Proceedings of the 2024 European Conference on Computing in Construction

Srećković, Marijana; Kassem, Mohamad; Soman, Ranjith; Chassiakos, Athanasios

DOI
10.35490/EC3.2024

Publication date
2024

Document Version
Final published version

Citation (APA)
Srećković, M., Kassem, M., Soman, R., & Chassiakos, A. (2024). *Proceedings of the 2024 European Conference on Computing in Construction*. (Proceedings of the European Conference on Computing in Construction; Vol. 2024). European Council on Computing in Construction (EC3). https://doi.org/10.35490/EC3.2024

Proceedings of the

2024 European Conference on Computing in Construction

Editors:

Dr. Marijana Srećković [1]
TU Wien (AT)

Prof. Mohamad Kassem [2]
Newcastle University (UK)

Dr. Ranjith Soman [3]
TU Delft (NL)

Dr. Athanasios Chassiakos [4]
University of Patras (GR)

[1] *Programme Chair*
[2] *Conference Chair*
[3] *Dissemination Chair*
[4] *Publications Chair*

14-17 JULY 2024
Chania, Crete, Greece

2024 EC³

Proceedings of 2024 European Conference on Computing in Construction

July 14–17, 2024
Chania, Crete, Greece

Edited by

Dr. Marijana Srećković, TU Wien (AT)

Prof. Mohamad Kassem, Newcastle University (UK)

Dr. Ranjith Soman, TU Delft (NL)

Dr. Athanasios Chassiakos, University of Patras (GR)

Published at Newcastle University, Newcastle, UK.

To cite this publication:

M. Srećković, M. Kassem, R. Soman and A. Chassiakos, (2024). *Proceedings of the 2024 Conference of European Council for Computing in Construction (EC3)*. July 14–17, 2024; Chania, Crete, Greece: European Council for Computing in Construction (Computing in Construction, 2684–1150). DOI: 10.35490/EC3.2024

ISBN: 9789083451305
ISSN: 2684-1150

PREFACE

The 2024 European Conference on Computing in Construction was held as a mixed Conference from July 14 – July 17, 2024. The conference had 170 attendees that presented work and exchanged ideas in the areas of the conference. This book contains the papers that were submitted to the conference and were accepted after a rigorous peer-review process.

The EC^3 conference will be a core activity of EC^3. The conference organisation is managed independently by a group of volunteers. The present conference is the 6th annual event of EC^3.

The EC^3 2024 proceedings include an illustrated review of the program, the names of organizations and persons who contributed to the technical program. The peer review process consisted of two phases. Firstly, we received 55 optional abstracts that were reviewed by the respective area chairs. Following a rigorous full paper peer review process (with each full paper being reviewed by at least two reviewers drawn from the scientific committee of international experts, and final decisions being made collectively by the corresponding track and programme chairs, for two rounds of review), 131 outstanding full papers were ultimately included in the proceedings and presentation at the conference. The manuscripts were presented during 2 plenary sessions and 22 technical sessions, including topics that focused on:

- Blockchain & Distributed Ledger Technology
- Data Analysis, Simulation, & Resilience
- Data Integration Methods
- Data Sensing & Acquisition
- Education, Policy, and Standardisation
- Energy Modelling & Monitoring
- Product and Process Modeling
- Virtual and Augmented Reality

Please note: All EC^3 proceedings and session recordings are available at no cost from https://ec-3.org/publications/conference/. *All conference papers now have a unique DOI and are comprehensively meta-tagged to ensure easy discovery by commonly used search and indexing services.*

Additionally, the conference included four sessions dedicated to the technical committee work of the EC^3 and these included:

- Data Sensing and Analysis (DSA) Committee
- Modelling and Standards Technical (M&S) Committee
- Education (EDU) Committee
- Human Digital Interaction (HDI) Technical Committee

One further track allowed participants to communicate embryonic or thesis-related work through sessions dedicated to:

- Thesis-related work in the form of a Thesis-in-3 competition for students.

A day-long practical workshop for junior researchers on how to become an effective author of highly-ranked, peer-reviewed scholarly journal papers has been held by Prof. Mirosław J. Skibniewski. It conveyed the practical experience of a 30-year-long veteran editor-in-chief of a top-ranking international research journal focused on construction IT and related technologies.

During a dedicated session, PhD/Junior Researchers-Mentor round tables were conducted, with the scope of developing mentees' writing strength, clear goals, and personal growth.

In addition to the technical content, the conference also provided opportunities for fellowship and networking in informal events.

We would like to thank the EC^3 scientific community, including both academic and industry members for their contributions and support; the scientific committee (see specific acknowledgements below); Wahib Saif, Marta Boscariol, Dominik Hartmann for their invaluable support during the conference and the preparatory stages, and Sobia Bano for maintaining the conference website.

To all EC^3 2024 attendees: we sincerely appreciate your participation and involvement in this conference. We hope this experience provided opportunities to renew friendships and professional relationships, forge new ones, spark exciting new research ideas, and enjoy the scenery and surroundings in a beautiful setting!

Dr. Marijana Srećković, TU Wien, EC^3 Programme Chair, and Prof. Mohamad Kassem, Newcastle University, EC^3 Conference Chair.

EC3 SPONSORS:

We would like to thank our sponsors and academic partners for their generous contributions to the congress.

Organizer:

European Council on Computing in Construction

Gold sponsor:

ESTECO

Silver sponsor:

Bentley®
Advancing Infrastructure

ACADEMIC PARTNERS:

EUROPEAN COUNCIL ON COMPUTING IN CONSTRUCTION

The **European Council on Computing in Construction (EC^3)** is a recently established society of construction professionals, academics, researchers and national Professional Bodies, aspiring to become the leading European forum in the area of information technology in construction engineering and management.

Role of EC^3

The **European Council on Computing in Construction (EC^3)** advances professional knowledge and improves engineering practice in the built environment by fostering research, education and policy in current and emerging computing and information technologies.

EC^3 is founded on the following four pillars and corresponding Technical Committees:

- DSA: Data Sensing & Analysis
- M&S: Modelling & Standards
- HDI: Human Data Interaction
- EDU: Education

EC^3 interacts strongly with other Architecture, Engineering, Construction and Facility Management (AEC/FM) societies in related areas, strengthen the collaborations between academia and industry in topics related to **EC^3**'s mission, spearhead research on such topics, identify and promote effective ways to advance the state of knowledge and the level of education and practice in these topics, assist in the making of policy, and support existing and new related specialty conferences and publications.

Technical Committees

The role of all technical committees is to:

- Gather, maintain and disseminate information on the application of the committee's area to AEC/FM
- Organise and support joint research activities in the committee's area
- Organise and support conference sessions, workshops, and meetings in the committee's area
- Disseminate innovation in the committee's area through position papers, white papers, grand challenges reports and policy work throughout the European spectrum
- Keep the SPAs aware of developments in the committee's area

Conference Committee:

Prof. Mohamad Kassem, Newcastle University (General Chair)
Dr. Marijana Srećković, TU Wien (Programme Chair)
Dr. Ranjith Soman, TU Delft (Dissemination Chair)
Dr. Athanasios Chassiakos, University of Patras (Publications Chair)

Acknowledgements

Conference host:

The European Council on Computing in Construction

The organisers would like to thank the Area Chairs and the Scientific Committee members for their contributions to the paper review and selection process. We would also like to thank the Technical Committee members for their invaluable contributions and hard work to deliver the other activities that took place in addition to the mainstream paper presentations.

Area Chairs:

Blockchain & Distributed Ledger Technology:

Dr. Jennifer Li, *Northumbria University, UK*

Data Analysis, Simulation, & Resilience:

Dr. Alessandro Carbonari, *Polytechnic University of Marche, IT*

Data Integration Methods:

Dr. Pieter Pauwels, *Eindhoven University of Technology, NL*

Data Sensing & Acquisition:

Dr. Thomas Linner, *OTH Regensburg, DE*

Education, Policy & Standardisation:

Dr. Yiannis Xenidis, *Aristotle University of Thessaloniki, GR*

Energy Modelling & Monitoring:

Dr. Dimitrios Rovas, *University College London, UK*

Product and Process Modelling:

Dr. Moretti Nicola, *University College London, UK*

Virtual and Augmented Reality:

Dr. Iro Armeni, *Stanford University, USA*

CONFERENCE SCIENTIFIC COMMITTEE

Dr.	Akbarieh, Arghavan	*Eindhoven University of Technology*
Ms.	Albrecht, Sophie	*OTH Regensburg*
Dr.	Ali, Usman	*University College Dublin*
Prof.	Amor, Robert	*University of Auckland*
Dr.	Ariyachandra, Mahendrini	*University of Cambridge*
Dr.	Armeni, Iro	*Stanford University*
Dr.	Bloch, Tanya	*Technion Israel Institute of Technology*
Dr.	Boje, Calin	*Luxembourg Institute of Science and Technology*
Prof.	Borrmann, André	*Technical University of Munich*
Dr.	Braun, Alex	*Technical University of Munich*
Mr.	Bruttini, Alessandro	*University of Florence*
Dr.	Cadena, Juan Diegno Blanco	*Politecnico di Milano*
Mr.	Cao, Jianpeng	*Delft University of Technology*
Dr.	Castronovo, Fadi	*University of Brighton*
Ms.	Chamari, Lasitha	*Eindhoven University of Technology*
Dr.	Chan, Yin-Chi	*University of Cambridge*
Dr.	Chen, Weiwei	*University of Cambridge*
Mr.	Chen, Zhe	*University of Hong Kong*
Mr.	Chung, Sehwan	*Seoul National University*
Dr.	Corneli, Alessandra	*Polytechnic University of Marche*
Dr.	Costin, Aaron	*University of Florida*
Ms.	De Andrade Pereira, Flavia	*University College Dublin*
Dr.	Demian, Peter	*Loughborough University*
Mr.	Donkers, Alex	*Eindhoven University of Technology*
Mr.	Doukari, Omar	*Northumbria University*
Ms.	Durmus, Dilan	*Università Politecnica delle Marche*
Prof.	Ergen Pehlevan, Esin	*Istanbul Technical University*
Dr.	Farghaly, Karim	*University College London*
Dr.	Fauth, Judith	*University of Cambridge*
Dr.	Fierro, Gabe	*Colorado School of Mines*
Dr.	Garcia de Soto, Borja	*New York University Abu Dhabi*

Mr.	Getuli, Vito	*University of Florence*
Prof.	Guldur Erkal, Burcu	*Hacettepe University*
Dr.	Hadjidemetriou, Georgios	*Loughborough University*
Mr.	Hagedorn, Philipp	*Ruhr-Universität Bochum*
Dr.	Hall, Daniel	*TU Delft*
Dr.	Hamdan, Al-Hakam	*A+S Consult GmbH*
Prof.	Hartmann, Timo	*TU Berlin*
Dr.	Hofmeyer, Hèrm	*Eindhoven University of Technology*
Dr.	Hu, Rongbo	*Technical University of Munich*
Dr.	Ilhan Jones, Bahriye	*Northumbria University*
Dr.	Iturralde, Kepa	*Technical University of Munich*
Ms.	Jankovic Rakic, Sanja	*OTH Regensburg*
Prof.	Jurelionis, Andrius	*Kaunas University of Technology*
Dr.	Kalemi Vakaj, Edlira	*Birmingham City University*
Ms.	Karamara, Merve	*OTH Regensburg*
Prof.	Kassem, Mohamad	*Newcastle University*
Dr.	Korolija, Ivan	*University College London*
Prof.	Lee, Ghang	*Yonsei University*
Dr.	Li, Jennifer	*Northumbria University*
Ms.	Li, Yitong	*George Mason University*
Dr.	Lilis, Georgios	*University College London*
Dr.	Linner, Thomas	*OTH Regensburg*
Ms.	Liu, Liu	*Ruhr-Universität Bochum*
Prof.	Maile, Tobias	*Augsburg University of Applied Sciences*
Dr.	Maltese, Sebastiano	*University of Applied Sciences and Arts of Southern Switzerland*
Dr.	Marincioni, Valentina	*University College London*
Dr.	Mastrolembo Ventura, Silvia	*University of Brescia*
Dr.	Matoseiro Dinis, Fábio	*University of Porto*
Mr.	Mavrokapnidis, Dimitris	*University College London*
Mr.	Mêda, Pedro	*Instituto para a Construção Sustentável*
Dr.	Merino, Jorge	*University of Cambridge*

Dr.	Meschini, Silvia	*University of Turin*
Mr.	Messi, Leonardo	*Polytechnic University of Marche*
Dr.	Mirarchi, Claudio	*Politecnico di Milano*
Dr.	Moretti, Nicola	*University College London*
Mr.	Msawil, Mahir	*Northumbria University*
Dr.	Mukherjee, Anandarup	*University of Cambridge*
Dr.	Murguia, Danny	*University of Cambridge*
Prof.	Ng, Ming Shan	*Kyoto Institute of Technology*
Dr.	O'Donnell, James Thomas	*University College Dublin*
Dr.	Pauwels, Pieter	*Eindhoven University of Technology*
Dr.	Petrova, Ekaterina	*Eindhoven University of Technology*
Prof.	Ploennigs, Joern	*University of Rostock*
Prof.	Poças Martins, João	*University of Porto*
Dr.	Poveda-Villalon, Maria	*Universidad Politécnica de MaDr.id*
Prof.	Puust, Raido	*Tallinn University of Technology*
Dr.	Rezaei Rad, Aryan	*École Polytechnique Fédérale de Lausanne*
Prof.	Rinaldi, Stefano	*University of Brescia*
Dr.	Rovas, Dimitrios	*University College London*
Prof.	Sacks, Rafael	*Technion Research & Development Foundation LTD.*
Mr.	Sheikhkhoshkar, Moslem	*University of Lorraine*
Dr.	Simeone, Davide	*Webuild Group*
Prof.	Singh, Vishal	*Indian Institute of Science*
Prof.	Smallwood, John Julian	*Nelson Mandela University*
Prof.	Smarsly, Kay	*Hamburg University of Technology*
Dr.	Soman, Ranjith	*TU Delft*
Dr.	Srećković, Marijana	*TU Wien*
Prof.	Tagliabue, Lavinia Chiara	*University of Turin*
Dr.	Tang, Rui	*University College London*
Mr.	Tzanakakis, Iason	*Sweco AB*
Dr.	Wang, Tong	*TU Delft*
Prof.	Weininger, Florian	*OTH Regensburg*
Dr.	Wen, Ya	*University of Cambridge*

Prof.	Xenidis, Yiannis	*Aristotle University of Thessaloniki*
Dr.	Xie, Xiang	*Newcastle University*
Dr.	Yilmaz, Gokcen	*University of Cambridge*
Prof.	Yitmen, Ibrahim	*Jönköping University*
Dr.	Zhu, Aiyu	*Southeast University*

TABLE OF CONTENTS

Blockchain and Distributed Ledger Technology

BLOCKCHAIN-BASED COMMON DATA ENVIRONMENTS TO ADDRESS DATA AND INFORMATION FRAGMENTATION IN THE DUTCH CONSTRUCTION INDUSTRY

Rowan Kloosterman[1], Martijn Smits[1], Wasim Haji[1], Jennifer Li[2]
[1]Amsterdam University of Applied Sciences, Amsterdam, The Netherlands
[2]Northumbria University, Amsterdam, The Netherlands

Abstract

In the Dutch construction industry, the demand for advanced information storage and sharing is growing due to the complexity of construction projects. Limitations of traditional methods include lack of transparency and inefficient communication. Blockchain offers a promising solution by enabling decentralized storage and immutable recording of data increasing transparency and efficiency in the construction supply chain. Combining a Common Data Environment with the InterPlanetary File System – decentralized file storage and exchange, and a powerful tool for secure, efficient and reliable data management in construction – can emerge to improve cooperation between parties increasing effectiveness of projects.

Introduction

The construction industry is not known for being highly innovative. In general, construction companies struggle to keep up with the speed of technological developments and innovations (van Sante, 2016). A major obstacle to innovation and change in construction is the high degree of project-driven working and fragmentation (Nawi et al., 2014; Li, 2023). The different construction phases of a structure (vertical fragmentation), the specialized stakeholders (horizontal fragmentation) and the different, often complex projects create many challenges within the industry (Fergusson, 1993; Adriaanse, 2014).

While the application of Building Information Modelling (BIM) has become increasingly accepted in construction, making optimal use of it appears difficult (Canon, 2019). According to Koutsogiannis and Berntsen (2019), 95% of the data generated is lost between the parties involved and the different construction process phases. Construction information is rarely reliable and available, which is largely due to the limited connection between parties as a result of fragmentation (Adriaanse, 2014; Volker, 2019).

In the Netherlands, strong fragmentation and specialization perpetuate, with project-based approaches being the norm (Adriaanse et al., 2020). These characteristics present significant challenges, especially considering the growing amount of information generated throughout the life cycle of projects (Koutsogiannis and Berntsen, 2019). In practice, there often appear to be problems related to the availability and reliability of construction information, which has significant implications for the efficiency and effectiveness of construction projects (Chong et al., 2014). The transfer and preservation of construction information is a crucial part of the construction process and will become even more important in the future. Making and keeping information available and reliable is something that is a hot topic in many sectors. A development which is often associated with transparent and reliable exchange of data is Blockchain Technology (Puthal *et al.*, 2018).

The overarching objective of this study is to assess the viability of integrating blockchain with Common Data Environments (CDEs) to enhance the accessibility and dependability of building information within the Dutch construction sector. This focus is crucial as building information plays a pivotal role in streamlining construction processes across the life cycle of buildings in the Netherlands, ultimately fostering opportunities for long-term material reuse initiatives (Jaskula, et al., 2023).

The research question for this study is: "*How can the integration of blockchain and CDEs be leveraged to enhance the availability and reliability of building information in the Netherlands?*"

This paper reports on preliminary research focused on the exploration of information management within the Dutch construction sector. By examining local practices and challenges, this study allows researchers to contextualize and compare the findings in the Dutch context with research from other countries regarding information management in construction. This adds insights to the body of knowledge and facilitates the Netherlands in learning from other contexts, thus contributing to a broader understanding of information management practices in the construction sector.

In next section, the literature review introduces blockchain technology, and identifies challenges to current information management and CDEs in construction. The methodology explains the approach for the study centred on interviews and a use case evaluation, the results of which are then presented. The final two chapters bring discussion and conclusions of the study.

Literature Review

Blockchain technology

Many current database systems operate within a centralized architecture, where users have permission to modify data stored on the central server. A central authority manages the entire database, including access control and user authentication (Sarmah, 2018). Blockchain, often associated with cryptocurrencies such as Bitcoin, offers a decentralized approach. It is a type of Distributed Ledger Technology (DLT) that has the potential to spark a new industrial revolution (Perez, 2009). Unlike centralized systems, blockchain is managed and validated by a global network of peer-to-peer connected computers where no one individual or organisation has control over the data (Jena and Dash, 2021). This creates a digital database that is built, shared, validated and synchronized by participants eliminating the need for a central authority where everyone has access to the latest version of the ledger (Lashkari and Musilek, 2021). Transactions are sent to the network, validated by computer algorithms, and linked to previous transactions, creating an immutable chain. The distributed structure of blockchain and the confirmed guarantees of nodes make information almost impossible to manipulate, which increases trust between participants. This provides a solution to the limitations of traditional centralized systems (Bodkhe *et al.*, 2020).

Blockchains can be public or private. Public blockchains are open systems without authority; anyone can join and be granted full rights to participate (Guégan, 2017). The verifiability and transparency of information are central features as there are no access restrictions imposed. Public blockchains are not recommended for large or energy-sensitive domains. Private blockchains have limited access to the network and are shared privately between specific participants with communication protocols. Only pre-verified individuals meeting certain requirements are allowed to perform certain operations on the blockchain (Yang et al., 2020). Between public and private sits the consortium blockchain where only organizations or groups with the same goals can join (Jabbar et al., 2022). The degree of decentralization varies in DLT networks from full decentralization (e.g., Bitcoin) to lower decentralization with some form of hierarchy for direction and control (e.g., Hyperledger), depending on the specific application and parties involved (Hamilton, 2019).

Information management

Fragmentation in information management

Due to the dynamic and complex nature of construction, different professions and specializations have emerged, roughly divided into the well-known construction phases: initiators, designers, builders, maintenance parties (Nawi et al., 2014). In this subdivision, a clear structure is visible with strong separation between the different construction phases (Adriaanse, 2014). Within these phases, there is a high degree of specialization leading to many companies becoming involved in a construction project all having their own contributions and interests. In many cases, this leads to difficulties in communication and coordinating all the parties in the construction process (Chong et al., 2014; Di Giuda et al., 2020). Despite this high degree of specialization often presenting challenges, construction has proven to be a key piece in the overall development of society. Without the necessary infrastructure, many developments would have been much more difficult and society would be less developed (van Breugel, 2019). The construction industry exhibits considerable fragmentation as an inherent feature of its structure and mode of operation. It is often compared to an ‘archipelago’, consisting of several ‘islands’ that are highly distributed across three levels (Adriaanse, 2014) – vertical, horizontal and longitudinal. This complex structure leads to communication challenges between stakeholders, construction process phases and projects themselves (Dave and Koskela, 2009). This division results in significant challenges in communication and coordination among stakeholders (Fergusson, 1993).

CDEs for information management

The increasing complexity of construction projects, coupled with the need for real-time information sharing, underscores the necessity for digitalization in the construction industry (Agarwal et al., 2016). A CDE serves as a comprehensive data store for all project-related information, including geometric and semantic data, as well as documentation, throughout the construction project life cycle (British Standards Institution, 2013). By consolidating all information into one accessible location, the CDE acts as a central information management tool, facilitating efficient collaboration and communication among stakeholders (Preidel *et al.*, 2017). We consider whether the CDE can play a central role in integrating and managing information among stakeholders.

Efficient information management is paramount for project success due to the complexity and abundance of data generated during construction projects (Di Giuda et al., 2020). Defined by ISO 19650-1:2018, a CDE functions as an agreed-upon information resource for a project or asset, enabling the collection, management, and dissemination of information through a controlled process (ISO, 2018). It encompasses both a CDE solution, typically cloud- or server-based technology with database management features, and a structured CDE workflow that organizes information flow and management throughout the asset life cycle (BIM Dictionary, 2022).

Jaskula et al. (2022) outlines three levels of CDE maturity based on document management, life cycle functionality, security, and BIM integration. At its core, the CDE enhances information management by providing stakeholders with a unified platform for accessing, sharing, and collaborating on project data. This centralized approach fosters transparency, reduces errors, and improves decision-making, ultimately contributing to the overall success of construction projects. Cloud storage services such as Dropbox are categorized as Level 1 CDE due to limited integration with BIM technology and do not offer advanced management capabilities nor a high level of security. More advanced CDE tools such as Viewpoint, Asite, Procore, Deltek, or ProjectWise are considered Level 2 CDE for BIM because they offer BIM visualization and communication based on BIM formats such as BIM Collaboration Format (BCF) and Industry Foundation Classes (IFC) in addition to document management. Additionally, integrating technology such as the InterPlanetary File System (IPFS) could enhance the capabilities of CDE tools, potentially elevating them to Level 3 BIM integration. IPFS, with its decentralized and distributed file system, offers advantages such as increased data integrity, redundancy, and resilience against censorship. Platforms like Autodesk's BIM 360 or BIMcollab, integrated with IPFS, could offer Level 3 functionality for multidisciplinary real-time collaboration and document management, serving as a more comprehensive single source of truth throughout the life cycle of construction projects. However, this integration may still require further development and evaluation to ensure seamless compatibility and functionality within existing CDE frameworks (Jaskula et al., 2023).

IPFS as a solution to information management

While blockchain can store information, it is not a realistic solution for storing large amounts of building information (Kloosterman and Smits, 2023) for being prohibitively expensive and demands significant computational power. Therefore, current iterations of blockchain are not suitable for the collaboration and sharing of large amounts of data. This challenge can be addressed by integrating blockchain with IPFS (Steichen *et al.*, 2018), which allows participants to directly connect to each other without the need for a central server (Muralidharan and Heedong, 2019; Bennet, 2020). A major advantage is that IPFS allows flexible storage of large files. Each file uploaded to the IPFS network is encrypted into a cryptographic hash value referred to as a content identifier (CID). IPFS is considered a valuable addition because it can solve the problem of inefficient, bulky data storage in blockchain (Steichen *et al.*, 2018; Nyaletey *et al.*, 2019). In other words, users can choose to store design files or documents in IPFS and place only CIDs in blockchain transactions.

Implementation of the blockchain-IPFS CDE solution faces two challenges: (1) establishing the collaboration workflow of a CDE that integrates blockchain and IPFS, and (2) exchanging design information in such an environment. Regarding (1), limited studies have examined which CDE container(s) should use blockchain and the logic of collaborating in such a distributed environment (Ye and König, 2021). Regarding (2), in a blockchain network, users exchange information by proposing transactions and using smart contracts. But the data model of transactions and smart contracts that meets the requirements of CDE is not yet developed, which complicates communication between designers (Tao *et al.*, 2021). It is expected that these challenges will be addressed in due course within the research community; proposing solutions is outside the scope of this paper.

Methodology

To answer the research question, this study adopted an interpretivist philosophy and an inductive approach to data collection and analysis. The methodology can be seen graphically in Figure 1 below.

First, a literature review was conducted to understand the issues surrounding fragmentation in the Dutch construction industry, the nature of construction information and the opportunities offered by Blockchain.

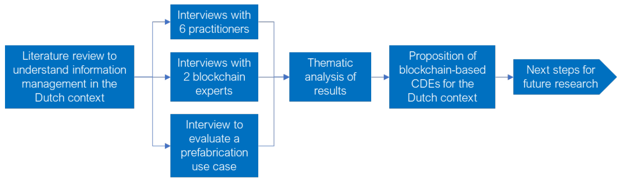


Figure 1: Methodology for the study

The findings of the literature review formed the basis of interviews with six industry practitioners (#1 to #6 in Table 1) and two blockchain experts (#7 and #8 in Table 1) contrasted with evaluation of a use case with a BIM Coordinator of a prefabrication project (#9 in Table 1) in the Netherlands. Finally, a validation interview took place with an academic with extensive knowledge of blockchain in construction (#10 in Table 1). The interviews with industry practitioners centred on investigating opportunities for enhancing communication and optimizing information utilization. Both positive and negative feedback were documented regarding various aspects including the availability and reliability of information, communication exchanges, and data storage. These insights are crucial for identifying areas for improvement and pinpointing specific pain points within current practices.

Table 1 presents the profiles of the participants, who were chosen based on their roles to ensure representation across different phases of construction projects. Conducted in Dutch, semi-structured interviews allowed for the organic evolution of topics and the emergence of new insights, with subsequent translation into English for analysis.

Table 1: Profile of Participants

ID	Function	Type of business	Subject
#1	Deputy director	Developer	
#2	Project leader	Architect	
#3	Managing director	BIM model elaboration	Data exchange and life cycle optimisation
#4	Project leader	Contractor	
#5	Project leader	Housing corporation	
#6	Transfer team leader	Housing corporation	
#7	Blockchain expert	Software developer	
#8	Blockchain expert	Knowledge organization	Blockchain
#9	BIM Coordinator	Prefabrication	Data generation and transfer
#10	Lecturer	University	Validation

Subsequent to the interviews, a use case was evaluated to examine the impact of using a single central source of truth on the availability and reliability of construction information with specific relevance to the Dutch context. The interviews with the blockchain experts also addressed these topics, but with a specific focus on the potential role of blockchain technology in transforming data exchange. The central question here was whether blockchain can add value to an alternative approach to data communication, addressing aspects such as information availability, reliability and storage in an innovative way. The use case was used to gain insights into whether working within a single network – such as a blockchain – has benefits for information exchange and availability. In fact, different departments within the company use the same information sharing system.

Upon collection of the data, thematic analysis was applied following Williams and Moser's (2019) three-step coding process. First open codes were identified, then refined into selective codes and finally themes were identified across the data. Table 2 shows the resultant themes along with analysis of positive or negative comments being made across the three sets of interviews.

Table 2: Evaluation of comments from the interviews

Themes	Current		Blockchain		Use Case	
	-ve	+ve	-ve	+ve	-ve	+ve
Availability	58	56	5	18	2	23
Reliability-Trust	60	67	6	44	3	25
Exchange-Communication	83	95	2	12	4	11
Storage	29	19	9	16	0	9
Construction information	133	130	23	68	1	19
Collaboration	51	76	5	16	4	25
BIM	58	87	2	16	1	6

Results of the interviews

The interviews were designed to explore three main aspects: the current situation, the potential added value of blockchain and a use case. Three sets of interviews were conducted with respondents from the construction industry, blockchain experts and as part of a case study.

The results, analysed thematically and summarized in Table 2, show the current situation around these topics is fairly even, with roughly equal numbers of positive and negative comments. However, when asked whether blockchain can add value, more positive statements emerged. This suggests that blockchain could be beneficial for these aspects. Furthermore, it was found that working within one network and with one source of truth produced more positive as well as negative expressions on the various topics. This indicates the potential of a blockchain-based structure to improve collaboration and information sharing.

Current situation with information exchange

The results of examining the current situation with information exchange show there is a clear need for available and reliable construction information. One crucial aspect that emerged from the interviews is the location where information is stored. Several stakeholders emphasized that construction work information is often stored on their own servers or systems, resulting in scattered and fragmented data storage. The interviews also revealed that just a shared storage location is not enough. Participant #2 said, "*That is the place where everyone leaves their information. In several projects, we have found this often acts as a kind of dumping ground for information. While anyone can put information there, there is often a lack of adequate version control and determination of relevance. Managing older data or directing that information is often neglected*". This makes it difficult to quickly access the information needed and forces parties to exchange information through external channels as indicated by participant #6: "*I get a very large information portion from the contractor through a USB stick*" (#6).

An adverse consequence of this information sharing is the challenge of identifying the most recent version of documents (e.g., drawings). "*The more information you have to share, the more mistakes are made*" (#2). This creates the risk of working with outdated or incorrect information, which can damage the reliability of construction work information. "*You will always have the customer or contractor working with a different system*" (#9). To address these issues, some stakeholders are already experimenting with software solutions that bring together construction work information in one central location. These CDEs are welcomed by respondents who recognize the benefits of centralized access to and exchange of information: "*You really need to have a place where the information is centralized, because it disappears everywhere. Everyone has different information and then what's the right version? You just lose that overview*" (#3). They believe such CDEs will improve the availability and reliability of construction information in the future: "*Of course, it would be much nicer if something from the contractor went directly into our system*" (#6).

These statements represent that the exchange and storage of construction information in construction could be significantly improved. Often, information is distributed to stakeholders in different ways, leading to ambiguity about the correct version. This leads to mutual consultation among stakeholders without everyone being aware, often resulting in parallel efforts and different versions of the 'truth'. According to respondents, the demand for a central source of information can yield significant gains in information management. Participant #6 indicated that a lot information is lost during construction and after handover. As a result, valuable data are missing during operation. Participants also indicated that there is no seamless connection between business information, with participant #6 indicating that it would be more efficient if the contractor could add information directly to the server. When asked whether applying blockchain in the Dutch construction industry can help improve information management and collaboration, there were mixed responses. Mainly, the lack of knowledge about blockchain and overall market adoption are mentioned as obstacles. Proper education about the possibilities can promote broader implementation. Participant #9 indicated that they are open to new technologies if they are beneficial for overall construction and information management in the different phases.

Added value through blockchain technology

The interviews with blockchain experts aimed to explore how blockchain can contribute to achieving more reliable and accessible information. A key focus was to identify the most appropriate form of blockchain for collaboration between different stakeholders. This involved asking about different types of blockchains (e.g., public, private, consortium), and examining the features most conducive to effective collaboration. In addition, the research assessed the impact of using blockchain in businesses by examining the potential benefits of blockchain adoption and identifying the challenges and adjustments needed when transitioning to a blockchain-based business model. Finally, they explored why blockchain, despite several years of existence, has not yet been fully utilized. Through these interviews, it was hoped the potential obstacles that may hinder implementation of blockchain in the Dutch construction industry could be better understood.

The results indicate that while storing large volumes of information on a blockchain is possible, it is also be prohibitively expensive and demands significant computational power. Therefore, it may be more practical to consider recording only the transactions associated with that information on the blockchain. *"The bitcoin-blockchain is extremely slow and poor to use for storing information. But it is very good at transferring values because it is so secure"* (#8). This approach can enhance the reliability of information without incurring excessive costs or energy consumption. Additionally, it is worth noting that the information is hashed, and the resulting hash code is processed through the blockchain, ensuring data integrity and security. Blockchain expert #8 expressed that "*you have to see the blockchain as the party that fixes the information that parties exchange with each other. A kind of notary*". This makes the implementation of a CDE a considerable necessity, as it promotes the availability of information for all parties involved. It ensures transparency and accessibility of information, which enhances collaboration and information sharing. The findings also highlight that blockchain is an additional layer within the ICT infrastructure because "*you actually have to really separate it from the database, it's just a different kind of technology*" (#8). It provides a secure and decentralized infrastructure for storing and verifying data, which promotes trust and transparency.

For collaborations involving stakeholders with hierarchical structures, private (consortium) blockchains are recommended by participant #8. This allows participants to work together within a shared network where specific roles and authority can be assigned as well as structured collaboration and decision-making giving parties "*the ability to share data among themselves. Moreover, it is configurable, so you can configure which parties specifically share data with each other. You have full control over this configuration*" (#7).

Based on these interviews, it can be said that the demand for structured and reliable information management is strong in the industry. The large-scale implementation of CDEs is lagging partly because of the lack of structure and control. This is where blockchain technology can add value. By capturing construction information using blockchain, it is possible to keep track of changes and other important issues. The ultimate goal with this is to roll out the large-scale application of central data management in construction.

Use Case: Prefabricated wall panels

To test an 'ideal situation' and focus on information exchange between cooperating departments within a single organization, a use case was evaluated. A factory specializing in the production of precast elements was chosen. As a modern method of construction, prefabrication has a more integrated supply chain than that of traditional construction sites (Ocheoha and Moselhi, 2018) and is, therefore, potentially more susceptible to adoption of blockchain in the near-term (Olawumi *et al.*, 2022).

The aim of the use case evaluation was to gain insight into how information is exchanged when internal stakeholders work together within the same network. This study focused specifically on the methods and processes used to share and communicate information among the departments involved and whether there is a 'Single Source of Truth' for all stakeholders to reliably access. The evaluation produced several results that highlight the benefits of working with a single source of truth. It was found to improve the reliability of information because all necessary data are stored in one central location, eliminating incorrect or contradictory information. This contributes to a higher degree of reliability and accuracy in the production process such that "*IFC is the digital source model of the prototype we make and with that we control the production*" (#9). In addition, working with a single source of truth adds value to the optimization and automation process. It streamlines information exchange and creates an efficient workflow. "*We have software that does our planning and also ultimately controls production. And through that system we have files for the lasers, for the plotter; files are generated*" (#9). By working with a centralized resource, processes can be automated and more emphasis can be placed on optimization and improvement of production. An additional advantage is that all employees and departments have access to the same data "*so we are working with the same source information*" (#9). Having the source information stored in one place with easy accessibility promotes communication and collaboration within the organization. Everyone can have access to the most recent and relevant data at any time ensuring better coordination and handling of tasks and processes. Information is no longer distributed through emails between internal stakeholders, but everyone knows where to find the information. This prevents duplication of information and ensures that the 'production model' is and remains the only source of truth. The downside of intensive use of the source model is that any changes made during the process must be updated in the model to avoid production errors, as many processes are based on the source information and automation. "*We aim to integrate as much information as possible into the production model that drives our production process. We want this to be the source of information*" (#9).

During the interview, it was noticed that company employees showed remarkably more mutual engagement than is common in the construction industry. This result can be attributed to the fact that they work together within one network or all work for the same company, which fosters a sense of belonging. As a result, all employees have a common interest in the final product, which differs from the usual fragmented nature of the construction industry in which individuals perform their own tasks.

Evaluation of this use case has shown that working with a single source of truth offers significant benefits, including improved reliability of information, availability of information, optimization of processes and a sense of belonging among employees and departments as they advance toward a common goal.

Discussion

This paper aimed to evaluate the use of blockchain in combination with CDEs to improve the accessibility and reliability of construction information in the Netherlands. The Dutch construction industry is known for its fragmentation and specialization, which poses challenges for efficient information exchange. Integrating blockchain with CDEs could help solve problems of information management essential for optimizing construction processes. The research question being answered was, "*How can the integration of blockchain and CDEs be used to improve the availability and reliability of construction information in the Netherlands?*"

The literature review highlighted the benefits of blockchain and CDEs, particularly in terms of reliability, accessibility and efficiency of information management. Blockchain is recognized for its ability to secure information and provide transparency without the need for trusted intermediaries (Lashkari and Musilek, 2021). The concept of CDEs is considered an effective way to centralize construction information and facilitate collaboration among stakeholders (Preidel *et al.*, 2017). The interviews showed current practices in construction often result in fragmented and piecemeal storage of information, leading to problems such as finding the most recent versions of documents and miscommunication between different parties. This highlights the need for centralized solutions such as CDEs to address these challenges.

The results show combining blockchain and CDEs offers a potential solution to the limitations of traditional centralized systems in the construction industry by making manipulation of information virtually impossible increasing trust between participants. This is valuable due to the complex and specialized nature of the construction industry, where different professional groups are involved in different phases of the construction process (Nawi et al., 2014). A centralized information management system can help consolidate information and facilitate collaboration and communication between stakeholders. However, it was found that storing information on a blockchain is not effective due to its latency and poor storage abilities. The combination of blockchain and a CDE was considered a potential solution to this with IPFS suggested to complement blockchain to help solve problems with inefficient data storage (Steichen *et al.*, 2018; Nyaletey *et al.*, 2019). This was supported by the findings from the interviews, though practical considerations are still to be addressed, such as the perceived high cost and energy requirements of storing large amounts of information on the blockchain requiring a targeted approach. A solution being, for example, capturing only transactions involving information on the blockchain with the actual information stored on IPFS.

The promising applications of blockchain in construction, particularly in improving efficiency, transparency and time savings, are a key finding of this research. Optimizing CDEs by further exploring and developing these areas for the Netherlands, innovative solutions can be developed that lead to more efficient, transparent and sustainable construction processes and projects. Moreover, it is important to research similar situations in other countries. This helps not only to understand the applicability and effectiveness of blockchain in different contexts, but also to identify global best practices. By making international comparisons and drawing lessons from different experiences, we can gain a deeper understanding of the potential impact of blockchain on the construction industry worldwide. This can ultimately contribute to a more informed and inclusive approach to the implementation of blockchain technology in the construction industry.

Conclusion

The Dutch construction industry, characterized by fragmentation, experiences challenges in communication and coordination of construction work information, with loss of valuable data. This research answered the question, "*How can the integration of blockchain and CDEs be leveraged to enhance the availability and reliability of building information in the Netherlands?*" Qualitative methods were used to examine construction information in the chain, as well as opportunities and limitations of blockchain.

The research findings emphasized the importance and issues surrounding availability and reliability of information in the construction supply chain, often associated with negative experiences. The lack of connection between companies, operating mainly on their own 'islands', leads to manual dissemination of information across stakeholders. This results in conflicting versions, miscommunication and lack of a clear source of truth.

Research on blockchain integrated with a CDE has provided insights on its potential and limitations. Although further research is required to explicitly answer the research question, it is suggested that integration of a CDE with blockchain could be a valuable approach. Blockchain is not suggested for storing large amounts of information, rather integration with IPFS for a CDE that acts as a central repository for all construction information, could see increased security and prevent manipulation of data. This enables reliable and immutable storage of information, which is crucial for the efficient and effective execution of construction projects.

The evaluated use case showed that using a shared network environment and working with a central source of truth brings several benefits in terms of optimization, efficiency and trust. This is because information is stored, managed and shared more efficiently within the network. The availability and reliability of construction information increases greatly when using a single source of truth. Because stakeholders know where information is stored, individual sharing is no longer necessary, which benefits information reliability and reduces the risk of miscommunication and duplicate versions.

One notable result is that stakeholders have more mutual trust and cooperate better when operating within the same network environment. This improvement comes from relying on each other's information within the same system, allowing individual goals to converge into a common interest. This highlights the importance of a shared platform and the potential of blockchain to enhance communication, trust and collaboration in the supply chain. In short, the study concludes that construction stakeholders benefit from available and reliable construction information, with blockchain serving as a valuable tool to achieve these goals.

As preliminary research, this study was limited regarding the number of stakeholders engaged and the use case evaluated. While the study gained insights from relevant perspectives, it is acknowledged that more engagement is needed. A wider range of stakeholders could provide a broader spectrum of insights and allow for a more comprehensive representation of diverse viewpoints in the Dutch construction industry. Extending the study to engage with a wider range of stakeholders across the project life cycle to gain a deeper understanding of the complex dynamics and challenges within the Dutch construction industry. The use case chosen considered information management at an intraorganizational level that arguably faces less challenges than at an interorganizational level, the latter will be included in the extended study. In addition, blockchain's technological elaboration and application alongside a central database will be explored.

References

Adriaanse, A. (2014) *Bruggen bouwen met ICT*. Universiteit Twente. Available at: https://ris.utwente.nl/ws/portalfiles/portal/5119421/oratieboekje-Adriaanse.pdf.

Adriaanse, A., Borsboom, W. and Roef, R. (2020) *Naar netwerken van predictive twins van de gebouwde omgeving*. Available at: https://www.tno.nl/nl/newsroom/2020/11/predictive-twins-oplossing-uitdagingen/.

Agarwal, R., Chandrasekaran, S. and Srigrid, M. (2016) *Imagining construction's digital future*. Available at: https://www.mckinsey.com/industries/capital-projects-and-infrastructure/our-insights/imagining-constructions-digital-future2/28.

Bennet, J. (2020) *InterPlanetary File System (IPFS)*. Available at: https://ipfs.tech/.

BIM Dictionary (2022) *Common Data Environment (CDE)*. Available at: https://bimdictionary.com/en/common-data-environment/2.

Bodkhe, U. *et al.* (2020) *Blockchain for Industry 4.0: A comprehensive review*, *IEEE Access*. Available at: https://doi.org/10.1109/ACCESS.2020.2988579.

van Breugel, K. (2019) *Het model: vehikel voor glorie en schande*. Available at: https://pure.tudelft.nl/ws/portalfiles/portal/96293783/Afscheidsrede_Prof._K._van_Breugel_27.09.2019_.pdf.

British Standards Institution. (2013) *PAS 1192-2:2013 : specification for information management for the capital/delivery phase of construction projects using building information modelling.* Available at: https://www.hfms.org.hu/joomla/images/stories/PAS/PAS1192-2-BIM.pdf.

Canon (2019) *De nationale benchmark Digitalisering in de Bouw*. Available at: https://www.canon.nl/business/digitalisering-in-de-bouw/.

Chong, H.Y., Wong, J.S. and Wang, X. (2014) 'An explanatory case study on cloud computing applications in the built environment', *Automation in Construction* [Preprint]. Available at: https://doi.org/10.1016/j.autcon.2014.04.010.

Dave, B. and Koskela, L. (2009) 'Collaborative knowledge management—A construction case study', *Automation in Construction*, 18(7), pp. 894–902. Available at: https://doi.org/10.1016/J.AUTCON.2009.03.015.

Fergusson, K.J. (1993) *Impact of integration on industrial facility quality. Unpublished Dissertation*. Palo Alto. Available at: https://stacks.stanford.edu/file/druid:xj721fn3242/TR084.pdf.

Di Giuda, G.M., Giana, P.E. and Pattini, G. (2020) 'The shortening and the automation of payments: The potentiality of smart contract in the aeco sector', in *Proceedings of International Structural Engineering and Construction*. ISEC Press, p. CON-12-1-CON-12-6. Available at: https://doi.org/10.14455/ISEC.2020.7(2).CON-12.

Guégan, D. (2017) *Public Blockchain versus Private blockchain*. Paris. Available at: http://centredeconomiesorbonne.univ-paris1.fr/.

Hamilton, M. (2019) *Blockchain distributed ledger technology: An introduction and focus on smart contracts*. University of Alabama. Available at: https://doi.org/10.1002/jcaf.22421.

ISO (2018) *ISO, Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information* *modelling*, *Part 1: Concepts and principles*. Available at: https://www.iso.org/standard/68078.html.

Jabbar, R. *et al.* (2022) *Blockchain Technology for Intelligent Transportation Systems: A Systematic Literature Review*, *IEEE Access*. Available at: https://doi.org/10.1109/ACCESS.2022.3149958.

Jaskula, K. *et al.* (2022) *Common Data Environments in construction: State-of-the-art and challenges for practical*. The Bartlett School of Sustainable Construction. Available at: https://dx.doi.org/10.2139/ssrn.4249458.

Jaskula, K., Papadonikolaki, E. and Rovas, D. (2023) *Comparison of current Common Data Environment tools in the construction industry*. University College London. Available at: https://ec-3.org/publications/conference/paper/?id=EC32023_315.

Jena, A.K. and Dash, S.P. (2021) 'Blockchain Technology: Introduction, Applications, Challenges', in *Blockchain Technology: Applications and Challenges*, pp. 1–11. Available at: http://www.springer.com/series/8578.

Kloosterman, R. and Smits, M. (2023) *Blockchain Technologie in de bouw*. Amsterdam University of Applied Sciences.

Koutsogiannis, A. and Berntsen, N. (2019) *Blockchain and construction: the how, why and when*. Available at: https://www.bimplus.co.uk/blockchain-and-construction-how-why-and-when/.

Lashkari, B. and Musilek, P. (2021) *A Comprehensive Review of Blockchain Consensus Mechanisms*, *IEEE Access*. Available at: https://doi.org/10.1109/ACCESS.2021.3065880.

Li, J.J. (2023) *A socio-technical framework to guide implementation and value realisation of distributed ledger technologies (DLT) in the construction sector*. Available at: https://nrl.northumbria.ac.uk/id/eprint/51601/.

Muralidharan, S. and Heedong, K. (2019) *An InterPlanetary File System (IPFS) based IoT framework*. Available at: https://doi.org/10.1109/ICCE.2019.8662002.

Nawi, M.N.M., Baluch, N. and Bahauddin, A.Y. (2014) *Impact of Fragmentation Issue in Construction Industry: An Overview*. School of Technology Management and Logistics, University Utara Malaysia. Available at: https://doi.org/https://doi.org/10.1051/matecconf/20141501009.

Nyaletey, E. *et al.* (2019) *BlockIPFS - Blockchain-enabled interplanetary file system for forensic and trusted data traceability*, *Proceedings - 2019 2nd IEEE International Conference on Blockchain, Blockchain 2019*. Available at: https://doi.org/10.1109/Blockchain.2019.00012.

Olawumi, T.O. *et al.* (2022) *Automating the modular construction process: A review of digital technologies and future directions with blockchain technology*, *Journal of Building Engineering*. Available at: https://doi.org/10.1016/J.JOBE.2021.103720.

Perez, C. (2009) *Technological revolutions and techno-economic paradigms*. Technological University of Tallinn. Available at: www.carlotaperez.org.

Preidel, C. *et al.* (2017) *Seamless integration of common data environment access into BIM authoring applications the BIM integration framework*, *eWork and eBusiness in Architecture, Enginering and Construction*. Available at: https://mediatum.ub.tum.de/doc/1306961/654ocv4bit32uh73pjj1z1vjs.pdf.

Puthal, D. *et al.* (2018) *Everything You Wanted to Know about the Blockchain: Its Promise, Components, Processes, and Problems*, *IEEE Consumer Electronics Magazine*. Available at: https://doi.org/10.1109/MCE.2018.2816299.

van Sante, M. (2016) *Technologie in de bouw*. Amsterdam. Available at: https://assets.ing.com/m/b9c7308fea908f98/original/Technologie-in-de-bouw.pdf.

Sarmah, S.S. (2018) *Understanding Blockchain Technology*, *Computer Science and Engineering*. Available at: https://doi.org/10.5923/j.computer.20180802.02.

Steichen, M. *et al.* (2018) *Blockchain-Based, Decentralized Access Control for IPFS*. Available at: https://doi.org/https://doi.org/10.1109/Cybermatics_2018.2018.00253.

Tao, X. *et al.* (2021) *Distributed common data environment using blockchain and Interplanetary File System for secure BIM-based collaborative design*, *Automation in Construction*. Available at: https://doi.org/10.1016/j.autcon.2021.103851.

Volker, L. (2019) *Just a little of that human touch: Towards a value-based ecosystem for delivering infrastructure services*. University of Twente . Available at: https://ris.utwente.nl/ws/portalfiles/portal/201586728/Inaugural_lecture_L_Volker.pdf.

Williams, M. and Moser, T. (2019) *The Art of Coding and Thematic Exploration in Qualitative Research*, *International Management Review*. 15(1), pp. 45-55.

Yang, R. *et al.* (2020) *Public and private blockchain in construction business process and information integration*, *Automation in Construction*. Available at: https://doi.org/10.1016/J.AUTCON.2020.103276.

Ye, X. and König, M. (2021) *Framework for Automated Billing in the Construction Industry Using BIM and Smart Contracts*, *Lecture Notes in Civil Engineering*. Available at: https://doi.org/10.1007/978-3-030-51295-8_57.

A BLOCKCHAIN-BASED APPROACH FOR EMBODIED CARBON MANAGEMENT ALONG THE CONSTRUCTION SUPPLY CHAIN

K.L. Kuan[1], Yuqing Xu[1], Chengliang Zheng[2], Jessica S.M. Chan[3], Jack C.P. Cheng[1], and Alexis K.H. Lau[1]
[1]The Hong Kong University of Science and Technology, Hong Kong, China
[2]Wuhan University, Wuhan, China
[3]MTR Corporation Limited, Hong Kong, China

Abstract

Blockchain offers a potential solution to enhance transparency and trackability in carbon management, but limited research exists on its implementation in the construction industry. This study proposes a blockchain-based approach for transparent and trackable carbon management in the construction supply chain to address concerns regarding data quality and confidentiality. A system was developed comprising off-chain carbon management and on-chain (public blockchain) storage, with technical highlights of: role-based access control for functionality, hierarchical hashing strategy for validation, and selective data disclosure before on-chain. The research's contributions include enhancing data quality and traceability for carbon assessments.

Introduction

The construction industry significantly influences global carbon emissions due to its energy-intensive processes, use of materials, and the operation of buildings (Sizirici et al., 2021). Carbon emissions from the construction supply chain, referred to as upfront embodied carbon, are estimated to account for around 30% of construction projects' whole life cycle carbon emissions (Gan et al., 2017). Therefore, how to manage and reduce carbon emissions from construction supply chains has become an urgent task in the construction industry. Although many practices and studies have been implemented to manage carbon emissions at the specific stage of the construction supply chains, it is still difficult to track carbon footprints along the supply chain. For one thing, the fragmentation of supply chains makes it difficult to gather consistent and accurate data on carbon emissions from each stage of the supply chain among multiple tiers of suppliers and subcontractors (Hijazi et al., 2021). For another, considering the wide scope of carbon emission data required during carbon assessment and tracking, stakeholders may find it difficult to access (Lai et al., 2023). In that case, the data transparency is very limited both for internal carbon management along the construction supply chain and external supervision for sustainability disclosure, which hampers the ability to track carbon footprints effectively.

As a promising technology, blockchain offers a possible and effective solution for the insufficient transparency and low trackability problems during carbon management along the construction supply chain (Turk et al., 2017). Blockchain technology is a decentralized and distributed ledger system that allows multiple parties to maintain a shared database without the need for a central authority. It is a chain of blocks, where each block contains a list of transactions or data (Hunhevicz et al., 2020). These blocks are linked together using cryptographic hashes, creating an immutable and transparent record of all transactions or data entries (Tao et al., 2022). With the help of blockchain, problems such as fragmentation, inefficient data-sharing, insufficient data transparency, and poor data trackability have been solved in many construction scenarios. Existing studies have explored blockchain's capability and feasibility in various construction scenarios, including construction payment transactions, construction information-sharing and collaboration during multiple project stages, and governmental supervision of construction work (Xu et al., 2023). However, there is limited research in the field of blockchain-based carbon management along the construction supply chain, especially focusing on the whole supply chain lifecycle and how to motivate all stakeholders to participate in such an innovative approach. During the interviews, it was noted that other challenges include concerns regarding data quality and sensitive information. Therefore, the two research questions in this research are:

(1) How can blockchain technology integrate with carbon management along the whole construction supply chain?
(2) How can carbon data quality and privacy be enhanced in the blockchain-based carbon management system?

This study aims to propose a blockchain-based approach for transparent and trackable carbon management along the construction supply chain for all stakeholders (suppliers, contractors, architects, and developers of the construction project) by developing a blockchain-based carbon management system. This approach covers all stages that will produce carbon emissions in construction supply chains (from material processing, transportation, and on-site construction activities), which is consistent with the A1 to A5 stages from the life cycle assessment identified in international standards.

Related work

Carbon management in the construction supply chain

Carbon emissions from the construction supply chain, referred to as upfront embodied carbon, are estimated to account for around 30% of construction projects' whole life cycle carbon emissions. Under the urgent trends of global carbon reduction, the importance of carbon management in construction supply chains is increasingly recognized.

A range of international standards have been developed and can be applied to assess and manage carbon emissions in the construction industry, including carbon assessment and management in construction supply chains. The European standard, EN 15978, is commonly recognized for the assessment of the environmental performance of new and existing buildings based on a life cycle approach. This standard specifies the calculation method and multiple carbon calculation stages based on Life Cycle Assessment (LCA) and other quantified environmental information based on the building level. Especially, A1 to A5 stages of product sourcing and construction stage identified in EN 15978 help assess carbon emissions from the construction supply chains, including sourcing, transportation, fabrication, and construction of all materials and products (British Standards Institution, 2011).

Moreover, multiple carbon certification schemes for construction materials and products have also been implemented in the construction industry, targeting to control and reduce carbon emissions at the start of construction supply chain management (A1 to A3 stages). For instance, the carbon reduction label managed by the Carbon Trust in the UK is one of the earliest certification schemes in the world. Under this scheme, organizations are required to report environmental claims and detailed carbon-proof data to get the certified label of products. Similar practices can also be found in The Singapore Green Building Product Certification (SGBPC), and the Construction Industry Council (CIC) Green Product Certification scheme in Hong Kong.

In academia, how to manage and reduce carbon emissions in construction supply chains to realize a green supply chain and improve its sustainability is also a hot topic. There are various research categories in the existing papers about carbon management in construction supply chains, including (1) green purchasing and procurement, (2) low-carbon design and manufacturing of construction materials and products, (3) green logistics, and (4) construction waste management. Green purchasing and procurement mainly focus on evaluation schemes and strategies for material and product selection along the construction supply chains. Bagul et al. developed a sustainable sourcing strategy for mega-construction projects using the analytic hierarchy process (AHP) technique and a multi-objective Goal Programming (GP) model, which has also been verified using a single construction megaproject case (Bagul et al., 2023). In the field of low-carbon design and manufacturing of construction materials and products, studies are focusing on how to design environmentally friendly construction materials and limit resource consumption during the production process of construction materials and products. For instance, Yang et al. developed a low-carbon design of an Ultra-High Performance Concrete (UHPC) by incorporating high-volume phosphorous slag (PS), which shows a promising approach to developing a cleaner building material with lower carbon footprints (Yang et al., 2019). Zheng et al. proposed a knowledge-based integrated product design framework to support low-carbon product development. Green logistics is another research field in carbon management along the construction supply chains. It aims to reduce carbon emissions during the delivery process of construction materials and products (Zheng et al., 2021). Chen et al. integrated building information modeling (BIM) and web map service (WMS) for the source selection of sustainable construction materials to reduce carbon emissions during the delivery (Chen et al., 2019). The carbon management of the recycling process of the construction supply chain is also studied in existing studies. Wibowo et al. proposed a measurement model to analyze the carbon reduction performance of recycled materials to support sustainable construction (Wibowo et al., 2017).

Blockchain applications in the construction industry

Blockchain is one of the most prominent types of distributed ledger technology (DLT) that allows all transactions to broadcast and operate on a distributed peer-to-peer (P2P) network without a centralized administrator. It enables the secure and transparent recording of transactions and data. The blockchain transactions are first grouped together, validated by particular consensus mechanisms (e.g., Proof of Work, Proof of Stake) among all nodes, and finally added to a block in a specific order. Since the newly generated block links to the previous block by unique a hashing index, it is difficult to modify once a block is added (needs to modify all previous block indexes and value if want to change one block), which ensures the integrity and trackability of the recorded data (Scott et al., 2021). Currently, blockchain's capability and feasibility have been discussed and validated in various construction scenarios, including construction payment and procurement, information-sharing and collaboration, construction supply chain management (including construction carbon management), regulations and compliance, and contract management (Li et al., 2021 a).

For construction payment scenarios, blockchain-based payment systems can enable faster payments by eliminating intermediaries and self-executing smart contracts. A distributed blockchain-based framework that does not require trust to automatically enforce the terms and conditions related to interim payments was proposed by Das et al., aiming to facilitate payment transparency, enforce conditions of interim payments, and automatically record payment cycles (Das et al., 2020). Similarly, to avoid the risks and disputes caused by slow payments among construction stakeholders, an autonomous payment administration solution integrating blockchain-enabled smart contracts and robotic reality capture technologies was proposed by Hamledari et al., which can automatically transfer cryptocurrencies by smart contracts after finishing phased work (Hamledari et al., 2021).

Blockchain technology has also been applied in the single stage or multiple stages of construction projects, such as design, construction, and maintenance, for information sharing and stakeholder collaboration to enhance construction data transparency, reliability, and trackability. For example, a blockchain-based prototype system was developed and evaluated to address the challenges of design liability control and information security during the construction design process (Pradeep et al., 2021). For the construction stage, a blockchain-based verification framework of adequate scaffolding was proposed for onsite inspections, aiming to make the onsite operations safer (Baek et al., 2020).

Among the above blockchain applications in the construction industry, blockchain-integrated construction supply chain management is one of the most popular research directions across multiple project stages. Recent studies have illustrated blockchain-enabled construction supply chain management via case studies. Wang et al. proposed a blockchain-based information management framework for precast supply chain information-sharing, real-time controlling, and status tracking, and a case study was used to validate the performance of the proposed framework (Wang et al., 2020). Moreover, Lu et al. developed a smart construction objects-enabled blockchain oracles framework to bridge the on-chain and off-chain worlds, which was examined in the context of off-site logistics and on-site assembly services (Lu et al., 2021). Some studies in this field also integrated blockchain with other digital technologies such as BIM and IoT to improve the performance of construction supply chains. For example, Wu et al. linked a permissioned blockchain to the Internet of Things (IoT)-BIM platform for off-site production management in modular construction by providing better information visibility, traceability, and a more collaborative working environment (Wu et al., 2022).

For construction carbon management, blockchain's potential to facilitate carbon data transparency and traceability has been examined in both carbon quantification and assessment. Rodrigo et al. explored the potential application of blockchain for accurate embodied carbon estimation in construction supply chains and developed a data model for the blockchain-based embodied carbon estimator for construction (Rodrigo et al., 2020). Recent literature has developed several theoretical frameworks for blockchain-enabled carbon management. For instance, Wang et al. proposed a conceptual framework integrating blockchain, supply chain, and environmental performance, which suggests using blockchain to enhance supply chain management and reduce carbon emissions (Wang et al., 2020). A blockchain-based identification and coordination framework was designed by Wang et al. based on a specific multi-tier supply chain for sustainable supply chain management, aiming to ensure compliance with sustainability standards in construction supply chains (Wang et al., 2023). Liu et al. proposed a conceptual framework integrating blockchain as a carbon management tool to achieve transparent carbon footprint disclosure during product certification and supply chain management (Liu et al., 2019). Moreover, Xu et al., developed a more detailed and implementable blockchain-based framework for the embodied carbon certification of construction materials and products, which was validated via real certification cases (Xu et al., 2024).

Some other blockchain-based applications in the construction field can also be found in construction supervision (Li et al., 2021 c), and contract management for solving contract disputes and claims (Li et al., 2021 b). In summary, blockchain technology has already been implemented in many construction cases with different objectives, such as increasing data transparency, enhancing mutual trust, and facilitating tracking functionalities. However, existing blockchain-integrated research mainly based on less decentralized solutions without transparency to the public, and there is limited research about sufficient transparency and traceability of upfront embodied carbon management among all stakeholders along construction supply chains.

Methodology

The research methodology employed in this study is based on the Design Science Research (DSR) method, which facilitates problem-solving activities as depicted in Figure 1. The research problems and objectives were identified through an extensive literature review of existing studies on carbon management in construction supply chains and blockchain-based applications, as well as through interviews conducted with stakeholders across the construction supply chain. Once the research problems were identified, the system framework and workflow of the blockchain-based system were developed. This entailed creating a model diagram that showcases the key components and a process diagram that illustrates the system workflows among different stakeholders. Subsequently, a system prototype was developed, tested, and demonstrated through a real-life case study, enabling the recording and tracking of carbon footprints along the construction supply chain. The final two steps of this study involve the evaluation of the system and its practical assessment through communication with relevant stakeholders.

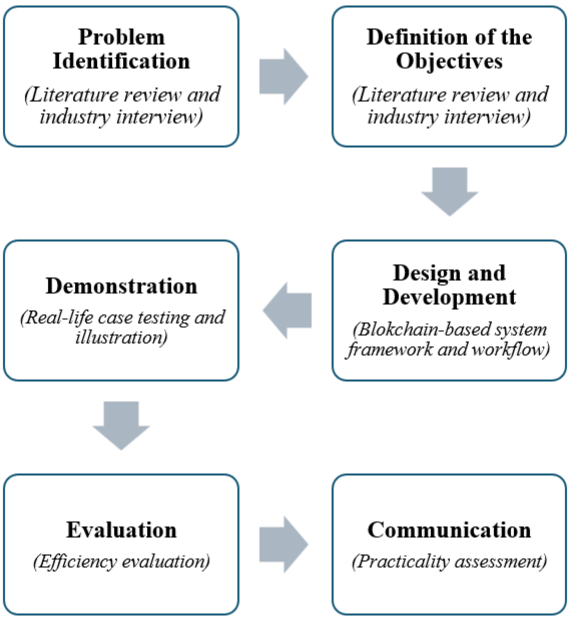


Figure 1: Research methodology in this study

System requirement analysis

Interview results for system requirements

We conducted interviews with one supplier, two contractors and one developer to gain insights into the industry practices related to embodied carbon practices. We also sought their opinions on the proposed blockchain-based carbon management system. The interview questions are provided below:

1. How does your company collect and store carbon-related information/data for a project? For instance, what types of data are collected, how is data collected from different parties, who is responsible for inputting the data, and how is the data stored?
2. Does your company have any concerns about using a blockchain-based carbon management system for tracking carbon-related data in the construction supply chain?

Based on the interviews, it is noted that only a few stakeholders have a practice of collecting and tracking embodied carbon data in their projects, such as having a centralized procurement system and an online centralized database for storing data. The stakeholders also expressed concerns regarding data quality and sensitive information (e.g., the design formula of the product, total quantity and inventory sources). Besides, they expressed expectations for the system to support data sharing throughout the project and industry (e.g., project emission factor), and facilitate the establishment of industry benchmarks.

System design

The research primarily focuses on the integration of blockchain technology with carbon management across the entire construction supply chain to enhance the quality and traceability of embodied carbon data for more accurate carbon assessments. In our blockchain-based carbon management system, the stakeholders involved are developers, architects, consultants (such as structural consultants and carbon consultants), contractors (such as foundation contractors and main contractors), and product suppliers (such as concrete and steel suppliers). The contractor and supplier are further categorized into general staff and senior staff for different roles in the system. The system framework is referenced in Figure 2.

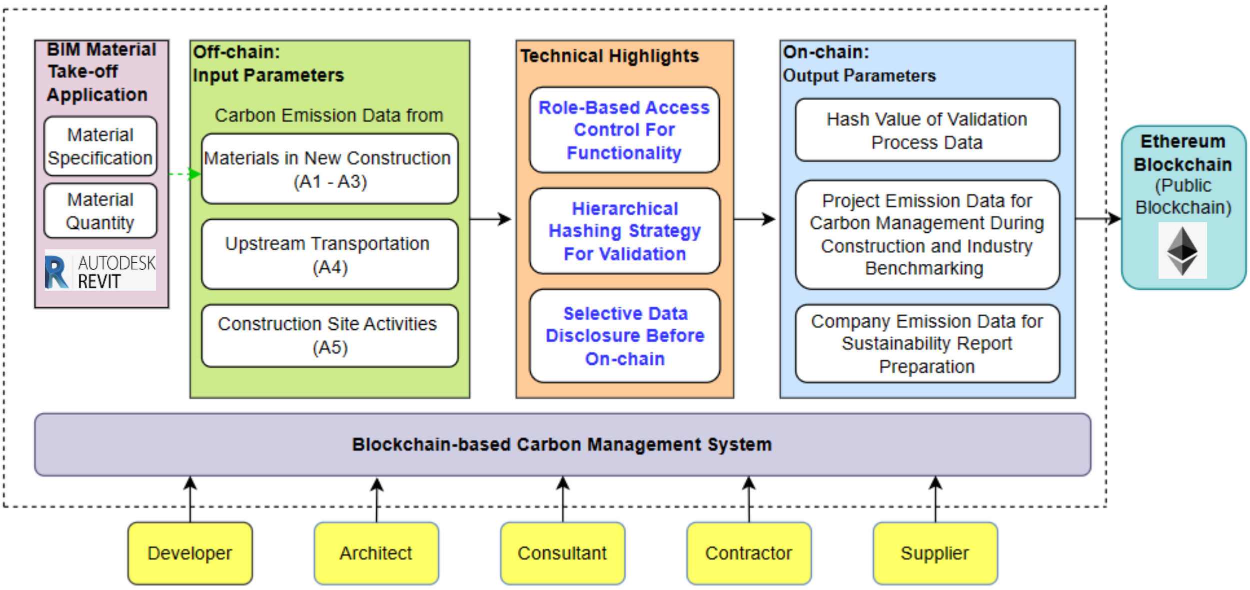


Figure 2: System framework

The proposed system framework provides the traditional off-chain carbon management functions, encompassing user registration and data input. Within this system, all stakeholders should register a unique account in the blockchain and carbon management system with their respective identities. Additionally, the architect creates a project account in the system. Under normal circumstances, the general staff of the contractors and product suppliers will be responsible for data input based on their respective roles in the project. Carbon data input parameters are divided into three groups based on the building life cycle: Materials in New Construction (A1 - A3), Upstream Transportation (A4), and Construction Site Activities (A5). These parameters include raw materials, origin, emission factors, material quantities, and transportation modes and distances. The BIM material take-off function is utilized to extract information for the input of A1 – A3, which covers material specification and quantity.

To make carbon footprints transparent, reliable, and trackable, the proposed blockchain-based system framework entails an on-chain smart contract and public blockchain storage for data recording and output. A decentralized database with self-executing smart contracts is utilized to store significant project information and carbon emissions data. In addition to the previously mentioned hash value of validation process data, the system also records and generates output data on a project basis and company basis. This includes details such as the project address, product emission factor, product quantity, monthly carbon emissions, total emissions in Materials in New Construction (A1 - A3), Upstream Transportation (A4), Construction Site Activities (A5), and the overall project emissions. These data serve various purposes, such as facilitating carbon management throughout the construction period, preparing sustainability reports, and establishing benchmarks. By storing data on the blockchain, it becomes immutable and can be easily shared with the project team and the public for enhanced traceability and transparency.

Technical Specification

There are three technical highlights of this carbon management system:

- *Role-Based Access Control For Functionality*: Every stakeholder has a clearly defined role and functionality within the carbon management system. The levels of power within the system, listed in descending order, are as follows: The Developer holds the highest level, followed by the Architect, Consultant, Contractor, and Supplier, while within each level, the Senior Staff holds a higher position than the General Staff. For example, in the context of the carbon management system being applicable to the industry rather than a single company, ensuring proper access control for individual and project accounts becomes crucial. Therefore, during user registration and project initialization, only the Architect has the authority to create projects in the system and to add, edit, or delete the accounts of Developers, Contractors, and Consultants. The Senior Staff, on the other hand, can add, edit, or delete the accounts of General Staff members. Furthermore, in the system view, the Contractor has the right to create and view all inputted carbon information, while the Developer is limited to viewing only the total carbon emissions data.
- *Hierarchical Hashing Strategy For Validation:* Data quality is a significant concern for stakeholders. In order to enhance the quality of carbon data, a validation process consisting of three hash values is implemented within both a company and a project team. The first value represents the original carbon data, the second value corresponds to the validation process executed by the first designated stakeholder (e.g., senior staff of the supplier reviewing and validating the input data provided by the general staff of the supplier), and the third value represents the validation process executed by the second designated stakeholder (e.g., after the senior staff of the supplier validates the input, the senior staff of the contractor further reviews and endorses the data). These three hash values, associated with the validation process data, are then stored in the blockchain.
- *Selective Data Disclosure Before On-chain:* Once the data has been validated, stakeholders who are concerned about sensitive data disclosure have the right to choose whether to disclose or encrypt the data before uploading it to the blockchain and sharing it with the public. For example, if the supplier has reservations about disclosing the product design formula, the senior staff of the supplier can opt to encrypt this information while disclosing other carbon data such as product specifications and emission factors.

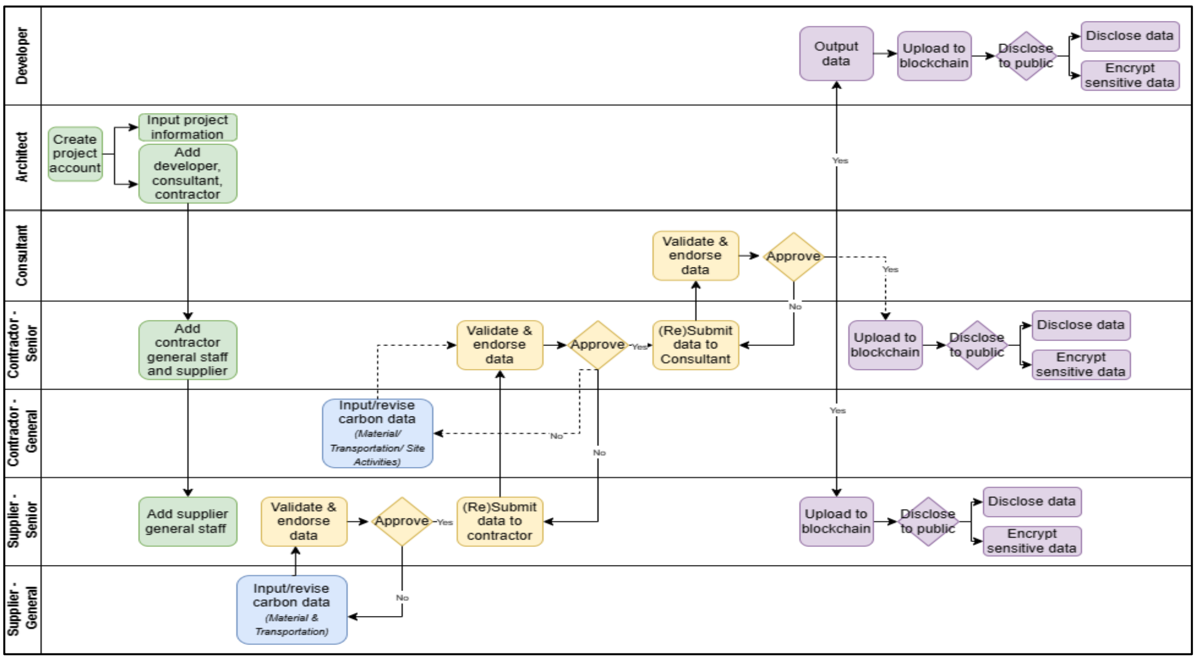


Figure 3: Operational Workflow (Green - User/Project Initialization; Blue – Data Input; Orange – Validation Process; Purple - Smart Contracts and Public Blockchain)

Operational Flow

In order to have a comprehensive understanding of the system, an operational workflow process is presented in Figure 3. Firstly, the architect initiates the creation of a project account and inputs the project information, as well as adds the relevant stakeholders (e.g., developers and contractors) within the project account. Additionally, the contractor adds the product suppliers, while the senior staff of the supplier/contractor includes their respective general staff in the project. Secondly, the general staff of product supplier and/or contractor input the carbon data into the information management system. Once the data is ready, their senior staff will review and validate the data before submitting it to the upper-level party for further validation. Thirdly, following the data validation process, the senior staff of the product supplier/contractor can choose specific data for encryption and subsequently upload it to the blockchain. This upload includes the three hash values associated with the validation process data, as elaborated in the previous section. Lastly, the developer possesses the final authority to select data pertaining to the total project and company emission for encryption and subsequently uploads it to the blockchain.

Implementation and results

System development

Figure 4 shows the system development details. First, a front end was developed for different users to collect carbon emission data via the web browser. A blockchain wallet tool called Metamask was used in this study to manage user blockchain accounts and provide methods to sign blockchain transactions during token transfer. At the backend, we use Alchemy to build the blockchain environment and provide the blockchain infrastructure. Considering this study aims to make carbon management along construction supply chains more transparent and directly assist with carbon data disclosure, Ethereum, a public blockchain platform, is thus used in this study to realize blockchain functions with smart contracts. For data storage, we use the integration of traditional databases and blockchain databases for safety and transparency. The traditional database is used to store original carbon emission data provided by all users. The blockchain transactions store and disclose all public carbon data transparently. At the same time, confidential carbon data for each party will also be stored on the blockchain after hash processing to satisfy the user's need to track their own carbon data through the blockchain-based carbon management system.

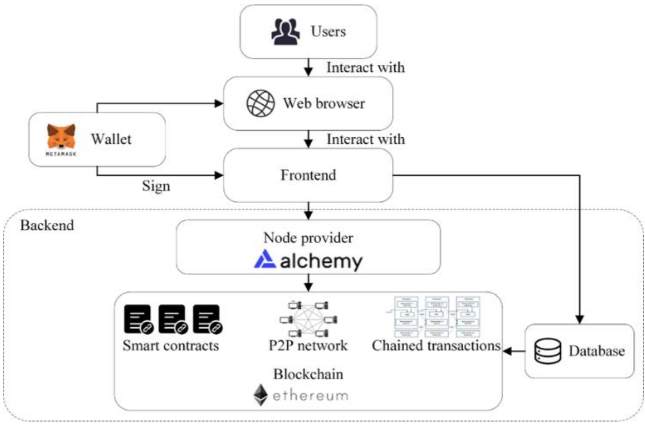


Figure 4: System development details

Result demonstration – User interface of input parameters

As mentioned previously, the DSR method is used as the research methodology in this study, consisting of six steps covering 1) problem identification, 2) definition of objectives, 3) design and development, 4) demonstration, 5) evaluation, and 6) communication. Based on the research progress, the study is currently at the stage of system design and development, with the user interface of input parameters being presented in this section. The input parameters consist of two main categories. The first category is the project information (e.g., site address, construction floor area) and project team information (e.g., names of stakeholders and their corresponding blockchain accounts), which are input by the architect. The second category is the upfront embodied carbon-related data, which covers the A1 - A5 modules of the building life cycle (from raw material extraction to the construction stage). This category is further divided into three groups based on the building life cycle and the parties responsible for inputting the data (i.e., product suppliers and contractors) for the substructure and superstructure stages. Figure 5 displays the extracted user interface of the input parameters.

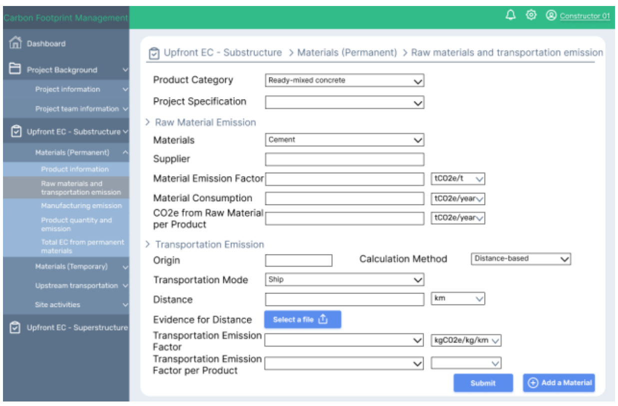


Figure 5: User interface diagram – carbon data related to raw materials and transportation emission

1. *Materials for constructing new buildings (permanent and temporary)*: This group includes emissions from raw material extraction (A1), transport to manufacturing (A2), and manufacturing (A3). The data includes product

information (e.g., product category and specifications), raw material emission data (e.g., emission factors, material consumption), transportation emission data (e.g., origin, transportation mode, fuel emission factors, and consumption), and manufacturing emission data (e.g., manufacturing process type and the fuel used).

2. *Upstream transportation and distribution*: This group includes emissions from A4 (transport to site) only, which refers to emissions from transporting products and materials from manufacturing plants to the project site. The input data is similar to the transportation emissions data in group 1 (e.g., transportation mode and distance).
3. *Construction site activities*: This group covers emissions during the construction of the building(s). The data includes emissions from fuel and energy-related activities (e.g., electricity, town gas, refrigerant, welding, and flame cutting) and waste generated during operations (e.g., landfill, public fill, metal, and timber recycling).

Conclusions

Carbon footprint management, especially carbon tracking, is complex along the construction supply chain due to the lack of accurate data collection and sufficient data transparency. Blockchain provides a potential and effective solution to enhance transparency and trackability issues in carbon management across the construction supply chain. However, there is limited research in the field of blockchain-based carbon management along the construction supply chain, especially focusing on the whole supply chain lifecycle and how to motivate all stakeholders to participate in such an innovative approach. Therefore, this study proposes a blockchain-based approach for transparent and trackable carbon management along the construction supply chain for all stakeholders by developing a blockchain-based carbon management system. The research methodology employed in this study is based on the Design Science Research (DSR) method based on (1) the literature review results of existing research about carbon management in construction supply chains and blockchain-based applications and (2) industry interview feedback from stakeholders along the construction supply chain. Based on the system requirement analysis results, a blockchain-based carbon management system is designed and developed, which consists of an off-chain traditional carbon management method and an on-chain (public blockchain) storage. There are three technical highlights of this carbon management system: role-based access control for functionality, hierarchical hashing strategy for validation, and selective data disclosure before on-chain. This research mainly contributes to enhancing the quality and traceability of embodied carbon data for more accurate carbon assessments by using blockchain during construction supply chain management. In the future, we intend to improve the system development and test the system by inviting more industry stakeholders to provide feedback.

Acknowledgments

The authors would like to thank Gammon Construction and Limited, China State Hong Kong for their support during the interviews. Also, the work described in this paper was substantially supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region (Project No. T31-603/21-N). Any opinions and findings are those of the authors and do not necessarily reflect the views of the Research Grants Council of the Hong Kong Special Administrative Region.

References

Baek, Chan-woo, Do-Yeop Lee, and Chan-Sik Park. "Blockchain based framework for verifying the adequacy of scaffolding installation." ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction. Vol. 37. IAARC Publications, 2020.

Bagul, A., Balon, V., & Mathur, A. (2023). Sustainable sourcing strategy for mega construction projects: An empirical analysis. Business Strategy & Development, 6(3), 411-419.

British Standards Institution. (2011) BS EN 15978 Sustainability of construction works -assessment of the environmental performance of buildings -calculation method, British Standards Institution, London.

Chen, P. H., & Nguyen, T. C. (2019). A BIM-WMS integrated decision support tool for supply chain management in construction. Automation in Construction, 98, 289-301.

Das, M., Luo, H., & Cheng, J. C. (2020). Securing interim payments in construction projects through a blockchain-based framework. Automation in construction, 118, 103284.

Gan, V. J., Chan, C. M., Tse, K. T., Lo, I. M., & Cheng, J. C. (2017). A comparative analysis of embodied carbon in high-rise buildings regarding different design parameters. Journal of Cleaner Production, 161, 663-675.

Hamledari, H., & Fischer, M. (2021). Construction payment automation using blockchain-enabled smart contracts and robotic reality capture technologies. Automation in Construction, 132, 103926.

Hijazi, A. A., Perera, S., Calheiros, R. N., & Alashwal, A. (2021). Rationale for the integration of BIM and blockchain for the construction supply chain data delivery: A systematic literature review and validation through focus group. Journal of construction engineering and management, 147(10), 03121005.

Hunhevicz, J. J., & Hall, D. M. (2020). Do you need a blockchain in construction? Use case categories and decision framework for DLT design options. Advanced Engineering Informatics, 45, 101094.

En, L. K., Rahiman, N. A., Othman, N., Ali, K. N., Wah, L. Y., Moayedi, F., & Dzahir, M. A. M. (2023). Quantification process of carbon emissions in the construction industry. Energy and Buildings, 113025.

Li, Jennifer, and Mohamad Kassem. (2021 a). Applications of distributed ledger technology (DLT) and Blockchain-enabled smart contracts in construction. Automation in construction, 132, 103955.

Li, W., Duan, P., & Su, J. (2021 b). The effectiveness of project management construction with data mining and blockchain consensus. Journal of Ambient Intelligence and Humanized Computing, 1-10.

Li, X., Wu, L., Zhao, R., Lu, W., & Xue, F. (2021 c). Two-layer adaptive blockchain-based supervision model for off-site modular housing production. Computers in Industry, 128, 103437.

Liu, K. H., Chang, S. F., Huang, W. H., & Lu, I. C. (2019). The framework of the integration of carbon footprint and blockchain: using blockchain as a carbon emission management tool. Technologies and eco-innovation towards sustainability I: Eco-design of products and services, 15-22.

Lu, W., Li, X., Xue, F., Zhao, R., Wu, L., & Yeh, A. G. (2021). Exploring smart construction objects as blockchain oracles in construction supply chain management. Automation in construction, 129, 103816.

Pradeep, A. S. E., Yiu, T. W., Zou, Y., & Amor, R. (2021). Blockchain-aided information exchange records for design liability control and improved security. Automation in Construction, 126, 103667.

Rodrigo, M. N. N., Perera, S., Senaratne, S., & Jin, X. (2020). Potential application of blockchain technology for embodied carbon estimating in construction supply chains. Buildings, 10(8), 140.

Scott, D. J., Broyd, T., & Ma, L. (2021). Exploratory literature review of blockchain in the construction industry. Automation in construction, 132, 103914.

Sizirici, B., Fseha, Y., Cho, C. S., Yildiz, I., & Byon, Y. J. (2021). A review of carbon footprint reduction in construction industry, from design to operation. Materials, 14(20), 6094.

Tao, X., Liu, Y., Wong, P. K. Y., Chen, K., Das, M., & Cheng, J. C. (2022). Confidentiality-minded framework for blockchain-based BIM design collaboration. Automation in Construction, 136, 104172.

Turk, Ž., & Klinc, R. (2017). Potentials of blockchain technology for construction management. Procedia engineering, 196, 638-645.

Wang, B., Lin, Z., Wang, M., Wang, F., Xiangli, P., & Li, Z. (2023). Applying blockchain technology to ensure compliance with sustainability standards in the PPE multi-tier supply chain. International Journal of Production Research, 61(14), 4934-4950.

Wang, M., Wang, B., & Abareshi, A. (2020). Blockchain technology and its role in enhancing supply chain integration capability and reducing carbon emission: A conceptual framework. Sustainability, 12(24), 10550.Wang, Z., Wang, T., Hu, H., Gong, J., Ren, X., & Xiao, Q. (2020). Blockchain-based framework for improving supply chain traceability and information sharing in precast construction. Automation in construction, 111, 103063.

Wibowo, M. A., Sholeh, M. N., & Adji, H. S. (2017). Supply chain management strategy for recycled materials to support sustainable construction. Procedia engineering, 171, 185-190.

Wu, L., Lu, W., Xue, F., Li, X., Zhao, R., & Tang, M. (2022). Linking permissioned blockchain to Internet of Things (IoT)-BIM platform for off-site production management in modular construction. Computers in Industry, 135, 103573.

Xu, Y., Tao, X., Das, M., Kwok, H. H., Liu, H., Wang, G., & Cheng, J. C. (2023). Suitability analysis of consensus protocols for blockchain-based applications in the construction industry. Automation in Construction, 145, 104638.

Xu, Y., Tao, X., Das, M., Kwok, H. H., Liu, H., Kuan, K. K., ... & Cheng, J. C. (2023). A blockchain-based framework for carbon management towards construction material and product certification. Advanced Engineering Informatics, 61, 102242.

Yang, R., Yu, R., Shui, Z., Gao, X., Xiao, X., Zhang, X., ... & He, Y. (2019). Low carbon design of an Ultra-High Performance Concrete (UHPC) incorporating phosphorous slag. Journal of Cleaner Production, 240, 118157.

Zheng, H., Yang, S., Lou, S., Gao, Y., & Feng, Y. (2021). Knowledge-based integrated product design framework towards sustainable low-carbon manufacturing. Advanced Engineering Informatics, 48, 101258.

DIRECTIONS FOR RESEARCH ON INCENTIVISING AND GOVERNING THE INCREASE OF BIODIVERSITY IN CITIES WITH BLOCKCHAIN TECHNOLOGIES

Jennifer Li[1], and Theodoros Dounas[2]
[1]Northumbria University, Amsterdam, The Netherlands
[2]Antwerpen University, Antwerp, Belgium

Abstract

Biodiversity is crucial in sustaining all life on earth. However, extinction of species critical to the functioning of our global ecosystem is a direct threat to biodiversity. The built environment in its urban forms is one of the biggest causes of biodiversity loss through the exploitation of finite resources and land use on which it sits. In this paper we consider the application of blockchain to create new forms of incentivisation and governance for maintaining and increasing biodiversity in cities at the intersection of co-design, blockchain, biodiversity, and design and planning, and propose directions for future research on this topic.

Introduction

Biodiversity is "*the variety of life on Earth and the natural patterns it forms*" (Convention on Biological Diversity, 2009, para. 1). Such variety of species is critical to the functioning of the earth's ecosystem (Duffy, 2009). However, extinction of species is accelerating faster than ever in human history (WEF, 2020) threatening this variety. Biodiversity loss is reported as fourth in the top five global risks to humanity regarding likelihood and third regarding impact (WEF, 2020). Rockström *et al.* (2009) estimated three planetary boundaries (biodiversity loss, climate change, the nitrogen cycle) that sustain the Holocene and safe operating space for humanity are already exceeded. Global urbanisation, "*the process of anthropogenic transformation of wildlands to the built environment where people live and work*" (2021, p. 125), has accelerated these changes; since 1950, human activity has resulted in more energy use than the entire 11,000 years of the previous Holocene (Elmqvist *et al.*, 2021). As the planet deals with several crises centred on climate change, pollution, biodiversity loss (UN, 2023), and a clash between finite resources and an increasing population (Scheel *et al.*, 2020), a fundamental shift is essential in the relationship between humanity and nature to solve these problems (UN, no date). Being more sustainable is part of the solution, yet not enough is being done to prevent the disastrous impacts they bring (Abbass *et al.*, 2022).

The aim of this paper is to establish directions for research to investigate the extent to which blockchain technologies and smart contracts can be used to create new forms of incentivisation and governance for maintaining and increasing biodiversity in cities. The project will focus specifically on cities as the areas causing most of the biodiversity loss across the planet through urbanisation (Kirk *et al.*, 2021) and with a spotlight on the urban design and planning phase of the construction lifecycle where the decisions about biodiversity would have the most impact.

In the next section, the methodology for conducting the research for this paper is outlined, then literature on the role of design and planning in constructing the built environment is reviewed. Blockchain as a socio-technical system is presented with consideration to its possible use to govern and incentivise the increase of biodiversity in cities. This is done by analysing literature at the intersection of biodiversity, blockchain and construction through a socio-technical lens. Next, future directions for this research project are discussed before concluding the paper in the final section.

Methodology

In Figure 1 below, the three areas of research interest for this project are shown. There are existing studies at the intersections for two components (e.g., UB and BCT; UB and BE, BCT and BE). However, there is very little research sitting at the centre where all three interest areas intersect (i.e., UB, BCT *and* BE).

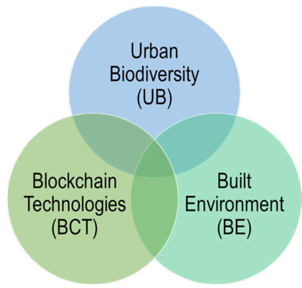


Figure 1: Intersection of research interests

To fill this gap, this paper took an inductive approach to conduct a traditional literature review to establish what is known at the intersections between two or three of the themes above. Given that there is little research at the intersection of all three, the paper suggests directions to fill this gap based on a synthesis of what is currently known.

Literature Review

The built environment's impact on biodiversity

The construction sector is responsible for constructing and maintaining the built environment necessary to maintain human life. However, it is also "*known to be the highest contributor*" to climate change (Oguntona and Aigbavboa, 2019, p. 513). Cities are essential to withstand the increase in population density but rapid expansion of urban areas are associated with resultant biodiversity loss (Uchida *et al.*, 2021). Through excessive resource use and adaptation of the land, construction can damage natural habitats jeopardising wildlife and plants. Thus, construction has a responsibility to protect sites and minimise damage to ecology. However, understanding of biodiversity in the sector is poor (Woodall and Crowhurst, 2003) and "*is often viewed as an afterthought or final addition once an urban development nears completion*" (Kirk *et al.*, 2021, p. 1). Biodiversity loss has critical consequences for the human race caused by current methods of production and overconsumption that could see the failure of food systems, collapse of healthcare systems and see whole supply chains disrupted (WEF, 2020). This creates an imbalance of the ecosystems that sustain clean air, water and all life on earth (WEF, 2020). The sector uses a substantial amount the world's resources from extracted materials, to global water and energy supplies, whilst being one of the highest contributors of greenhouse gasses (GHG) (Wieser *et al.*, 2019). It contributes heavily to the generation of waste and reduction of global ecological integrity, as well as buildings and infrastructure having a long-term environmental footprint (Oguntona and Aigbavboa, 2019). Cities are of strategic importance to the "*global governance of climate change*" and limiting global warning to only 1.5°C (Bulkeley, 2022, p. 266). It hinges on the ability to redress planetary limits and boundaries through reshaping urbanisation (Elmqvist *et al.*, 2021). An estimate puts global temperature increase at 4 degrees by 2100 if governments and the construction sector continue current policies and practices, thus, municipal level decisions are critical to limit the temperature rise to 1.5°C (Miller, 2020).

The role of design and planning in construction

Construction needs to be more sustainable utilising "*a sociotechnical approach, [that] underpins the principle of sustainable development*" (Medaglia and Damsgaard, 2020, p. 3). This can be done in part by moving toward a co-design approach that involves citizens in municipality-level decisions that feed into the design and planning phase of the project lifecycle. Such a concept gaining attention in the creation of the built environment is regenerative design. This positions citizens as both co-creators and contributors of earth's ecosystems moving beyond the current mindset of resource exploitation without consideration of the consequences (Wang *et al.*, 2023) where quality of life is put before economic growth (Axinte *et al.*, 2019). Camrass (2022) states that urban planners and designers are in a unique position to drive forward the regenerative agenda.

Miller (2020) writes about true democracy taking place in Canadian cities where citizens are regularly engaged with and are part of decision-making at municipality levels regarding urban planning. Joachain and Klopfert (2011) discuss two complementary currency programmes as policy instruments as follows: NU-Spaarpas was a loyalty card scheme in partnership with independent shops in Rotterdam, The Netherlands to promote greener consumption and behaviour from citizens. E-portemonnee based in Overpelt, Belgium focused on promoting sustainable behaviours by rewarding sustainable actions of citizens. This demonstrates the appetite and necessity for community engagement in promoting more sustainable practices in cities.

Blockchain

Blockchain is rooted in a desire to challenge central elites and associated power structures by offering a transparent, traceable, decentralised and distributed ledger (Ekblaw *et al.*, 2016) that can record anything of value (Mathews *et al.*, 2017). Its decentralised network of computers stores a common ledger and agrees on the truth of transactions recorded on the chain via an algorithmic consensus mechanism. Transactions are recorded on a chain of blocks where each one engulfs the previous blocks' cryptographic hash. It offers security and privacy based on its append-only nature while its immutability provides a single source of truth (Hijazi *et al.*, 2022).

Blockchain has been identified as a socio-technical system by several scholars (Shin *et al.*, 2022; Li, 2023; Selvanesan and Satanarachchi, 2023) who acknowledge society as central to its utility; consideration of this dimension key to its successful adoption. It gave rise to cryptoeconomics – applying cryptography to economic systems for new forms of organisational business models and commons frameworks (Brekke and Alsindi, 2021) – comprised of computer engineering, economics and game theory (Brekke, 2021). Smart contracts are programmable pieces of code that can run on a blockchain and self-execute based on pre-defined logic offering automation to create efficiencies and increase productivity without human intervention. They can facilitate the creation of powerful, domain-specific decentralised applications (Ye *et al.*, 2022). There is concern of the high energy consumption of blockchains, however, this is a misconception for it is applicable only to early blockchains that use the proof-of-work (PoW) consensus mechanism. Alternatives such as proof-of-authority, proof-of-burn, proof-of-capacity and others are very energy efficient (Lashkari and Musilek, 2021). For example, Ethereum originally ran PoW but changed to proof-of-stake in 2022 to consume 99.9% less energy equating to a megaton of carbon weekly (Kessler, 2022).

Blockchain as an incentive and governance tool

Several studies have considered blockchain for incentive and governance systems. Wang *et al.* (2023) discuss blockchain's potential to govern information, procedures, ownership and values of a regenerative built environment. Zhao *et al.* (2023) present a blockchain-based token incentive mechanism for environmental, social, and governance (ESG) performance to be used by organisations to attract investors, issue dividends, grant access to products and services, and allocate voting rights. The integration of blockchain, BIM, artificial intelligence, machine learning, building energy management systems and life cycle assessment is proposed by Desmond and Salama (2023) to manage energy usage of buildings. Incentives lie in the trading of stored energy, and smart contracts facilitate tax relief to buildings meeting sustainability objectives. Dounas *et al.* (2022) present a cryptoeconomics-based incentive system centred on architectural design governance connecting building performance and carbon reduction with financial incentives. Incentivising better than net-zero energy buildings are proposed by O'Reilly and Mathews (2019) by storing excess energy generated by solar panels and selling it as a commodity. This incentivises designers to optimise building performance and occupants to behave energy efficiently.

Beyond sustainability-related research, blockchain tokenisation can incentivise changes in human behaviour through decentralisation and offering multidimensional payment systems against today's one-dimensional monetary system (Hunhevicz and Hall, 2020). Hunhevicz *et al*. (2021) employ cryptoeconomics to incentivise lifecycle performance rewarding parties based on performance-based contracts. Green Coin, a blockchain-powered social currency, incentivises citizens to exchange urban waste via a tradable cryptocurrency that can be used in local shops in a Brazilian municipality (França *et al.*, 2020). It helps low-income families, supports awareness of correct ecological behaviours of citizens, reduces landfill waste, and improves health through reduction of fatal diseases such as dengue and zika. A survey on blockchain-incentivised applications across transport, energy and recycling in smart cities is presented in Kahya *et al.* (2021). Finally, in Marsal-Llacuna (2018), blockchain forms the basis of a governance and incentive system to facilitate prioritisation and submission of citizens needs to municipalities leading to the development of bottom-up, citizen-centric masterplans connected to urban codes (e.g., polices, planning, regulations).

Blockchain and biodiversity in research

There are few studies that focus on blockchain *and* biodiversity and none that intersects with construction. A blockchain-based biodiversity value chain is being developed to manage the unregulated exploitation of indigenous natural habitats (Bose *et al.*, 2019). Blockchain is used in food supply chains to conserve biodiversity through incentivising sustainable practices in agriculture and food waste (Kafi *et al.*, 2023). A literature review finds blockchain can support sustainable forestry, minimise illegal logging and help conserve biodiversity (He and Turner, 2022). The co-design of inter-species relations based on programmable cryptocurrency creates "*cities' transition towards Post-Anthropocene for cross-species bio-digital coliving*" facilitated by blockchain (Davidova and McMeel, 2020, p. 338). Beyond academic literature, Greener Tokens (2023) is an initiative that aims to connect companies to restorative finance (ReFi) to protect the Brazilian rainforest. They offer carbon emissions offsetting, ESG project management, and environmental consulting. They avoid greenwashing through their governance framework to ensure true protection of biodiversity. Funded by the Inter-American Development Bank (IDB), the design of digital tokens for a biodiversity innovation challenge aims to preserve natural capital and biodiversity in Colombia, Ecuador, Peru, and Trinidad and Tobago (IDB, 2023). Crypto Altruism (2023) discusses ways in which preservation of natural capital can be incentivised through tokenisation (e.g., carbon offsetting, fractional ownership, preservation of endangered species and the surrounding habitats). Token Kitchen is exploring the preservation of endangered African Forest elephants that support carbon capture and sequestration in the surrounding forests that incentivises locals by issuing tokens upon proof they are maintaining natural habitats (Voshmgir, 2021). Czura (no date) presents three applications: issuing Green Bonds to give investors more control over how their money is invested; TreeCycle that promotes eucalyptus tree planting in Paraguay where investors receive 40% of profits at harvest, 10% goes to local charities and 50% is reinvested into replanting; and the promotion of more sustainable supply chains (e.g., in aquaculture) with a blockchain-based trusted chain of provenance data.

Giving agency to nature, individuals and 'things'

Can blockchain give agency to nature, individuals and 'things' where a thing could be an inanimate object, a system or whole ecosystem? In Athens, Georgia in the United States, there is a 'Tree That Owns Itself' where it was gifted its ownership by the owner of the land upon his death (Figure 2).

Figure 2: A plaque to commemorate The Tree That Owns Itself, "For and in consideration of the great love I bear this tree and the great desire I have for its protection for all time. I convey entire possession of itself and all land within eight feet of the tree on all sides – William H. Jackson" (The Treeographer, 2018)

In a legal sense, the tree does not own itself but nobody disputed the self-ownership and so it was accepted. Upon the death of the Tree That Owns Itself, an acorn from the original tree was planted in its place where a new oak tree now stands; it is known as 'The Son of the Tree That Owns Itself' (The Treeographer, 2018). While this might be difficult to comprehend on a boarder scale, there is research ongoing that allows assets to own themselves in the form of no1s1, a blockchain-based project of a self-owned meditation pod at ETH Zurich facilitated by a decentralised autonomous organisation (DAO) (J. Hunhevicz *et al.*, 2021). The University of Edinburgh's BitBarista project is a coffee machine that accepts payment in bitcoin before administering a coffee to the user and pays people in bitcoin for servicing the machine (e.g., refilling water, cleaning). Individuals buying coffee answer questions about what they want such as high quality, low price, sustainable coffee and so on, the data for which determine the type of coffee BitBarista then automatically orders for itself (McMeel and Sims, 2020).

The challenge of giving agency to nature is that nature is limited in its ability to demonstrate intent. According to Schlosser (2019, p. 1), "*an agent is a being with the capacity to act, and 'agency' denotes the exercise or manifestation of this capacity*". Equally, an inanimate object such as a meditation pod is also limited in its ability to demonstrate intent given its lack of consciousness. The DAO is encoded with a set of rules determined by the creators of the pod (i.e., researchers at ETH Zurich). This is of course different to, for example, plants and animals that have inherent intent within the natural order of the ecological system around us, whereas with no1s1 and BitBarista, the logic is decided by the humans who created it, whose inherent intent is arguably different.

BitBarista is an example of "*devolved decision-making*" (McMeel and Sims, 2020, p. 215) that we could also call co-design where decisions are made based directly on users' wants, needs and means. This new form of business model can challenge existing hierarchical decision-making systems that typically benefit the few profiting from those decisions. Smart contracts and be explored for their ability to facilitate agency for nature and things; the challenge will lie in who is given the task and responsibility for establishing intent that is for the greater good and not the profiteering few.

Consolidation of the literature

Consolidating the above literature, the built environment is one of the biggest exploiters of the world's resources causing substantial biodiversity loss. Current efforts to reduce the negative impacts are insufficient. Therefore, new governance and incentive mechanisms are required to achieve lasting positive impact. Regenerative design aims to bring together different components of the urban ecosystem across the environmental, social and cultural domains with a view to leave the built environment in a better state than it was found. Blockchain could be a solution to bring these elements together as demonstrated by research and initiatives looking to leverage cryptoeconomics for behavioural change. The current efforts of applying blockchain to tackle issues of biodiversity focus on indigenous forests and habitats, and financial incentives, with very little focus on the urban complex of cities. The subsequent sections consider what can be done through this research project to address some of these challenges.

Future research directions

In this section, future directions for this research project are considered through the lens of the socio-technical framework as set out below. Then it considers initial gaps identified in the literature, and challenges and constraints that will form the focus of the future research before discussing the proposed outcomes.

A socio-technical perspective

A socio-technical framework for implementation of blockchain in construction will form the basis of the theoretical underpinnings for this research project. The framework (Li, 2023) consists of a set of tools and models to support analysis, understanding and implementation of blockchain in construction applications. Its four dimensions of technology, process, policy and society acknowledge the importance of technology whilst giving equal attention to the surrounding ecosystem ensuring any application is developed with a focus on meeting the needs at the social level where benefits are realised.

Technology considers the hardware, software, networks, infrastructure and technological integrations of blockchain-based systems; *process* considers how blockchain will integrate into existing organisational processes and the changes required to adopt blockchain at an enterprise level; *policy* focuses on the regulatory environment surrounding blockchain-based applications regarding standards, laws, compliance etc.; and *society* addresses the impact of blockchain applications on its users and the social environment where benefits will be realised.

Research on blockchain in construction has accelerated in recent years (Li and Kassem, 2021) to propositions of use cases across the built environment with plentiful proof-of-concepts and a handful of applications demonstrated in the 'real-world'. It is now time to focus the framework on achieving real change not just for the built environment, but for the betterment of humanity. In this paper, an initial evaluation of blockchain's ability to address the state of urban biodiversity is made. In later work, the framework in its entirety will be applied to better understand the challenges and provide solutions.

Technology: The literature reviewed in this paper has shown the potential of blockchain technologies to facilitate incentive and governance mechanisms that can be directed toward biodiversity. This has been shown by conservation of forests and biodiversity outside of the urban complex (e.g., He and Turner, 2022; Kafi *et al.*, 2023) plus establishment of the several tokens focused on biodiversity in the previous section. The next step is to consider the added complexities of biodiversity maintenance and increase in cities and evaluate how the use of blockchain and smart contracts can facilitate the integration of co-creation in the design and planning phase of constructing the built environment.

Process: From a process perspective, such a re-conceptualisation as co-creation requires re-thinking at several levels including citizen, municipal, and government as well as an understanding of incentives and governance models for urban planners, architects and designers. The organisations integrating blockchain technologies and implementing new business models will need to be aware of the implications blockchain has on existing systems and consider how existing processes will need to adapt or be replaced by new ones. The financial implications of new technologies and business models requires consideration so it is not a barrier to adoption, for this will be a challenge at all scales from micro to meso and macro.

Policy: Policy will be central to the blockchain-based maintenance and increase of biodiversity at the city scale. It is hoped the future findings of this research will inform policymakers on the benefits the application of blockchain can bring to the table. Part of the philosophy of an underpinning blockchain system is the rationale that policy and regulatory frameworks are partly shifted on-chain (i.e., to the computing protocol). As such, diligent care needs to be taken on forming the policies that drive the design of the blockchain and related smart contracts. A key question that would be interesting in the formulation of such policy is the governance rights that might be attributed to nature itself. Other researchers have raised the question as well. However, the use of a particular question as a case study might help us make inroads into forming such policy, even outside the blockchain frame. Beyond policy, of course, there is also the issue on how nature might be able to exercise such rights, so it is not only a question of agency but also of practical execution of such agency.

Society: Biodiversity has implications for humans that make up society. Unfortunately, the activities that we as humans have become accustomed to, to live comfortably and enjoy day-to-day living have negative effects on the environment and result in biodiversity loss. For blockchain as an initiative to incentivise and govern biodiversity maintenance and increase requires: a) acceptance of blockchain as a facilitator, and b) comprehensive participation of citizens in the co-design process to achieve this. This requires establishing what the role of citizens will be in maintaining and increasing biodiversity at the city scale. It is a departure from today's limited participation in municipal governance so will require adjustments from both citizens and municipalities.

Initial gaps in the literature

While there is literature on blockchain-based incentives and blockchain-based governance, the biggest gap lies in a general lack of research on biodiversity at the city scale facilitated by blockchain. While the initial motivation for this research lies with the increase of biodiversity without a specified utility, there might also be another unexpected shift that can directly benefit the economy and production of the built environment. This is the fostering of an environment where bio-based materials and bio-based processes for construction are being adopted within frameworks of the circular economy. An increase in biodiversity could be coupled with an increase of more readily available bio-based construction materials, directly contributing to the circular economy of the built environment, not only in the financial sense but also in the ecological sense (i.e., in a systemic change).

Challenges and constraints

While the challenges and constraints of blockchain-based governance and incentive mechanisms to maintain and increase biodiversity at the city scale are scattered throughout this paper, this section brings them together to concretise the focus of the wider research project going forward.

- Acceptance of blockchain as a viable solution is required, acknowledging that selling 'blockchain' as the underpinning technology is not the way to do this. The benefits brought about by building on blockchain platforms will offer new propositions, business models, and revenue streams that can be reaped not by marketing a blockchain-based application, but by promoting the benefits of the 'killer app' that will be the tipping point for blockchain applications.
- A change in current business models and levels of responsibility at government and municipal levels is required to embrace co-creation with citizens as a genuine and integrated piece of the solutions puzzle.
- A critical mass of participants is needed to make utilising blockchain technologies worthwhile where decentralisation of power (in this instance away from national and local governments) allows true democracy and a focus on the long-term impacts of the built environment.
- There is limited understanding of the importance of biodiversity across the general public. If citizens are asked to co-create new and rethink existing built environments, they must be aware of the importance of biodiversity, what it looks like and the different ways in which it can be achieved. This requires educating the public (i.e., through participatory action research) to highlight the benefits of bringing back biodiversity to cities such as improved physical and mental health, lower levels of pollution, better use of resources, and more sustainable urban environments.

Proposed outcomes

This research will create a framework to facilitate increasing biodiversity at the city scale. It will examine case studies from regions around the world that have high levels of biodiversity (e.g., Singapore, Brazil) either through protecting the natural environment or through rewilding back into cities. A set of detailed use cases for governance and incentive mechanisms of blockchain in construction will support the framework and an IT prototype will be created to demonstrate how blockchain technologies can enable this.

The outputs of this research will contribute knowledge to a new area of interdisciplinary research at the intersection of co-design, blockchain, biodiversity and design and planning in construction with a humanitarian focus.

Conclusions

Biodiversity is crucial to the survival of humanity and is central to the ecological systems in which the built environment exists. However, the built environment has evolved in recent decades to one that exploits natural resources and damages the environment without sufficient regard to the long-term effects. The aim of this paper was to consider the potential of emerging blockchain technologies to mitigate the negative impacts of the built environment at the urban complex where most of the damage is seen. As such our motivation lies in improving the performance of the built environment in issues of biodiversity, via the exploration of blockchain incentives and governance. We view biodiversity increase in the built environment as a public good, hence it has a very good alignment with the application of blockchain technologies as a socio-technical system for facilitation of governance and incentivisation.

Existing research has been conducted on blockchain as a facilitator of governance and incentive mechanisms showing its potential in changing user behaviours and organisational processes to create efficiencies and promote real democracy. This paper presents early plans for a research project that will create a framework and an IT prototype to utilise blockchain in maintaining and increasing biodiversity at the city scale. It will do this by conducing interdisciplinary research at the intersection of co-design, blockchain, biodiversity, and design and planning of the built environment.

References

Abbass, K. *et al.* (2022) 'A review of the global climate change impacts, adaptation, and sustainable mitigation measures', *Environmental Science and Pollution Research*, 29(28), pp. 42539–42559. Available at: https://doi.org/10.1007/s11356-022-19718-6.

Axinte, L.F. *et al.* (2019) 'Regenerative city-regions: a new conceptual framework', *Regional Studies, Regional Science*, 6(1), pp. 117–129. Available at: https://doi.org/10.1080/21681376.2019.1584542.

Bose, R.P.J.C. *et al.* (2019) 'A Decentralized Application for Fostering Biodiversity: Opportunities and Challenges', in *2019 IEEE/ACM 41st International Conference on Software Engineering: Companion Proceedings (ICSE-Companion)*, pp. 284–285. Available at: https://doi.org/10.1109/ICSE-Companion.2019.00116.

Brekke, J.K. (2021) 'Hacker-engineers and Their Economies: The Political Economy of Decentralised Networks and "Cryptoeconomics"', *New Political Economy*, 26(4), pp. 646–659. Available at: https://doi.org/10.1080/13563467.2020.1806223.

Brekke, J.K. and Alsindi, W.Z. (2021) 'Cryptoeconomics', *Internet Policy Review*, 10(2). Available at: https://policyreview.info/glossary/cryptoeconomics (Accessed: 18 November 2023).

Bulkeley, H. (2022) 'Climate changed urban futures: environmental politics in the Anthropocene city', in *Trajectories in Environmental Politics*. Routledge.

Camrass, K. (2022) 'Urban regenerative thinking and practice: a systematic literature review', *Building Research & Information*, 50(3), pp. 339–350. Available at: https://doi.org/10.1080/09613218.2021.1922266.

Convention on Biological Diversity (2009) *Sustaining Life on Earth*. Secretariat of the Convention on Biological Diversity. Available at: https://www.cbd.int/convention/guide/ (Accessed: 5 November 2023).

Crypto Altruism (2023) *Natural Capital Backed Assets: Incentivizing conservation and regeneration with asset tokenization*, *Crypto Altruism*. Available at: https://www.cryptoaltruism.org/blog/natural-capital-backed-assets-incentivizing-conservation-and-regeneration-with-asset-tokenization (Accessed: 21 November 2023).

Czura, C. (no date) *Blockchain for Biodiversity: An overview of various blockchain applications to help increase biodiversity funding*. UNEP.

Davidova, M. and McMeel, D. (2020) 'Codesigning with blockchain for synergetic landscapes: The cocreation of blockchain circular economy through systemic design', in D. Holzer et al. (eds). *CAADRIA 2020: Re:Anthropocene - Design in the Age of Humans*, Bangkok: Association for Computer Aided Architectural Design in Asia, pp. 333–342. Available at: https://orca.cardiff.ac.uk/id/eprint/134033/ (Accessed: 18 November 2023).

Desmomd, L. and Salama, M. (2023) 'Integrating Blockchain & Emerging Technologies for Sustainability Assurance in the Built Environment', in *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*. *2023 IEEE International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings)*, pp. 1–5. Available at: https://doi.org/10.1109/AIBThings58340.2023.10292472.

Dounas, T., Lombardi, D. and Jabi, W. (2022) 'Collective Digital Factories for Buildings: Stigmergic Collaboration Through Cryptoeconomics', in T. Dounas and D. Lombardi (eds) *Blockchain for Construction*. Singapore: Springer Nature (Blockchain Technologies), pp. 207–228. Available at: https://doi.org/10.1007/978-981-19-3759-0_11.

Duffy, J.E. (2009) 'Why biodiversity is important to the functioning of real-world ecosystems', *Frontiers in Ecology and the Environment*, 7(8), pp. 437–444. Available at: https://doi.org/10.1890/070195.

Ekblaw, A. *et al.* (2016) 'Bitcoin and the Myth of Decentralization: Socio-technical Proposals for

Restoring Network Integrity', in. *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, Augsburg, Germany, pp. 18–23. Available at: https://doi.org/10.1109/FAS-W.2016.18.

Elmqvist, T. *et al.* (2021) 'Urbanization in and for the Anthropocene', *npj Urban Sustainability*, 1(1), pp. 1–6. Available at: https://doi.org/10.1038/s42949-021-00018-w.

França, A.S.L. *et al.* (2020) 'Proposing the use of blockchain to improve the solid waste management in small municipalities', *Journal of Cleaner Production*, 244, p. 118529. Available at: https://doi.org/10.1016/j.jclepro.2019.118529.

Greener Tokens (2023) *Greener Tokens*. Available at: https://www.greenertokens.com/ (Accessed: 20 November 2023).

He, Z. and Turner, P. (2022) 'Blockchain Applications in Forestry: A Systematic Literature Review', *Applied Sciences*, 12(8), p. 3723. Available at: https://doi.org/10.3390/app12083723.

Hijazi, A.A. *et al.* (2022) 'A data model for integrating BIM and blockchain to enable a single source of truth for the construction supply chain data delivery', *Engineering, Construction and Architectural Management* [Preprint]. Available at: https://doi.org/10.1108/ECAM-03-2022-0209.

Hunhevicz, J. *et al.* (2021) 'no1s1 - a blockchain-based DAO prototype for autonomous space', in. *2021 European Conference on Computing in Construction*.

Hunhevicz, J.J. and Hall, D. (2020) 'Crypto-Economic Incentives in the Construction Industry', in *Proceedings of the ARCOM Doctoral Workshop: Exploring the mutual role of BIM, Blockchain and IoT in changing the design, construction and operation of built assets*, pp. 1–2. Available at: https://doi.org/10.3929/ethz-b-000420837.

Hunhevicz, J.J., Motie, M. and Hall, D.M. (2021) *Digital Building Twins and Blockchain for Performance-Based (Smart) Contracts*. Available at: http://arxiv.org/abs/2105.05192.

IDB (2023) *IDB | IDB Lab Announces the Results of its Digital Tokens for Biodiversity Challenge*. Available at: https://www.iadb.org/en/news/idb-lab-announces-results-its-digital-tokens-biodiversity-challenge (Accessed: 20 November 2023).

Joachain, H. and Klopfert, F. (2011) 'Emerging trend of complementary currencies systems as policy instrument for environmental purposes: changes ahead?', *Working Papers CEB* [Preprint]. Available at: https://ideas.repec.org//p/sol/wpaper/2013-101209.html (Accessed: 19 November 2023).

Kafi, A. *et al.* (2023) 'Meta-analysis of food supply chain: pre, during and post COVID-19 pandemic', *Agriculture & Food Security*, 12(1), p. 27. Available at: https://doi.org/10.1186/s40066-023-00425-5.

Kahya, A. *et al.* (2021) 'Blockchain-enabled Personalized Incentives for Sustainable Behavior in Smart Cities', in *2021 International Conference on Computer Communications and Networks (ICCCN)*. *2021 International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–6. Available at: https://doi.org/10.1109/ICCCN52240.2021.9522340.

Kessler, S. (2022) *The Ethereum Merge Is Done, Opening a New Era for the Second-Biggest Blockchain*. Available at: https://www.coindesk.com/tech/2022/09/15/the-ethereum-merge-is-done-did-it-work/ (Accessed: 28 October 2022).

Kirk, H. *et al.* (2021) 'Building biodiversity into the urban fabric: A case study in applying Biodiversity Sensitive Urban Design (BSUD)', *Urban Forestry & Urban Greening*, 62, p. 127176. Available at: https://doi.org/10.1016/j.ufug.2021.127176.

Lashkari, B. and Musilek, P. (2021) 'A Comprehensive Review of Blockchain Consensus Mechanisms', *IEEE Access*, 9, pp. 43620–43652. Available at: https://doi.org/10.1109/ACCESS.2021.3065880.

Li, J. (2023) *A socio-technical framework to guide implementation and value realisation of distributed ledger technologies (DLT) in the construction sector*. Doctoral Thesis. Northumbria University. Available at: https://nrl.northumbria.ac.uk/id/eprint/51601/.

Li, J. and Kassem, M. (2021) 'Applications of distributed ledger technology (DLT) and Blockchain-enabled smart contracts in construction', *Automation in Construction*, 132, p. 103955. Available at: https://doi.org/10.1016/j.autcon.2021.103955.

Marsal-Llacuna, M.-L. (2018) 'Future living framework: Is blockchain the next enabling network?', *Technological Forecasting and Social Change*, 128, pp. 226–234. Available at: https://doi.org/10.1016/j.techfore.2017.12.005.

Mathews, M., Robles, D. and Bowe, B. (2017) 'BIM+Blockchain: A solution to the trust problem in collaboration?', in *CITA BIM Gathering 2017*. Dublin, Ireland, pp. 1–11. Available at: https://arrow.tudublin.ie/bescharcon.

McMeel, D. and Sims, A. (2020) 'Blockchain: a new building block for the built environment?', in *Imaginable Futures: Design Thinking, and the Scientific Method. 54th International Conference of the Architectural Science Association 2020*. Aukland, New Zealand: The Architectural Science Association, pp. 211–291. Available at: https://anzasca.net/wp-content/uploads/2021/03/22-Blockchain-a-new-building-block-for-the-built-environment.pdf.

Medaglia, R. and Damsgaard, J. (2020) 'Blockchain and the United Nations Sustainable Development Goals:

Towards an Agenda for IS Research', in. *Twenty-Fourth Pacific Asia Conference on Information Systems*, Dubai, UAE. Available at: https://blockchainacademy.dk/wp-content/uploads/2021/01/PACIS2020_paper_613-4.pdf.

Miller, D. (2020) *Solved: How the World's Great Cities Are Fixing the Climate Crisis*. University of Toronto Press. Available at: https://doi.org/10.3138/9781487534905.

Oguntona, O.A. and Aigbavboa, C.O. (2019) 'Benefits of Biomimicry Adoption and Implementation in the Construction Industry', in J. Charytonowicz and C. Falcão (eds) *Advances in Human Factors, Sustainable Urban Planning and Infrastructure*. Cham: Springer International Publishing (Advances in Intelligent Systems and Computing), pp. 506–514. Available at: https://doi.org/10.1007/978-3-319-94199-8_49.

O'Reilly, A. and Mathews, M. (2019) 'Incentivising Multidisciplinary Teams with new methods of Procurement using BIM + Blockchain', in *CITA BIM Gathering*. Gallway, Ireland. Available at: https://doi.org/10.21427/14aq-jn02.

Rockström, J. *et al.* (2009) 'A safe operating space for humanity', *Nature*, 461(7263), pp. 472–475. Available at: https://doi.org/10.1038/461472a.

Scheel, C., Aguiñaga, E. and Bello, B. (2020) 'Decoupling Economic Development from the Consumption of Finite Resources Using Circular Economy. A Model for Developing Countries', *Sustainability*, 12(4), p. 1291. Available at: https://doi.org/10.3390/su12041291.

Schlosser, M. (2019) 'Agency', *Stanford Encyclopedia of Philosophy*, Winter 2019 Edition. Available at: https://plato.stanford.edu/entries/agency/?trk=public_post_comment-text (Accessed: 7 November 2023).

Selvanesan, H. and Satanarachchi, N. (2023) 'Potential for synergetic integration of Building Information Modelling, Blockchain and Supply Chain Management in construction industry', *Journal of Information Technology in Construction*, 28, pp. 662–691. Available at: https://doi.org/10.36680/j.itcon.2023.035.

Shin, D. *et al.* (2022) 'How does the blockchain find its way in the UAE The blockchain as a sociotechnical system'. Available at: https://doi.org/10.1504/IJTM.2022.124616.

The Treeographer (2018) 'The Tree That Owns Itself – Can a Tree Have Rights?', *The Treeographer*, 22 March. Available at: https://thetreeographer.com/2018/03/22/the-tree-that-owns-itself/ (Accessed: 7 November 2023).

Uchida, K. *et al.* (2021) 'Urban Biodiversity and the Importance of Scale', *Trends in Ecology & Evolution*, 36(2), pp. 123–131. Available at: https://doi.org/10.1016/j.tree.2020.10.011.

UN (2023) *The Sustainable Development Goals Report 2023: Special Edition*. Available at: https://unstats.un.org/sdgs/report/2023/ (Accessed: 7 November 2023).

UN (no date) *Goal 15 | Department of Economic and Social Affairs*. Available at: https://sdgs.un.org/goals/goal15#overview (Accessed: 7 November 2023).

Voshmgir, S. (2021) *Biodiversity Tokens: Rebalance Earth*, *Token Kitchen*. Available at: https://token.kitchen/biodiversity-tokens-rebalance-earth (Accessed: 21 November 2023).

Wang, H. *et al.* (2023) 'Blockchain for regenerative built environment governance', *Journal of Physics: Conference Series*, 2600(18), p. 182001. Available at: https://doi.org/10.1088/1742-6596/2600/18/182001.

WEF (2020) *Global Risk Report 2020*. Available at: https://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf (Accessed: 5 November 2023).

Wieser, A.A. *et al.* (2019) 'Implementation of Sustainable Development Goals in construction industry - a systemic consideration of synergies and trade-offs', in *IOP Conference Series: Earth and Environmental Science*. Graz, Austria: IOP Publishing, p. 012177. Available at: https://doi.org/10.1088/1755-1315/323/1/012177.

Woodall, R. and Crowhurst, D. (2003) *Biodiversity indicators for construction projects*. W005 RP662. CIRIA. Available at: http://216.92.66.137/pdf/w005.pdf.

Ye, X., Zeng, N. and König, M. (2022) 'Systematic literature review on smart contracts in the construction industry: Potentials, benefits, and challenges', *Frontiers of Engineering Management*, 9(2), pp. 196–213. Available at: https://doi.org/10.1007/s42524-022-0188-2.

Zhao, R., Wang, J. and Xue, F. (2023) 'A blockchain-based token economic model for incentivizing ESG in the construction industry', in. *EC3 Conference 2023*, European Council on Computing in Construction (Computing in Construction), pp. 0–0. Available at: https://doi.org/10.35490/EC3.2023.235.

CONTRACT THEORY-BASED BLOCKCHAIN REWARD AND PENALTY MECHANISM FOR CONSTRUCTION TENDERING AMONG MULTIPLE STAKEHOLDERS

Zijun Zhan[1], Yaxian Dong[2], Daniel Mawunyo Doe[1], Yuqing Hu[2], Shuai Li[3], Shaohua Cao[4], and Zhu Han[1]
[1]University of Houston, Houston, USA
[2]The Pennsylvania State University, University Park, USA
[3]University of Tennessee, Knoxville, USA
[4]China University of Petroleum (East China), Qingdao, China.

Abstract

With the prosperity of the Web 3.0 era, the construction sector is turning to blockchain-based tendering to address stakeholder conflicts. However, bid rigging among suppliers and information asymmetry between suppliers and the general contractor (GC) potentially undermine the GC utility. To this end, we introduce contract theory into the blockchain-based tendering framework to augment the utility of GC. Simulation results indicate a potential 20.7% profit increase for the GC, due to fostering competition among suppliers, such as material delivery time, compared to traditional blockchain-enabled bidding processes.

Introduction

Construction projects typically require the procurement of materials, a process commonly facilitated via tendering. To select competent and qualified suppliers, the lowest price (Enshassi and Modough, 2012) and comprehensive bid evaluation (van der Meer et al., 2022) methods are primarily employed for their bid evaluation. Meanwhile, the mainstreaming tendering procedures are manual and e-tendering, respectively. However, several drawbacks, such as extremely time-consuming (Santoso and Bourpanus, 2019) and heavily reliant on the third party (Chen et al., 2016), remain in current tendering procedures, thereby posing motivation to augment the accuracy, automation, transparency, security, fairness, and efficiency in terms of the tendering procedures.

To this end, scholars have become increasingly interested in developing efficient and automatic tendering systems (Dong et al., 2023). In addition, in light of the tendering process being mainly used by governments and companies as a dominant procurement method (Mali et al., 2020), security issues are imperative to be resolved. Notably, given the legal and security issues and lack of transparency in the e-tendering process, the authors in (Torkanfar et al., 2023) developed a distributed e-tendering system via the integration of blockchain, public key infrastructure (PKI), and interplanetary file system (IPFS). Meanwhile, with the prosperity of the Web 3.0 era (Zhan et al., 2023), some scholars in the construction domain have attempted to implement blockchain technology into the tendering system to achieve a secure and fair tendering process (Zhang et al., 2023; Yutia and Rahardjo, 2019). In the recent past, a blockchain-based tendering system for construction projects has been proposed, which employs decentralized smart contracts to execute the tendering process, a decentralized storage system to exchange and store relevant off-chain documents, and a decentralized application to enhance interaction between the GC and suppliers (Ahmadisheykhsarmast et al., 2023).

Nevertheless, certain shortcomings persist within the blockchain-based tendering framework. In this framework, the selection of competent suppliers heavily relies on their submitted Request for Proposals (RFPs). In this way, the bidding initiative is not in the hands of GC, thus causing severe consequences, such as prevalent misconduct of bid rigging. In addition, suppliers' confidential information may also compromise the utility[1] of GC. For instance, in the United States, most participants reported experiencing or anticipating delays in the delivery of materials (Alsharef et al., 2021). Apart from this, some truck companies prioritize proximity orders and are reluctant to transport across state lines to avoid the risk of material delivery delays due to policy changes (Ren et al., 2023). Therefore, when facing concurrent material delivery orders, the priority level for delivering GC's material that GC is unknown. Since construction work requires the support of these upstream manufacturing plants and trucking companies in the supply chain, such information asymmetry issues between GC and suppliers might lead to delayed delivery and financial losses. Although penalty clauses are deployed when designing contracts to avoid material delivery delays, such as mandatory fines per day (Bergantiños and Lorenzo, 2019), the GC remains unable to control the risk of delayed material delivery during the supplier screening phase. Therefore, we aim to enhance existing blockchain-based tendering by considering two primary issues, ***i) GC does not completely possess the bidding initiative***, and ***ii) the information asymmetry between the GC and suppliers***.

Methodology

Regarding aforesaid issues, Nobel-winning contract theory (Li et al., 2024) excels in resolving information asymmetry and can transfer the initiative from suppliers to GC ingeniously. To this end, we propose a novel tendering framework that leverages blockchain technology and contract theory together, thereby ensuring a fair and transparent tendering process and optimizing the GC's utility. The proposed framework comprises several key Steps, as depicted in Figure 1. First, the GC will initiate a prequal-

[1]It is a measure of the benefit or value that an individual places on a particular choice.

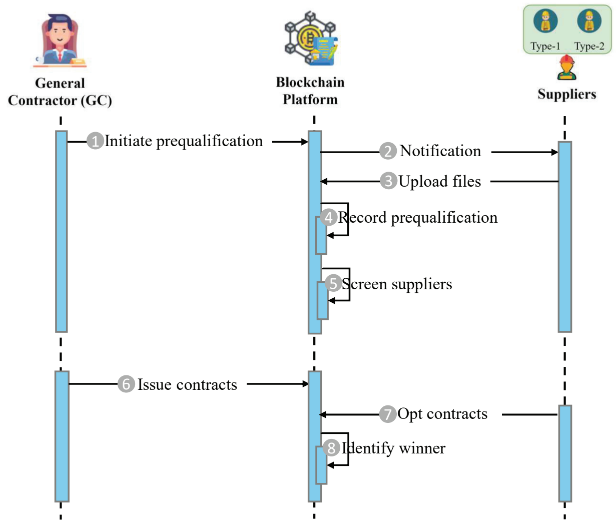


Figure 1: Sequence diagram of blockchain-based and contract theory-enabled tendering.

ification process to screen honest and qualified suppliers for participation in the tendering process. Suppliers will then upload their qualification documents as required by the GC for evaluation. Next, compiled smart contracts will be employed to screen suppliers according to predefined conditions. After that, a set of contract bundles will be designed and issued by the GC based on the contract theory for suppliers. Notably, the contract bundles will be selected by suppliers based on their actual confidential information (Bolton and Dewatripont, 2004). Subsequently, the smart contract will automatically identify the winning supplier based on previously defined metrics. Finally, the GC will sign with the winning supplier. In summary, our main contributions are summarized as follows:

1. To address the bid-rigging issue, we propose a blockchain and contract theory-based tendering framework. This framework enables the GC to take ownership of the tendering process while ensuring a fair and efficient tendering process for suppliers.
2. Under information asymmetry, we systematically model the tendering process of a blockchain-based construction project and formulate an optimization problem with the supplier's delivery time as the optimization variable. To address this problem, we employ contract theory to design a set of contract bundles that compel competition among suppliers under their confidential information.
3. The simulation and related numerical analysis results of the contract theory-based tendering model are presented. We compare the impact of different key parameters on the results and demonstrate the feasibility and effectiveness of the proposed model.

Blockchain-based contract theory procedure

To illustrate the blockchain and contract theory-based tendering framework, especially the connection between the blockchain and contract theory, we will delve into the details of our proposed tendering framework.

Algorithm 1: Blockchain-based contract theory

Input: timestamp, supplierAddress
Output: winning supplier, S_w
1 **if** *Invoke* ***DataCollectorContract*** **then**
    `// Obtain historical data`
2   data ⇐requestHistoricalData(timestamp, supplierAddress);
    `// Provide obtained data to GC`
3     $\mathscr{D}$ ⇐HistoricalDataCollected(timestamp, data);
4 **end**
5 As per $\mathscr{D}$ to partition suppliers into different types, $\theta = \{\theta_H, \theta_L\}$;
`// Obtain contract bundles`
6 **T** ⇐ContractTheory(θ);
7 **if** *Invoke* ***SupplierSelectionContract*** **then**
    `// Use contract bundles in smart contract`
8   ContractBundle ⇐**T**;
9   Shortlist qualified suppliers;
10   Record selection results of suppliers;
11   Identify the winning supplier, S_w;
12 **end**

First of all, our proposed tendering framework, as depicted in Figure 1, which mechanism is analogous to the blockchain-based tendering framework proposed in (Ahmadisheykhsarmast et al., 2023). The critical difference is that we have integrated the contract theory into the blockchain-based tendering system, Steps 5-8 in Figure 1, and the detailed procedure is illustrated in Algorithm 1. Concretely, GC will first utilize blockchain to collect trustworthy data to evaluate the potential delivery time that suppliers can do (Lines 1-4 in Algorithm 1). Subsequently, as per evaluation results, GC will design contract bundles with the utilization of contract theory (Lines 5-6 in Algorithm 1). Then, the GC will code smart contracts as per calculated contract bundles (Lines 7-12 in Algorithm 1). Eventually, the GC will deploy the smart contract on the blockchain, thus enabling an autonomous and transparent tendering process, i.e., award the material delivery order to the supplier who opts for the strictest contract bundle, thereby effectively reducing the risk of delayed delivery.
It is worth noting that we utilize immutability and transparency two crucial properties of blockchain when designing the tendering framework. As for immutability, GC will trace the historical delivery data of the supplier by using smart contract ***DataCollectorContract***. Regarding transparency, our proposed framework will enable GC and suppliers to oversee the whole tendering process, i.e., the running process of smart contract ***SupplierSelectionContract***, including contract bundle selection and winning supplier identification. In addition, in light of our proposed tendering framework is enhanced as per (Ahmadisheykhsarmast

et al., 2023), thus it can be deployed on the Ethereum [2] as well.

In the remainder of this paper, we will present the mathematical model of blockchain and contract theory-enabled tendering process and numerical simulation comparison results to demonstrate the effectiveness of our proposed framework.

Contract theory model

Contract theory background

We consider K types of construction materials, collectively denoted by $\mathscr{M} = \{M_1, \cdots, M_k, \cdots, M_K\}$, to be tendered in the construction project. The execution sequence and interdependencies of various materials in the construction project implementation plan give rise to varying degrees of importance for different materials. We use δ_k to represent the importance of material M_k. Additionally, each material has a specific delivery time requirement to satisfy construction scheduling needs. While timely delivery of M_k is preferable, advance or delayed delivery may occur due to unforeseen circumstances. For ease of modeling, we assume that regardless of whether M_k is delivered early or late, the difference between the on-time delivery and both the earliest and the latest delivery times are identical.

In this paper, since the tendering process is similar for each material, we only consider the tendering process for M_k. As for tendering, the GC will solicit several suppliers to bid publicly for M_k. Without loss of generality from a mathematical perspective, we assume that GC calls only 2 suppliers to bid for M_k tendering, denoted as $\mathscr{S}_k = \{S_k^1, S_k^2\}$. Akin to (Doe et al., 2023), we can easily extend to a practical scenario where multiple suppliers together tender for M_k.

For simplicity, we will use S_1 and S_2 to represent two suppliers in the remainder of this paper. Each supplier can be denoted as S_n[3] and has a unique identity type, θ_n, corresponding to the material delivery time, which may influence by the count of concurrent delivery and delivery priority level for GC's material that suppliers possess. Due to privacy concerns, the exact identity type of the suppliers is unknown to the GC, resulting in information asymmetry. Thus, the GC can only stipulate the delivery time of M_k based on the public information submitted by suppliers, which we denote as T_k^{max}. However, setting the delivery time as T_k^{max} may not maximize the GC's utility, and arbitrary delivery time settings can result in supplier dissatisfaction. In a nutshell, under information asymmetry, it becomes challenging to maximize the utility of GC. Therefore, we have integrated contract theory with the blockchain-based tendering scheme proposed in (Ahmadisheykhsarmast et al., 2023).

[2]https://ethereum.org/whitepaper

[3]$n \in \{1,2\}$

Supplier model

For clarity exposition, we assume S_n with only two possible identity types, θ_H and θ_L. We define $\theta_H > \theta_L$[4], and each identity type corresponds to a different extent of advance delivery requirement with respect to T_k^{max}, represented by T_k^H and T_k^L, in which $T_k^H > T_k^L$. For the sake of modeling, we will drop the subscript k and use T_i to indicate advance delivery requirement, i.e., $i \in \{H, L\}$.
In addition, we assume that the tendering will use the comprehensive tendering method, a commonplace bidding model in the construction industry (van der Meer et al., 2022). As per our initial assumption, the GC follows Steps 1-4 of the framework depicted in Figure 1 to comprehensively evaluate and shortlist several suppliers. Subsequently, the GC implements Steps 5-8 to identify the ideal supplier using a multilateral contract theory-based (Bolton and Dewatripont, 2004)[5] tendering scheme. We assume that the average bid of selected suppliers in Figure 1 is $\bar{R}$. The utility function of S_n can be defined as follows:

$$U_{S_n} = \omega(\theta, \bar{R}) - \phi(T_i) - p(\delta_k) + r(\bar{R}), \qquad (1)$$

where $\omega(\cdot)$ represents the profit that S_n can earn for this project. $\phi(\cdot)$ denotes the cost function for the supplier to transport the construction materials. $p(\cdot)$ represents the penalty imposed on the supplier if they fail to deliver within the delivery time specified in the signed contract. Finally, $r(\cdot)$ represents the revenue generated by the blockchain platform for S_n.
As this paper adopts the comprehensive tendering model, the GC only needs to pay the successful bidder with $\bar{R}$. The difference in profit among suppliers is primarily determined by their identity type θ. Specifically, the better the delivery, the more intangible wealth the supplier will gain, such as a relatively high probability of getting the next order. Therefore, we define the profit function that S_n can obtain from the project as:

$$\omega(\theta_i, \bar{R}) = \theta_i \bar{R}. \qquad (2)$$

In light of the stricter delivery time requirement, i.e., T_i, the supplier will possess a lower concurrent delivery order or necessitate the supplier to set GC as a high priority, potentially impacting its profit. Therefore, we define the cost of S_n is related to T_i:

$$\phi(T_i) = \alpha(\bar{T} + T_i), \qquad (3)$$

where α is the coefficient of suppliers' cost, and $\bar{T}$ is the initial cost when suppliers' delivery time requirement is T_k^{max}.
The penalty function for construction material not delivered before the delivery time requirement imposed on S_n is related to the importance of the materials and the punctuality of delivery, which is defined as:

[4]Due to the θ_H type supplier with fewer concurrent delivery orders and will set GC's delivery order as the high priority.

[5]The multilateral contract theory is akin to the bidding process, where the winner takes all.

$$p(\delta_k) = P\delta_k \times \tau, \tag{4}$$

where $P\delta_k$ represents the penalty charged for the breach, in which P is a constant coefficient, and τ denotes the probability of a supplier breaching the contract.

The payoff function of operating a blockchain platform includes two main components: the revenue generated by the blockchain platform and the expenses required to operate it (e.g., human resources and hardware equipment). According to (Griffiths et al., 2017), the benefits derived from using a blockchain platform are typically associated with the project's bid $\bar{R}$. As a result, we define the blockchain payoff function as:

$$r(\bar{R}) = \Delta\bar{R} - C, \tag{5}$$

where $\Delta\bar{R}$ represents the reduced transaction cost generated by the blockchain, and C denotes the cost of operating the blockchain.

Following the above analysis, the utility function of S_n can be summarized as:

$$U_{S_n} = \theta_i\bar{R} - \alpha(\bar{T} + T_i) - \varphi + \rho, \tag{6}$$

For clarity, we use φ to denote $p(\delta_k)$, and ρ to represent $r(\bar{R})$.

General Contractor Model

The utility function of the GC is defined as follows:

$$U_{GC} = \pi(T_i) - \xi(\bar{R}) + p(\delta_k) + r(\bar{R}), \tag{7}$$

where $\pi(\cdot)$ represents the revenue function of the GC, and $\xi(\cdot)$ denotes the project funds that the GC pays to the supplier.

The revenue function of the GC is defined by

$$\pi(T_i) = \alpha' \ln(D + \varepsilon T_i), \tag{8}$$

where α' is a coefficient and D represents the revenue earned when the material delivery time is T_k^{max}. Also, GC can proactively control the delivery time of M_k earlier than the default delivery time setting T_k^{max}, so as to effectively reduce the risks of delayed delivery, thus we utilize εT_i to denote the additional revenue earned by GC.

The remuneration function paid by the GC to the supplier is defined as follows:

$$\xi(\bar{R}) = \beta\bar{R}, \tag{9}$$

where β is the project remuneration coefficient due to market fluctuations. In this article, assuming that the supply and demand in the market are in equilibrium, we set $\beta = 1$. In summary, the utility function of the GC can be expressed as:

$$U_{GC} = \alpha' \ln(D + \varepsilon T_i) - \bar{R} + \varphi + \rho. \tag{10}$$

Contract definition and problem formulation

Definitions of contract theory

Analogous to (Li et al., 2024), we assume that the GC faces competition from two risk-neutral suppliers[6] for an indivisible material supply order, without loss of generality and for the sake of simplicity. The cases with multiple risk-neutral suppliers can be studied similarly. As discussed earlier, each supplier S_n may have two identity type for supplying M_k, given by:

$$\theta = \begin{cases} \theta_H, & \lambda_n, \\ \theta_L, & 1 - \lambda_n, \end{cases} \tag{11}$$

where λ_n denotes the probability that the GC evaluates the identity type of S_n as θ_H, which can be determined by using empirical-based methods or data mining techniques. Consistent with (Bolton and Dewatripont, 2004), we assume that $\lambda_n = \lambda$, and all suppliers are identical ex-ante. Without loss of generality, we can restrict the GC to symmetric auctions where suppliers are treated equally.

During the tendering process, suppliers compete with each other to secure a contract and maximize their profits. In such a competitive scenario, suppliers take into account not only the contract bundles they choose but also the contract bundles that their competitors may opt for. To reflect this scenario, we present the types of contracts that the GC will offer during the tendering process in the following Definition 1.

Definition 1 (*Competed Delivery Time Requirement*)**.** *Suppliers' delivery time requirements will depend on their respective potential identity type, denoted by* $\mathbf{T} = \{T_{HH}, T_{HL}, T_{LH}, T_{LL}\}$.

To illustrate, T_{HH} denotes the contract variable designed by GC for S_n when the identity types of both suppliers are θ_H. Nevertheless, T_{HH} is a temporary variable set by GC during contract design, and the actual contracts issued are based on the potential identity type of each supplier. Consequently, in this study, only two types of contracts are issued by the GC, namely, T_H and T_L.

During the tendering process, suppliers have the freedom to choose either the same or different contracts. However, it is worth noting that their bid success rate may differ depending on the contracts they opt for. To provide clarity on this matter, Definition 2 is presented:

Definition 2 (*Bidding Success Rate*)**.** *Suppliers' probability of winning the material delivery order depends on their potential identity type, denoted by* $\mathbf{x} = \{x_{HH}, x_{HL}, x_{LH}, x_{LL}\}$.

For instance, x_{HH} represents the probability that supplier S_n wins the order when the identity type of both suppliers is θ_H.

[6] The supplier will be indifferent to risk when making decisions, i.e., the supplier is rational.

As there is only one indivisible construction project, the GC can only select one successful bidder among all suppliers. As a result, it is essential to ensure that the probability of successful bidding for all suppliers, who choose different types of contracts, must be less than or equal to 1. To formalize this requirement, Definition 3 is proposed:

Definition 3 (*Feasibility*)**.** *Supplier S_n has feasibility constraints on the probability of winning the order, i.e.,* $2x_{HH} \leq 1; 2x_{LL} \leq 1; x_{HL} + x_{LH} \leq 1$.

For example, x_{HH} represents the identity type of both suppliers as θ_H. In this case, x_{HH} is the bid success rate for both suppliers, and only one supplier will ultimately succeed in the tendering process, leading to the constraint $2x_{HH} \leq 1$.

With the previous definitions, we have established a mathematical framework for tendering scenarios. However, it is important to note that a rational supplier will only participate in a tendering project if it is profitable for them. Therefore, GC needs to design contracts in a way that ensures positive utility for suppliers. This idea is represented mathematically through Definition 4.

Definition 4 (*Individual Rationality, IR*)**.** *IR denotes that the GC designs the contract in a manner that S_n agrees to it only if there are guaranteed positive benefits.*

Considering the competition among suppliers, we redefined the utility function of S_n as:

$$U_{S(\theta_i,x_i,T_i)} = \sum_{j=H,L} \lambda_j(\theta_i \bar{R} x_{ij} - \alpha(\bar{T} + T_{ij}) - \varphi + \rho), \quad (12)$$

where $\lambda_H = \lambda$, and $\lambda_L = 1 - \lambda$. The physical meaning of (12) is that since there is only one indivisible material delivery order, the probability of S_n bidding success will depend on the contract chosen by other suppliers. Concretely, if the other supplier chooses the H-type contract with probability λ_H and the L-type contract with probability λ_L, S_n will have different probabilities of bid success, as defined in Definition 2. It is important to note that the effect of $\mathbf{x}$ is only on the supplier reward function. This is because the probability of receiving a reward for a successful bid is dependent on another supplier, even if they select the contract of their type and devote their best efforts.

Therefore, we can express the *IR* constraint as:

$$U_{S(\theta_i,x_i,T_i)} \geq 0, \quad i = H, L. \quad (13)$$

While the *IR* constraint ensures supplier participation in tendering, it does not guarantee that each supplier will choose the contract specifically designed for them, which would not maximize GC's utility. To ensure that each supplier selects only contracts suitable for their identity type, we introduce Definition 5.

Definition 5 (*Incentive Compatibility, IC*)**.** *IC denotes that suppliers can only maximize their benefits if they choose the contract that matches their identity type, i.e.,*

$$U_{S(\theta_i,x_i,T_i)} > U_{S(\theta_i,x_j,T_j)}, \quad i = H, L; j \neq i. \quad (14)$$

From (14), we can intuitively see that when a supplier chooses a contract that does not match its identity type, it will result in lower revenue. Therefore, a rational supplier will choose only the contract that meets its identity type, maximizing its benefits and satisfying GC's requirements.

Problem formulation

This paper proposes that the GC maximizes its profit by formulating a series of contracts. Specifically, the optimization problem of the GC is formulated as **P1**:

$$\textbf{P1} \quad \max_{\mathbf{T}} \quad \Pi_{GC} = 2 \sum_{i=H,L} \sum_{j=H,L} \lambda_i(\lambda_j U_{GC(T_{ij})}) \quad (15a)$$

$$s.t. \quad U_{S(\theta_i,x_i,T_i)} \geq 0, \quad i = H, L, \quad (15b)$$

$$U_{S(\theta_i,x_i,T_i)} > U_{S(\theta_i,x_j,T_j)}, \quad i = H, L; j \neq i, \quad (15c)$$

$$2x_{HH} \leq 1, 2x_{LL} \leq 1, x_{HL} + x_{LH} \leq 1, \quad (15d)$$

$$0 \leq T_{ij} < T_k^{max}, \quad i = H, L; j \neq i. \quad (15e)$$

The optimization objective of **P1** is to maximize the expected benefit of the GC. Constraint (15b) is the *IR* constraint on the supplier. Constraint (15c) represents the *IC* constraint on the supplier. Constraint (15d) is the feasibility constraint in the tendering process, and finally, constraint (15e) sets a limit on the material delivery time requirements of the supplier that should be met by the GC.

Optimal contract

Due to the presence of multiple non-convex constraints, solving **P1** is computationally intractable (Zhan et al., 2023). To overcome this issue, we aim to simplify **P1** by leveraging Propositions 1, 2, and 3, the detailed proof can be referred to (Bolton and Dewatripont, 2004).

Proposition 1. $x_{HH} = x_{LL} = \frac{1}{2}, x_{HL} = 1, x_{LH} = 0.$

Proposition 2. *If IR constraint of type θ_L hold, IR constraint of type θ_H will automatically hold.*

$$U_{S(\theta_L,x_L,T_L)} \geq 0 \Rightarrow U_{S(\theta_H,x_H,T_H)} \geq 0. \quad (16)$$

Proposition 3. *If IC constraint of type θ_H is binding, then IC constraint of type θ_L will automatically hold.*

$$\begin{aligned} U_{S(\theta_H,x_H,T_H)} &= U_{S(\theta_H,x_L,T_L)} \\ &\Rightarrow U_{S(\theta_L,x_L,T_L)} > U_{S(\theta_L,x_H,T_H)}. \end{aligned} \quad (17)$$

By using Propositions 1, 2, and 3, and temporarily relaxing constraint (15e), we can transform the intractable problem **P1** into a tractable one, denoted as **P2**.

$$\textbf{P2} \quad \max_{T_{ij}} \quad \Pi_{GC} = 2 \sum_{i=H,L} \sum_{j=H,L} \lambda_i(\lambda_j U_{GC(T_{ij})}) \quad (18a)$$

$$s.t. \quad U_{S(\theta_L,x_L,T_L)} = 0, \quad (18b)$$

$$U_{S(\theta_H,x_H,T_H)} = U_{S(\theta_H,x_L,T_L)}. \quad (18c)$$

By observing the two equality constraints in **P2**, we can obtain expressions for T_{LH} and T_{HH} in terms of T_{LL} and T_{HL}, respectively. Specifically, we have

$$T_{LH} = \frac{(1-\lambda)(\frac{1}{2}\theta_L\bar{R} - \alpha T_{LL}) - \alpha\bar{T} - \varphi + \rho}{\lambda\alpha}, \quad (19)$$

$$T_{HH} = \frac{\frac{1}{2}\theta_H\bar{R} + \frac{1-\lambda}{2}\theta_L\bar{R} - (1-\lambda)\alpha T_{HL} - \alpha\bar{T} - \varphi + \rho}{\lambda\alpha}. \quad (20)$$

For convenience, we define $T_{LH} = g(T_{LL})$ and $T_{HH} = f(T_{HL})$ and substitute these expressions into the objective function of the optimization problem. Hence, the optimal contract can be obtained by solving the following optimization problem, given by (21):

$$\begin{aligned} T_{ij}^* = arg\max_{T_{ij}} 2[&\lambda_H(\lambda_H U_{GC(f(T_{HL}))} + \lambda_L U_{GC(T_{HL})}) \\ &+ \lambda_L(\lambda_H U_{GC(g(T_{LL}))} + \lambda_L U_{GC(T_{LL})})]. \end{aligned} \quad (21)$$

First, we obtain the second-order derivatives of the optimization function with respect to T_{HL} and T_{LL}, respectively. Specifically, we have

$$\frac{\partial^2\Pi}{\partial(T_{LL})^2} = -\left[p_1\frac{(\varepsilon g'(T_{LL}))^2}{(D+\varepsilon g(T_{LL}))^2} + p_2\frac{\varepsilon^2}{(D+\varepsilon T_{LL})^2}\right], \quad (22)$$

$$\frac{\partial^2\Pi}{\partial(T_{HL})^2} = -\left[p_3\frac{(\varepsilon f'(T_{HL}))^2}{(D+\varepsilon f(T_{HL}))^2} + p_1\frac{\varepsilon^2}{(D+\varepsilon T_{HL})^2}\right], \quad (23)$$

where p_1, p_2, and p_3 are defined as $(1-\lambda)\lambda$, $(1-\lambda)^2$, and λ^2, respectively.

Then, since both second-order derivatives are negative, we can determine T_{ij}^* by taking the first-order derivatives of T_{HL} and T_{LL}. Specifically, we have

$$T_{LL}^* = T_{LH}^* = \frac{\frac{1}{2}(1-\lambda)\theta_L\bar{R} - \alpha\bar{T} - \varphi + \rho}{\alpha}, \quad (24)$$

$$T_{HL}^* = T_{HH}^* = \frac{\frac{1}{2}\theta_H\bar{R}}{\alpha} + T_{LL}^*. \quad (25)$$

Finally, we should also check whether T_{ij}^* satisfies constraint (15e) in **P1**. If this constraint is not satisfied, we can use various convex optimization tools such as Gurobi or CVX in Matlab to obtain the optimal solution.

Results

Comparison methods

To provide a comprehensive evaluation of our proposed blockchain-enabled and multilateral contract-based tendering scheme, we compare it with the following benchmark methods.

1. *Multilateral Contract in Complete Information Scenario* (Bolton and Dewatripont, 2004): This method assumes that the GC possesses complete information about the supplier's identity type and, as a result, designs contracts that can extract all the profits of the supplier. As such, we employ this method as an upper bound for performance evaluation.
2. *Bilateral Contract in Incomplete Information Scenario* (Doe et al., 2023): This method assumes that multiple material shipping orders are simultaneously available for the GC, enabling them to issue exclusive shipping order contracts to each supplier.
3. *Multilateral Contract without blockchain in Incomplete Information Scenario*: This scheme differs from the one proposed in this paper only in that it excludes blockchain; all other parameters remain identical.
4. *Traditional Tendering Scheme in Incomplete Information Scenario* (Ahmadisheykhsarmast et al., 2023): This scheme leverages the conventional tendering method employed in the construction sector to identify material suppliers for the GC on the blockchain platform.

For ease of reference in subsequent discussions, we refer to our proposed scheme and the four benchmark comparison methods as *Optimal contract*, *Complete contract*, *Bilateral contract*, *WoBC contract*, and *No contract*, respectively.

Experiment parameters

To indicate the high and low identity type of S_n with respect to M_k, we set θ_h and θ_L as 0.1 and 0.02, respectively. We categorize the importance levels of M_k into five categories, namely extremely low, low, medium, high, and extremely high, and represent them using $\{1, 2, 3, 4, 5\}$. We use key parameters with reference to studies involving blockchain-based construction payment management and blockchain-based tendering scheme (Ahmadisheykhsarmast et al., 2023; Griffiths et al., 2017; Bidhive, 2018; Brown et al., 2006).

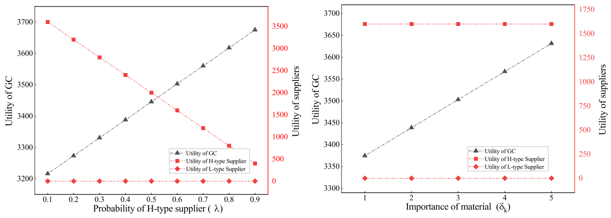
Discussion and result analysis

Contract analysis

In Figure 2, we depict the trend of utility changes for the GC, θ_L type supplier, and θ_H type supplier when varying key parameters. Notably, the utility of the θ_L type supplier remains at 0 for all four parameter variations due to the contract theory extracting the entire profit of the lowest-rated supplier, which aligns with previous research (Zhang and Han, 2017; Li et al., 2022; Doe et al., 2023).

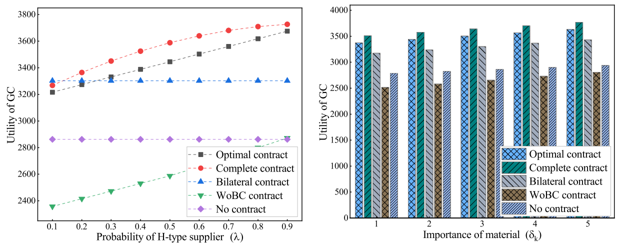
As shown in Figure 2a, when the identity types of θ_H increases, the utility of the GC also increases while that of θ_H type suppliers decreases. This result is intuitive as a higher λ value causes the GC to tailor contracts exclusively to θ_H type suppliers, extracting more profit.

Figure 2b reveals that when material importance δ_k increases, the GC's utility gradually increases. This outcome arises because heightened δ_k values necessitate a stricter delivery time requirement for materials, enabling the GC



(a) (b)

Figure 2: Contract-based tendering scheme for comparison of GC, θ_H and θ_L type suppliers' utilities as parameters vary.



(a) (b)

Figure 3: Contract-based tendering scheme for comparison with other four benchmark methods as parameters vary.

to extract more profit and thereby reduce risk. Notably, the utility of θ_H type suppliers remains constant across δ_k variations. This result stems primarily from the supplier's lead time increasing alongside δ_k. In other words, although the supplier's cost coefficient increases, the GC also obtains a more lenient delivery requirement, thereby reducing supplier delivery costs and leaving the utility of θ_H type suppliers unchanged.

Utility of GC

In Figure 3a, it is evident that the *Optimal contract*, *Complete contract*, and *WoBC contract* lead to an increasing trend for the GC's utility as λ increases. This result arises from varying the supplier's delivery time requirement as λ increases. In contrast, the *Bilateral contract* and *No contract* exhibit no change when λ increases since these methods do not affect the supplier's delivery time requirement. Specifically, in the *Bilateral contract*, the GC formulates designated contracts for suppliers with different identity types. Therefore, varying λ does not impact the supplier's delivery time requirement or the GC's utility. As for the *No contract*, it is intuitive that the supplier's delivery time requirement will not change irrespective of λ.

The reasons for the increase in GC's utility depicted in Figure 3b are similar to our analysis of Figure 2b, where an increase in δ_k leads to the extraction of more profit from the supplier by the GC. This increased profit stems from tailoring the contract to mitigate GC's own risk, resulting in an increase in the GC's utility as δ_k increase.

Based on the findings presented in Figure 3, we can determine the effectiveness ranking of the different contracts for GC's utility as follows: *Complete contract* > *Optimal contract* > *Bilateral contract* > *No contract* > *WoBC contract*. The *Complete contract* outperforms all other methods because the GC is aware of the supplier's identity type before contract formulation and can extract the maximum profit possible. Next, the *Optimal contract* generates more profit for the GC than the *Bilateral contract* by incorporating a competition mechanism during the contract formulation. Information asymmetry is the reason why the *No contract* method is inferior to the other three contract theory-enabled methods. Finally, the reason why the *WoBC contract* ranks last is that conventional tendering incurs an excessive amount of administration cost compared to the blockchain-based approach. Typically, administration costs range from 2% to 3% of the contract value (Bidhive, 2018; Brown et al., 2006), whereas blockchain cost remains fixed, which typically amounts to around \$750, including the cost of smart contract deployment and invocation, prequalification files submission, and evaluation and contract awarding, etc., for completing a tendering project on Ethereum. (Ahmadisheykhsarmast et al., 2023). Consequently, blockchain will generate more revenue as the contract value increases.

Conclusion

Considering information asymmetry and bid-rigging undermining the GC's utility, this paper proposed a blockchain and multilateral contract theory-based construction tendering framework. To promise the effectiveness of multilateral contract theory, the GC first utilized the blockchain to acquire trustworthy data. Then, with the utilization of contract theory, the GC can provide precise material delivery time via the smart contract for suppliers to opt for. The simulation results demonstrated that the proposed multilateral contract theory-based method improves the GC's profit by approximately 20.7% in comparison with the traditional bidding method, with only a 5% profit difference displayed compared to the performance-optimized upper bound complete information scenario, the ideal situation. Furthermore, the proposed framework benefits the scenarios where information asymmetry is prevalent, such as construction payment management and material transportation processes. In future work, we will take the continuity of different materials into consideration, and will not assume the tolerance for advanced and delayed delivery are identical since delayed delivery will cause more severe consequences.

Acknowledgments

This research was funded by the U.S. National Science Foundation (NSF) via Grants 2222730, 2222670, and 2222810. In addition, this work is partially supported by NSF CNS-2107216, CNS-2128368, CMMI-2222810, ECCS-2302469, US Department of Transportation, Toyota. Amazon and JST ASPIRE JPMJAP2326.

References

Ahmadisheykhsarmast, S., Senji, S. G., and Sonmez, R. (2023). Decentralized tendering of construction projects using blockchain-based smart contracts and storage systems. Automation in Construction, 151:104900.

Alsharef, A., Banerjee, S., Uddin, S. J., Albert, A., and Jaselskis, E. (2021). Early impacts of the covid-19 pandemic on the united states construction industry. International journal of environmental research and public health, 18(4):1559.

Bergantiños, G. and Lorenzo, L. (2019). How to apply penalties to avoid delays in projects. European Journal of Operational Research, 275:608–620.

Bidhive (2018). The human cost of bidding and tendering. `https://bidhive.com/the-human-cost-of-bidding-and-tendering/`.

Bolton, P. and Dewatripont, M. (2004). Contract theory. MIT press.

Brown, K., Hampson, K., and Brandon, P. (2006). Clients driving construction innovation: Moving ideas into practice. CRC for Construction Innovation.

Chen, Z., Chen, L., Huang, L., and Zhong, H. (2016). Towards secure spectrum auction: both bids and bidder locations matter: poster. In MobiHoc: Mobile and Ad Hoc Networking and Computing, New York, United States.

Doe, D., Li, J., Dusit, N., Gao, Z., Li, J., and Han, Z. (2023). Promoting the sustainability of blockchain in web 3.0 and the metaverse through diversified incentive mechanism design. IEEE Open Journal of the Computer Society, 2023:1–12.

Dong, Y., Hu, Y., Li, S., Cai, J., and Han, Z. (2023). Bim and blockchain-based automatic asset tracking in digital twins for modular construction. In Computing in Civil Engineering 2023, Corvallis, Oregon. American Society of Civil Engineers.

Enshassi, A. and Modough, Z. (2012). Case studies in awarding the lowest bid price in construction projects. IUG Journal of Natural and Engineering Studies, 20:113–137.

Griffiths, R., Lord, W., and Coggins, J. (2017). Project bank accounts: The second wave of security of payment? Journal of Financial Management of Property and Construction, 22:322–338.

Li, J., Liu, T., Niyato, D., Li, J., and Han, Z. (2022). On sidechain-assisted transaction service management for internet of things: A random contract approach. IEEE Transactions on Network Science and Engineering, 9:3437–3453.

Li, J., Niyato, D., and Han, Z. (2024). Cryptoeconomics: Economic Mechanisms behind Blockchains. Cambridge University Press.

Mali, D., Mogaveera, D., Kitawat, P., and Jawwad, M. (2020). Blockchain-based e-tendering system. In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pages 357–362. IEEE.

Ren, R., Li, H., Han, T., Tian, C., Zhang, C., Zhang, J., Proctor, R. W., Chen, Y., and Feng, Y. (2023). Vehicle crash simulations for safety: Introduction of connected and automated vehicles on the roadways. Accident Analysis & Prevention, 186:107021.

Santoso, D. S. and Bourpanus, N. (2019). Moving to e-bidding: Examining the changes in the bidding process and the bid mark-up decisions of thai contractors. Journal of Financial Management of Property and Construction, 24:2–18.

Torkanfar, N., Azar, E. R., and McCabe, B. (2023). Bidchain: A blockchain-based decentralized application for transparent and secure competitive tendering in public construction projects. Journal of Construction Engineering and Management, 149:04023050.

van der Meer, J., Hartmann, A., van der Horst, A., and Dewulf, G. (2022). Raising risk awareness in multi-criteria design decisions for integrated design and construction tenders. Construction Management and Economics, 40:296–312.

Yutia, S. N. and Rahardjo, B. (2019). Design of a blockchain-based e-tendering system: A case study in lpse. In ICISS, Tokyo, Japan.

Zhan, Z., Dong, Y., Doe, D. M., Hu, Y., Li, S., Cao, S., Li, W., and Han, Z. (2023). Mitigate gender bias in construction: Fusion of deep reinforcement learning-based contract theory and blockchain. In 2023 IEEE International Conference on Blockchain (Blockchain), Oceasion Island, China.

Zhang, X., Liu, T., Rahman, A., and Zhou, L. (2023). Blockchain applications for construction contract management: a systematic literature review. Journal of Construction Engineering and Management, 149:03122011.

Zhang, Y. and Han, Z. (2017). Contract Theory for Wireless Networks. Springer.

EXPLORING TOKENIZED PRODUCT PASSPORT FOR CIRCULAR CONSTRUCTION SUPPLY CHAINS

Brandon S. Byers[1], Jens Hunhevicz[12], Meliha Honic[1], and Catherine De Wolf[1]
[1]ETH Zurich, Zurich, Switzerland
[2]Empa, Dübendorf, Switzerland

Abstract

A token is a cryptoeconomic entity on a blockchain that can be used to digitally secure, represent, and trade assets. Existing research does not sufficiently explore the use of tokenization as digital representations of physical construction assets within the context of circular supply chains. Thus, this research explores why tokenization for circular construction may be helpful by employing a mixed methods approach via quantitative and qualitative analysis of expert surveys and a technical review of tokens. The contribution proposes scenarios of tokenization's potential for blockchain-based product tracking and product passports for AEC.

Introduction

The built environment is facing two major transformations: a circular economy transformation and a digital transformation, and both are critical paradigms that can target enhancing sustainability (Çetin et al., 2021). Digitization for the architecture, engineering, and construction (AEC) industry is known to be slow because of general resistance to changing processes and adopting new technologies. Changing our built environment from a linear to a circular model is also inhibited by resistance to changing processes, resistance to upfront added costs, and information deficits (Byers et al., 2024).

In the built environment, implementing a circular economy primarily deals with the complexity of the supply chain. This includes the procurement and specification of materials, involving numerous participants, and the certification of material quality. Circular construction supply chain (CCSC) compared to typical construction supply chains also include connecting the materials from disassembly in one life cycle to a use case for additional life cycles as well as preserving information on the materials over time (Wijewickrama et al., 2021). Additional challenges include the presumed added costs of component reuse, though it appears this depends primarily on local regulation, culture, and precedent (Mollaei et al., 2023). Other challenges for reusing materials include product liability and the producer's responsibility for the circulation of the materials at the end of use, particularly if the original information on them is unknown (Farooque et al., 2019). Therefore, research is growing on how to track construction products over time, what information to track, and where to store it. Blockchain technology is considered useful for tracking and tracing information in supply chains and construction (Wang et al., 2020; Qian and Papadonikolaki, 2020). Though there is an overlap of tracking product information with the concept of Digital Product Passports (DPPs) European Commission (2022), there is insufficient work on their application for AEC. In particular, the use of tokenization to represent circular supply chain assets with a tokenized product passport (TPP) seems a promising concept. TPPs may help ensure data continuity and availability through multiple stakeholders and life cycles. Yet, this application is nascent and there is a gap in research and practice. Therefore, this research addresses the question: *why would tokenization be useful for product passports for a circular built environment?* The paper begins with a background on relevant concepts, states the methods used, analyzes survey responses, provides a technical review of token features, proposes resulting applications of tokens, and finally discusses existing challenges and future work.

Background

To contextualize the research objective of this work, some background is provided on product passports, blockchain tokens, and existing applications in both academia and industry.

Product Passports

Recent legislation in the EU pushes for Digital Product Passports (DPPs), which will be used to inform product supply chains and environmental impact (Çetin et al., 2023; Honic et al., 2024; European Commission, 2022). Within AEC, Material Passports (MP) are similar to DPPs and describe characteristics of materials in products regarding their constituents, recyclability, and reuse potential, which have been found to support circularity (Honic et al., 2021) and reuse (Byers and De Wolf, 2023). While MPs and building logbooks apply to the built environment, DPPs are a cross-sectoral concept shaped by the regulatory framework "setting eco-design requirements for sustainable products" of the European Commission (2022). A report from the Wuppertal Institute by Jansen et al. (2022) reviews 76 corporate, policy, and research initiatives for DPPs. Though several of these initiatives do look at buildings, many are focused on batteries, textiles, machinery, and automotive products and discuss the challenges of heterogeneous industries and stakeholders. Interestingly, nine of these implemented blockchains for immutability, transparency, and decentralization along the supply chain.

Blockchain Tokenization

Blockchain technology is a type of distributed ledger technology (DLT) that is run over a decentralized set of computer nodes and ensures secure and transparent peer-to-peer transactions without intermediaries. Blockchain's potential application for construction is well-researched (Li et al., 2019), including supply chain and lifecycle applications (Scott et al., 2021).

This paper concentrates on Ethereum, which is among the most widely used blockchains, chosen for its capabilities in handling smart contracts, offering distributed computing solutions, and having a robust developer community. Tokens can be programmed through smart contracts, which are executable code stored on the blockchain that can interact with and define transaction logic. Smart contracts are used to specify token features (e.g., destroyable, ownable, metadata, etc.), supply count, and transfer mechanisms. On a blockchain, tokens provide a critical role as a medium of exchange and a way to represent digital or real-world asset ownership. Tokens have already been integrated into various digital ecosystems (e.g., gaming, decentralized markets of digital goods, and real estate fractionalization) reshaping industries through innovative business models on new forms of ownership and transactions (Sazandrishvili, 2020).

Because tokens are produced through smart contracts and are customizable, token standards are used to ensure interoperability and secure implementation through standardized building blocks in the network. In the Ethereum ecosystem, this standardization happens through an Ethereum Request for Comment (ERC). An ERC is a proposed protocol for suggesting a new standard defining rules and specifications for tokens or smart contracts to be adopted by individual developers and projects. The most common token standards are ERC-20 and ERC-721. ERC-20 establishes a fungible token that is tradable and dividable into discretized shares. ERC-721 is a non-fungible token (NFT) that is tradeable, non-divisible, and used to represent unique values. Tokens often consist of links within the metadata to the storage of data too large to be stored on-chain as shown in Figure 1. A token, often constituted as a JSON-LD entity, combines principles of the semantic web with interoperability and is readable by both humans and machines.

The general benefits of tokenization are greater liquidity, global access, reduced intermediaries, faster transactions, increased transparency, and immutability. The associated challenges are legal, regulatory compliance and uncertainty, cybersecurity, and absence of public sector involvement (Tian et al., 2020; Sazandrishvili, 2020). This is similar to the report from McKinsey & Company that discusses the impact of tokenization including improved capital efficiency; democratization of access; operational cost savings; enhanced compliance, auditability, and transparency; and, cheaper infrastructure (Banerjee et al., 2023).

```json
"name": "DT4C2 MP NFT",
"symbol": "NFTPro",
"description": "NFT of a material passport of a truss for exchange in
the course Digital Transformation for Circular Construction at ETH",
"seller_fee_basis_points": 100,
"image": "https://www.arweave.net/
E3o1gs0MN450vaFzDJ1lgaYLvuJW-4AfyGiwyzs-rU8?ext=PNG",
"attributes": [
  {
    "trait_type": "wood truss",
    "value": "physical asset"
  },
  {
    "trait_type": "material passport",
    "value": "circular construction"
  }
],
```

Figure 1: Example of a Material Passport NFT as a JSON

Tokenization Applications

In the context of using tokens within markets, Ferrara et al. (2022) states, "tokens of digital assets are indeed defined as digital objects in which the relevant information are stored to guarantee a single, unique matching between asset and token." The work proposed two processes for using blockchain in digital markets of physical assets: tokenization and legitimacy of ownership. Additional research from Weingärtner (2019) states the advantage of tokens is they facilitate the self-sovereignty of data and proposes three types: payment tokens (currency), utility tokens (rights or services), and assets or security tokens (represents a share).

A report from the consulting firm EY discusses tokenization of real-world assets for the transparency and improved traceability and trust in the chain of ownership. (EY, 2020). Tokens can be considered software with a unique asset reference connected to properties or legal rights. The report discusses three token standards and their ideal implementation: ERC-20 for transferring value between users, ERC-1400 for transferring asset or security tokens requiring a certificate, and ERC-721 for transferring ownership of a specific asset.

A recent position paper on the value of tokens for AEC provides a slightly modified token classification noting that the applications are either economic or technical: utility (governance), security (object representation), pegged (stablecoins), and direct (payments) (Kifokeris et al., 2023). The paper provides recommendations for a classification of tokenization opportunities by thematic area (e.g., circularity), and the associated benefits, challenges, issues, and opportunities.

Tezel et al. (2021) explored blockchain implementation in construction and suggested three primary supply chain management models: project bank accounts, reverse auction-based tendering, and asset tokenization for financing (crowdfunding). Wu et al. (2023) developed a prototype of an NFT-based construction waste material passport using a design science research approach. The prototype was built using Hyperledger, a private and permissioned blockchain. Additional research overlaps the concepts of smart construction objects and blockchain oracles for recording information around the construction processes on a blockchain (Lu et al., 2021).

Within the context of AEC, tokens have often been

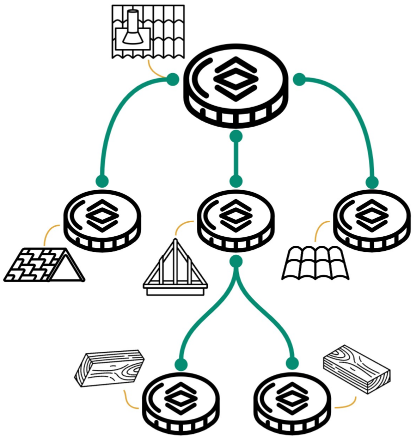


Figure 2: Visualization of Nested Token Hierarchy Representing Nested Component Assembly

proposed as an incentivization mechanism between stakeholders (Kouhizadeh et al., 2022). However, work from Dounas et al. (2021) proposed a system of organizing topology graph representation of a building as Ethereum tokens. Research from Hunhevicz et al. (2023) demonstrates a prototype of using tokens for Web3-based data access roles for material passports. Earlier work from 2019 proposes a framework for supply chain traceability based on tokenizing the bill of materials and component IDs (Dasaklis et al., 2019). Different tokens for source materials, elemental tokens, and compound tokens can be used to build up a tree structure of the data, emulating the similar hierarchy of assemblies in construction. Figure 2 symbolically represents this relationship of nested tokens for an assembly.

Boston Consulting Group and Arianee (2023) created a report on DPP tokenization that provided five main utilities: access to product information, a certificate of authenticity and ownership, a product lifecycle management tool, a customer relationship management tool, and a virtual replica of a physical object. Additionally, the three main proposed architectures are centralized DPPs, permissioned blockchain DPPs, and tokenized DPPs. Though in the short term, and for low-value products, tokenized DPP adoption is not as advantageous as other solutions, it is expected to provide much greater value in the long-term and for high-value products.

Methods

This paper is an exploratory paper on the overlap between tokenization and CCSC, therefore a mixed-methods approach was used including the exploratory data analysis of a survey of researchers and practitioners and a technical review of token features relevant to CCSC. The survey was issued to explore interest and elicit context on current challenges, perceptions, and suggestions for blockchain in CCSC. A more detailed description of the survey distribution and analysis is found in the respective section. In addition, technical standards were reviewed to better understand what features tokens can offer for product passports within AEC. Lastly, the mixed-methods approach combines the findings from each step into proposed scenario applications for tokenized product passports.

Survey on Exploring TPPs

Survey Approach and Overview

The survey distributed had 30 questions to gauge proximal knowledge on product passports for CCSC, technical requirements, and blockchain and tokenization applications. The survey was built online using Qualtrics and distributed via email through a convenience sampling of experts known by the authors with relevant experience in at least two of the three topics: sustainable construction, blockchain technologies, and circular economy. The submissions were analyzed anonymously using the language *R*. A total of 30 respondents completed the survey, though not all respondents answered every question. This paper shows the analyses of the most relevant questions for this work. The anonymized responses are a published dataset (https://doi.org/10.3929/ethz-b-000656624) and the code for the analysis can be found on github: https://github.com/cea-eth/TPP-EDA.

The format of the questions included five-point Likert-response, multi-select multiple-choice questions, and open-response questions. The open-response questions were analyzed using a natural language processing technique called latent Dirichlet allocation (LDA) for topic modeling. LDA is a generative probabilistic model for extracting similar topics over multiple documents (in this case, the documents are individual survey responses). The set of words for each extracted topic and the frequency of the topic are plotted and then interpreted by the authors for their meaning.

The respondents were asked to report in which countries they have projects. The highest reported country was Switzerland with 24% (nine responses), then the US tied with the UK with 13% (five responses), two responses in the Middle East, one response in Asia-Pacific, and the rest of the responses (n=16) were across Europe. About 40% of the respondents have 1-5 years of experience, about 23% have 5-10 years of experience, and 37% have greater than 10 years of experience. Of all the respondents, twelve work in academia, two for a government entity, and the rest within the private industry (in a combination of AEC firms, startups, and consultants).

Responses on DPP and Product Tracking

Of the 30 responses, four stated they were slightly familiar or not familiar with product passports, and two were left blank. Yet, 40% stated it is "Extremely Important" to track construction and building products over time, 40% stated it's "Very Important", and 13% stated it's "Moderately Important". The other two responses were blank.

Some of the questions asked were on potential issues related to data in the respondent's field and are displayed in Figure 3. The most striking results show a 100% positive

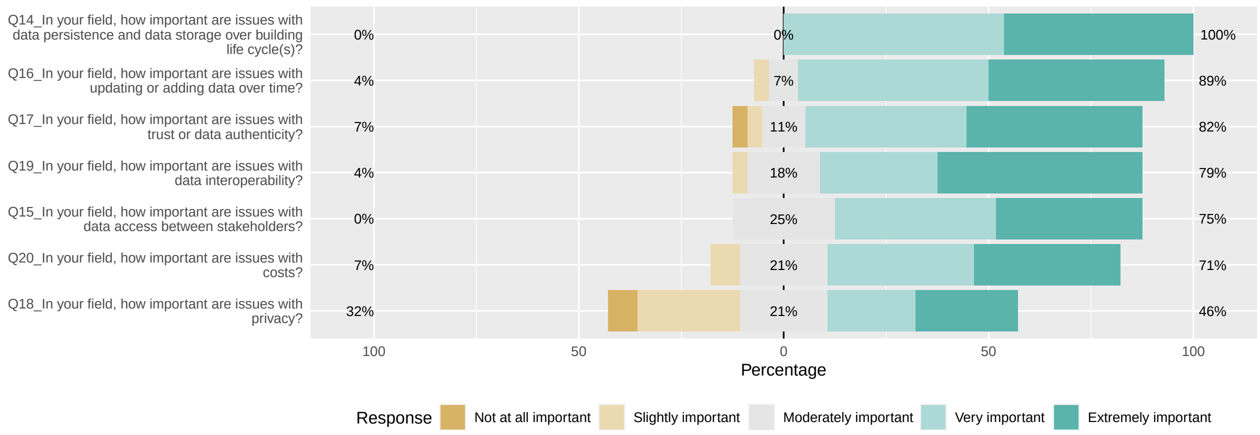


Figure 3: Responses on Importance of Issues on Data

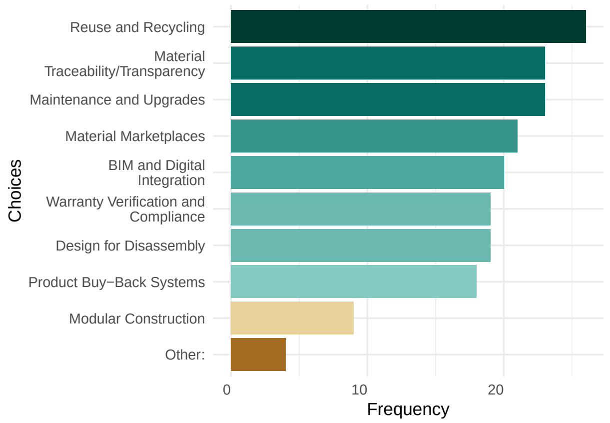


Figure 4: Responses on Ideal Use Cases for Product Passports

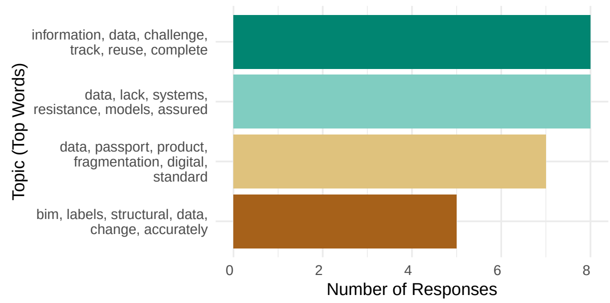


Figure 5: LDA Topics on the Challenges Inhibiting Product Passports for AEC

response for the importance of data persistence and storage over the building life cycle, and about one-third state issues with data privacy are only slightly or not important at all.

Among the various phases of the traditional product lifecycle, the three selected as having the greatest potential for implementing a product passport are B1-B3: Use, Maintenance, and Repair; B4-B5: Replacement and Refurbishment; and C1-C4: End of Life Stage. These responses are validated by another question on potential use cases for product passports shown in Figure 4.

The results from the LDA topic modeling on the problems and challenges inhibiting the use of product passports in AEC are shown in Figure 5. The y-axis shows the words in the topics produced from the model and the x-axis is their frequency found in responses. There are two topics on the challenges found eight times from the responses. The first topic is on the challenge of having complete information for tracking and reuse, the other major topic is on the resistance to data models and lack of assured data. Additional topics allude to challenges of product passport standardization and data fragmentation.

Responses on the Use of Blockchain and Tokenization

Figure 6 is a color-based correlation matrix from the Likert-based responses in the survey. The individual questions have been simplified to their topic and shown in the axes. A question has a perfect positive correlation with itself and is illustrated as dark green and a negative correlation is colored brown.

One of the questions asked the respondent how useful tokenizing product passports would be. The results from Figure 6 show that there is a slight positive correlation between tokenizing product passports and the questions asking about the applicability of blockchain and tokenization.

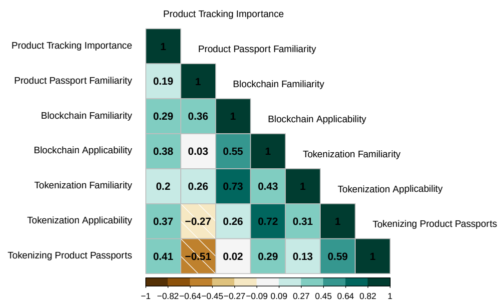


Figure 6: Correlation Matrix between Blockchain, Tokenization, Product Passports, and Product Tracking

Table 1: *Features of Tokens Relevant for Designing CCSC Systems*

Category	Token Features	Feature Description
Product Identification	Metadata Capacity	Amount of storage to hold product information
	Asset Linkage	Connecting the token to the product it represents
	Data Authenticity	Verification of unaltered product data
	Metadata Flexibility	Different types of data to be stored
Supply Chain Integration	Interoperability	Integration across platforms, actors, and file type
	Data Update Mechanism	Modifying data as the product moves or changes
	Queryable	Ease in searching for and retrieving token data
Lifecycle Management	Metadata Storage location	Considering on-chain or off-chain storage
	Composability	Combining and nesting tokens for assemblies or sub-products
Compliance & Privacy	Regulatory Compliance	Adhering to reporting regulations and other norms
	Data Privacy	Public or private chain for sensitive information
Business Proposition	Tradeability	Facilitating B2B/B2C transactions by token transfer
	Cost Efficiency	Evaluating overheads against its ROI
Product End-of-Life	Data Handling	Handling or burning of data at the end of its lifecycle
	Asset Linkage	Connecting the token to the product it represents

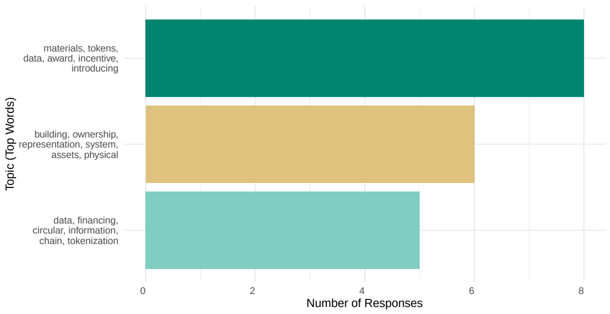


Figure 7: LDA Topics from the Potential Use Cases of Tokenization for AEC

Nevertheless, there is a slight negative correlation between the usefulness of tokenizing product passports and the familiarity level with product passports.

From the question of how might blockchain be used, the largest topic (n=13) extracted from the LDA was as the underlying technology for data transparency and tracking technology for building materials. From the question of how might tokenization be used for circular construction or circular supply chain management, only 19 responses were given as shown in Figure 7. Of those, the most frequent topic (n=8) discussed an application to introduce a system to award incentives for material information. Six responses discussed the topic of using tokens for ownership or representation of building assets. The third largest topic with five responses was on financing data for information in circular supply chains.

Review of Technical Token Features for CCSC

This section explores token standards for Ethereum and provides a table constructed from examined token standards that can be used to guide the applicability of a specific token for CCSC. Beyond ERC-20 and ERC-721, additional token standards have added functionality and potential relevance for this domain. Some are briefly listed below but this list is not exhaustive, which inspired the abstraction of token attributes from these ERC standards in Table 1.

- ERC-1155: Enables fungible and non-fungible tokens in one contract
- ERC-998: Composable NFTs that can own other NFTs and ERC-20 tokens
- ERC-6960: Dual Layer Tokens for asset classification, aiding in fractional and versatile ownership
- ERC-994: Delegated NFTs that register a geo-location of land and property
- ERC-5791: Physical-Backed Tokens representing physical item ownership and authenticity
- ERC-6551: Token-Bound Accounts are NFTs with unique Ethereum accounts for asset ownership and multi-chain application interaction
- ERC-5114: Soul-Bound Token, a non-transferable token permanently attached to a *soul* or address

Table 1 proposes features from various token standards that are relevant for applications to CCSC. The categories of token features should be considered by stakeholders interested in designing and implementing token-based systems. A similar classification approach was previously conducted on designing blockchain oracles for AEC (Dounas et al., 2023).

Resulting Scenarios

This section offers detailed scenarios for tokenized product passports, derived from a mixed-methods approach. The approach is a result from the analysis of token features, the review of relevant literature, and insights from the survey on blockchain's utility and use cases for CCSC. Tokens have broadly been shown to be used for asset representation, payment and incentive systems, or utilitarian and governance systems as defined in Kifokeris et al. (2023); Weingärtner (2019); Boston Consulting Group and Arianee (2023). The proposed scenarios examine the potential utility of tokens in CCSC by focusing on how tokens can serve as a means for representing objects. Proposing payment, incentive, or governance systems, although expressed from the survey results in Figure 7, is beyond the interest of this study.

These scenarios are not direct responses from the questionnaire but a synthesis of all results. Additionally, some token features are universally relevant for CCSC applications including Asset Linkage, Metadata Storage Location, Cost Efficiency, Data Authenticity, Composability, and Metadata Capacity.

Scenario 1) *Tokenize the Material DPP*: Drawing from industry proposals and the results from the survey, this scenario proposes a TPP system where each token acts as a representation for the DPP of the construction assets. The TPP can be connected directly to tracking hardware installed into the asset and used for streamlining compliance and transparency along the supply chain. Tokens could serve to create a unique digital identity for the provenance of each component, similar to Wu et al. (2023); Boston Consulting Group and Arianee (2023). According to DPP regulation, these tokens could facilitate accurate tracking and verification of the origin, composition, and recycling credentials of materials, which is essential for the integrity of a circular supply chain. Specific token features from Table 1 include: Regulatory Compliance, Queryable, and Interoperability.

Scenario 2) *Tokenize the Material Asset*: In this scenario, an asset tokenization platform for the CCSC can be developed for tokenizing the asset and supported by the results shown in Figure 7. Tokenizing a real-world asset reflects ownership within the real world and economic markets. Tokenized assets could allow the transparent traceability and management of material (EY, 2020) as well as their physical relationship to other materials (Dasaklis et al., 2019). An example of a preliminary implementation of this scenario and its integration with building modeling is proposed by Dounas et al. (2021). The token features from Table 1 important for this architecture include Tradeability, Metadata Flexibility, and Interoperability.

Scenario 3) *Tokenize the Material Ownership*: While the first scenario tokenizes the material information, the second scenario tokenizes the asset, and this third scenario utilizes tokens as a security for the ownership of the asset. This token security is a subset of asset tokenization that uses tokens to represent its value contractually, which can be reflected along the CCSC, particularly for material exchange, leasing, or purchasing in decentralized marketplaces (Hunhevicz et al., 2023; Ferrara et al., 2022). Though similar to the other scenarios, the financial and business considerations are more of an emphasis EY (2020); Boston Consulting Group and Arianee (2023). Using tokens as a mechanism for material asset securities also introduces new sustainable finance mechanisms for investing in circular economy solutions. Because of the legal and financial implications, the specific token features in the scenario include Regulatory Compliance, Tradeability, and Data Privacy from Table 1.

Discussion

One of the most interesting results from the survey is found in Figure 6, which shows a negative correlation between familiarity with product passports and the utility of tokenizing product passports. This implies either that those very familiar with product passports do not think there is much utility with TPPs, or that those who think there is high utility with TPP aren't very familiar with product passports. Nevertheless, there is a stronger positive correlation between TPP utility and those who believe tokenization has high potential for CCSC. These results align with Figure 7, where the most frequently mentioned topic was using tokens as incentives, not as representations of the asset itself. Although this supports the *direct* and *utility* cases found in Kifokeris et al. (2023), those cases are out of scope as this paper focuses on the *object representation* case.

Storing product passports on-chain ensures their accessibility and permanence for any stakeholder through time. This approach secures the data and also potentially enables new ecosystem interactions, such as unique wallets assigned to physical assets and the possibility for tokens to be held by non-human entities. The proposed scenarios for token usage are directional but grounded in existing research and survey findings. They require further development and verification through technical implementation and stakeholder engagement. Although this research primarily examines tokens on the Ethereum blockchain, other blockchains also employ tokens, which could be leveraged for DPP applications.

Challenges and Limitations

Blockchain in construction, in addition to product life cycle tracking, is a cyber-physical-social systems problem. To abate this challenge, the researched scenarios should be able to integrate first with existing systems for business-as-usual. Thus the use of a common data format and metadata schema is critical. Additionally, linking data in tokens to external servers is often critiqued because of the challenges associated with the maintenance of hosting externally.

This study is limited by not yet testing implementations of the proposed scenarios. For the survey, the respondents primarily work across Europe, with some in the UK and the US. Responses may vary based on geographical region. There is a natural selection bias from the method of approaching the respondents and cannot be taken as a generalizable state of knowledge. There are inherent limitations to using a questionnaire as it is more likely respondents will provide brief or no answers at all. The use of more exhaustive methods, like interviews or the Delphi method, may produce more informed responses.

Given the heterogeneity of experiences of the survey respondents, a deeper statistical analysis can be used for analyzing results. For instance, the results can be partitioned and analyzed separately based on the categories of responses (e.g., familiarity level with product passports, familiarity level with tokenization, or country of work).

Future Work

The survey included additional open-ended questions that were not covered in this study. Additional work can explore these other inputs from the respondents on hardware approaches, ideal stakeholders, and challenges and use cases. A few of the respondents noted the importance of analyzing the value-add of tokenizing product passports and an understanding of the environmental tradeoffs. These are valid concerns and though they are out of the scope of this study, stand as points for future work.

This work sets the exploratory and theoretical underpinning for future hypothesis development and research on qualitative implementation and quantitative testing of these new processes. Future efforts by the researchers will address the *how-to* for technical implementation and aim to quantify impacts. Further work will use a design science research approach for an exploration of TPP in a real building. The survey findings in this study will inform the goals and design of such a system.

Conclusions

This research investigated the application of tokenization for product passports in CCSC. The authors employ a mixed methods approach via a technical review and survey of experts to frame the potential value of tokenization in circular construction. The results extend the understanding of blockchain-based product tracking for AEC by focusing on the potential of using tokens for object representation of a product passport.

The academic literature confirmed general interest in blockchain and tokenization for product tracking, although with limited CCSC applications tested. Survey results supported the need for product tracking and blockchain for data transparency, while also noting challenges such as data gaps and system integration. Features extracted from various Ethereum tokens were explored for their potential relevance for product passports. Lastly, these results were combined for three proposed scenarios of tokenization within the CCSC: tokenizing the material DPP, tokenizing the material asset, and tokenizing the material ownership. Future research intends to expand upon this theoretical foundation by testing the specified scenarios. This study's findings suggest that tokens could serve as a comprehensive information repository, providing innovative application possibilities for CCSC.

References

Banerjee, A., De Bode, I., de Vergnes, M., Higginson, M., and Sevillano, J. (2023). Tokenization: A digital-asset déjà vu. Technical report, McKinsey.

Boston Consulting Group and Arianee (2023). The case for native Digital Product Passport tokenization. Technical report.

Byers, B. and De Wolf, C. (2023). QR Code-Based Material Passports for Component Reuse Across Life Cycle Stages in Small-Scale Construction. Circular Economy, 1(1). Accepted: 2023-11-06T08:45:37Z Publisher: DSRPT UG.

Byers, B. S., Raghu, D., Olumo, A., De Wolf, C., and Haas, C. (2024). From research to practice: A review on technologies for addressing the information gap for building material reuse in circular construction. Sustainable Production and Consumption, 45:177–191.

Dasaklis, T. K., Casino, F., Patsakis, C., and Douligeris, C. (2019). A Framework for Supply Chain Traceability Based on Blockchain Tokens. In Di Francescomarino, C., Dijkman, R., and Zdun, U., editors, Business Process Management Workshops, Lecture Notes in Business Information Processing, pages 704–716, Cham. Springer International Publishing.

Dounas, T., Hunhevicz, J., and Byers, B. (2023). Design dimensions for blockchain oracles in the AEC industry. volume 4 of Computing in Construction, pages 0–0. European Council on Computing in Construction.

Dounas, T., Jabi, W., and Lombardi, D. (2021). Topology generated non-fungible tokens: blockchain as infrastructure for a circular economy in architectural design. 2. Publisher: Association for Computer-Aided Architectural Design Research in Asia.

European Commission (2022). Proposal for a REGULATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL establishing a framework for setting ecodesign requirements for sustainable products and repealing Directive 2009/125/EC.

EY (2020). Tokenization of Assets, Decentralized Finance (DeFi). Technical report, Switzerland.

Farooque, M., Zhang, A., Thürer, M., Qu, T., and Huisingh, D. (2019). Circular supply chain management: A definition and structured literature review. Journal of Cleaner Production, 228:882–900.

Ferrara, G., Messina, F., De Benedetti, M., and Santoro, C. (2022). Physical Assets Tokenization for Blockchain Market. In Camacho, D., Rosaci, D., Sarné, G. M. L., and Versaci, M., editors, Intelligent Distributed Computing XIV, Studies in Computational Intelligence, pages 273–282, Cham. Springer International Publishing.

Honic, M., Kovacic, I., Aschenbrenner, P., and Ragossnig, A. (2021). Material Passports for the end-of-life stage of buildings: Challenges and potentials. Journal of Cleaner Production, 319:128702.

Honic, M., Magalhães, P. M., and Van den Bosch, P. (2024). From Data Templates to Material Passports and Digital Product Passports. In De Wolf, C., Çetin, S., and Bocken, N. M. P., editors, A Circular Built Environment in the Digital Age, Circular Economy and Sustainability, pages 79–94. Springer International Publishing, Cham.

Hunhevicz, J. J., Bucher, D. F., Soman, R. K., Honic, M., Hall, D. M., and De Wolf, C. (2023). Web3-based role and token data access: the case of building material passports. volume 4 of Computing in Construction, pages 0–0. European Council on Computing in Construction.

Jansen, M., Gerstenberger, B., Bitter-Krahe, J., Berg, H., Sebestyén, J., and Schneider, J. (2022). Current approaches to the digital product passport for a circular economy: An overview of projects and initiatives. Working Paper 198, Wuppertal Papers.

Kifokeris, D., Dounas, T., Tezel, A., and Moon, S. (2023). What is the potential value of tokens and token engineering for the architecture, engineering, and construction industry? A positional paper. volume 4 of Computing in Construction, pages 0–0. European Council on Computing in Construction.

Kouhizadeh, M., Zhu, Q., and Sarkis, J. (2022). Circular economy performance measurements and blockchain technology: an examination of relationships. The International Journal of Logistics Management, 34(3):720–743. Publisher: Emerald Publishing Limited.

Li, J., Greenwood, D., and Kassem, M. (2019). Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. Automation in Construction, 102:288–307.

Lu, W., Li, X., Xue, F., Zhao, R., Wu, L., and Yeh, A. G. O. (2021). Exploring smart construction objects as blockchain oracles in construction supply chain management. Automation in Construction, 129:103816.

Mollaei, A., Byers, B., Christovan, C., Olumo, A., De Wolf, C., Bachmann, C., and Haas, C. (2023). "A global perspective on building material recovery incorporating the impact of regional factors". Journal of Cleaner Production, 429:139525.

Qian, X. A. and Papadonikolaki, E. (2020). Shifting trust in construction supply chains through blockchain technology. Engineering, Construction and Architectural Management, 28(2):584–602. Publisher: Emerald Publishing Limited.

Sazandrishvili, G. (2020). Asset tokenization in plain English. Journal of Corporate Accounting & Finance, 31(2):68–73. _eprint: https://onlinelibrary.wiley.com/-doi/pdf/10.1002/jcaf.22432.

Scott, D. J., Broyd, T., and Ma, L. (2021). Exploratory literature review of blockchain in the construction industry. Automation in Construction, 132:103914.

Tezel, A., Febrero, P., Papadonikolaki, E., and Yitmen, I. (2021). Insights into Blockchain Implementation in Construction: Models for Supply Chain Management. Journal of Management in Engineering, 37(4):04021038. Publisher: American Society of Civil Engineers.

Tian, Y., Lu, Z., Adriaens, P., Minchin, R. E., Caithness, A., and Woo, J. (2020). Finance infrastructure through blockchain-based tokenization. Frontiers of Engineering Management, 7(4):485–499.

Wang, Z., Wang, T., Hu, H., Gong, J., Ren, X., and Xiao, Q. (2020). Blockchain-based framework for improving supply chain traceability and information sharing in precast construction. Automation in Construction, 111:103063.

Weingärtner, T. (2019). Tokenization of physical assets and the impact of IoT and AI. Technical report, European Union Blockchain Observatory and Forum.

Wijewickrama, M. K. C. S., Rameezdeen, R., and Chileshe, N. (2021). Information brokerage for circular economy in the construction industry: A systematic literature review. Journal of Cleaner Production, 313:127938.

Wu, L., Lu, W., Peng, Z., and Webster, C. (2023). A blockchain non-fungible token-enabled 'passport' for construction waste material cross-jurisdictional trading. Automation in Construction, 149:104783.

Çetin, S., De Wolf, C., and Bocken, N. (2021). Circular Digital Built Environment: An Emerging Framework. Sustainability, 13(11):6348. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.

Çetin, S., Raghu, D., Honic, M., Straub, A., and Gruis, V. (2023). Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. Sustainable Production and Consumption, 40:422–437.

CHEAPER SMART CONTRACTS FOR THE BUILT ENVIRONMENT? LINKING ON-CHAIN AND OFF-CHAIN IN A BLOCKCHAIN-GOVERNED APPROACH

Jens J. Hunhevicz[1,2], David F. Bucher[2], Kepeng Hong[3], Daniel M. Hall[4], and Catherine De Wolf[2]
[1]EMPA, Dübendorf, Switzerland
[2]ETH Zurich, Zurich, Switzerland
[3]DTU, Kongens Lyngby, Denmark
[4]TU Delft, Delft, The Netherlands

Abstract

The digitization and automation of contracts within the built environment through blockchain has demonstrated potential. Nonetheless, the use of on-chain smart contracts can amount to substantial costs. This study proposes a blockchain-governed approach to individually assess whether and how to use blockchain for different components of contracts. We explain the rationale behind the concept and implement a pilot prototype of a performance-based, blockchain-governed contract. The results promise an alternative and more cost-effective approach to smart or intelligent contracts in the built environment, while still allowing for trusted verification of key contract parts through blockchain.

Introduction

Contracts are important to the collaborative built environment and for interactions between its actors. With the ongoing digital transformation, the digitization and automation of contracts is increasingly being discussed. In particular, blockchain technology promises to enable "smart legal", sometimes also termed "intelligent" contracts (Mason, 2017; McNamara and Sepasgozar, 2021; Allen and Hunn, 2022). Independent of the term, we refer here to the idea to encode contract terms in so-called blockchain smart contracts, which are scripts deployed on a blockchain that can then enforce interaction logic with any blockchain transactions, for example to execute a payment. Blockchain as an immutable ledger of peer-to-peer transactions ensures that contract terms are executed in a transparent and trustworthy manner without the need for a third-party institution. The applications of smart contracts for the built environment are diverse (Li and Kassem, 2021). Examples include the automatic execution of coded terms for contract management (Msawil et al., 2022), automated contracting for construction projects (Gupta and Jha, 2023), or performance-based contracts for lifecycle services based on IoT data (Hunhevicz et al., 2022).

Although a very promising technology for contracting, the use of blockchain smart contracts also creates challenges and risks (Mezquita et al., 2019). Most relevant to this study, their use can amount to substantial costs (Zou et al., 2021; Hunhevicz et al., 2022). As discussed later in this paper, potential solutions are to use a blockchain with low transaction costs or permissioned blockchains. However, we suggest that a more unexplored option is to decouple the functionality of smart contracts and use blockchain only for parts of a contract.

Therefore, we introduce the concept of "blockchain-governed" contracts in the built environment, and why it could be a way to reduce the costs associated with blockchain-based contract management with reasonable trade-offs. To do so, we first cover the necessary background in the point of departure, then introduce the concept of the chosen approach, and show an exemplary implementation of a performance-based contract. Finally, we discuss our contribution by comparing it to a previous on-chain implementation of a similar smart contract and point out limitations and further research.

Point of Departure

Smart Contract Components

As already mentioned in the introduction, smart contracts encode interaction logic with blockchain transactions. In simple terms, a smart contract performs tasks like *"if this happens on the blockchain, then execute this transaction logic"*. Smart contract can be abstracted into four key components (Hunhevicz et al., 2022), as shown in the upper part of Figure 1 and further described below. Although they can be combined into complex contract constructs, for the scope of this paper, we can focus on these components individually.

Identity: Blockchain identifies users with a pair of public / private keys. The public key serves as an address to identify a blockchain user on the network. The private key acts as a password to sign transaction to prove control of the respective address. Smart contracts require information about which public address is allowed to interact or is affected by the smart contract functionality.

Payment: Smart contracts can initiate payment transactions in native cryptocurrency or another token. They can also hold funds and act as bank accounts. Even if the functionality of a smart contract is not related to finance, smart contract functionality still involves a monetary fee to execute transactions. Therefore, for the scope of this paper, we refer to payments as any use of a token or cryptocurrency, both for contract payments or fees.

Data: Smart contracts can store and act on data. Depending on the blockchain, data is stored through available data primitives, for example, in the case of Ethereum, state vari-

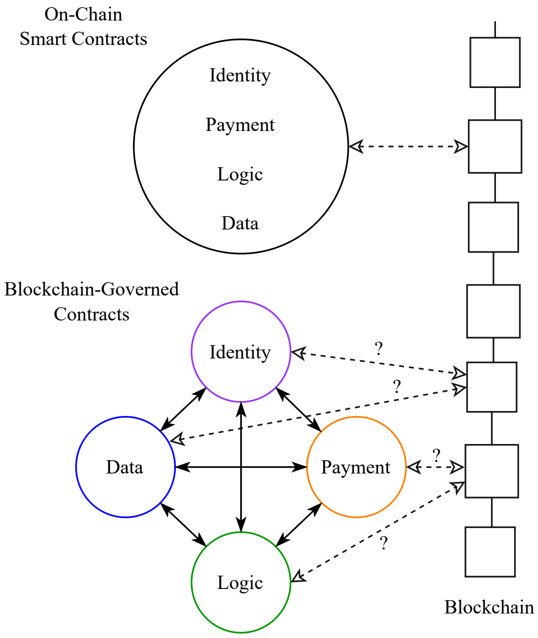


Figure 1: Smart contracts involve identity, payment, logic, and data; a blockchain-governed contract assesses individually whether and how to use blockchain for each category.

ables, arrays, structs, or mappings[1]. To add data to a smart contract, executing a transaction is required to pay the network for storing the data.

Logic: The key part of a smart contract is its logic related to transaction execution and interaction, possibly interacting with and connecting the other described components of identity, payment, and data.

Current Approaches to Smart Contracts

This section reviews current approaches to smart contracts and, where applicable, provides examples of research in the context of the built environment.

On-Chain Smart Contracts

In this paper, we refer to on-chain smart contracts as what is most often meant with the concept of blockchain smart contracts. They include all the above-introduced components in one construct. Once deployed on the blockchain, the logic is deterministic and fully transparent. Ethereum piloted the use of such Turing-complete smart contracts (Buterin, 2014). Several studies in the built environment use on-chain smart contracts, e.g. on Ethereum (Yang et al., 2020; Hamledari and Fischer, 2021; Hunhevicz et al., 2022).

Depending on the used blockchain, the speed of transaction execution can be slow and the associated transaction costs can amount to large sums (Zou et al., 2021). To reduce the costs associated with on-chain contracts, we identified two main options.

First, a blockchain with cheaper transaction costs can be used. For example, new generations of blockchain networks optimize and market themselves for high transaction throughput and low transaction cost, e.g. the Solana blockchain[2]. Moreover, existing blockchain networks, such as Bitcoin and Ethereum, promote the use of second-layer protocols to process transactions faster and cheaper (Gangwal et al., 2023). A potential trade-off to faster transaction can concern the security of the blockchain (Kiayias, 2016). Furthermore, transaction costs can also increase as a function of growing network usage. For now, we are not aware of much research discussing this approach in the built environment. Examples include Naderi et al. (2023) and Scott et al. (2024).

The second option is a permissioned blockchain, also known as a consortium blockchain, e.g. Hyperledger Fabric[3]. This type of blockchain is run by a set of trusted actors who operate the blockchain nodes. Permissioned blockchains offer high throughput and no transaction fees, as the consensus mechanisms are fast and overhead costs are typically paid and shared by the consortium. As a potential downside, permissioned blockchains rely on trusted actors to run the nodes, who have the ability to exclude transactions and users and even shut down the network at their discretion. Furthermore, ensuring the availability of nodes can be challenging over long time periods and changing stakeholders. Nevertheless, due to its project-based nature, a consortium approach is often considered in the built environment (Yang et al., 2020).

Off-chain Approaches to Contracts

A second way to reduce the costs associated with smart contracts involves moving parts of the smart contract off-chain, meaning not stored or implemented on the blockchain. One reason is that the transaction cost of an on-chain smart contract is typically determined by the transaction size and involved computation of a smart contract execution. Especially in the context of networks like Ethereum, moving data off-chain e.g. to external data networks like IPFS[4], is becoming increasingly common, also for applications explored in the built environment (Tao et al., 2021; Adel et al., 2023).

Another reason to research off-chain contracts is that some networks, e.g. Bitcoin, do not support Turing complete smart contracts. Therefore, different approaches are suggested to move data and logic off-chain, e.g. outsourcing to service providers (Wüst et al., 2019), trusted execution environments (TEEs) (Das et al., 2019), or off-chain schemas using blockchain as a state commitment layer and ownership control[5]. We are not yet aware of research in the built environment exploring such contract approaches.

Finally, governance platforms can facilitate trusted process execution through a blockchain-verified and secure semi-automated process (Dursun and Üstündağ, 2021). "Politea"[6] is an example for such a platform used as a pro-

[1] `https://docs.soliditylang.org/en/latest/internals/layout_in_storage.html`, accessed 30.01.2024

[2] `https://solana.com/`
[3] `https://www.hyperledger.org/projects/fabric`
[4] `https://ipfs.tech/`
[5] `https://rgb-org.github.io/`
[6] `https://proposals.decred.org/`

posal and payment system for the Decred blockchain contractors. Although no formal on-chain logic is utilized, it makes use of timestamps to allow for checking and re-evaluating the process in case of disputes. Timestamping is inherent to blockchains; each block and transaction has a timestamp for clear recognition. Using this mechanisms, data can be attached to a transaction to create a fingerprint for off-chain data as a proof of existence.

Motivation and Scope of This Study

The previous sections showed that there are different approaches to smart contracts. An alternative to on-chain smart contracts is moving parts or all of the contract components off-chain. We see these approaches as a viable alternative for smart or intelligent contracts in the built environment for the following reasons.

Known actors: Construction and the built environment rely on established processes. The decision to use blockchain involves a trade-off between trusting the technical system vs. trusting the involved actors (Hunhevicz and Hall, 2020). If actors are known and accountable, the cost premium of a fully decentralized on-chain smart contract may not be justified.

Time spans: Construction contracts usually span duration of several years of individual project phases, e.g., for design and construction. Even though such contracts can benefit from a blockchain approach, it might not be justified to pay for an on-chain contract that lasts as long as a blockchain exists.

Physical nature: Blockchain-based contracts need to rely on trusted feedback loops of physical processes and products. Although the logic of an on-chain smart contract is trustworthy, wrong input data could jeopardize a correct execution. A partial off-chain approach could be a practical middle ground with reasonable trust and transparency.

Trust in existing systems: The construction industry is known for slow technology adoption. In many cases, the possibility of traditional identity verification, FIAT payments, or the possibility for more data privacy has priority. A blockchain-governed approach could better consider these aspects than a pure on-chain smart contract.

Therefore, this study proposes the concept of ”blockchain-governed” contracts for the built environment, and pilots an off-chain governance system for an exemplary performance-based contract.

Blockchain-Governed Contracts

The term ”blockchain-governed” contract is proposed in this paper for contracts that use blockchain in one of the four components introduced (identity, payment, logic, or data), but at the same time also do not use blockchain with an off-chain approach in at least one category. As shown in Figure 1, such a blockchain-governed approach allows to individually evaluate for use cases whether and how blockchain is used for each component. Table 1 provides non-exhaustive examples of the differences between an on-chain and an off-chain approach for each component.

Proposed Off-Chain Governance System

To illustrate the concept of blockchain-governed contracts in the built environment, we propose an off-chain governance system for a contract in the built environment in Figure 2. The figure is organized in a matrix structure, with the four components of identity, payment, and logic arranged vertically, and the physical built environment, the off-chain governance system, and blockchain (on-chain) arranged horizontally. The proposed system is inspired by the aforementioned Decred governance system, which has been operational since 2017. Therefore, this system is already proven in the context of managing the development

Table 1: Exemplary differences in choosing an on-chain vs. off-chain approach for the four contract components. The proposed concept of blockchain-governed contracts could then use either option, but at least one on-chain, and one off-chain component.

	On-Chain	Off-Chain
Identity	To execute an on-chain transaction, actors need a blockchain address, e.g. for payments, logic execution, or storing data. Each transaction needs to be signed with the corresponding private key.	Traditional credentials such as username and password can be used as means of identification for off-chain contract applications, timestamping, or read-only functionality of blockchain state.
Payment	Cryptocurrency and tokens can be used as a means of payment. Execution guarantee, censorship resistance, and customization of the monetary asset can be advantages of using on-chain payments. Costs refer to network transaction costs.	Traditional third-party payment methods in FIAT currencies such as EUR can be linked to the contract, e.g. via application interfaces (APIs). Benefits include compliance with traditional systems and regulations. Costs relate to service fees.
Logic	On-chain logic provides several benefits, including direct interaction with the blockchain state (such as executing a payment), transparent and verifiable logic, and censorship resistant logic that cannot be changed unnoticed.	Off-chain logic can be transparent and verifiable by using accessible (public or distributed local) code repositories combined with timestamping. This enables the logic to be recalculated and verified. It may also be possible to use other approaches (e.g. TEEs).
Data	Data stored on a blockchain is transparent, immutable, and available as long as the blockchain exists. These characteristics ensure the data cannot be altered or deleted unnoticed.	Moving data to external storage locations can reduce on-chain computation costs. The smart contract can reference the data's location and timestamp to verify its authenticity.

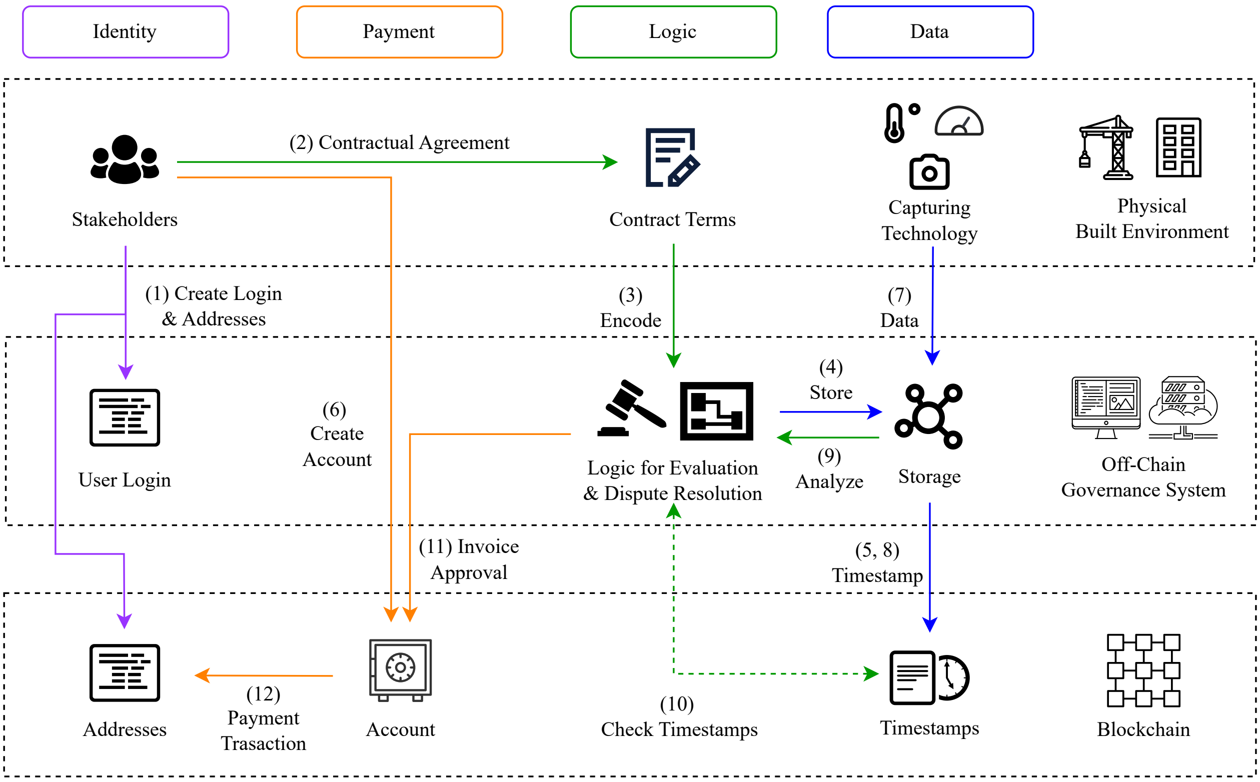


Figure 2: Exemplary blockchain-governed contract system for the built environment.

of the Decred ecosystem, making it a good starting point for investigating blockchain-governed systems in the built environment. In addition, it demonstrates the interaction between the different components using both on- and off-chain approaches.

A typical process would work as follows (see Figure 2): (1) All stakeholders create both login credentials for the governance platform, as well as an address so that the contract payout terms can be defined and encoded. (2) The parties negotiate the contractual agreement. (3) The contract terms are encoded. (4) The contract is stored in the off-chain governance database. (5) The contract is timestamped so that it can later be verified as authentic. (6) The responsible stakeholders create and fund the project account from which the payments will be released. (7) The required data is captured, streamed, and stored off-chain in the project database. All stakeholders should have a local copy, or at least access to the data. (8) The data is periodically timestamped so that it can be verified at a later time based on the locally stored copies. (9) At defined intervals, performance is automatically evaluated based on the input data and contract terms. All results are published transparently to stakeholders. (10) Stakeholders can confirm the evaluation. Only in the case of an error or disagreement, a dispute resolution process would start to verify the correct performance logic and data based on timestamps (dotted line in Figure 2). (11) If there is no disagreement, or after the dispute is resolved, the generated invoices are confirmed and signed by the necessary stakeholders. (12) Payouts are made in cryptocurrency to the defined addresses.

Implementation

We developed a pilot prototype to obtain first insights into the feasibility and challenges of the proposed blockchain-governed contract approach. After introducing the tested process, we outline for each component (identity, payment, logic, and data) the chosen approach for our prototype, in line with the proposed governance system in Figure 2.

Tested Process

The example follows the use case of a performance-based contract that rewards stakeholders for meeting the energy performance targets of operating a building. A similar process was implemented and evaluated in a previous study (Hunhevicz et al., 2022), allowing for a good comparison regarding process, cost, and efficiency. For the interested reader, the referenced study also gives more details on the rationale of performance based contracting.

Figure 3 shows the interaction between the technical parts implemented, as well as the stakeholders. The general idea is that a building owner initiates the use of a blockchain-governed contract to issue automatic payments to a contractor responsible for operating and managing the energy system. The exact logic of the contract is not important for the scope of this work; in this example we used historic temperature data from the NEST building at EMPA Dübendorf, Switzerland[7], issuing payments when the managed temperature stays within a defined range around the set-point temperature.

[7]https://www.empa.ch/web/nest/

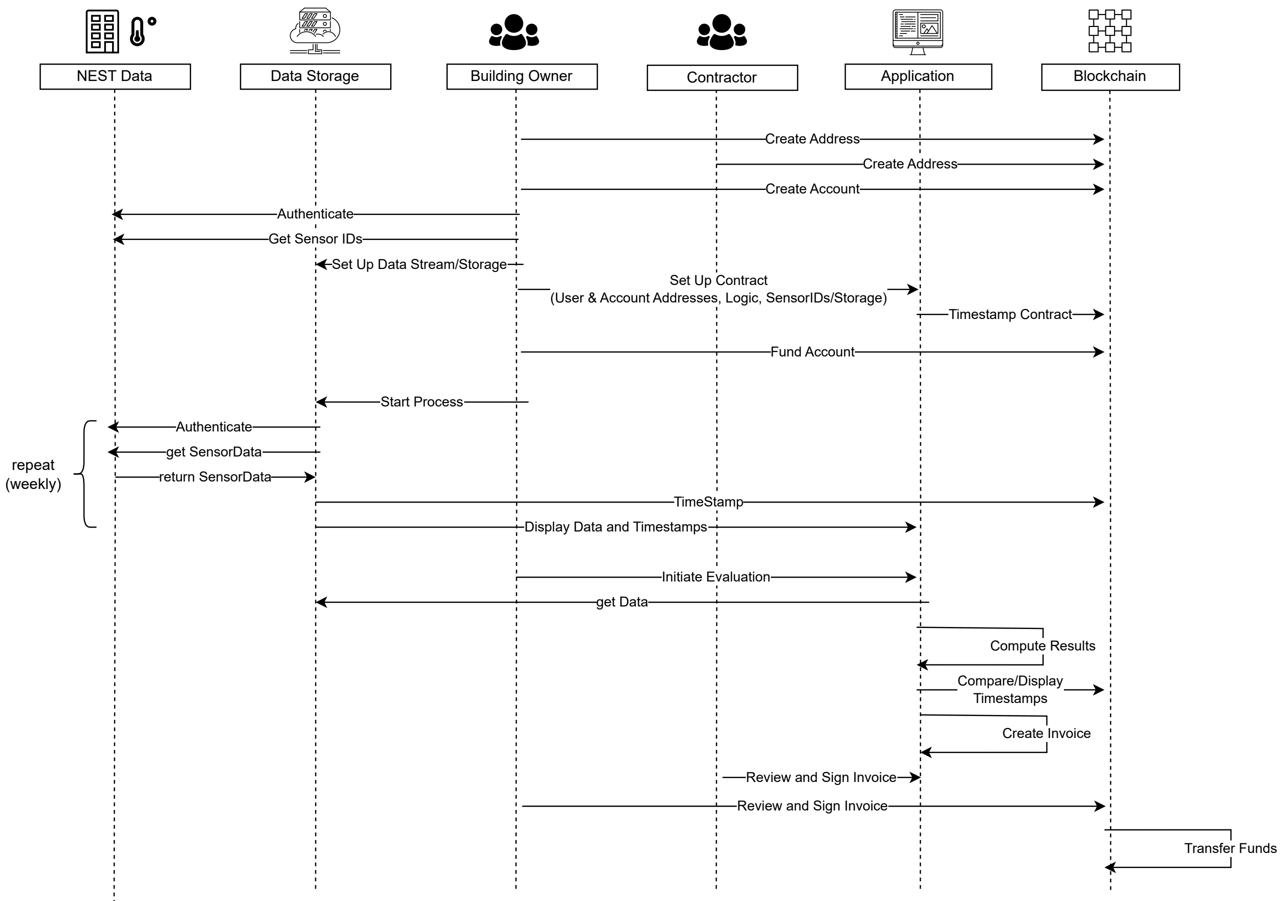


Figure 3: Implemented process of the temperature performance contract. In the tested process, both the owner and the contractor accept the evaluation and sign the invoice transaction.

Logic

The core of the prototype is a web application built with Next.js[8] that manages the main contract governance process. The stakeholders can define a new data set, as well as the contract logic with the main parameters of the contract (see Figure 4). In our case, we defined the sensor data set, the payout address, the set point temperature, the allowed deviation and threshold for penalties, and the payout amount of the performance based contract.

The contract will be time-stamped after its creation, anchoring the hash of the JSON contract file to the Decred blockchain (see Figure 4). We chose dcrtime[9], because the service is open source and freely accessible via API. Dcrtime timestamps approximately every hour, called anchoring. To indicate the status of the process, the application changes from "not timestamped", to "waiting for anchoring time", to "timestamped". Once timestamped, the digest can be retrieved from the application and checked for the timestamp and transaction hash on Timestamply[10]. In Decred's blockchain explorer[11], the transaction details can be displayed.

After a defined evaluation period, the results can be computed for the specified data set and defined contract logic. To make this evaluation trustworthy, the data and results are visualized in the application (see Figure 5). In case there are concerns about the validity of the process, the code and data can be checked against local copies or copies stored in an open source repository, and if needed, even redeployed and recalculated with the timestamped data and contract logic to check for authenticity.

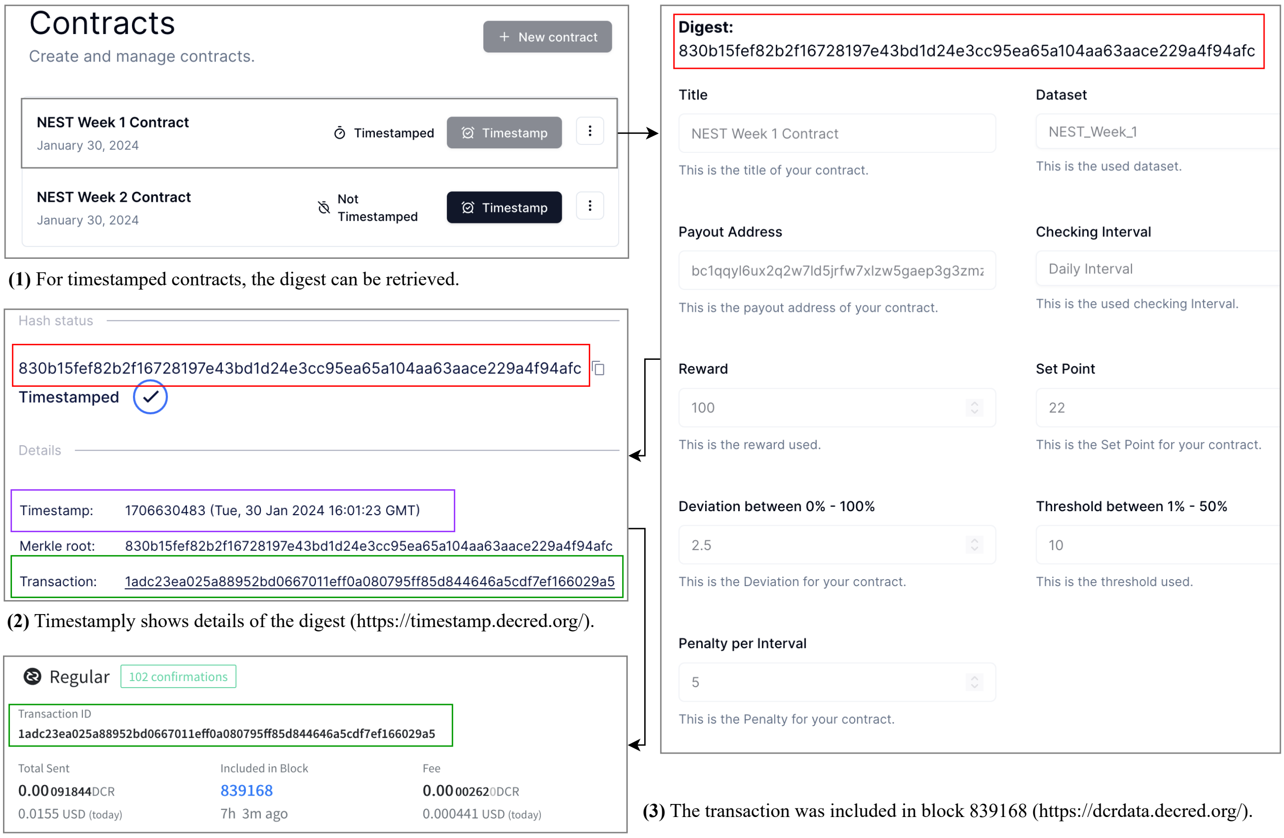
Payment

Payouts are released from a separate project account, which is set up as a multi-sig wallet by the owner and contractor, meaning that both the owner's and contractor's blockchain signatures are required to move funds. This was implemented so that neither party could spend the funds without the other's approval. After calculation of the payouts, the invoice to credit the project account is automatically created based on the specified payout addresses. If both parties agree, they can import the proposed transaction to their respective wallets and sign for execution.

Identity

The above described web application requires a standard username and password for identity verification. However, the application could also have used blockchain-

[8] https://nextjs.org/
[9] https://github.com/decred/dcrtime
[10] https://timestamp.decred.org
[11] https://dcrdata.decred.org



(1) For timestamped contracts, the digest can be retrieved.

(2) Timestamply shows details of the digest (https://timestamp.decred.org/).

(3) The transaction was included in block 839168 (https://dcrdata.decred.org/).

Figure 4: Contracts can be created in the application with the needed parameters for the evaluation and payout. After creation, they can be timestamped to ensure verifiability at a later point.

based access verification (Hunhevicz et al., 2023). This demonstrates that there can be different identity procedures within one blockchain-governed contract.

For the payouts, a blockchain wallet is needed. Since we chose the Bitcoin blockchain for payments, we used the Electrum[12] wallet to create the public/private key pairs, since it allows for easy multi-sig capabilities. The owner and the contractor each need their own wallet.

Data

Data is stored off-chain in a relational SQL database accessed by the application. In our example, the application used historical temperature tiime series of the NEST building from the ehub platform[13] through the available APIs. In addition to the main off-chain storage, the data is timestamped in weekly intervals, as introduced above for the contract logic (see also Figure 4). Therefore, the hashes of the data are stored on-chain on the Decred blockchain.

Discussion and Conclusion

This study is a first attempt to find practical alternatives to pure on-chain smart contracts in the built environment. The proposed decoupling and individual evaluation of whether and how to use blockchain for identity, payment, logic, and data was termed ”blockchain-governed contract”. The then proposed and evaluated exemplary contract governance system combines on- and off-chain approach, mainly using blockchain for proofing authenticity of data through timestamping, as well as payouts.

One of the most interesting consequences of this approach is that it forces an implementer to think about the individual parts of a contract and whether blockchain makes sense. With a complete on-chain approach, the entire contract is usually implemented on-chain, although only a subset would benefit from it. This means that blockchain could be used for only parts of a contract and that different parts of the contract may use different features of blockchain, possibly even different blockchains. For example, timestamping could benefit from the free services offered by e.g. the Decred dcrtime[14] library or Bitcoin's opentimestamps[15], while other features that require on-chain smart contracts could use a fast blockchain layer for high throughput and reasonable transaction costs. A more extensive and systematic analysis on when to use which blockchain would be interesting future research.

In addition, more work is needed to complete an implementation and validate it in a real-world contracting scenario. One of the missing pieces of the here tested system is the dispute resolution process, in case the parties do not accept the evaluation and payments performed.

[12]https://electrum.org/
[13]https://info.nestcollaboration.ch/wikipediapublic/
[14]https://github.com/decred/dcrtime
[15]https://opentimestamps.org/

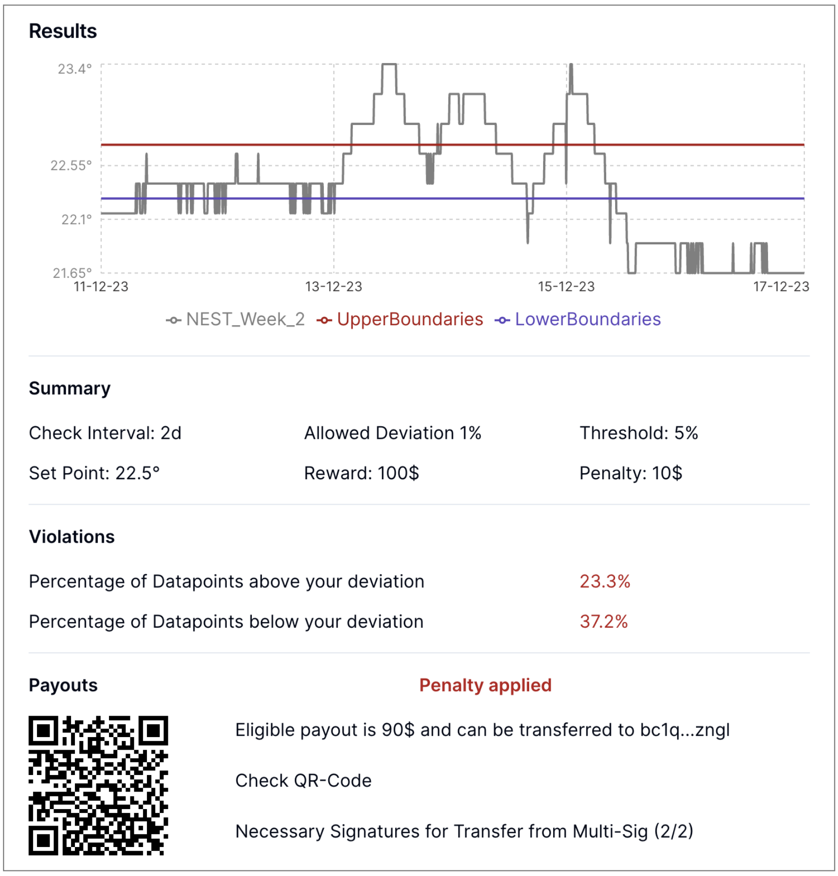


Figure 5: Evaluation of data in the application based on the contract terms and data set to calculate payouts. Authenticity of the data and contract can be checked against the timestamps. The invoice is created and needs to be signed by both the contractor and owner.

Other works have already proposed blockchain-based dispute resolution platforms for the built environment (Saygili et al., 2021; Son and Lien, 2022), which could serve as inspiration. We are confident that a suitable mechanism could be designed and implemented. Furthermore, while the proposed system seems like a reasonable approach, it is only one of many possible combinations between on- and off-chain for the different contract components as described in the departure section. Other combinations should be designed and tested for comparison.

A blockchain-governed contract approach appears to offer more flexibility and the potential for cost savings compared to an on-chain smart contract. In addition, we believe it could also improve the usability of smart contracts, since only parts of the contract require a blockchain, which is currently a mostly new technology unfamiliar to most stakeholders in the built environment. Complicated smart contract applications could be avoided. In our case, the proposed governance platform is similar to current web applications, except that payment execution requires a wallet. Another interesting observation is that a blockchain-governed approach requires more manual and human input than a full on-chain smart contract. In this sense, it can be considered less automated then previously proposed smart or intelligent contracts, but is potentially simpler and closer to existing practice. However, the verification process is likely to take longer to resolve than a fully transparent and deterministic on-chain approach, particularly in the event of a dispute. More detailed research is needed to evaluate user experience aspects, the industry requirements and readiness for different blockchain-governed smart contract platforms, as well as a quantitative comparison of cost and performance, e.g. under what circumstances the cost of on-chain smart contracts compared to a blockchain-governed approach is justified. Overall, a blockchain-governed contract approach seems a solution worth exploring in the built environment.

Acknowledgments

We thank Empa for access to the ehub NEST time series data.

References

Adel, K., Elhakeem, A., and Marzouk, M. (2023). Decentralized system for construction projects data management using blockchain and IPFS. Journal of Civil Engineering and Management, 29(4):342–359.

Allen, J. and Hunn, P., editors (2022). Smart Legal Contracts: Computable Law in Theory and Practice. Oxford University Press, 1 edition.

Buterin, V. (2014). Ethereum: A Next-Generation Smart Contract and Decentralized Application Platform.

Das, P., Eckey, L., Frassetto, T., Gens, D., Hostáková, K., Jauernig, P., Faust, S., and Sadeghi, A.-R. (2019). FastKitten: Practical Smart Contracts on Bitcoin.

Dursun, T. and Üstündağ, B. B. (2021). A novel framework for policy based on-chain governance of blockchain networks. Information Processing & Management, 58(4):102556.

Gangwal, A., Gangavalli, H. R., and Thirupathi, A. (2023). A survey of Layer-two blockchain protocols. Journal of Network and Computer Applications, 209:103539.

Gupta, P. and Jha, K. N. (2023). A Decentralized and Automated Contracting System Using a Blockchain-Enabled Network of Stakeholders in Construction Megaprojects. Journal of Management in Engineering, 39(4):04023021.

Hamledari, H. and Fischer, M. (2021). Construction payment automation using blockchain-enabled smart contracts and robotic reality capture technologies. Automation in Construction, 132:103926.

Hunhevicz, J. J., Bucher, D. F., Soman, R. K., Honic, M., Hall, D. M., and De Wolf, C. (2023). Web3-based role and token data access: The case of building material passports. In 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference.

Hunhevicz, J. J. and Hall, D. M. (2020). Do you need a blockchain in construction? Use case categories and decision framework for DLT design options. Advanced Engineering Informatics, 45(February):101094.

Hunhevicz, J. J., Motie, M., and Hall, D. M. (2022). Digital building twins and blockchain for performance-based (smart) contracts. Automation in Construction, 133:103981.

Kiayias, A. (2016). Speed-Security Tradeoffs in Blockchain Protocols.

Li, J. and Kassem, M. (2021). Applications of distributed ledger technology (DLT) and Blockchain-enabled smart contracts in construction. Automation in Construction, 132:103955.

Mason, J. (2017). Intelligent Contracts and the Construction Industry. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 9(3):04517012.

McNamara, A. J. and Sepasgozar, S. M. (2021). Intelligent contract adoption in the construction industry: Concept development. Automation in Construction, 122:103452.

Mezquita, Y., Valdeolmillos, D., González-Briones, A., Prieto, J., and Corchado, J. M. (2019). Legal Aspects and Emerging Risks in the Use of Smart Contracts Based on Blockchain. In Uden, L., Ting, I.-H., and Corchado, J. M., editors, Knowledge Management in Organizations, Communications in Computer and Information Science, pages 525–535, Cham. Springer International Publishing.

Msawil, M., Greenwood, D., and Kassem, M. (2022). A Systematic evaluation of blockchain-enabled contract administration in construction projects. Automation in Construction, 143:104553.

Naderi, H., Shojaei, A., and Ly, R. (2023). Autonomous construction safety incentive mechanism using blockchain-enabled tokens and vision-based techniques. Automation in Construction, 153:104959.

Saygili, M., Mert, I. E., and Tokdemir, O. B. (2021). A decentralized structure to reduce and resolve construction disputes in a hybrid blockchain network. Automation in Construction, page 104056.

Scott, D., Ma, L., and Broyd, T. (2024). Project bank account (PBA) decentralised application for the construction industry. Construction Innovation, ahead-of-print(ahead-of-print).

Son, P. V. H. and Lien, P. N. (2022). Blockchain crowdsourced arbitration in construction project delay resolution. Journal of Science and Technology in Civil Engineering (STCE) - HUCE.

Tao, X., Das, M., Liu, Y., and Cheng, J. C. P. (2021). Distributed common data environment using blockchain and Interplanetary File System for secure BIM-based collaborative design. Automation in Construction, 130:103851.

Wüst, K., Diana, L., Kostiainen, K., Karame, G., Matetic, S., and Capkun, S. (2019). Bitcontracts: Supporting Smart Contracts in Legacy Blockchains.

Yang, R., Wakefield, R., Lyu, S., Jayasuriya, S., Han, F., Yi, X., Yang, X., Amarasinghe, G., and Chen, S. (2020). Public and private blockchain in construction business process and information integration. Automation in Construction, 118:103276.

Zou, W., Lo, D., Kochhar, P. S., Le, X.-B. D., Xia, X., Feng, Y., Chen, Z., and Xu, B. (2021). Smart Contract Development: Challenges and Opportunities. IEEE Transactions on Software Engineering, 47(10):2084–2106.

THE ROLE OF BLOCKCHAIN IN ENHANCING TRUST: A CONSTRUCTION PROJECT GOVERNANCE APPROACH

Seongha Hwang[1], Karen Blay[1], Mohamed Osmani[1], and Mingzhu Wang[2]
[1]School of Architecture, Building and Civil Engineering, Loughborough University, UK
[2]Department of Architecture and Civil Engineering, City University of Hong Kong, China

Abstract

Construction project governance (CPG) offers a strategic framework to guide participants through complex construction projects. Emphasizing the trust mechanism within CPG is crucial for managing inter-organizational relationships. Blockchain (BC) can serve as an enabler for trust-based CPG by facilitating and regulating relationships at the project level. Accordingly, this paper aims to identify CPG's trust requirements and assess the relationship between CPG's trust requirements and BC's characteristics through a state-of-the-art review. The results indicate that integrating BC's distinctive characteristics across project delivery can enhance CPG's trust requirements. This study presents comprehensive guidelines for developing a trust-based CPG model.

Introduction

Construction projects face temporal constraints, including limited duration, unique tasks, ambiguous hierarchies, diversity, and ad-hoc coordination (Sergeeva, 2019). These temporal aspects negatively impact long-term engagement, crucial for establishing trust among project stakeholders (Yan and Zhang, 2020). Trust describes relational aspects such as shared confidence and positive expectations that each organization will act in a mutually beneficial way (Li *et al.*, 2021). Trust between project stakeholders is one of the most critical factors for successful construction projects (Qian and Papadonikolaki, 2020). To establish trust, a comprehensive construction contract should clearly outline the requirements, obligations, and specifications of all project stakeholders. However, incomplete contracts arise during construction projects due to project stakeholders' bounded rationality and asymmetric information (Winch, 2010). As a result, continuous adjustment between project stakeholders is necessary throughout the project life cycle (Mansor and Rashid, 2016). As such, incomplete formal contracts with transactional relationships and informal agreements governing complex construction projects failed to ensure high trust at the project level, leading to issues like time delays, cost overruns, and compromised project quality (Liu *et al.*, 2022).

Construction project governance (CPG) offers new insights into the trust challenge from a broader perspective (Lin *et al.*, 2020). It aims to guide construction projects toward meeting stakeholders' goals by inducing collective behaviors through efficient regulation and principles (Müller and Martinsuo, 2015). Providing structures, processes, decision-making frameworks, and project management tools, CPG aligns objectives with each stakeholder's organizational governance models (PMI, 2017). However, the current CPG's rigid monitoring, strict contractual management, and rigorous audit mechanisms decrease communication, transparency, cooperation, and motivation, leading to low trust in construction projects (Lin *et al.*, 2020).

Blockchain (BC) has the potential to establish a trust-based CPG (Xu *et al.*, 2022). BC can facilitate the trust mechanism of CPG by directing network participants' behaviors through autonomous services and effectively governing relationships among BC network participants by enabling reliable information sharing (Lumineau *et al.*, 2021). As a digital ledger technology, BC chronologically records transactions across a peer-to-peer network (Das *et al.*, 2020). It provides real-time transmission for synchronized and immutable data through predetermined consensus algorithms, securing block-linked databases via cryptography and hashing algorithms (Perera *et al.*, 2020). However, existing research has not revealed the relationship between CPG's trust requirements and BC (Xu *et al.*, 2022). Therefore, this research aims to explore the trust requirements of CPG and assess the relationship between these requirements and BC characteristics through a state-of-the-art review.

Research Methodology

This study employed a state-of-the-art literature review to provide insights into current scientific advancements (Barry et al., 2022) and fresh viewpoints on the association between BC characteristics and trust requirements in CPG.

Table 1: Database, inclusion criteria, and search terms for a state-of-the-art review

Category	Details
Database	Scopus and Web of Science
Inclusion criteria	Academic journal papers
Search terms	(“trust” OR “governance” OR “relational governance” OR “relational norm” OR “construction”) AND “blockchain”

The Scopus and Web of Science academic databases were used to source high-quality scholarly articles. Details regarding inclusion criteria and specific search terms are outlined in Table 1.

Figure 1 outlines the methodology of the comprehensive review carried out in this study.

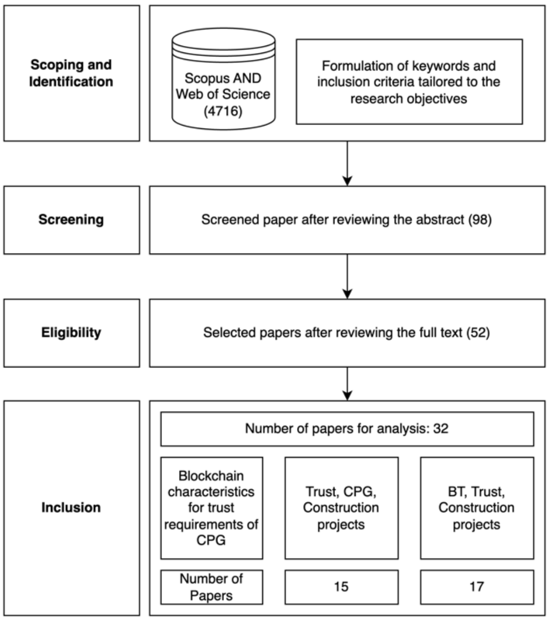


Figure 1: The state-of-the-art review process (devised by authors)

Specifically, the initial step involved a preliminary examination and identification of pertinent literature using specific keywords and inclusion criteria within the chosen database. The second step entailed a screening phase conducted through an abstract analysis. The final selection of 32 papers for in-depth analysis was made after an eligibility assessment based on a full-text review. Hence, the thorough review identified a gap in the existing literature, particularly its lack of exploration into the connection between BC characteristics and trust requirements of CPG.

Trust in Construction Project Governance

A common-pool resource is a shared resource that can be used without causing any harm under favorable conditions (Ostrom, 2019). A construction project can be regarded as a common resource scenario with a shared resource pool, decision-making rights, and risk/reward. Common-pool resource scenarios governed by a top-down governance approach exhibit multiple systematic failures (Hall and Bonanomi, 2021). However, the current CPG is biased towards top-down project governance using contractual project governance rather than bottom-up project governance. This top-down governance approach using rigorous monitoring processes and strict contract terms and conditions decreases trust at the project level (Sergeeva, 2019).

Additionally, according to agency theory and transaction cost economics, construction project stakeholders can be opportunistic, selfish, and possess bounded rationality in the common-pool resource contexts (Winch, 2010; Hall and Bonanomi, 2021). Trust is critical to limit these negative features of legally intertwined entities by facilitating mutual benefits, knowledge sharing, and emotional bonds (Laan *et al.*, 2012). As such, trust can improve project performance by achieving unique objectives within a set period and controlling uncertainty and fragmentation in complicated construction processes (Lu *et al.*, 2021b). In particular, trust is essential when uncertain situations and issues bearing risk occur. It can lead to a reliable relationship based on practical cooperation (Lakusic, 2021). However, fragmented construction processes, the temporality of projects, and project stakeholders' opportunistic behaviors hinder the establishment of trust in the construction project environments (Ke *et al.*, 2015).

Furthermore, self-organization explains the phenomenon where many agents interact with each other in a disordered complex situation but increase order and regularity through the interaction between spontaneous agents. Due to the uncertainties of the construction contract system, the agents participating in the construction project voluntarily form a self-organizing network to address these uncertainties (Steen *et al.*, 2018). Specifically, stigmergy is an example of self-organization involving indirect communication and learning processes through non-linear rules and methods. This provides a basis for understanding social agent interaction that induces complex patterns (Ramos and Abraham, 2004). Stigmergic principles need more flexible project governance structures and play an important role in building trust in construction projects by providing reliable coordination solutions and inducing collective actions within limited information-sharing environments (e.g., construction projects) (Dounas *et al.*, 2022).

Trust Requirements of Construction Project Governance

While agency theory and transaction cost economics as project governance theories provide the contexts of low trust in construction projects involving multiple project stakeholders, contingency and network theories provide the foundation for conceptualizing and understanding CPG's three trust requirements (mutuality, flexibility, and solidarity) derived from relational norms (Lu et al., 2015; Musawir et al., 2020):

- Mutuality: Willingness to mutually improve the current situation compared to the previous situation.

- Flexibility: Willingness to accept adjustments and modifications according to changed circumstances.
- Solidarity: Willingness to maintain and stabilize partner relationships.

Table 2 collates trust requirements underpinned by contingency and network theories, from the 15 reviewed papers on Trust CPG, Construction projects.

Table 2: Research on trust requirements of CPG contexts

Authors	Trust requirements		
	Mutuality	Flexibility	Solidarity
(Tian *et al.*, 2023)		✓	✓
(Paswan *et al.*, 2017)	✓		✓
(Liu *et al.*, 2022)		✓	✓
(Benítez-Ávila *et al.*, 2019)		✓	✓
(Zheng *et al.*, 2008)		✓	✓
(Chakkol *et al.*, 2018)		✓	
(Xu *et al.*, 2022)	✓	✓	✓
(Lu et al., 2015)	✓	✓	✓
(Yang *et al.*, 2022)	✓	✓	✓
(Lin *et al.*, 2020)	✓		
(Haq *et al.*, 2019)		✓	✓
(Müller and Martinsuo, 2015)		✓	✓
(Benítez-Ávila *et al.*, 2018)		✓	✓
(Bonatto *et al.*, 2020)		✓	✓
(Cao and Lumineau, 2015)		✓	✓

Contingency theory focuses on organizational effectiveness, described as the degree of alignment between the organizational characteristics and contingencies reflecting an organization's internal and external environment (Lizarralde *et al.*, 2011). Contingency theory posits that contingency variables (such as governance, size, culture, strategies, stakeholder motivations, or legal frames) align with the changing internal and external environment to improve organizational performance (such as time, cost, quality, scope, and benefit) and thus facilitate mutuality and solidarity (Deng and Smyth, 2013). The dynamic and flexible nature of contingency theory through its continual response to changing environments, suits the examination of various types of construction projects from diverse and complex backgrounds (Hanisch and Wald, 2012). However, construction project's temporality including limited duration, featuring ambiguous hierarchies, diversity, and informal coordination hinders the establishment of trust at the project level (Deng and Smyth, 2013). According to the contingency, CPG should highlight three relational norms: mutuality, flexibility, and solidarity among construction project stakeholders. These relational norms enhance trust by increasing responsiveness to the internal and external environments at the project level (Musawir *et al.*, 2020; Xu *et al.*, 2022). In addition to contingency theory, network theory emphasizes the efficiency of a network comprising multiple stakeholders participating in a construction project (Wang and Yin, 2023). Like the contingency theory, the network theory emphasizes flexibility, solidarity, and mutuality of relational norms as trust requirements through networking and a network's efficiency (Musawir *et al.*, 2020). Network theory is a subset of graph theory, which is the study of the properties of graphs and the mathematical definition of networks. In general terms, a graph is composed of nodes (objects that constitute the graph) and edges, representing the relationship between the nodes. Graph theory is used to analyze connected node data through various graph algorithms, such as basic statistics, graph data queries, visual exploration, and machine learning (Needham and Hodler, 2019). Similarly, network theory focuses on research into the connection patterns, network structures, node positions, and node outcomes of actors corresponding to the nodes in the graph. Accordingly, network theory is used in social science to study the characteristics of human society, such as group phenomena and human communications (Lu *et al.*, 2015b). Hence, network theory is widely used in project governance, requiring network thinking to study complex construction project networks (Li *et al.*, 2020). However, the non-linear execution of the construction project network, self-organization of multiple stakeholders within that network, and various emergent project situations can decrease network efficiency.

In response to these trust requirements, BC has been focused on as a potential solution to meet CPG's trust requirements (Lumineau *et al.*, 2021; Xu *et al.*, 2022).

Fundamentals of Blockchain Technology

BC consists of four fundamental technologies (Nawari and Ravindran, 2019; Perera *et al.*, 2020):

- P2P networks enable nodes to efficiently store, share, and manage files in a decentralized manner, enhancing speed and security in various digital services without requiring a central server.
- Hashing algorithms assure consistent outputs for identical inputs and resistance to reverse calculation, linking multiple transactions within a BC.

- Cryptography secures data in BC networks, employing either symmetric methods (fast but less secure) or asymmetric methods (more secure but computationally intensive).
- Consensus algorithms aim to maintain transaction orders even in adversarial environments by ensuring agreement among nodes on a single piece of data. These algorithms prevent malicious actions, addressing the Byzantine Generals' Problem, describing how Byzantine troops attacked a completely enclosed city.

Characteristics of Blockchain Technology in Construction Project Governance Contexts

A permissioned blockchain, known as a consortium blockchain, is appropriate for governance contexts in construction projects (Zhong *et al.*, 2020). A synthesis of literature from the 17 reviewed papers on BC, trust and CPG revealed the fundamentals of BC lead to six significant BC characteristics (see Table 3) in the permissioned BC contexts:

Table 3: The characteristics of blockchain in the construction project governance contexts

Characteristics	Details
Autonomy	Blockchain offers a range of autonomous services using smart contracts, which function as computerized protocols for business logics (Nawari and Ravindran, 2019).
Decentralization	Blockchain, as a digital ledger system, operates without a central administrator and a centralized data storage framework (Perera *et al.*, 2020).
Immutability	Once data are added in the blockchain networks, transactions on the blockchain cannot be canceled or altered (Das *et al.*, 2020).
Security	Blockchain protects data and prevents fraud by using public keys for network transactions and private keys for managing confidential information (Ciotta *et al.*, 2021).
Traceability	Every transaction on the blockchain is accurately recorded and timestamped, allowing network participants to access and track these records on any node (Msawil *et al.*, 2022).
Transparency	Authorized participants in the blockchain networks or channels have access to the same version of data (Teisserenc and Sepasgozar, 2021).

The Relationship Between Trust Requirements of Construction Project Governance and Characteristics of Blockchain Technology

There are strong relationships between CPG trust requirements and BC characteristics in construction project contexts characterized by a common pool resource and stigmergy (see Figure 2).

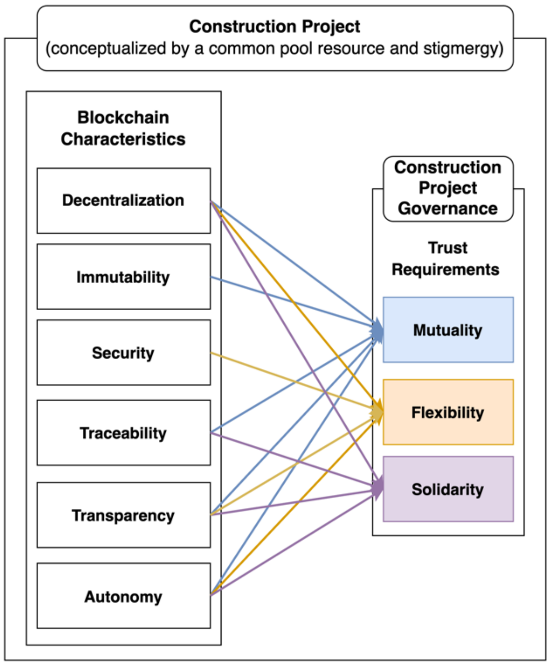


Figure 2: Blockchain characteristics and trust requirements of construction project governance (devised by authors)

First, decentralization, transparency, traceability, immutability, and autonomy are key factors that can enhance mutual relationships among project stakeholders. Decentralization can promote mutuality by creating interconnected networks that are directly operated by distributed entities without any central authority at the project level (Ciotta *et al.*, 2021). Transparency and traceability allow authorized project stakeholders to verify the actions of other network participants, thus promoting mutual trust (Das *et al.*, 2022). Immutability is also vital in mutual relationships, as it ensures that data is not altered or tampered with (Lu *et al.*, 2021b). Autonomy can positively impact mutuality by enabling transparent obligations and increasing the commitment of network participants through smart contracts with coded terms (Xu *et al.*, 2022). These transparent obligations and increased commitments align with the principle of mutuality (Nawari and Ravindran, 2019). Thus, blockchain technology can help establish mutual relationships that are based on a willingness to improve the current situation at the project level (Perera *et al.*, 2020). Second, decentralization, transparency, security, and autonomy can significantly improve flexibility in the context of CPG (Ciotta *et al.*, 2021). Compared to centralized systems, BC's decentralized architecture enhances flexibility in terms of response time against the risk of system failures or attacks (Fiorentino and Bartolucci, 2021). On the other hand, transparency and security enable reliable and secure information sharing, which positively impacts behavioral flexibility by allowing project stakeholders to respond quickly and efficiently to internal and external changes (Wang and Yang, 2022). Likewise, autonomy increases flexibility for data integration into a secure network by governing data access control within the construction project network (Msawil *et al.*, 2022). As a result, blockchain-enabled flexibility can help construction project stakeholders adapt effectively to any adjustments against rapidly changing external environments (Perera *et al.*, 2020).

Finally, decentralization, transparency, traceability, and autonomy are important factors that contribute to solidarity among stakeholders in construction projects (Xu *et al.*, 2022). Decentralization and autonomy enable the democratization of control in the CPG context, which promotes equal and participatory decision-making processes (Das *et al.*, 2022). This, in turn, leads to better solidarity through collaboration among project stakeholders (Perera *et al.*, 2020). For example, decentralized autonomous organizations and voting applications through BC ensure that stakeholders work together through a fair democratic system for collective decision-making (Nawari and Ravindran, 2019). Transparency and traceability are essential for promoting ethical practices and establishing an accountable cultural system in construction projects. This positive culture promotes solidarity from the project stakeholders at the project level (Li *et al.*, 2019). As a result, the strengthened solidarity positively impacts stakeholders' willingness to maintain and stabilize relationships (Yang *et al.*, 2022).

Hence, focusing on CPG trust requirements (mutuality, flexibility, and solidarity) through BC enables flexible bottom-up rather than top-down governance, enhancing trust at the project level (Xu *et al.*, 2022). Accordingly, the network-based bottom-up project governance through BC aligns with the stigmergy of construction projects as a common pool resource (Dounas *et al.*, 2022; Hall and Bonanomi, 2021).

Discussion

The first objective of this study was to explore CPG's trust requirements. The findings of this research indicated that mutuality, flexibility, and solidarity derived from relational norms are significant trust requirements in the CPG context. These results are aligned with those of previous studies about relational norms in the CPG context. For instance, Benítez-Ávila et al. (2018) argued that the CPG trust can be enhanced by mediating the effect of formal CPG designs and increasing project coalitions' capacity to coordinate tasks and reach high levels of cooperation. Similarly, mutual association through established normative practices and expectations in construction projects facilitates trust. The relational norms underpinning trust requirements yield positive outcomes and foster a high degree of trust by curbing opportunistic behaviors in CPG environments (Xu *et al.*, 2021). However, previous studies have failed to define CPG trust requirements and the theories underpinning them. The study highlights the importance of utilizing a flexible bottom-up project governance approach for construction projects conceptualized as a common-pool resource and stigmergy instead of top-down governance approaches. Additionally, the examination of contingency theory as a governance theory revealed that the current CPG system diminishes trust in construction projects due to its inability to swiftly adapt to rapid changes in the construction project's internal and external environments. Likewise, network theory suggests that traditional CPG systems fall short in facilitating effective communication among project stakeholders in the intricate networks of construction projects, which erodes trust. In particular, project clients and main contractors (Li *et al.*, 2020), dealing with intricate interests involving multiple stakeholders, prioritize efficient and rapid decision-making in CPG. They should also emphasize network-based project management, scrutinizing network connection patterns, structures, node positions, and outcomes within the multifaceted networks of construction projects.

Nevertheless, previous research on blockchain-based CPG tended to focus on the control aspects of CPG, which emphasizes the relationship between BC and the tools of control-based CPG: contract management (Kim *et al.*, 2022), procurement management (Gupta and Jha, 2023), project assurance (Lu *et al.*, 2021a; Das *et al.*, 2022) and quality assurance (Lu *et al.*, 2022) – to limit opportunism in the construction network. In response to the limitations of previous studies, this study aimed to assess the relationship between CPG trust requirements and BC characteristics. This state-of-the-art review introduces new knowledge to CPG research by identifying that BC's six distinctive characteristics can increase CPG's three trust requirements. According to Ostrom (2019), there are many instances of common-pool resource situations where top-down governance has led to significant and consistent failures. In response to this project governance challenge, the result theoretically presents a basis for establishing a BC-aided relational CPG that can guide the reliable coordination and collective behaviors of construction project stakeholders through the bottom-up approach. The relational CPG through BC leads to structural relationships between network participants, enhancing trust through collaborative project networks under uncertain project environments (Liu *et al.*, 2022).

However, developing relational governance under complex construction environments demands significant time investment and resource-intensive social processes from project stakeholders (Xu *et al.*, 2022). In response to these challenges, this study provided the practical trust-based CPG model through BC (see Figure 3). The

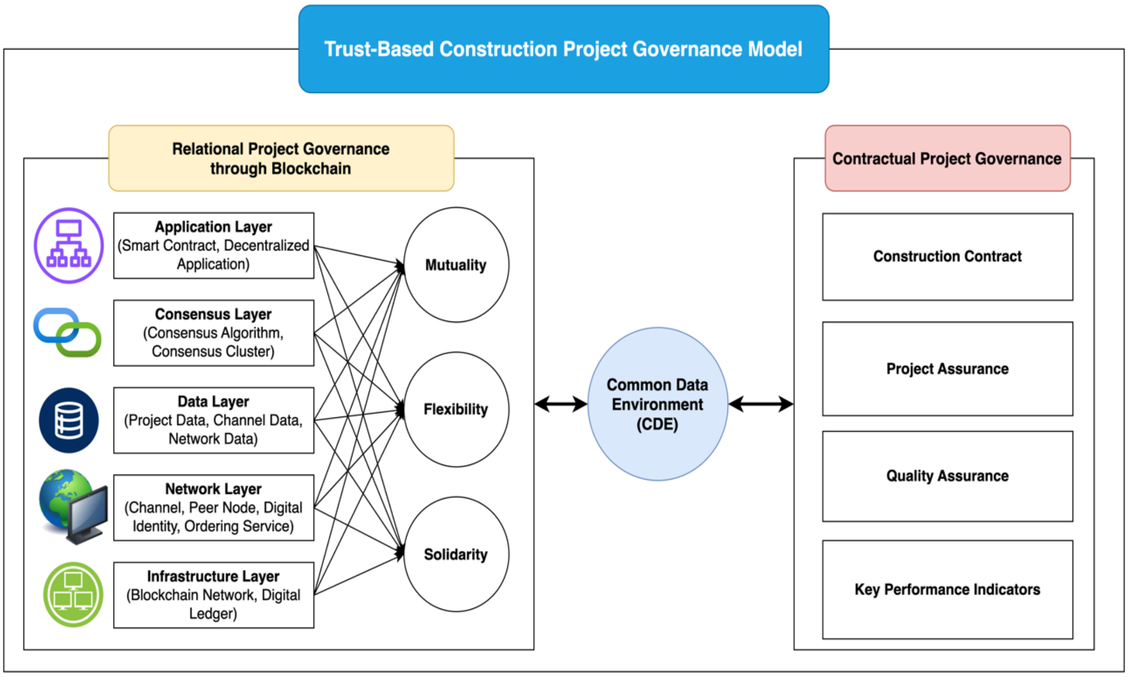


Figure 3: Trust-Based Construction Project Governance Model

findings indicate the possibility of developing a practical trust-based CPG model. Specifically, BC channels, nodes, orderers, and ledgers (Zhong *et al.*, 2020) can be configured, taking into account the unique context of each construction project. Additionally, a common data environment (CDE) can be used to connect existing CPG systems using construction contracts, project assurance, quality assurance and key performance indicators (KPIs). Trust among project stakeholders can be improved by securely storing, validating, and accessing project information using smart contracts. Ultimately, this trust-based CPG with BC can satisfy three trust requirements (mutuality, flexibility, and solidarity) of CPG. Therefore, this blockchain-enabled CPG, which is characterized by reliable decentralization, can efficiently govern and direct construction projects by enabling systematic self-organization and stigmagy environments. In particular, this system can provide novel insights allowing key project stakeholders (clients, main contractors, and consultants) to establish a trust CPG system while maintaining the existing governance system, including construction contracts and project assurance (Sergeeva, 2019).

Conclusions

This study, which adopted a state-of-the-art review, aimed to explore CPG's trust requirements and assess the relationship between CPG's trust requirements and BC's characteristics in construction projects distinguished by a common pool resource and stigmergy. The study revealed that CPG has three trust requirements: mutuality, flexibility, and solidarity. The distinctive six characteristics of BC; autonomy, decentralization, immutability, security, traceability and transparency can improve these trust requirements. Consequently, the findings bridge the knowledge gap concerning the trust requirements of CPG and their association with BC. The result theoretically lays the groundwork for establishing a BC-supported relational CPG, guiding stakeholder behavior in construction projects. The findings also indicate the potential for a practical, trust-based CPG model by using BC. This research serves as a foundational step toward achieving fully automated, trust-based CPG systems.

References

Barry, E.S. et al. (2022) State-of-the-art literature review methodology: six-step approach for knowledge synthesis. *Perspectives on Medical Education*, 11(5), 1–8.

Benítez-Ávila, C. et al. (2018) Interplay of relational and contractual governance in public-private partnerships: The mediating role of relational norms, trust and partners' contribution. *International Journal of Project Management*, 36(3), 429–443.

Benítez-Ávila, C. et al. (2019) Contractual and Relational Governance as Positioned-Practices in Ongoing Public–Private Partnership Projects. *Project Management Journal*, 50(6), 716–733.

Bonatto, F. et al. (2020) Relational governance in supply chain: a systematic literature review. *Benchmarking: An International Journal*, 27(6), 1711–1741.

Cao, Z. & Lumineau, F. (2015) Revisiting the interplay between contractual and relational governance: A qualitative and meta-analytic investigation. *Journal of Operations Management*, 33–34(1), 15–42.

Chakkol, M. et al. (2018) The governance of collaboration in complex projects. *International Journal of Operations & Production Management*, 38(4), 997–1019.

Ciotta, V. et al. (2021) Integration of blockchains and smart contracts into construction information flows: Proof-of-concept. *Automation in Construction*, 132, 103925.

Das, M. et al. (2020) Securing interim payments in construction projects through a blockchain-based framework. *Automation in Construction*, 118, 103284.

Das, M. et al. (2022) A blockchain-based integrated document management framework for construction applications. *Automation in Construction*, 133, 104001.

Deng, F. & Smyth, H. (2013) Contingency-Based Approach to Firm Performance in Construction: Critical Review of Empirical Research. *Journal of Construction Engineering and Management*, 139(10), 04013004.

Dounas, T. et al. (2022) *The Architecture Decentralised Autonomous Organisation - A stigmergic exploration in architectural collaboration*. In: eCAADe 2022: Co-creating the Future - Inclusion in and through Design. Ghent, Belgium, 2022. Ghent, Belgium. Available at: doi:10.52842/conf.ecaade.2022.1.567 [Accessed: 28 March 2024].

Fiorentino, S. & Bartolucci, S. (2021) Blockchain-based smart contracts as new governance tools for the sharing economy. *Cities*, 117, 103325.

Gupta, P. & Jha, K.N. (2023) A Decentralized and Automated Contracting System Using a Blockchain-Enabled Network of Stakeholders in Construction Megaprojects. *Journal of Management in Engineering*, 39(4), 04023021.

Hall, D.M. & Bonanomi, M.M. (2021) Governing Collaborative Project Delivery as a Common-Pool Resource Scenario. *Project Management Journal*, 52(3), 250–263.

Hanisch, B. & Wald, A. (2012) A Bibliometric View on the Use of Contingency Theory in Project Management Research. *Project Management Journal*, 43(3), 4–23.

Haq, S.U. et al. (2019) Project governance mechanisms and the performance of software development projects: Moderating role of requirements risk. *International Journal of Project Management*, 37(4), 533–548.

Ke, H. et al. (2015) The Impact of Contractual Governance and Trust on EPC projects in Construction Supply Chain Performance. *Engineering Economics*, 26(4), 349–363.

Kim, E.W. et al. (2022) Blockchain-Based Automatic Tracking and Extracting Construction Document for Claim and Dispute Support. *KSCE Journal of Civil Engineering*, 26(9), 3707–3724.

Laan, A. et al. (2012) Levels of Interorganizational Trust in Construction Projects: Empirical Evidence. *Journal of Construction Engineering and Management*, 138(7), 821–831.

Lakusic, S. (ed.) (2021) Reducing information asymmetry and building trust in projects using blockchain technology. *Journal of the Croatian Association of Civil Engineers*, 73(10), 967–978.

Li, J. et al. (2019) Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. *Automation in Construction*, 102, 288–307.

Li, J. et al. (2020) The effects of trust network among project participants on project performance based on SNA approach: a case study in China. *International Journal of Construction Management*, 20(8), 837–847.

Li, Y. et al. (2021) Influencing factors on inter-organizational trust asymmetry behavior in construction projects Evidence from China. *Engineering, Construction and Architectural Management*, 28(1), 308–331.

Lin, Y.-H. et al. (2020) The impact of relational governance on the adaptability of international contractors: a comparative study between China and Korea. *Engineering, Construction and Architectural Management*, 27(10), 3235–3259.

Liu, T. et al. (2022) Effects of Contractual and Relational Governances on BIM Collaboration and Implementation for Project Performance Improvement. *Journal of Construction Engineering and Management*, 148(6), 04022029. American Society of Civil Engineers.

Lizarralde, G. et al. (2011) Structuring of Temporary Multi-Organizations: Contingency Theory in the Building Sector. *Project Management Journal*, 42(4), 19–36.

Lu, P. et al. (2015a) The effectiveness of contractual and relational governances in construction projects in China. *International Journal of Project Management*, 33(1), 212–222.

Lu, W. et al. (2021a) Blockchain Technology for Governmental Supervision of Construction Work: Learning from Digital Currency Electronic Payment Systems. *Journal of Construction Engineering and Management*, 147(10), 04021122.

Lu, W. et al. (2021b) Rebuilding trust in the construction industry: a blockchain-based deployment framework. *International Journal of Construction Management*, 1–12.

Lu, W. et al. (2022) Construction E-Inspection 2.0 in the COVID-19 Pandemic Era: A Blockchain-Based Technical Solution. *Journal of Management in Engineering*, 38(4), 04022032.

Lu, Y. et al. (2015b) Organizational Network Evolution and Governance Strategies in Megaprojects. *Construction Economics and Building*, 15(3), 19–33.

Lumineau, F. et al. (2021) Blockchain Governance—A New Way of Organizing Collaborations? *Organization Science*, 32(2), 500–521.

Mansor, N.S. & Rashid, K.A. (2016) Incomplete Contract in Private Finance Initiative (PFI) Contracts: Causes, Implications and Strategies. *Procedia - Social and Behavioral Sciences*, 222, 93–102.

Msawil, M. et al. (2022) A Systematic evaluation of blockchain-enabled contract administration in construction projects. *Automation in Construction*, 143, 104553.

Müller, R. & Martinsuo, M. (2015) The impact of relational norms on information technology project success and its moderation through project governance. *International Journal of Managing Projects in Business*, 8(1), 154–176.

Musawir, A. et al. (2020) Project governance and its role in enabling organizational strategy implementation: A systematic literature review. *International Journal of Project Management*, 38(1), 1–16.

Nawari, N.O. & Ravindran, S. (2019) Blockchain and the built environment: Potentials and limitations. *Journal of Building Engineering*, 25, 100832.

Needham, M. & Hodler, A.E. (2019) *Graph algorithms: practical examples in Apache Spark and Neo4j*. First edition. Sebastopol, California: O'Reilly Media.

Ostrom, E. (2019) *Governing the commons: the evolution of institutions for collective action*. 10th printing. Cambridge: Cambridge University Press.

Paswan, A.K. et al. (2017) Opportunism, governance structure and relational norms: An interactive perspective. *Journal of Business Research*, 77, 131–139.

Perera, S. et al. (2020) Blockchain technology: Is it hype or real in the construction industry? *Journal of Industrial Information Integration*, 17, 100125.

PMI (ed.) (2017) *A guide to the project management body of knowledge*. Sixth edition. Newtown Square, PA: Project Management Institute.

Qian, X. (Alice) & Papadonikolaki, E. (2020) Shifting trust in construction supply chains through blockchain technology. *Engineering, Construction and Architectural Management*, 28(2), 584–602.

Ramos, V. & Abraham, A. (2004) *Evolving a Stigmergic Self-Organized Data-Mining*. [object Object]. [Accessed: 28 March 2024].

Sergeeva, N. (2019) Towards more flexible approach to governance to allow innovation: the case of UK infrastructure. *International Journal of Managing Projects in Business*, 13(1), 1–19.

Steen, J. et al. (2018) Projects and Networks: Understanding Resource Flows and Governance of Temporary Organizations with Quantitative and Qualitative Research Methods. *Project Management Journal*, 49(2), 3–17.

Teisserenc, B. & Sepasgozar, S. (2021) Adoption of Blockchain Technology through Digital Twins in the Construction Industry 4.0: A PESTELS Approach. *Buildings*, 11(12), 670.

Tian, B. et al. (2023) Can relational governance improve sustainability in public-private partnership infrastructure projects? An empirical study based on structural equation modeling. *Engineering, Construction and Architectural Management*, 30(1), 19–40.

Wang, M. & Yang, Y. (2022) An empirical analysis of the supply chain flexibility using blockchain technology. *Frontiers in Psychology*, 13, 1004007.

Wang, X. & Yin, Y. (2023) Structural Dimensions and Measurement of Trust Networks among Construction Project Participants. *Sustainability*, 15(5), 4112.

Winch, G. (2010) *Managing construction projects: an information processing approach*. 2nd ed. Chichester ; Ames, Iowa: Blackwell Pub.

Xu, D. et al. (2022) Leveraging digital and relational governance mechanisms in developing trusting supply chain relationships: the interplay between blockchain and norm of solidarity. *International Journal of Operations & Production Management*, 42(12), 1878–1904.

Xu, J. et al. (2021) Towards the dynamics of trust in the relationship between project-based firms and suppliers. *International Journal of Project Management*, 39(1), 32–44.

Yan, L. & Zhang, L. (2020) Interplay of Contractual Governance and Trust in Improving Construction Project Performance: Dynamic Perspective. *Journal of Management in Engineering*, 36(4).

Yang, X. et al. (2022) Prior and governed stakeholder relationships: The key to resilience of inter-organizational projects. *International Journal of Project Management*, 40(1), 64–75.

Zheng, J. et al. (2008) The dynamics of contractual and relational governance: Evidence from long-term public–private procurement arrangements. *Journal of Purchasing and Supply Management*, 14(1), 43–54.

Zhong, B. et al. (2020) Hyperledger fabric-based consortium blockchain for construction quality information management. *Frontiers of Engineering Management*, 7(4), 512–527.

COULD GREEN BUILDING CERTIFICATION SYSTEMS BENEFIT FROM THE INTRODUCTION OF BLOCKCHAIN? A SYSTEMATIC REVIEW

Akman, Veysel Hüseyin*[1] ; Dounas, Theodoros [2]
[1]University of Warsaw, Poland; [2]University of Antwerp, Belgium

Abstract

We investigate the Green Building Incentive mechanisms to overcome identified barriers in adoption. With the aim of increasing the overall efficiency of the certification system, integration with the blockchain is proposed. Further, the blockchain concept and its applications are introduced. Blockchain-related incentive mechanisms and their intricate relationship with design and governance are also highlighted, through a systematic review. We conclude with the need for a stakeholder survey for the most suitable blockchain design and governance.

Introduction

The construction industry's impact on sustainability highlights the need for green buildings Jayakody and Vaz (2023). Measuring a building's sustainability performance becomes crucial as we strive to mitigate negative consequences. In response, several versions of the Green Building Certification System (GBCS) have evolved to measure a building's sustainability performance Braulio-Gonzalo et al. (2022). Despite its importance, barriers hinder the widespread adoption of green building certification Agyekum et al. (2019). In response, incentive mechanisms have been introduced, but their effectiveness still needs to be improved Olubunmi et al. (2016). Lack of sufficient motivation for green building certification drives exploration into innovative solutions, with blockchain emerging as a promising option.

Blockchain technology introduces a decentralized and secure way to manage records Turk and Klinc (2017). With their ability to represent various incentives, blockchain applications expand to the construction industry, offering viable solutions to the barriers encountered by GBCS . However, an effective application for GBCS requires careful consideration of BC design and governance Fan and Wu (2020).

This article explores the relationship between blockchain, governance, and incentives, examining their roles in potentially improving green building certification adoption. Investigating stakeholder perspectives becomes decisive because their preferences are the main driver for a tailored blockchain design in the green building certification area. This study reveals blockchain's potential to advance GBCS by promoting sustainability through efficient governance and alternative incentive mechanisms.

The article's structure is as follows: in the first section, we delve into green building certification and examine barriers to its adoption. Moving to the second section, we introduce blockchain technology and its application requirements. In the third section, we investigate the capabilities of blockchain in GBCS. Finally, we conclude by harmonizing key insights and proposing future research.

Methodology

To conduct a literature review and define the knowledge gap, we researched relevant articles in Google Scholar. We utilized Boolean operations to filter our search with the following keywords: "blockchain AND green building certification AND incentive mechanisms", "blockchain AND sustainable building AND financial incentives OR non-financial incentives", and "blockchain AND green building certification AND financial incentives OR non-financial incentives". These keywords were determined based on their relevance and incorporation with the key topics of research; blockchain, green building certification, and incentive mechanisms. We also applied the snowballing technique to improve the literature review's comprehensiveness. During the primary research, we identified two key articles and then extended the literature pool by systematically examining the references cited in these articles. This iterative process helped to collect broader relevant literature. Eventually, in our literature review process, we identified 82 relevant papers. These papers are examined to identify their alignment with the research objectives and key topics. Only articles published in English are taken into account. Articles published before 2000 are excluded from this study to provide a contemporary and relevant analysis. Relevant doctorate and masters' theses are also not included in the scope to focus and prioritize peer-reviewed research publications. For purposes of this research, we selected to present papers discussing the intersection of blockchain and green building certification systems, primarily focusing on the incentive mechanisms and their impact on sustainability and innovation in construction.

Why Green Buildings are Important?

The building sector is criticized due to its high energy and resource consumption. A recent study by Jayakody and Vaz (2023) stated that buildings constitute nearly 70% of the total energy consumption in the United States and around 35% globally. In addition to energy usage, buildings have a significant impact on other natural resources, contributing to 35% of greenhouse gas emissions, 50% of extracted materials, and approximately one-third of both water consumption and waste worldwide, as highlighted in another recent study by Braulio-Gonzalo et al. (2022).

The concept of green building has emerged as a response to the adverse impacts of construction on the three pillars of sustainability: environmental, social, and economic. It is a primary tool for advancing sustainability objectives within the construction industry as stated by Karji et al. (2021). Green buildings, alternatively called eco-friendly or low-energy buildings, are intentionally designed to alleviate stress on resources by limiting negative impacts on human health and addressing challenges like resource scarcity.

Green building contribution extends beyond environmental considerations; it is inherently connected to social sustainability. Green buildings support creating a healthy environment, fostering community, and boosting human productivity. As stated in Goh et al. (2020), they "complement their environmental impact, creating sustainable and livable spaces that promote a higher quality of life for individuals and communities."

Embracing green building practices has a substantial impact on all three sustainability pillars. Beyond the environmental and social perspectives, green buildings present economic advantages. Effective implementations result in cost savings throughout their lifespan and offer stakeholders the chance for increased profits. In essence, green building practices provide noticeable benefits both environmentally, socially and economically Olubunmi et al. (2016).

What is the GBCS?

Measuring the performance of development is fundamental for ensuring its contribution by establishing standards, aligning goals, and identifying successes and challenges. Braulio-Gonzalo et al. (2022) details that performance measurement optimizes resource allocation, supports the decision-making process, and exhibits the tangible impact of development initiatives.

Standards and frameworks used in the sustainability assessment are typically established by internationally recognized organizations such as ISO, GRI, etc. These institutions develop and define global standards that become laws or serve as the foundation for industry. Standards specify requirements in two ways: prescriptive standards provide methods of achieving objectives, while performance-based standards outline expectations for the desired outcomes Vierra (2016). To illustrate with energy consumption, prescriptive standards require using methods like efficient LEDs. In contrast, performance-based standards set a goal for energy deduction without specifying methods, allowing flexibility in solution.

GBCS emerged to establish standardized practices in construction, offering a common framework to evaluate and differentiate buildings that follow sustainability standards. Over time, these systems expanded to address not only environmental concerns but also social and economic, reflecting a broader understanding of sustainability as stated by Awadh (2017). While implementing these global standards locally, adjustments can be made according to local requirements. But yet, the overarching goal is to encourage and recognize practices that contribute to a holistic sustainability approach in building.

The Building Research Establishment Environmental Assessment Method (BREEAM) arose in the early 1990s as one of the first comprehensive green building certification systems in the United Kingdom. Following the success of BREEAM, other countries and regions established their own green building certification processes: LEED (Leadership in Energy and Environmental Design) in the United States, Green Star in Australia, and Estidama in the United Arab Emirates are a few examples which are extended their usage beyond the countries. All these initiatives have developed throughout time, adding social and economic aspects into their criteria to give a comprehensive approach to building sector sustainability Braulio-Gonzalo et al. (2022).

The functional benefits of green building begin with enhancing companies' competitiveness in the market. Tan et al. (2011) explored how green building certificates enhance contractors' competitiveness. The study revealed that "good corporate governance of environmental and social issues not only enhances companies' shareholder value but also safeguards their highly valuable reputations". It's crucial to emphasize that the advantages of these certificates go beyond contractors, benefiting all stakeholders at different levels. In addition to being a valuable marketing tool for developers, Vierra (2016) highlights the market desire for green building certification.

As mentioned, requirements of green building certification systems aren't universally applicable due to project dynamics, making one more suitable than another based on factors like location, budget, and goals. Selecting an appropriate certification system involves considering essential aspects such as cost, ease of submission, and building performance to determine the most relevant rating system and achievable certification level Vierra (2016). It's crucial to note that each green building certificate comes with distinct specifications, which may vary based on project requirements, whether for refurbishment, new construction, or specific needs. Obtaining green certification requires teamwork across various functions such as supply chain, planning, design, and manufacturing. Yet, the unique aspects of green construction and certification pose challenges, with cost as the main obstacle hindering broader adoption of these innovative practices, as highlighted by Agyekum et al. (2019).

Barriers to GBCS Adoption

Barriers to adopting sustainable practices exist across various dimensions, including economic, social, cultural, political, and systemic factors, as identified by Horner and Ryan (2019). Moreover, the type and intensity of these barriers vary among countries, contributing to distinctions in their quantity and severity Agyekum et al. (2019). Yet, it is essential to carefully analyze the requirements that provide guidance and an implementation plan for stakeholders. When examining current GBCS' characteristics, key requirements centered around energy efficiency, water preservation, minimum negative environmental impact, reliance on renewables and sustainable materials, and efficient space use.

Several studies have investigated the barriers hindering the widespread adoption of green building certificates. Fan and Wu (2020) identified obstacles as high upfront costs, limited access to capital, social and psychological barriers, and a lack of information, experience, and effective incentives. Another study by Karji et al. (2021) indicated common obstacles, including a lack of green building knowledge, high costs, insufficient incentives, lack of motivation, and insufficiency in policies and regulations. Similarly, Agyekum et al. (2019) highlighted various barriers highlighting regulatory deficiencies, inadequate training, developers' hesitation due to perceived additional costs, high investment cost barriers, and challenges in procurement and project-delivery mechanisms.

Considering local-level challenges, Shen et al. (2018) investigated barriers in Thailand using a Likert-scale survey among different stakeholders. The findings show social barriers, with the lack of owner motivation ranking as the major obstacle among all participants. Collectively, these studies underscore the multifaceted challenges impeding the widespread adoption of green building certificates, encompassing knowledge gaps, financial considerations, policy shortcomings, and social factors.

Furthermore, it is important to mention that traditional construction practices have been in place for many years. Transitioning from conventional to sustainable and digital methods will require significant time and changes across individual, organizational, and industry levels. For this reason, this transformation will not be without its hurdles Diyana et al. (2013) Li et al. (2019).

All stakeholders acknowledge the above-mentioned barriers, and as a response, various financial and non-financial incentive mechanisms have been introduced to motivate stakeholders. These incentives, typically offered by governmental entities, environmental agencies, and industry associations, are designed to encourage sustainable building practices and contribute to overarching sustainability objectives.

Incentive mechanism

An incentive mechanism is a structured strategy to encourage specific behaviors, actions, or outcomes by offering rewards or penalties. This concept has extensive application across several disciplines, leveraging components to drive individuals, organizations, or entities toward desired outcomes. In the context of green building, incentive mechanisms play a vital role in motivating developers to embrace sustainable construction Fan and Wu (2020). The eligibility of the incentive mechanism is mostly connected to the green building certification system because the incentive promoters are able to differentiate the project to award.

The Figure 1 illustrates the categorization of green building incentive mechanisms, highlighting their diverse nature and underscoring the necessity for a case-by-case approach. The literature created a spotlight on effectiveness. The reason for such a spotlight is that evaluating the efficiency of incentive mechanisms contributes valuable insights to understanding how different incentives influence stakeholders. The literature delves into the effectiveness of incentive mechanisms, often through case scenarios.

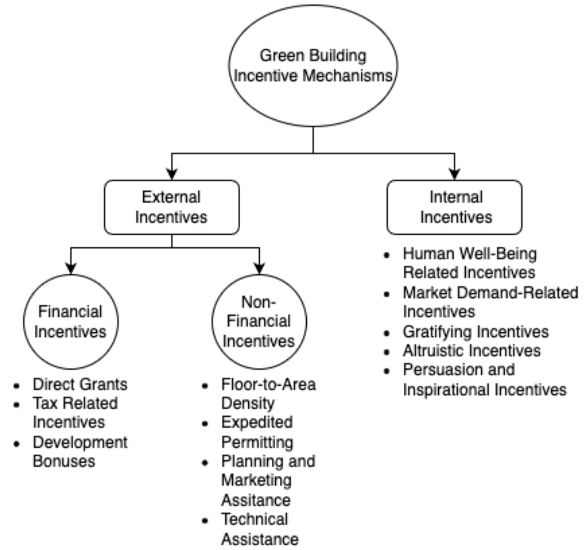


Figure 1: Categorization of Green Building Incentives adapted from Olubunmi et al. (2016)

The study conducted by Dounas et al. (2022) showcases the versatile usage of incentive mechanisms by proposing an innovative incentive model to reduce carbon and waste in the construction industry. It is highlighted that the proposed incentive model benefits environmental aspects and also enhances overall efficiency. Moreover, the research conducted by Sauer and Siddiqi (2009) compares the effectiveness of three different incentive mechanisms based on the production status offered by local government. By using regression analysis, the most preferred incentives are concluded as Gross Floor Area(GFA) concessions, administrative, and financial incentives. In another study, Agyekum et al. (2019) investigates the barriers to adoption in Ghana. The authors' findings show inadequate awareness of the benefits of green certification has the biggest impact on adoption in a developed country, Ghana. In a parallel study by Work (2007), the authors revealed that their findings point out that developers are aware of these incentives but don't always use them. The main reason is the timing of development decisions and the response time of local government don't always match together. Developers must take quick decisions, and governments prefer to move more slowly to see the outcomes.

Non-financial incentives have limited visibility in the literature compared to financial ones. Karji et al. (2021) connects this limitation to the poor impact of social incentives. Similar to the previously mentioned research, the root cause of the ineffectiveness of social incentives is the lack of information on the benefits of green building. It is possible to convey that recognition is elemental on non-financial incentives for governments and incentive receivers due to its ability to create public awareness.

On the other hand, Olubunmi et al. (2016) specifies weaknesses in the current green building incentive mechanisms, and in terms of the social effect of certificates, the authors criticize the close link between incentive mechanisms and third-party certification institutions due to the potential bias or inequity caused by economic relationship. Further criticisms of current green building incentive mechanisms are the lack of enforceability mechanisms, the absence of a mechanism to determine the optimal level and the non-transferability of incentives. But it is also mentioned that the production rate method gives limited information on real adoption rates therefore the outcomes are subject to discussion.

Fan and Wu (2020) conducted a cost-benefit analysis for incentive mechanism by using the Analytic Hierarchy Process. Based on outcomes, the authors proposed a hierarchy structure of cost and benefit criteria as actual and hidden. The high upfront cost of green building remains a major cost problem. Considering benefits, actual ones such as energy savings and GFA concession were deemed more important than hidden ones. However, it is crucial to note that stakeholders may not be aware of the proposed essence of hidden benefits like outdoor environment, productivity, and brand recognition, as mentioned, creating an opportunity for further research in the literature.

Considering the status of the construction industry, more adoption of green buildings is inevitably needed to increase sustainability performance. Incentive mechanisms play a vital on this as a main driver of stakeholder motivation. Research supports the feasibility of blockchain in incentive design, offering transparency and security to build stakeholder confidence. This decentralized and secure approach has the potential to generate economic benefits by cutting intermediaries, lowering transaction costs, and improving resource allocation within the incentive ecosystem, as stated by Fan and Wu (2020). Furthermore, side benefits due to the characteristics of blockchain may reveal new incentives and increase the overall effectiveness.

Blockchain

Definition and Basic Concept

Blockchain technology gained a reputation with the publication of "Bitcoin: A Peer-to-Peer Electronic Cash System" white-paper by an author under the name Satoshi Nakamoto Nakamoto and Bitcoin (2008). Fundamentally, blockchain is an unalterable distributed ledger allowing secure data transfer. Blockchain records data in groups on a peer-to-peer (P2P) network as detailed by Raj (2021). The name 'blockchain' is derived from two terms: 'block,' referring to the grouped data sets, and 'chain,' indicating that these data sets are securely linked using cryptographic principles.

In blockchain, transactions are a prominent feature that refers to an action that records the exchange of data or assets between parties. Once a transaction is completed, it leads to creating a new block with a time-stamp and linking to the preceding block to the end of the chain, as detailed by Natalia Maslova and CTP (2018). The information within each new block holds a secure, cryptographic summary of the one before it. This interconnectedness form enhances security naturally. Once data is added to the chain, it becomes immutable, making it impossible to alter or erase.

In a decentralized system like blockchain, consensus mechanisms are vital to ensure agreement among participants regarding transaction validity, thereby preventing system failures and keeping the whole system secure. Zhao et al. (2020) mentioned that besides security and continuity of the ecosystem, consensus mechanisms also significantly influence performance metrics such as scalability and block creation speed. In the consensus mechanism, the active participation of nodes in all roles is crucial to ensure system security, decentralization, and integrity Zhao et al. (2023). Roles and responsibilities within the ecosystem were outlined by Han et al. (2022) with broadcasting nodes participating in the verification, and spread of transaction records, while mining nodes actively execute the consensus mechanisms and support the block creation. Lastly, full nodes' responsibilities fall between the other 2 types.

It is important to highlight that executing consensus mechanisms comes with a cost, therefore, participants need to be compensated to maintain system continuity. Additionally, participants often prioritize actions that serve their own interests. In response to this challenge, incentive mechanisms are introduced to motivate cooperative behavior and maintain the security of the decentralized system, as elaborated further.

As we move towards the Blockchain 4.0 era, blockchain applications aim to enhance user experience and speed for creating advanced decentralized applications for all sectors. While blockchain technology has demonstrated transformative potential within the construction sector, its specific implications for green building certification systems warrant further investigation.

Blockchain Applications

Blockchain technology is widely adopted across numerous sectors, demonstrating its potential to transform traditional processes, promote sustainability, enhance effectiveness, and drive innovation. The finance sector was among the first to embrace blockchain technology, but its applications extend to sectors rapidly such as construction, energy, crowdsensing, and supply chain management, as detailed by Xu et al. (2017).

For instance Zhang et al. (2020) proposes a blockchain-based certification system to address inefficiencies regarding to centralization in China's renewable energy certificate process, emphasizing intelligent and automated processes facilitated by blockchain technology. The studies conducted by Marques et al. (2023) Delardas and Giannos (2023) also delve into the application of blockchain in the energy sector, aiming to enhance efficiency and effectiveness. The research conducted by Wei et al. (2020) explores decentralized crowdsensing architectures via blockchain, focusing on security enhancement, privacy protection, and incentivization mechanisms. Simulation of the proposed model confirms its feasibility, showing that the proposed incentive mechanism effectively encourages participation.

The research conducted by Wei et al. (2023) focuses on addressing technical information exchange barriers in the supply chain, which lead to substantial energy consumption. By leveraging blockchain for data management and government incentives, the paper investigates strategies to improve the sustainability of supply chains. The research proposes practical measures and introduces a data governance platform to foster eco-friendly supply chains, drawing insights from government involvement and behavioral theories.

In the built environment, distributed ledger technology offers promising applications across several domains. Li et al. (2019) categorized these domains as smart energy, smart cities and the sharing economy, smart government, smart homes, smart transport, Building Information Modelling (BIM) and construction management, and innovation in business models and organizational structures. The majority of the cited articles in this paper discuss the application of blockchain technology in the various domains of built environment. Also, studies like Smajgl and Schweik (2022) bring a further looking perspective and anticipate that blockchain technology will bring significant changes into the existing socio-economic framework, influencing the governance aspects and expanding the boundaries of solutions created in sustainability.

While numerous studies demonstrate the advantages of blockchain technology across different domains, its specific application in green building certification systems remains relatively unexplored. Understanding the governance and design principles of blockchain applications is crucial, particularly within green building certification systems, where tailored approaches are needed to address adoption and sustainability challenges effectively. The following chapter will delve into the governance and design considerations specific to blockchain technology in the context of green building certification, underscoring the dynamic and versatile nature of blockchain technology in diverse contexts.

Blockchain Governance and Design

The close connection between blockchain governance and its design arises from the foundational principles of technology. Design encompasses the underlying architecture, consensus mechanisms, and smart contracts that shape how the system operates. These design choices directly influence how governance is implemented within the blockchain network. Governance rules are often encoded in smart contracts, automating and enforcing predefined rules without the need for centralized control. As mentioned before these choices and rules also determine performance, including scalability, efficiency, etc. of the application, as further detailed in Han et al. (2022).

De Filippi et al. (2020) emphasizes that blockchain governance relies heavily on incentives coded into the protocol and game-theoretic principles align with community priorities. The goal of good governance is to align diverse stakeholders toward a common understanding and agreement. Effective governance can increase the reliability and overall success of domain by ensuring that created financial and non-financial incentives match desired behaviors and encouraging community adherence to shared standards.

Therefore, introducing blockchain into a specific domain requires a deep understanding of both design principles and governance requirements to achieve effective application. A holistic approach is essential to ensure that the blockchain solution aligns with the required outcome and goals of the application.

Blockchain Related Incentive Mechanisms

Incentives are vital in blockchain systems to motivate participants and align their interests with the collective goals of the system. Incentives support positive behavior, discourage malicious activities, and foster an active and engaged community within the decentralized networks. Well-designed incentive mechanisms not only drive the desired behavior but also play a pivotal role in shaping the governance structures that govern the decision-making processes of the blockchain protocol. Thanks to the inherited features of blockchain, participants can be motivated in several forms of incentives.

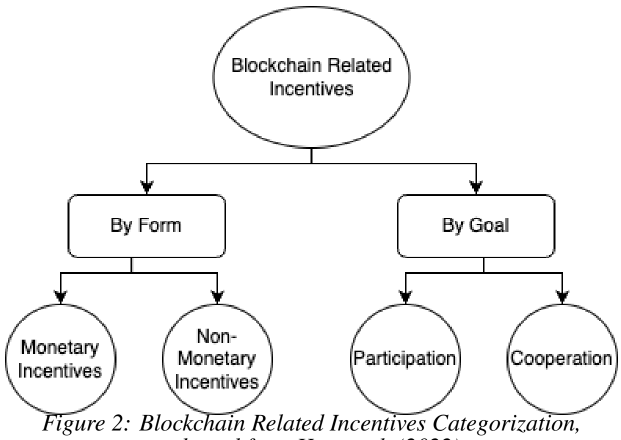


Figure 2: Blockchain Related Incentives Categorization, adapted from Han et al. (2022)

The Figure 2 illustrates the high-level classification of blockchain-related incentives based on the form and goal of the incentives in the literature. Monetary incentives are designed to regulate system entities economically, motivating participation with financial rewards. Non-monetary incentives, like credit-based and reputation-based systems, foster trust among entities, while gamified incentives leverage individuals' preferences for enjoyable experiences to guide their behavior within the system. Tokens and badges are commonly used in gamified mechanisms as incentives for system engagement as ticket.

In terms of goals, incentive mechanisms can be categorized as encouraging participation to maintain system sustainability and promoting cooperation among entities. The participation goal addresses nodes' selfish nature by providing incentives for their involvement, which is crucial for ensuring security and decentralization. On the other hand, the cooperation goal acknowledges the fact that selfish nodes may deviate from system design, posing threats to the system. To mitigate, introducing incentive mechanisms becomes essential, fostering cooperation through adherence to system protocols and discouragement from initiating attacks.

In conclusion, the multifaceted nature of blockchain, with its diverse array of incentive mechanisms, creates opportunities for various benefits within the ecosystem. By strategically aligning incentives with the goals and form, blockchain not only addresses the rational and profit-driven behavior of nodes but also promotes the safety, sustainability, and decentralization of the system. In essence, blockchain's innovative approach to incentives unlocks opportunities, driving positive contributions and reinforcing the resilience of various domains implemented.

Blockchain for GBCS

There is a growing need for a more comprehensive and digital green building certification system to address the challenge of inadequate motivations hindering widespread adoption. Blockchain, with its inherited features, has the potential to bring more integrity to certification practices. By leveraging its capabilities, blockchain has the potential to bolster the credibility of certificates while streamlining verification procedures.

Blockchain has been already implemented as a transformative factor in certification processes, providing a streamlined and reliable method to verify accomplishments. Fowler (2017) suggests that blockchain integration in certification processes holds promise for lowering costs associated with advancement in measurement methodologies to meet evolving quality standards. In the article by Woo et al. (2020), the authors support the idea of blockchain-enabled GBCS and target the high cost associated with the current systems. Their research proposes a blockchain-based framework for measurement, reporting, and verification as a cost-effective alternative.

Pu and Lam (2023) discuss the advantages of blockchain-enabled certification systems and highlights that GBCS is a particularly suitable domain to apply. Firstly, blockchain's ownership or identity management capabilities provide a secure way to verify the identity of individuals or entities involved in green building projects. Secondly, the provenance tracing capacities of blockchain ensure transparency in the supply chain, allowing stakeholders to trace the origin of materials used in construction. Lastly, ownership transfer mechanisms can streamline the transfer of green building certifications between different parties, enhancing overall efficiency in the certification ecosystem.

Introducing blockchain into the GBCS may enhance the certification system by providing both monetary and non-monetary benefits. Therefore, stakeholder dynamics play a vital role in blockchain design, particularly for incentive mechanisms and their effectiveness. Each stakeholder in the green building certification process has unique priorities and motivations as detailed in Mulligan et al. (2014). Building owners seek cost reduction and streamlined processes, while architects focus on sustainable building design principles. Certifying bodies prioritize integrity and credibility through incentive mechanisms. Regulatory bodies may enforce compliance with incentive requirements or promote adoption by implementing new incentives, depending on the ecosystem dynamics.

Dynamic Alliances, as described by Grover et al. (2021), refer to collaborations formed by entities with mutual interests and goals in a given domain. Dynamic Alliances aim to create a collaborative environment that allows parties to leverage strengths, optimize resources, and accomplish common objectives. The authors suggest that integrating blockchain into governance ecosystems of dynamic alliances would bring notable benefits. Integrated Project Delivery (IPD), a project management approach tailored to the construction industry, has governance characteristics and principles that resonate with Dynamic Alliances. In this context, the research by Hunhevicz et al. (2022) proposes using blockchain for IPD using in a common pool resources scenario. Likewise, the authors suggest that blockchain implementation will leverage the governance of the domain. This suggests that the governance structures and incentive mechanisms inherent in blockchain implementation can effectively align with the principles of collaboration, trust, and accountability within green building certification systems.

In conclusion, the alignment of blockchain's governance with those observed in Dynamic Alliances and IPD highlights the importance of carefully balancing stakeholder interests in design and governance decisions. Tailoring design and governance decisions to the diverse motivations of stakeholders ensures that incentive mechanisms effectively encourage collaboration, trust, and accountability within the green building certification ecosystem. In the context of GBCS, blockchain implementation can optimize collaboration, streamline processes, and foster innovation, thereby advancing sustainability goals within the construction industry.

Conclusion and Future Work

In overview, the construction industry encounters challenges including issues with information exchange, materials procurement, high costs, and trust, limiting its ability to achieve better sustainability performance. Blockchain technology holds promise in addressing these challenges by its decentralized and secure nature. Our literature review acknowledges the potential of blockchain in the construction industry to enhance sustainability performance. However, practical implementations are currently constrained, indicating a gap between theoretical recognition and widespread adoption.

To address this gap, we propose a comprehensive survey targeting diverse stakeholders involved in green building certification. This survey aims to explore stakeholders' motivations to participate in GBCS and assess their receptiveness to various incentive mechanisms and receptivity toward blockchain. The survey results are expected to close the gap between theoretical discussions and practical implementations of blockchain in green building certification. Also, these findings may support decision-making processes, driving refinements in certification practices toward greater sustainability.

Throughout the literature review, it is identified that the integration of blockchain technology promises to enhance transparency, efficiency, and trust in certification processes. Moreover, such a system has the potential to create a range of diverse incentives, including both financial and non-financial, which may be novel to the construction sector. In light of survey findings, this integration holds promise for promoting the adoption rates of GBCS by tailoring design and governance preferences to stakeholders. In conclusion, while the potential benefits of blockchain for GBCS are recognized, the dynamics of GBCS need to be thoroughly comprehended. The identified literature gap calls for a survey with GBCS stakeholders to investigate their motivations and preferred incentives. The outcomes of such a survey inform the governance and design choices of the best-suited blockchain technology for widespread adoption in green building certification.

References

Agyekum, K., Adinyira, E., Baiden, B., Ampratwum, G., and Duah, D. (2019). Barriers to the adoption of green certification of buildings: A thematic analysis of verbatim comments from built environment professionals. Journal of Engineering, Design and Technology, 17(5):1035–1055.

Awadh, O. (2017). Sustainability and green building rating systems: Leed, breeam, gsas and estidama critical analysis. Journal of Building Engineering, 11:25–29.

Braulio-Gonzalo, M., Jorge-Ortiz, A., and Bovea, M. D. (2022). How are indicators in green building rating systems addressing sustainability dimensions and life cycle frameworks in residential buildings? Environmental Impact Assessment Review, 95:106793.

De Filippi, P., Mannan, M., and Reijers, W. (2020). Blockchain as a confidence machine: The problem of trust challenges of governance. Technology in Society, 62:101284.

Delardas, O. and Giannos, P. (2023). Towards energy transition: Use of blockchain in renewable certificates to support sustainability commitments. Sustainability, 15(11):258.

Diyana, N., Abidin, Z., et al. (2013). Motivation and expectation of developers on green construction: a conceptual view. International Journal of Humanities and Social Sciences, 7(4):914–918.

Dounas, T., Lombardi, D., and Jabi, W. (2022). Collective digital factories for buildings: Stigmergic collaboration through cryptoeconomics. In Blockchain for Construction, pages 207–228. Springer.

Fan, K. and Wu, Z. (2020). Incentive mechanism design for promoting high-level green buildings. Building and Environment, 184:107230.

Fowler, M. D. (2017). Linking the public benefit to the corporation: Blockchain as a solution for certification in an age of do-good business. Vanderbilt Journal of Entertainment Technology Law, 20:881.

Goh, C. S., Chong, H.-Y., Jack, L., and Mohd Faris, A. F. (2020). Revisiting triple bottom line within the context of sustainable construction: A systematic review. Journal of Cleaner Production, 252:119884.

Grover, B. A., Chaudhary, B., Rajput, N. K., and Dukiya, O. (2021). Blockchain and Governance: Theory, Applications and Challenges, page 113–139. John Wiley Sons, Ltd.

Han, R., Yan, Z., Liang, X., and Yang, L. T. (2022). How can incentive mechanisms and blockchain benefit with each other? a survey. ACM Comput. Surv., 55(7).

Horner, J. and Ryan, P. (2019). Blockchain standards for sustainable development. Journal of ICT Standardization, 7(3):225–248.

Hunhevicz, J. J., Brasey, P.-A., Bonanomi, M. M., Hall, D. M., and Fischer, M. (2022). Applications of blockchain for the governance of integrated project delivery: A crypto commons approach. arXiv preprint arXiv:2207.07002.

Jayakody, T. A. C. H. and Vaz, A. (2023). Impact of green building certification on the rent of commercial properties: A review. Journal of Informatics and Web Engineering, 2(2):8–28.

Karji, A., Ghorbani, Z., Tafazzoli, M., and Rokoui, S. (2021). Revisiting social aspects of green buildings: Barriers, drivers, and benefits.

Li, J., Greenwood, D., and Kassem, M. (2019). Blockchain in the built environment and construction industry: A systematic review, conceptual models and practical use cases. Automation in Construction, 102:288–307.

Marques, N. L., Gomes, L. L., and Brandão, L. E. (2023). A blockchain-based model for token renewable energy certificate offers. Revista Contabilidade Finanças, 34:e1582.

Mulligan, T., Mollaoğlu-Korkmaz, S., Cotner, R., and Goldsberry, A. (2014). Public policy and impacts on adoption of sustainable built environments: Learning from the constuction industry playmakers. Journal of Green Building, 9:182–202.

Nakamoto, S. and Bitcoin, A. (2008). A peer-to-peer electronic cash system. Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf, 4:2.

Natalia Maslova, C. and CTP, P. (2018). Blockchain: Disruption and opportunity. Strategic Finance, 100(1):24–29.

Olubunmi, O. A., Xia, P. B., and Skitmore, M. (2016). Green building incentives: A review. Renewable and Sustainable Energy Reviews, 59:1611–1621.

Pu, S. and Lam, J. S. L. (2023). The benefits of blockchain for digital certificates: A multiple case study analysis. Technology in Society, 72:102176.

Raj, P. (2021). Demystifying the blockchain technology. In Advances in Computers, volume 121, pages 1–42. Elsevier.

Sauer, M. and Siddiqi, K. (2009). Incentives for Green Residential Construction.

Shen, W., Tang, W., Siripanan, A., Lei, Z., Duffield, C. F., and Hui, F. K. P. (2018). Understanding the green technical capabilities and barriers to green buildings in developing countries: A case study of thailand. Sustainability, 10(10).

Smajgl, A. and Schweik, C. M. (2022). Advancing sustainability with blockchain-based incentives and institutions. Frontiers in Blockchain, 5.

Tan, Y., Shen, L., and Yao, H. (2011). Sustainable construction practice and contractors' competitiveness: A preliminary study. Habitat International, 35(2):225–230.

Turk, and Klinc, R. (2017). Potentials of blockchain technology for construction management. Procedia Engineering, 196:638–645.

Vierra, S. (2016). Green building standards and certification systems. National Institute of Building Sciences, Washington, DC.

Wei, J., Yi, X., Yang, X., and Liu, Y. (2023). Blockchain-based design of a government incentive mechanism for manufacturing supply chain data governance. Sustainability, 15(88):6968.

Wei, L., Wu, J., and Long, C. (2020). A blockchain-based hybrid incentive model for crowdsensing. Electronics, 9(22):215.

Woo, J., Kibert, C. J., Newman, R., Kachi, A. S. K., Fatima, R., and Tian, Y. (2020). A new blockchain digital mrv (measurement, reporting, and verification) architecture for existing building energy performance. In 2020 2nd Conference on Blockchain Research Applications for Innovative Networks and Services (BRAINS), pages 222–226.

Work, G. B. I. T. (2007). A look at how local governments are incentivizing green development. Yudelson Associates. NAIOP (National Association of.

Xu, X., Weber, I., Staples, M., Zhu, L., Bosch, J., Bass, L., Pautasso, C., and Rimba, P. (2017). A taxonomy of blockchain-based systems for architecture design. In 2017 IEEE international conference on software architecture (ICSA), pages 243–252. IEEE.

Zhang, S., Xuan, J., Lyu, Z., and Fu, Y. (2020). Application prospect of blockchain in renewable energy certificates. In Proceedings of the 4th International Conference on Computer Science and Application Engineering, CSAE '20, page 1–5, New York, NY, USA. Association for Computing Machinery.

Zhao, F., Guo, X., and Chan, W. K. (2020). Individual green certificates on blockchain: A simulation approach. Sustainability, 12(9):3942.

Zhao, R., Wang, J., and Xue, F. (2023). A blockchain-based token economic model for incentivizing esg in the construction industry. volume 4 of Computing in Construction, page 0–0. European Council on Computing in Construction.

AUTOMATED TRACKING, INSPECTION AND COMMISSIONING OF WALL PANELS USING AN IOT-BLOCKCHAIN SOLUTION: THE CASE OF ETICS

Lavinia Chiara Tagliabue[1], Stefano Rinaldi[2], Jens Hunhevicz[3], Michele Melchiori[2], Daniel Hall[4], Maria Cairoli[5] and Angelo Luigi Camillo Cirbini[2]
[1]University of Turin, Turin, Italy
[2]University of Brescia, Brescia, Italy
[3]EMPA, Zurich, Switzerland
[4]Delft University of Technology, Delft, Netherlands
[5]Politecnico di Milano, Milan, Italy

Abstract

Poor installation and misaligned supply chain incentives can impact the energy performance and user comfort of External Thermal Insulation Composite System (ETICS) panels. This paper investigates a combined IoT-blockchain solution to create trusted traceability of the process, together with performance-based incentives based on sensor data about the panel state. The system is implemented on the University of Brescia prototype, an experimental test house. The results demonstrate the feasibility of an integrated system of smart contracts, IoT sensors and a digital twin for visualization. The work provides insights into an IoT-blockchain solution applied to a real-world use case for supply chain process improvement.

Introduction

Tracking of the state of health, inspection, and commissioning of installed components in the built environment has traditionally been a manual process. Tracking of progress depends on visual inspections, checklists, and logs to report deficiencies. Traditional practice relies heavily on inspectors' personal judgment and observational skills (Boschè et al., 2015). More recently, a huge body of work has emerged for automated monitoring of operation-level construction progress, using technologies such as laser scanning, computer vision and/or Building Information Modeling (BIM) for many applications, for example mechanical, electrical and plumbing systems or concrete floor flatness.

Another large field of research is the automated commissioning of buildings to show that they follow the promised energy performance (Panteli et al., 2020). Related research for facility management focuses on the real time monitoring through visualization at building and city scale, energy performance analysis, and energy benchmarking (Tomašević et al., 2015). Using these solutions, facility managers can retrieve energy-related information from the Building Management System (BMS) and view the BIM information using web-based energy management platforms (McGlinn et al., 2017). This allows facility managers to better visualize indoor environmental conditions such as temperature and humidity and make adjustments to improve user comfort (Marzouk and Abdelaty, 2014).

Other energy related solutions include the energy performance analysis based on real time data. Typically, energy analysis is based on mathematical simulation methods in the 3D environment; however, there is often a gap between predicted and actual building performance. Moreover, the likelihood for performance disparity increases in cases of complex, high-performance features which are difficult to model and simulate. However, one promising solution to improve accuracy of prediction versus performance is to use technologies to enable real time data as calculation input for the energy analysis (Yilmaz et al., 2023).

Despite advances in these research areas, solutions for progress tracking and commissioning technologies often mirror the fragmentation found in the greater industry (Hall et al., 2020). The proposed solutions tend to be oriented around individual project temporal phases: installation, inspection, and commissioning. Solutions are developed in isolation for tracking the installation, automating the inspection, or checking the commissioning. By contrast, much less research looks at integrating these activities with an object- or component-oriented perspective.

Monitoring these activities at the component level has several advantages. It avoids misaligned incentives where firms attempt to push risk down the supply chain. Better supply chain management of the digital asset can occur across the lifecycle and enable new incentive systems and business models for management of the asset, including opportunities for sustainability (e.g., circular economy), productivity (e.g., better collaboration, streamlined value chain) and/or security (e.g., the Hackett fire in London) (Watson et al., 2019).

Industry 4.0 for Lifecycle Traceability

The concept of "industry 4.0" is used to describe digital technologies and automation to create interconnected, intelligent, autonomous, and self-learning cyber-physical systems (Lasi at al, 2014). In the Architecture, Engineering, Construction, and Operations (AECO) industry, the term "Construction 4.0" is emerging (Klinc et al., 2019; FIEC 2015; Sawhney et al., 2020; Garcia de Soto et al., 2019). While Construction 4.0 encompasses many concepts and technologies, multiple sources mention the combination of BIM, Internet-of-Things (IoT) sensors, and distributed ledger technologies (DLT, also referred to as blockchain) as a favorable combination to enable supply chain traceability and data integration (Ye at al., 2018; Kinnaird and Geipel, 2017). BIM allows designers and builders to design, visualize, and coordinate construction systems with greater efficiency using three-dimensional digital models and processes.

BIM have proven that generation of large data sets in the AECO industry is possible. It has triggered a change towards digital working practices across stages of the design, construction, and operations process.

IoT refers to a network of sensors and connected devices that can be used for real-time feedback. To reduce the gap between the digital and physical world digital twins can act as (often) visual representations of information based on sensors for real-time representation of the physical world. For example, Zhai et al. (2019) finds that sensors in the construction process, in tandem with BIM, can help to address problems of inconvenient data collection, lack of automatic decision support, and incomplete information.

DLT and associated blockchain technologies can be an important enabler for more transparent and streamlined traceability in the AECO industry (Li et al, 2020). DLT enables direct peer-to-peer transactions of value and the immutable and trusted storage of data. These properties facilitate the building of trust between transacting parties and devices and can decrease the settlement time of transactions and reduction of costs associated with intermediaries (Viryasitavat at al., 2018). In addition, newer implementations of blockchain allow the execution of code protocols, often called smart contracts (SCs). SCs can encode business logic for value transactions and data handled on the blockchain to create automated workflows. Since these code protocols live on a blockchain, they benefit from the immutability and transparency of the system and hence it is ensured that they execute exactly as specified. Two studies critically analyzed blockchain for construction applications, concluding that the technology indeed has real world potential beyond the hype (Hunhevicz and Hall, 2020; Perera et al., 2020).

Supply chain traceability has been identified by several studies as one of the most promising ones for the application of DLT (Hunhevicz and Hall, 2020; Perera et al., 2020; Li at al., 2019). The opportunity is best described by Qian et al. (2020), who conclude that "blockchain technology is a key technology in Construction 4.0 that can bring the cyber (digital technologies) and physical (social capital) closer together by transforming trust to support various ecosystems of construction supply chains that shape and produce the built environment." However, overall there is a need to further investigate how the combined use of DLT, IoT and BIM should be implemented to create a system of trusted life cycle traceability and incentives.

Scope and Methodology

The goal of this research is to develop a tracing and performance monitoring system, leveraging combined use of new Industry 4.0 technologies, in particular BIM, IoT and blockchain. The innovation hereby lies in applying blockchain to the complete life cycle, beyond track and trace of just one supply chain stage. Furthermore, there is an opportunity to couple the performance evaluation through sensorized components with the benefits of trusted blockchain traceability and payments.

Research Setting

To achieve this goal, the study works with the supply chain for the retrofitting process using an External Thermal Insulation Composite System (ETICS). ETICS panels are widely used in retrofitting intervention to enhance the energy performance of the existing building. Correct installation of ETICS is important for a functioning insulation system to achieve the planned energy performance, having a direct impact on the cost of operation and comfort of inhabitants.

However, the current installation of ETICS is subject to repeated failure. ETICS represents an exemplary case for the often-present life cycle related issues in the construction supply chain. The success of the ETICS installation depends on different contractors that need to do their due diligence to deliver the final expected performance. The contractor is incentivized to maximize their profit during installation by taking "shortcuts" to save on material or time. Failures during production or installation manifest often, but they are not discovered until the operations phase because ETICS panels remain hidden behind the façade until the failure is propagated to the surface. This time discrepancy makes it very hard to investigate the exact cause of failure, as well as to hold the responsible parties accountable. The installing parties may pass blame to others, may have gone bankrupt, or the warranty time may have already expired.

A careful tracing and performance evaluation of ETICS panels over the course of their life cycle could benefit 1) the environment through emission reduction and recyclability of the material related to the concept of circular economy, 2) the operational efficiency through cost reduction both in heating or cooling losses and warranty claims, and 3) health and safety of both workers and occupants of the retrofitted building.

Implementation of a BIM-IoT-DLT approach is particularly suitable for the ETICS case study because: 1) it has a specific product focus that can be tracked, inspected, and commissioned; 2) the supply chain of ETICS panels has been repeatedly identified as a typical example where stakeholders misbehave at different working steps for their own benefit, to the detriment of the overall system performance. DLT can provide the trusted traceability to ensure accountability regarding correct production, commissioning, and installation, and incentivize correct installation through performance-based payments.

Case Study Description

This study uses the so-called "University of Brescia (UniBS) Prototype" as a research case study. The UniBS was built to investigate different improvement potentials around the lifecycle performance of ETICS. The physical prototype consists of a small demonstrator-house realized at the experimental facilities provided by Ente Sistema Edilizia (Building System Institute) Brescia, within a collaborative framework involving both the academia and industry. The research program was managed with the support of an industrial manufacturer Weber Saint-Gobain. The driving force from academia was the eLUX lab– "Energy Laboratory as University eXpo" at the smart campus of the University of Brescia in Italy, led by the Department of Civil, Environmental, Architectural Engineering and Mathematics and the Department of Information Engineering.

UniBS is equipped with detailed network of IoT to provide real time data on the ETICS performance. The UniBS prototype served as the test platform, applying the developed technical system throughout the life cycle of the prototype, from construction to operations.

The next section of the paper further specifies the role of the investigated ETICS supply chain workflow in the UniBS prototype, and the implemented blockchain process. The results section then gives details on 1) the transaction and SC design logic and implementation of the DLT-based system for the case study, and 2) the connection of the DLT SC with the digital twin and the IoT system for performance evaluation.

Blockchain for ETICS supply chain tracking

The main two value propositions of DLT in this prototype are 1) the transparency and immutability it can provide to track the ETICS panels and associated transactions in the supply chain, 2) workflow automation throughout the supply chain process for the execution of the performance-based incentive payments. The first one is important to establish a trusted database that can hold parties accountable for their actions, as well as assurance that data integrity is maintained. The second one can then create (semi-)autonomous workflows and automatic triggering of payments without any intermediary. Below we outline the investigated workflow of the ETICS panels, and afterwards give more details on the implemented blockchain process.

Investigated Workflow

Figure 1 pictures the overall workflow logic of the ETICS value supply chain for the UniBS Prototype with its involved stakeholders. To implement blockchain for the UniBS Prototype, the supply chain was simplified regarding both the number of actors and the workflow steps. The real supply chain is more extensive, nevertheless, this approach can showcase most important aspects to investigate a blockchain implementation. The considered actors involve the supplier of the ETICS panels, a contractor responsible to install the panels on the construction site, as well as the building owner.

Moreover, there is an actor introduced called "Oracle". The term oracle usually describes an external source of information that is needed to feed information to blockchain based SC. For the UniBS Prototype, the Oracle is a placeholder name for actors that will take the role of verifying various aspects regarding the energy performance-based contract. In the example of this paper, to feed the verification data of the installation steps (Figure 1; (7)) the Oracle is a trusted human third-party actor, while for the commissioning (Figure 1; (10)) the Oracle consists of the sensor network. To have both a human actor and sensors as an Oracle can also showcase these two possibilities in comparison. Having said that, not all SCs used in the prototype need input from external sources. Some of them only rely on the information saved on the blockchain itself. For example, accessibility features and users' roles, as well as the prescribed supply chain flow and the actual state of each panel can be described and checked on-chain. More information on the different SCs implemented in the prototype is given in the next section.

Implemented Blockchain Process

The blockchain-based process follows the defined workflow of the panels as defined in Figure 1. The smart contract logic controlling the access rules and the sequence of actions is then based on the "state" of panels. For example, the workflow starts with registering panels through the supplier once they are produced (Figure 1; (1)). The state is then "produced". This indicates to the SC that the next step needs to be selling the panels to the owner with a subsequent payment (Figure 1; (2), (3)). The workflow goes on to ship the panels to the construction site (Figure 1; (4)), where the contractor will receive and install them (Figure 1; (5), (6)). At this point, the SC only allows the payment and subsequent use of the panels in the state "installed" if all the installation checks are

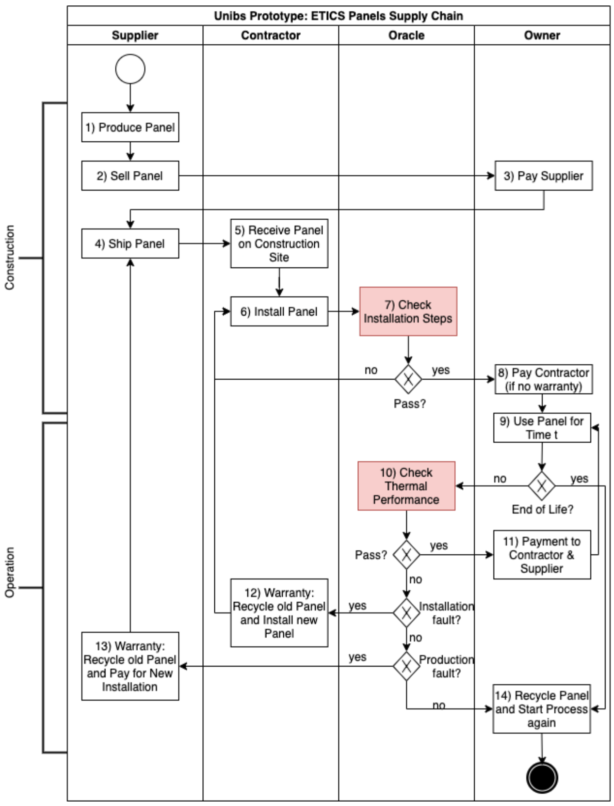


Figure 1: Flow chart of the ETICS Panels Supply Chain in the UniBS Prototype.

successful. This ensures verifiability and certification of the installation steps. For the use phase, the SC request for each time interval t the sensor data to check the thermal performance of the ETICS system (Figure 1; (9), (10)). If the situation is as expected, a payment to both the contractor and the supplier is automatically triggered (Figure 1; (11)) and the procedure repeats until the end of life of the panel is reached (Figure 1; (14)). In case the thermal performance of the ETICS is not as expected, the SC logic runs through the various causes, while considering the installation certification. In case it is an installation fault, the warranty and reinstallation of the panels needs to be performed through the contractor (Figure 1; (12)). In case of a production fault, the warranty to install a new ETICS panel needs to be paid by the supplier (Figure 1; (13)).

Overall, the proposed workflow acts as an energy performance smart contract (EPSC) and extends the liability of the supplier and contractor into the operational phase. Stakeholders get paid as usual to cover their immediate expenses, but only receive profit if the solution performs well during operation through an automated and continuous revenue stream. The blockchain smart contract gets funded up front by the owner and therefore acts as an escrow. If the solution does not perform well, the stakeholders will receive less or no profit. This should act as an incentive to collaborate and integrate across the life cycle phases. If behaving honestly, all participants can benefit from such a solution. The supplier has data available on where the panels are installed and how they perform. The owner gets assurance on the production an installation quality, as well as automatic warranty claims based on real-time sensor data. The contractor can benefit from the constant additional revenue streams but could also be held accountable for mistakes and warranty claims. Through the transparent installation certification, the contractor can also proof later that the installation was correct, but maybe a production error.

Implementation and Results

Transaction, SC Design and DLT-based system

The core of the implementation of the ETICS DLT-based system consists of a set of SCs supporting trust-demanding tasks developed for the Hyperledger Fabric permissioned blockchain infrastructure. Since not every task and data can benefit from on-chain execution, we defined a rule-based approach to identify these trust-demanding tasks and data objects. In general, these tasks identify critical activities in the workflows, with associated data objects, as explained above. A Private permissioned blockchain was chosen because it offers privacy and all parties are known (Hunhevicz and Hall, 2020). However, a permissionless blockchain could also be utilized, provided that appropriate encryption measures are implemented for the data written on it.

The Hyperledger Composer, a rapid prototyping environment running on top of Hyperledger Fabric, was used as development environment. The main concepts of a Hyperledger Composer project are *participant*, *asset*, *transaction* and *event*. An asset is anything, whether physical or virtual having some value in the considered domain. A participant is a role of the workflow who participates in the operations on the ledger. A transaction is an operation submitted by a participant to modify the ledger. It may consist of modifying the amount of an asset and it may raise events. A transaction typically modifies or creates some contents of the Hyperledger world state. Implementing a Hyperledger Composer project (M.M.S. et al., 2019) consists of creating three main files. A model definition (with extension .cto) file describes the business domain. It is written in a specific notation, called Composer Modeling Language, in which resources like participants, assets, transactions and events are declared. A JavaScript file, in which transactions declared in the .cto file are implemented as functions, describes the way assets are processed. Finally, a file containing a list of access control rules defines the rights (e.g., Read, Create, Update) assigned to each participant to operate on the ledger. These files form the so-called Business Network definition. For example, the following shows both the definitions of the Owner participant and the *Building* asset in the model definition file:

```
participant Owner extends Company {
o String name
        o String contactInfo optional
}
asset Building identified by buildingID{
o String buildingID
--> Design[] referencesToDesign
}
```

Here, the definition of the *Owner* participant extends the *Company* super-type. The resource has all properties and fields specified by the super-type and may add additional properties from its own definition. The definition of the *Building* asset includes the identified attribute *buildingID*. A resource definition specifies a set of relationships to other types that are not owned by the resource but that may be referenced from it. In this example, *Building* references a set of objects of *Design* type.

In our implementation, the participants can make an authentication on the ledger and perform the operations they are authorized to by the access control rules. The assets are the objects that the transactions operate on. We developed two kinds of transactions. The first one updates the ledger world state. The second one permits to create instances of assets (e.g., to insert a new panels installation on the ledger).

```
async function terminatePosing( tx ) {
    const NS = 'com.biz.eticssample';
    const factory = getFactory();
    const dateTime = new Date();
    const panRegistry = await getAssetRegistry('com.biz
.eticssample.installations');
    const instConfirmation =
factory.newResource(NS, 'instconfirmations', tx.id);
    instConfirmation.EndingDate = dateTime;
    instConfirmation.anchorsPattern = tx.anchorsPattern;
    instConfirmation.building = factory.newRelationship(
NS,'building',tx.building);
```

```
instConfirmation.contractor = factory.newRelationship(NS,'contractor',tx.contractor);
    await panRegistry.add(instConfirmation); }
```

The above function *terminatePosing* defined to process a transaction, exemplarily showing transaction execution. This transaction is declared in the model definition file and represents the corresponding workflow task. It permits to a contractor to create an *installationConfirmation* object sets its properties (i.e., date and time, pattern used for panel anchors, references to the building and to the contractor), and to add it to the ledger. The parameter *tx* is an instance of the *terminatePosing* function.

SC, IoT and DT for Commissioning of ETCIS Panels

The proposed system leverages IoT sensors to continually gather data on the condition of panels throughout their operational lifespan, enriching the Digital Twin (DT) model of the building. This integration allows for real-time monitoring of ETICS health. BIM is employed to facilitate intuitive visualization for building managers. Moreover, it serves as a proactive tool, triggering alarms promptly in response to any detected anomalies or critical events, with respect to the expected behavior. To assess the performance of the ETICS system during the operational phase, it is important to have a base for comparison (expected values). Therefore, an energy model was used to simulate the thermal behavior of the system. The purpose is to provide a predictive model of the energy flow that is transferred from inside (where a heating system in winter set up the temperature to 20°C) to the external, passing through the ETICS consisting of the materials EPS and glass wool. The model allows to calculate the expected temperature in the interface between the wall and the insulation layer defined by the boundary conditions through the heat transfer equations implemented in the dynamic simulation (Energyplus).

The temperature profile in the wall has been then defined in the calculation model and depicted in Figure 2. The sensor for measuring temperature and relative humidity should be properly installed in the interspace between the thermal layer and the wall. The calculated value can be then compared to the temperature measured by the sensors. Based on this method, it is possible to assess the deviation of the predicted value compared to the temperatures measured by the sensors on the control walls (north, west), as well as the temperatures on the failing walls (south, east).

The deployment of a distributed monitoring system could make the analysis of the energy efficiency and the maintenance state of an installed thermal insulation system. Such a system should be able to monitor in real-time several parameters from which can be inferred key performance indexes (KPIs) about the status of the building thermal coating. In particular, the monitoring system should be able to identify the following conditions: i) partial detachment of the insulation panel; ii) change in the energy efficiency of the coating detected by the temperature in the interface between the wall and the insulation panel and due to water infiltration for example. The monitoring system should be easy to install, low-cost and robust since it must be installed on the external wall.

Battery-operated devices are preferred to cabled ones, because they are easier to install. In addition, the monitoring system should have a lifetime comparable with the installation to monitor, i.e. approximately around 10/20 years. Thus, the system should be formed by devices with low power consumption, to guarantee a useful lifetime. Given these considerations, the monitoring system exploits the Wireless Sensor Network (WSN) approach for extensive monitoring of the parameters. Each of the monitoring points is composed of a wireless sensor, able to collect the parameters using transducers. Each of the sensors should monitor the temperature and the relative humidity (used to estimate the efficiency of the ETICS), and the 3-axis acceleration (used to estimate the partial detachment of the ETICS). The estimation of energy efficiency requires the estimation of the external and internal temperature and relative humidity. The block diagram of the sensor board is shown in Figure 3. Each sensor board is composed of three types of transducers: a temperature sensor, a 3-axis

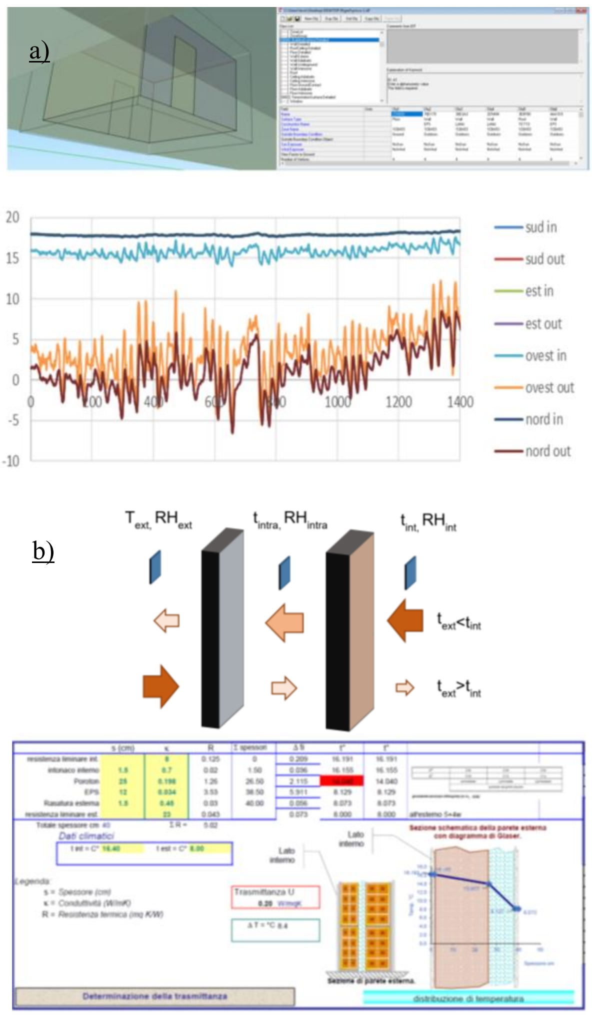


Figure 2: Calculation Method: a) Energy model of the test house b) Assessment of the expected sensor values.

accelerometer, and a relative humidity sensor. The data generated by the transducers are acquired and processed by the Micro Controller Unit (MCU). The data are temporarily stored in the MCU and, when possible, transmitted to a remote gateway through a Bluetooth Low Energy (BLE) modem. The MCU is also responsible for the optimization of the energy consumption of the entire board, through the adoption of proper low-power policy. All the peripherals of the board, including the transducers and the BLE modem, are turned off and the MCU itself remains in deep sleep mode until a data transmission is scheduled. The wake-up of the sensor board is obtained using the wake-up peripheral of the MCU. Each of the boards has been programmed to wake up every 3 hours, to reduce energy consumption. After the wake-up, the MCU acquires the data from the transducers and transmits the data through the BLE modem. The BLE protocol defines several communication modes, which require a data exchange with the local network coordinator. Such communication modes are not suitable for applications with strict requirements on power consumption. BLE provides a communication mode that does not require any interaction with the network coordinator: the advertising. Advertising packets are sent by a BLE node to advertise its presence to the coordinator. An advertise packet contains the Universally Unique Identifier (UUID), which identifies the sending node, and up to 37 bytes that can contain custom data. In the current case, the custom bytes are used to transmit the transducers' data, i.e. temperature, relative humidity, and 3-axis acceleration. The prototype of the sensor has been realized using the CC2650 Sensor Tag prototype board produced by Texas Instruments. Each of the boards is equipped with a high-performance ARM Cortex-M3 CC2650 wireless MCU, a 9-axis movement transducer (3-axis accelerometer, 3-axis gyroscope, 3-axis magnetometer), and a humidity transducer, able to monitor both relative humidity and temperature. The complete list of transducers used in the prototype and their main characteristics are summarized in Table 1.

Table 1: List of the transducers used by the prototype of the sensor board.

Sensor Type	Measurand	Bit resolution	Device
Movement	3-axis Acceleration, 3-axis Magnetometer, 3-axis gyroscope	16 bits	MPU9250
Humidity	Relative Humidity, temperature	16 bits	HDC1000

Each of the sensors installed in the interspace between the insulating layer and the wall represents a node of a Wireless Sensor Network (WSN). Each of the sensors transmits asynchronously the data acquired from the transducers one time every three hours toward the network coordinator, i.e. the GW in Figure 4. The WSN can be composed of a variable number of nodes. Since a connectionless transmission is used, there is no limit to the maximum number of nodes composing the WSN. The maximum number of nodes composing the WSN depends on the radio coverage and the computational resources of the GW itself. Approximately, is required a GW per each installation. The data acquired by the GW are then transmitted using HTTP RESTful services, through the public network, to a remote server for data processes (the data flow is shown in figure 4).

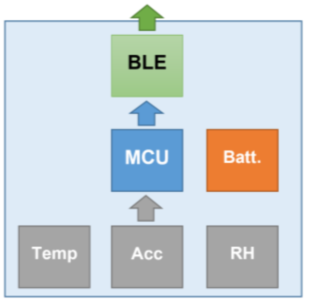


Figure 3: Block diagram of the sensor board. Relative Humidity (RH); Bluetooth Low Energy (BLE); Micro Controller Unit (MCU).

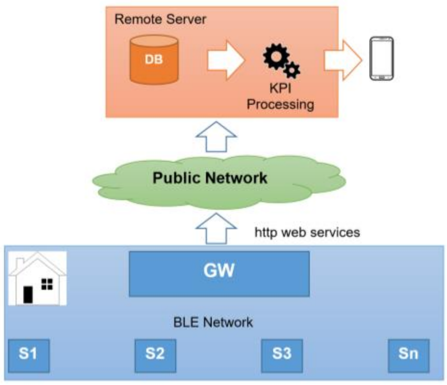


Figure 4: The architecture of the data analysis system.

A NoSQL DB (based on InfluxDB) is used to store the stream of data generated by the sensors. Computational resources are used to process the KPIs using the data stored in the DB. The data and the processed KPIs can be visualized using a dashboard, realized using the Grafana Framework. An example of the dashboard is shown in Figure 5. The proposed system can be easily scaled to process the data coming from different installations. The data from the sensors can be used to inform about the energy performance of the refurbished building and provide insight into discrepancies between the expected performance after the retrofit and the actual situation. The energy model of the building can be used to calculate the expected energy consumption and the temperature measured by the sensor can be used to calculate the energy flow for transmission.

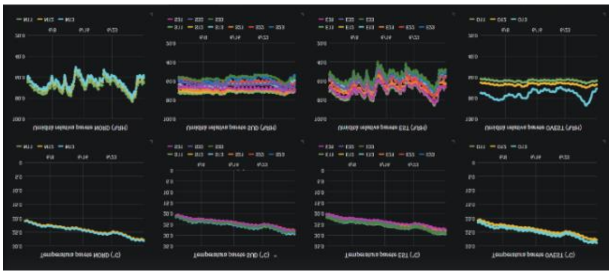


Figure 5: An example of the dashboard used to present the sensor data.

If the values are aligned, a discrepancy in consumption in winter could be related to the ventilation losses and to occupants' behavior for improving IAQ (indoor air quality) to the detriment of the energy saving. On the contrary, if the temperature is lower than expected and the ETICS is not properly insulating, it is possible to activate a detection procedure to understand if an issue is in progress for example for an infiltration which reduces drastically the insulation effect of the material. The experimental information model has been also used as an ideal Digital Twin of the test house working on the data mapping on the walls of the temperature distribution. Through VPL (Visual Programming Language) the sensors created into the information model for the different walls have been connected to the measures database of the value gathered from the sensors and the spread radius of the temperature variation have been visually approximated based on theoretical models since the empirical experience is not possible to exploit (Figure 6). However, the analysis of the data monitored showed that a difference of temperature due to the gradient distribution of the wall is detectable.

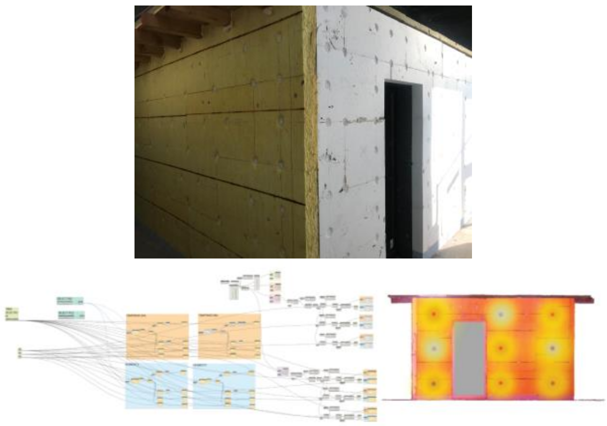


Figure 6: Test house before the installation of the finishing layers when it is possible to see the two target insulation materials, VPL model, and simulation of the spread radius of the temperature variation.

The data coming from the sensors, distributed on the walls, were collected and the difference of temperatures at different installation heights is used to define a sensor-based thermography of the situation of the walls (Figure 7).

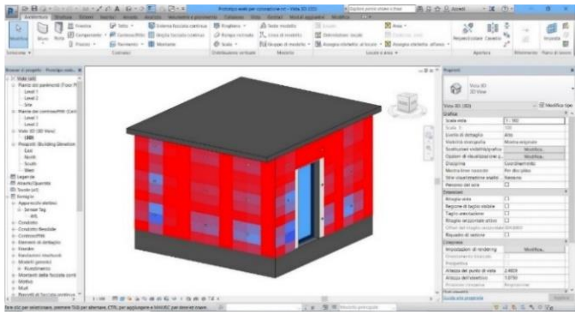


Figure 7: BIM based thermography related to the measured temperatures.

Conclusions

This study demonstrates the significant potential of blockchain technology in the AECO industry, particularly in addressing the challenges of traceability and transparency in supply chain scenarios where trust cannot be guaranteed. Our study focuses on the supply chain of ETICS panels for retrofitting existing buildings, which has been shown to suffer from misaligned incentives throughout the commissioning, installation, and operation of the panels.

Blockchain can provide trusted and transparent tracking of panels for accountability between supply chain actors and can use smart contracts to create performance-based incentive payments for a well-functioning system. For the necessary performance assessment, the sensor enabled IoT solution can provide real-time information over the entire building lifecycle, including the operations and maintenance (O&M) phase.

Our work explores an exemplary implementation of such a combined blockchain-IoT solution, focusing on a lifecycle perspective rather than just parts of the supply chain. We contribute to the emerging number of blockchain-IoT studies that implement energy performance smart contracts (e.g., Hunhevicz et al., 2022). Although the technical implementation and validation was performed on a test building, we believe it can provide valuable insights and inspiration for real-world implementation. Future research could extend to an actual construction project to further assess the cost and performance of the engineered system, compare between different available technical solutions, and evaluate the requirements and impact to industry beyond the focus on technology. Nonetheless, our findings demonstrate how the synergy between blockchain technology and sensor-based solutions can help shape the future of sustainable and resilient construction practices.

References

Bosché F., Ahmed M., Turkan Y., Haas C.T., Haas R., The value of integrating Scan-to-BIM and Scan-vs-BIM techniques for construction monitoring using laser scanning and BIM: The case of cylindrical MEP components, Automation in Construction, Volume 49, Part B, 2015, Pages 201-213, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2014.05.014.

FIEC. (2015). Construction 4.0. Eur. Constr. Ind. Fed. Retrieved from http://www.fiec.eu/en/themes-72/construction-40.aspx (accessed February, 2024).

García de Soto, B., Agustí-Juan, I., Joss, S., Hunhevicz, J. (2019). Implications of Construction 4.0 to the workforce and organizational structures. International Journal of Construction Management, 1–13. Retrieved from https://doi.org/10.1080/15623599.2019.1616414.

Hall, D. M., Whyte, J. K., & Lessing, J. (2020). Mirror-breaking strategies to enable digital manufacturing in Silicon Valley construction firms: a comparative case study. *Construction management and economics*, *38*(4), 322-339. https://doi.org/10.1080/01446193.2019.1656814

Hunhevicz, J. J., Motie, M., & Hall, D. M. (2022). Digital building twins and blockchain for performance-based (smart) contracts. Automation in Construction, 133, 103981. https://doi.org/10.1016/j.autcon.2021.103981

Hunhevicz, J.J., Hall, D.M., Do you need a blockchain in construction? Use case categories and decision framework for DLT design options, Adv. Eng. Informatics. 45 (2020) 101094. https://doi.org/10.1016/j.aei.2020.101094.

Kinnaird, C., Geipel, M. (2017). Blockchain Technology: How the Inventions Behind Bitcoin are Enabling a Network of Trust for the Built Environment. Arup. Retrieved from https://www.arup.com/perspectives/publications/research/section/blockchain-technology (accessed February 2024).

Klinc, R., Turk, Ž. (2019). Construction 4.0 – Digital Transformation of One of the Oldest Industries, 21, 292–410. Retrieved from https://doi.org/10.15458/ebr.92.

Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., Hoffmann, M., Industry 4.0, Bus. Inf. Syst. Eng. 6 (2014) 239–242. https://doi.org/10.1007/s12599-014-0334-4.

Li, J., Kassem, M., Watson, R., A Blockchain and Smart Contract-Based Framework to Increase Traceability of Built Assets, Proc. 37th CIB W78 Inf. Technol. Constr. Conf. (CIB W78). 2 (2020) 1–17.

M.S., M., Haritha, V., M.S., S., and M., T., "Comparison on Hyperledger Fabric and Hyperledger Composer of Block Chain Technology," 2019 International Conference on Intelligent Computing and Control Systems (ICCS), Madurai, India, 2019, pp. 1299-1305, doi: 10.1109/ICCS45141.2019.9065776.

Marzouk, M., Abdelaty, A., Monitoring thermal comfort in subways using building information modeling, Energy and Buildings, Volume 84, 2014, Pages 252-257, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2014.08.006.

McGlinn, K., Yuce, B., Wicaksono, H., Howell, S., Rezgui, Y., Usability evaluation of a web-based tool for supporting holistic building energy management, Automation in Construction, Volume 84, 2017, Pages 154-165, ISSN 0926-5805, https://doi.org/10.1016/j.autcon.2017.08.033.

Panteli, C., Kylili, A., Fokaides, P.A., Building information modelling applications in smart buildings: From design to commissioning and beyond A critical review, Journal of Cleaner Production, Volume 265, 2020, 121766, ISSN 0959-6526, https://doi.org/10.1016/j.jclepro.2020.121766.

Perera S., Nanayakkara S., Rodrigo M.N.N., Senaratne S., Weinand R., Blockchain technology: Is it hype or real in the construction industry?, J. Ind. Inf. Integr. 17 (2020) 100125. https://doi.org/10.1016/j.jii.2020.100125.

Sawhney, A., Riley, M., Irizarry, J. (2020). Construction 4.0. Routledge. Retrieved from https://doi.org/10.1201/9780429398100.

Tomašević, N.M., Batić, M.C., Blanes, L.M., Keane, M.M, Vraneš, S., Ontology-based facility data model for energy management, Advanced Engineering Informatics, Volume 29, Issue 4, 2015, Pages 971-984, ISSN 1474-0346, https://doi.org/10.1016/j.aei.2015.09.003

Viryasitavat, W., Xu, L., Da Bi, Z., Sapsomboon, A., Blockchain-based business process management (BPM) framework for service composition in industry 4.0, J. Intell. Manuf. (2018) 1–12. https://doi.org/10.1007/s10845-018-1422-y.

Watson, R., Kassem, M., Li, J. (2019). Traceability for Built Assets: Proposed Framework for a Digital Record. Retrieved from http://nrl.northumbria.ac.uk/39912/ and https://doi.org/10.3311/ccc2019-068

X. (Alice) Qian, E. Papadonikolaki, Shifting trust in construction supply chains through blockchain technology, Eng. Constr. Archit. Manag. ahead-of-p (2020). https://doi.org/10.1108/ECAM-12-2019-0676.

Ye, Z., Yin, M., Tang, L., Jiang, H. (2018). Cup-of-Water theory: A review on the interaction of BIM, IoT and blockchain during the whole building lifecycle, Pages 478-486 (2018 Proceedings of the 35th ISARC, Berlin, Germany, ISBN 978-3-00-060855-1, ISSN 2413-5844)

Yılmaz, D., Tanyer, A.M., Toker, I.D., A data-driven energy performance gap prediction model using machine learning, Renewable and Sustainable Energy Reviews, Volume 181, 2023, 113318, ISSN 1364-0321, https://doi.org/10.1016/j.rser.2023.113318.

Zhai Y., Chen K., Zhou J.X., Cao J., Lyu Z., Jin X., Shen G.Q.P., Lu W., Huang G.Q., An Internet of Things-enabled BIM platform for modular integrated construction: A case study in Hong Kong, Adv. Eng. Informatics. 42 (2019) 100997. https://doi.org/10.1016/j.aei.2019.100997.

ENHANCING CONTRACTUAL TIME PERFORMANCE MANAGEMENT: A HYPERLEDGER FABRIC-BASED CONCEPTUAL FRAMEWORK

Mahir Msawil[1], Jennifer Li[2], Faris Elghaish[3], David Greenwood[1], and Mohamad Kassem[4]
[1]Northumbria University, Newcastle, United Kingdom; [2]Northumbria University, Amsterdam, The Netherlands; [3]Queen's University Belfast, Belfast, United Kingdom; [4]Newcastle University, Newcastle, United Kingdom

Abstract

Time performance management is a core contractual mechanism that contributes to the success of construction projects. However, its interdependent processes face underlying challenges. Hence, this paper aims to build a conceptual framework of how Hyperledger Fabric can contribute to addressing challenges encountered in practice. The results reveal that existing digital solutions implemented on construction projects can act as software oracles for Hyperledger Fabric blockchain and the contractual logic can be converted into smart contract functions at the process level. The paper acts as a point of departure to develop a proof-of-concept for practical adoption within the context of FIDIC Red Book 2017.

Introduction

The delivery of large construction projects hinges on contractual governance (Zhang et al., 2023) and interdependent contractual mechanisms (Chen et al., 2018). The logical execution of those contractual mechanisms is operated by a construction contract administration (CCA) system. Among the various elements of the CCA system, time performance management is a core mechanism that contributes to the success of construction projects (Hetemi et al., 2020) and benefits all organisations involved (Maylor et al., 2023). Yet, time performance management continues to be perceived as fraught with challenges (Abdul-Malak and Mehdi, 2020). One main reason for this is that it is underpinned by interdependent processes with each process having its own underlying challenges (Farghaly et al., 2024).

To this effect, a host of research efforts have attempted to tackle an array of inherent challenges associated with each process through the adoption of advanced digital technologies. For example, some studies developed AI-based solutions to improve schedule generations (Soman and Molina-Solana, 2022). Another strand focused on enhancing progress reporting (El-Omari and Moselhi, 2011; Hamledari et al., 2017). In addition, attempts at improving the related claims management process were made (Palaneeswaran and Kumaraswamy, 2008; Ali et al., 2020). Furthermore, improving the analysis of schedule delays has occupied a considerable space in the extant literature (Bhih and Hegazy, 2021; Guévremont and Hammad, 2018). Despite their empirically validated benefits, those research efforts tend to treat their respective process as independent from the other processes. In addition, they tend to be rooted in the perspective of a single organisation (centralised) while overlooking the nexus of autonomous agents involved in the CCA process.

With the advent of Hyperledger Fabric (HLF, 2023), a permissioned blockchain, this overlooked perspective and its associated aspect can be tackled. A number of studies have proposed HLF-based solutions to tackle inherent challenges of contractual processes in construction projects while taking into account the inter-organisational perspective. Proof-of-concept-based studies have focused on financial management (Elghaish et al., 2022; Cheng et al., 2023), document and records management (Das et al., 2022), quality management (Sheng et al., 2020), and other relevant processes (Zhong et al., 2022). Yet, this emergent research line has not advanced the interplay between HLF and the processes associated with time performance management. In addition, while it encapsulates the inter-organisational perspective it tends to focus on developing the HLF-based solutions without rooting them in a defined contractual governance (Msawil et al., 2022). This is, of course, often due to a deliberately limited research scope. As a result, it can be argued that interested scholars and practitioners have remained uninformed of the practical and applicable alignment between this advanced technology and current contractual practices, on the one hand, and existing digital solutions implemented on construction projects, on the other hand. This argument is echoed by Çıdık and Boyd (2022).

Hence, this paper proposes an HLF-based conceptual framework for enhancing contractual time performance management. The framework attempts to integrate existing digital solutions of relevant processes into a single digital platform, while grounding the framework within the context of the FIDIC Red Book (RB) 2017 (FIDIC, 2017) - a standard contract model devised by the International Federation for Consulting Engineers (FIDIC) and it is widely adopted worldwide for infrastructure projects procured through the design-bid-build route.

Conceptual background

The conceptual background first formulates the research problem in the form of challenges. Subsequently, it summarises existing HLF-based solutions developed for various CCA processes before justifying the research need.

Challenges facing time performance management

Both scholars and practitioners continue to regard time performance management mechanisms to be fraught with challenges regardless of the form of contractual governance adopted. Table 1 consolidates relevant identified challenges based on analysing literature against the various processes associated with this contractual mechanism. From an inter-organisational perspective, it can be inferred that the persistence of such current challenges hinders efficient and effective management of time performance, which can result in financial losses to all involved organisations (Elazouni et al., 2023). In addition, such challenges damage the business relationship to varying degrees (e.g., due to unsettled time extension claims, legal disputes) (Jelodar et al., 2016). These challenges highlight the need for developing an innovative solution to achieve an efficient and effective time performance management from an inter-organisational perspective while respecting contractual governance. In this paper, it is argued that such an innovative solution can be realised through the adoption of Hyperledger Fabric (HLF), a permissioned blockchain protocol, and its associated smart contract technical functionality.

Table 1: Summary of analysed challenges against the relevant processes

Process	Reference (indicative)	Challenge
Programme production	Hong et al. (2021)	Ch-1: Incompleteness of programme
	Koc and Gurgun (2022)	Ch-2: Misapplication of related sub-clauses
	Hegazy and Kamarah (2022)	Ch-3: Unoptimised programmes
Progress reporting	al Qady and Kandil (2015)	Ch-4: Unspecified types of records
	Fan et al. (2015)	Ch-5: Absence of records collection tools
	Carson (2006)	Ch-6: Distortion of information
	Ibbs et al. (2017)	Ch-7: Lack of mitigation efforts records
Programme update	Carson (2006)	Ch-8: Concealment of schedule delays
	Demirel et al. (2019)	Ch-9: Lack of defined communication channels
	Seo and Kang (2020)	Ch-10: Information asymmetry at inter-organisational level
	Bhih and Hegazy (2021)	Ch-11: Inaccurate schedule and resource updates
Time extension claims management	Seo et al. (2021)	Ch-12: Missing contractual obligations
	Abdul-Malak and Mehdi (2020)	Ch-13: Inaction due to optimism bias
	Seneviratne and Michael (2020)	Ch-14: Missed time-bar
	Jelodar et al. (2016)	Ch-15: Avoidance of discussion and poor communication
	Francis et al. (2022)	Ch-16: Non-finality of decisions
Delay analysis	el Nemr (2021	Ch-17: Inaccessibility to information records
	Palaneeswaran and Kumaraswamy (2008)	Ch-18: Insufficient and contradictory information
	Guévremont and Hammad (2021)	Ch-19: Difficulty with establishing continuous reasoned causation
	el Nemr (2021)	Ch-20: Polarisation of views on delay analysis method selection

Hyperledger Fabric-based applications for CCA

HLF is a digital implementation protocol that was specifically devised to execute permissioned blockchain-based business processes. Accordingly, studies have justified the suitability/applicability of HLF as a permissioned blockchain to the managerial and contractual processes of construction projects (Tao et al., 2021; Elghaish et al., 2020). HLF blockchain-governed transactions are executed through a 'chaincode' that contains 'smart contract' functions (HLF, 2023). At its core, a blockchain-based 'smart contract' can be described as a logical structure of a programming code that implements the business logic of a contractual process (Mason, 2021). Feeding the required input data into smart contract functions can be achieved by several methods: it may be from previous transactions immutably stored in the HLF blockchain (Sheng et al.et al., 2020); from so-called 'oracles' based on hardware, software, or human agnets (Lu et al., 2021); or a combination of these methods (Mason, 2021).

In the research stream that concentrates on blockchain applications in the construction project management (CPM) sphere, a broad range of HLF-based solutions has emerged. Among this growing stream, only a limited number of studies can be mapped to the particular contractual processes/mechanisms of CCA (Msawil et al., 2022). These are presented in Table 2.

Table 2 is evidence that these identified CCA process challenges can be resolved by the HLF-based application's unique features (authorised accessibility, symmetric information, digital security, immutability, process streamlining, and simultaneous traceability of performed contractual obligations).Yet, the potential applicability of HLF to improve processes involved in the contractual time performance management in particular has not been unearthed, at least in the context of a defined contractual governance. In general, it is recognised that contextualising innovative digital solutions within existing practices can act as an enabler to their possible adoption in construction (Çıdık and Boyd, 2022).

Table 2: Relevant Hyperledger Fabric-based applications for CCA

Contractual Mechanism	Ref.	Focus of study	Oracle
Financial management	Cheng et al. (2023)	Developing a framework to facilitate construction cost management.	Human
	Wu et al. (2022)	Developing an HLF-based smart contract system for smart payment.	Human
	Elghaish et al. (2020)	Developing an financial system for Integrated Project Delivery procurement.	Software (BIM-based tools)
	Elghaish et al. (2022)	Developing a financial management solution for traditional procurement routes.	Software (BIM-based tools)
Performance reporting	Wang et al. (2020)	Building a framework for information sharing.	Human
Quality and acceptance management	Sheng et al. (2020)	Presenting a framework to support the management of quality information.	Human
	Lu et al. (2022)	Developing a technical solution to support remote E-inspection of building projects.	Human
Document and record management	Tao et al. (2021)	Proposing a framework for distributed common data environment	Human
	Das et al. (2022)	Developing a secure document management system to integrate project participants and document/record silos.	Human
Project 'legal' governance	Zhong et al. (2022)	Proposing an HLF framework to automatically monitor and record environmental pollution.	Hardware (Internet of Things (IoT) sensors)

Hence, this paper attempts to tackle the preceding formulated problematisation through the development of a conceptual framework. To this effect, it is conceptualised by converting the processes associated with time performance management under FIDIC RB 2017 into HLF-based technical architecture (on-chain) linked with available software implemented on construction projects (off-chain). Together with the analysed studies, the proposed conceptual framework paves the way forward to advancing the understanding of the potential applicability of HLF to improving time performance management as a practical/ particular instance of the wider CCA system.

Contractual derivation of the framework

This section derives the various processes and subprocesses associated with contractual time performance management. The derivation is based on analysing relevant contractual sub-clauses of FIDIC RB 2017 as an application context. Figure 1 shows the resulting analysis of the processes and subprocesses along with three forms of their contractual interdependencies: explicit, implicit, and absent. Due to limited space, the illustrated processes and subprocesses are briefly explained herein as deduced from their corresponding sub-clauses at the contractual process level.

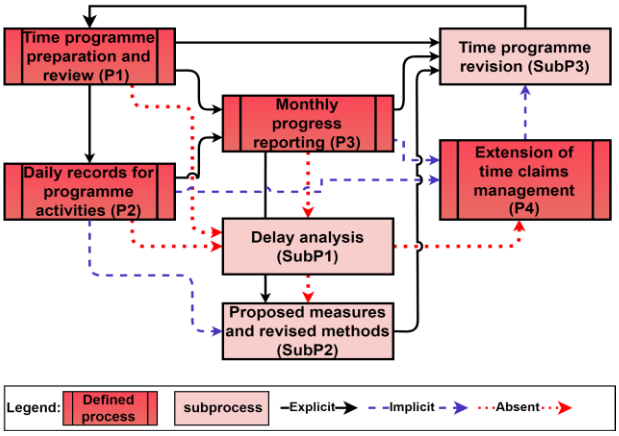


Figure 1: Key processes and subprocesses and their interdependencies

Time programme preparation and review (P1): this is derived from sub-clause 8.3 [Programme]. It places an obligation on the Contractor to prepare and submit the time programme to the Engineer, taking into account a wide range of factors and aspects.

Daily records for programme activities (P2): this is based on sub-clauses 6.10 [Contractor's Records] and 4.4.2 [As-Built Records]. It places an obligation on the Contractor to prepare and keep daily progress and as-built records for each activity shown in the programme to feed data into P3.

Monthly progress reports (P3): this is extracted from sub-clause 4.20 [Progress Reports] wherein a relatively long list of requirements is placed on the Contractor to prepare the monthly progress reports and submit them to the Engineer. Among other aspects, the requirements include a comparison of the actual and planned progress based on the programme update.

Under both P2 and P3, if a delay event that impacts the time performance occurs due to the Contractor's own cause, then the Contractor needs to propose [mitigation] measures to recover such time impacts. Such a Contractor's delay can be detected from the daily and monthly records and can be achieved through subprocess 1 (**SubP1**).

Delay analysis (SubP1): the contractual analysis of FIDIC RB 2017 did not reveal an explicit method of delay analysis. It recommends seeking advice from legal experts on the selection of an appropriate delay analysis method in line with the governing law of the Contract [project] and including it in the particular conditions.

Proposed measures and revised methods (SubP2): this can be considered as an undefined subprocess and it is inferred from sub-clauses 4.20 [Progress Reports] and 8.7 [Rate of Progress]. Any such measures and methods should be immediately incorporated into the programme to enable instant fulfilment of obligations and protection of rights of all involved organisations. Hence, it is logical to include this contractually inferred subprocess therein to place an emphasis on its role in closing the interdependence loop and linking the earlier processes to SubP3.

Time programme revision (SubP3): as a result of any changes to the initial programme that may result from (**P3**) and/or the adoption of measures/ methods under (**SubP2**), the Contractor shall revise the programme accordingly as a self-performed contractual obligation based on sub-clause 8.3 [Programme] or Engineer's instructions under 8.7 [Rate of Progress]. The revised programme becomes the programme for the subsequent period.

Under the aforementioned relevant processes and subprocesses, if a delay event occurs which is not a Contractor's responsibility, then the Contractor is entitled to an Extension of Time (EoT) as a contractual remedy.

Extension of time claims management (P4): this process is derived from sub-clauses 8.5 [Extension of Time for Completion], 20.2 [Claims for Payment and/or EoT], and 3.7 [Agreement or Determination]. It involves multi-sequential stages that hinge on the coordination and control-related obligations performed mainly by the Engineer and Contractor (with a minor involvement on the Employer's part). As evident in Figure 1, this process is contractually disconnected from the other processes, despite the very practical fact that its successful execution is influenced by information contained in (**P1, P2, P3, SubP1, Sub2, and SubP3**) and its outcome necessitates revising the programme. The authors emphasise here that this apparent 'disconnect' has been disregarded, as a critical review of the drafting of FIDIC RB 2017 is beyond the scope of this paper.

Based on the analysis above, it is self-evident that the means of improving contractual time performance management of FIDIC RB 2017 rests in addressing these processes and subprocesses. Yet, integrating them simultaneously while providing trust at an inter-organisational level and respecting the contractual governance is argued to have remained elusive. The following section presents an attempt to tackle this elusiveness.

The proposed HLF-based chained time performance management framework

This section presents the proposed conceptual framework by drawing on the tabulated HLF-based studies summarised in Table 2. As Figure 2 shows, the proposed concept connects the three involved organisations represented by human agents in a distributed manner around the notion of 'chained time performance management'. According to HLF design philosophy, each organisation is represented as a digital node and authorises its own human agents and software to act as peers on HLF network.

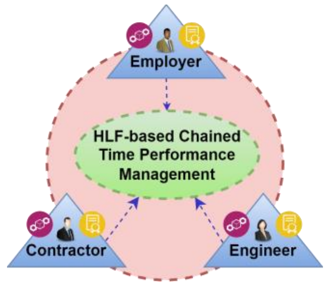


Figure 2: The multi-organisations connected to chained time management performance

Figure 3 shows the three key components of the framework and their digital relationships. In what follows, the key components along with their digital flow are described.

Off-chain: this component represents a construction contract environment where scheduled activities take place at a construction site, human agents perform contractual tasks, and software-based tools are implemented to aid human agents. In addition, it includes the defined contractual governance that regulates the interactions among the involved organisations. Of particular relevance to the current conceptual framework are human agents (human oracles) entrusted with performing a variety of contractual tasks and digital solutions (software oracles) implemented at the project level. The former includes project managers, project controls engineers, and quantity surveyors. The latter can include scheduling software (e.g., Microsoft Project and Oracle Primavera P6), BIM-based solutions, AI-based tools, and electronic document management systems. Both oracles interact with each other in a dynamic manner with human agents entrusted with data inputs while digital solutions entrusted with generating outputs. In this conceptual framework, it is proposed that both oracle types can act as dynamic data pipeline that feed required data variables of blockchain-based smart contracts through middleware.

Middleware: this component acts as a digital bridge between the off-chain oracles and the on-chain side. It performs a range of digital operations including invocation, queries, and updates. This component hinges on database servers and Application Programming Interfaces (APIs). It serves multiple technical functions. First, it enables human agents to interact with the HLF blockchain. This interaction can be to invoke a smart contract function by entering the required data variables or to query the blockchain ledger. Second, it enables available digital solutions (software oracles) to trigger a smart contract function in response to the occurrence of an event or query the blockchain ledger at regular coded intervals (e.g., each 24 hours) or invoke a specific function when pre-defined conditions are met based on programmed conditional statements. In addition, this set of middleware serves to return results retrieved from the ledger (on-chain) to human agents at both the project and inter-organisational levels. Furthermore, the database

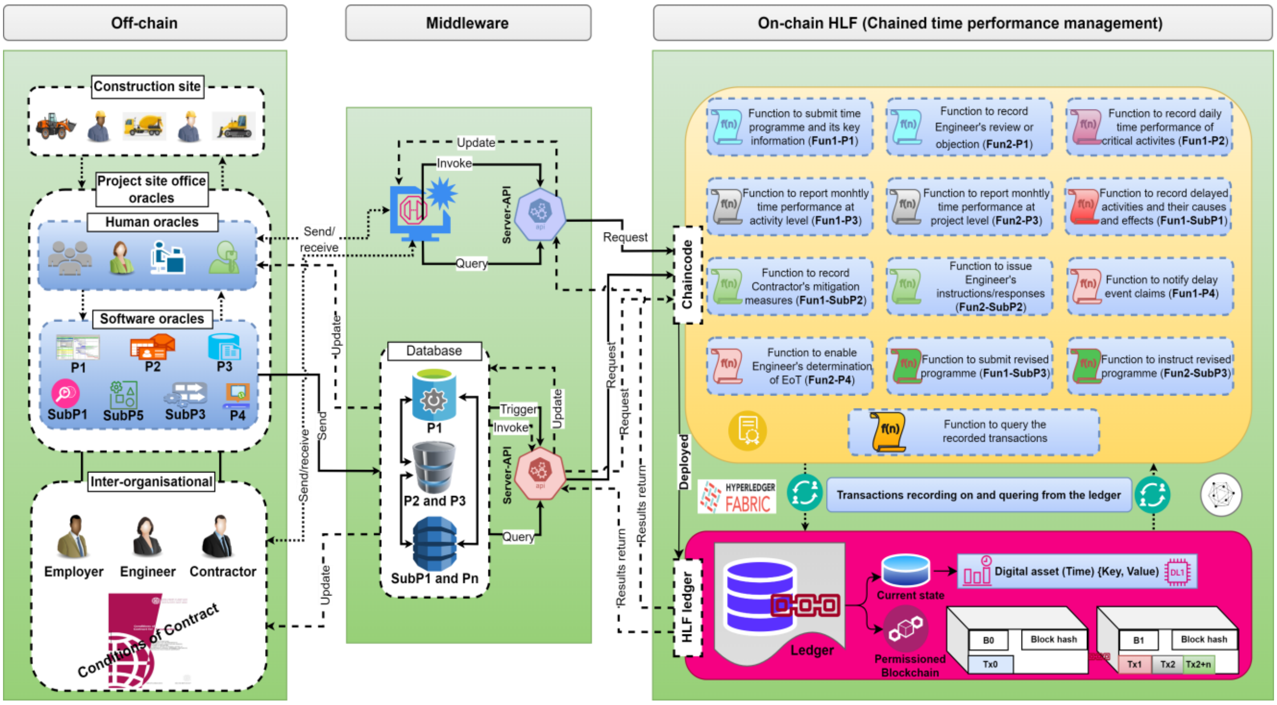


Figure 3: HLF-based chained time performance management framework

updates itself for subsequent digital operations (e.g., automatic invocation). Due to limited space, the solid and dotted arrows shown in Figure 3 conceptually show those digital operations. The detailed rules of the conceptualised operations and their interdependencies along with the server selection are currently under development.

On-chain HLF: the design of this component follows the HLF design philosophy. It includes a single channel that represents 'chained time performance management'. Authorised peer nodes at both the project and inter-organisational levels as well authorised software oracles are given access to the channel. On each authorised peer node, both the ledger and chaincode are deployed.

- **The chaincode**: the chaincode contains a set of the application-based smart contract functions coded to represent the contractual logic (see the previous contractual derivation of the framework) under each process or subprocess. In this paper, this set is called application-based smart contract functions which differs from system-based smart contract functions (e.g., initialise the ledger). Currently, the parameters that reflect the contractual logic of the relevant sub-clauses of FIDIC RB 2017 are being defined and analysed. To automatically execute a coded contractual task, the coded parameters inside a given smart contract function will receive data variables either from human agents or software agents or both, through middleware. If the received data variables meet the coded conditions, then the smart contract function executes a transaction that is immutability recorded on the blockchain. Subsequently, the various cumulative transactions can be queried by human agents or software agents to monitor the time performance and to automatically trigger specific action as coded.
- **The ledger:** the ledger consists of two interconnected parts: the blockchain and the current [world] state. The blockchain stores and processes all valid transactions whereas the current state shows the latest value of the defined asset (see the explanation below).

In addition to the preceding chaincode and ledger, the conceptual framework recognises the other digital HLF components (e.g., certificate authorities, endorsement policy, system-based chaincode, and identifying an ordering service for validating transactions and ordering blocks). For clarity and simplicity, those digital components are not shown in Figure 3.

Conceptual explanation of the proposed framework

This paper proposes a novel approach to managing time performance of construction projects in line with the objective ontology of FIDIC RB 2017. The proposed framework argues that time can be represented as a single digital asset in line with the philosophy of blockchain technology. To this effect, the contractual time is argued to be an asset owned by the Employer and it is allocated to the Contractor with the aim of using it to complete the construction project. The Engineer has a dual role in managing this asset: to monitor the Contractor's contractual state of time at defined intervals as stipulated in the Contract [project] data and to determine when the Contractor needs an increase in the allocated time (e.g., in case of a delay event which is not a Contractor's responsibility). This establishes an interdependent proactive approach to time performance management whereby all the processes and subprocesses are digitally connected through the adoption of an HLF-based solution.

Accordingly, this proposed approach chains and integrates the peers and their respective organisations enabling them to have complete status of time performance at any given moment with a history of fulfilled and remaining contractual obligations.

In addition, the human and software oracles providing input data are both identified in the recorded transactions based on HLF certificate authorities. On the part of human oracles, this increases the level of due-diligence and accountability in respect of data inputs. The need for increased due-diligence assessment fundamentally stems from the prospective awareness that a ‘smart contract’ function cannot be invoked without the input of all required data variables while taking into account that there will be neither an opportunity to change any input data nor to revoke recorded resultant transactions due to immutability. This awareness will substantially reduce erroneous data inputs and, hence, increase the likelihood of keeping the permissioned blockchain input error free. With this in mind, the data variables can be obtained from available digital solutions as software oracles.

It is envisaged that existing digital solutions can act as software oracles to feed smart contract functions with required data variables. However, an unintended limitation of current available digital solutions is that they tend to treat their digitalised processes as being independent which are in fact interdependent. Hence the proposed framework integrates those solutions (deployed off-chain) with their corresponding channel (on-chain ledger and smart contract functions) through APIs and databases that can technically be developed. In turn, smart contract functions are interconnected on the HLF platform with the capability of calling each other as well as calling APIs to obtain and feed data from the relevant processes of the ‘chained time performance management’.

The logical flow of this proposed approach follows contractual processes described in FIDIC RB 2017 with slight adjustments in terms of interdependence. This logical flow chains the entire associated processes from the commencement date of the Contract (i.e., project) to the date of completion in the form of a defined digital asset that represents the contractual time. The state changes of the defined asset are automatically and immutability recorded on the ledger through a series of digital transactions. As an example, the progress reporting process (**P3**) can be linked to the EoT claims management (**P4**) process through their respective smart contract functions supported by blockchain-governed API servers. With this linkage, as soon as the update of a given activity on the programme shows a negative float (i.e., delay) of a pre-defined numeric value while following a programmed delay analysis (**SubP1**) to detect the allocated delay responsibility, the ‘notice’ of EoT claim can be triggered through a smart contract function. As a result, the EoT claim management process automatically starts if that activity is not the Contractor’s responsibility. If, on the other hand, the delay is the Contractor’s responsibility then **SubP2** (see above) is invoked. With those interdependent processes and subprocesses being digitally executed, the relevant information flow coupled with the contractual obligations are captured and recorded in HLF-based transactions while being visible and immutably traceable at the inter-organisational level.

Discussion

In developing the notion and its conceptual framework, a number of findings emerged. In contrast with earlier HLF-studies, the conceptual framework highlights the possibility of enhancing the management of time performance by capitalising on HLF capabilities and connecting them with human and software oracles. It proposes a set of smart contract functions to enable the execution of various contractual processes and subprocesses as derived from FIDIC RB 2017.

Deploying and operating those prospective smart contract functions on an HLF-based platform can mitigate a number of inherent challenges. With reference to Table 1 above, examples at the project level include misapplication of related sub-clauses (Ch-2), missing contractual obligations (Ch-12), inaction due to optimism bias (Ch-13), and missed time-bar (Ch-14). Challenges addressed at the inter-organisational level include distortion of information (Ch-6), information asymmetry (Ch-10), avoidance of discussion and poor communication (Ch-15), and difficulty with establishing continuous reasoned causation (Ch-19). For both levels, it can be posited that the net foreseeable result is enhanced control and coordination of contractual obligations and information in respect of time performance management.

Evidently, with minor adjustments to the existing contractual logic of FIDIC RB 2017 at the contractual process level in terms of connection, the notion of ‘chained time performance management’ can be realised. For example, such adjustments may involve explicitly connecting the EoT claims management process with the other processes while defining a delay analysis protocol in the particular conditions of a construction contract. Yet, these adjustments should be coordinated and cross checked across the whole Contract to avoid unintended consequences. Prospective adopters can seek contractual and legal advice to verify and validate such adjustments before scaling the proposed framework into technical development and production.

Conclusions

This paper presents a conceptual framework for enhancing contractual time performance management from an inter-organisational perspective while respecting contractual governance. With available digital solutions linked to HLF as a digital core coupled with minor adjustments to contractual logic, a way to enhancing contractual time performance management can be built in a single digital environment.

This paper has implications for practice and research. In practical terms, it establishes an adaptable conversion path from practical problems associated with contractual time performance management to an HLF-based framework. This path may help interested practitioners adapt the approaches presented herein to other mechanisms of FIDIC RB 2017 and other standard contract models. Along the same line, it highlights the possibility of using existing digital solutions, implemented in practice, as software oracles to leverage advanced HLF-based solutions. Theoretical contribution to the body of knowledge was realised in two ways. First, in advancing the case of blockchain in CCA and construction at large by proposing an adaptable HLF-based framework for time performance management. Second, by offering an application reference within a defined contractual context that can be further validated by academic researchers.

Further work is currently being carried out to develop and refine the application framework. This includes: (*i*) defining the parameters of the proposed smart contract functions, (*ii*) coding the smart contract functions using a suitable programming language, and (*iii*) developing API rules and their technical requirements. Accordingly, the framework will be instantiated through the demonstration of a proof-of-concept on an infrastructure project to realise the notion of 'chained time performance management' for large construction projects.

Acknowledgments

The authors would like to acknowledge the funding from the International Centre for Connected construction (IC3) at Northumbria University under which this research was made possible.

References

Abdul-Malak, M.-A.U. and Mehdi, R.K. (2020) Investigation of the Reasonable Time Computation under Time-at-Large Construction Schedule Disputes. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 12(1).

Abdul-Malak, M.-A.U., Srour, A.H. and Demachkieh, F.S. (2020) Decision-Making Governance Platforms for the Progression of Construction Claims and Disputes. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 12(3).

Ali, B., Zahoor, H., Nasir, A.R., Maqsoom, A., Khan, R.W.A. and Mazher, K.M. (2020) BIM-based claims management system: A centralized information repository for extension of time claims. Automation in Construction, 110.

Bhih, M. and Hegazy, T. (2021) Enhanced Daily Windows Delay-Analysis Technique. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 13(3).

Carson, C. (2006) Claims Analysis Nested in Schedule Updates. AACE International Transactions, PS61.

Chen, Y., Wang, W., Zhang, S. and You, J. (2018) Understanding the multiple functions of construction contracts: the anatomy of FIDIC model contracts. Construction Management and Economics, 36(8) 472–485.

Cheng, J.C.P., Liu, H., Gan, V.J.L., Das, M., Tao, X. and Zhou, S. (2023) Construction cost management using blockchain and encryption. Automation in Construction, 152.

Çıdık, M.S. and Boyd, D. (2022) Value implication of digital transformation: the impact of the commodification of information. Construction Management and Economics, 40(11–12) 903–917.

Das, M., Tao, X., Liu, Y. and Cheng, J.C.P. (2022) A blockchain-based integrated document management framework for construction applications. Automation in Construction, 133.

Demirel, H.Ç., Volker, L., Leendertse, W. and Hertogh, M. (2019) Dealing with Contract Variations in PPPs: Social Mechanisms and Contract Management in Infrastructure Projects. Journal of Construction Engineering and Management, 145(11).

Elazouni, A., Gajpal, Y. and Fares, A. (2023) Negotiating win-win payment terms between construction contractors and subcontractors. Automation in Construction, 146.

Elghaish, F., Abrishami, S. and Hosseini, M.R. (2020) Integrated project delivery with blockchain: An automated financial system. Automation in Construction, 114.

Elghaish, F., Pour Rahimian, F., Hosseini, M.R., Edwards, D. and Shelbourn, M. (2022) Financial management of construction projects: Hyperledger fabric and chaincode solutions. Automation in Construction, 137.

El-Omari, S. and Moselhi, O. (2011) Integrating automated data acquisition technologies for progress reporting of construction projects. In: Automation in Construction. October 2011 699–705.

Fan, H., Xue, F. and Li, H. (2015) Project-Based As-Needed Information Retrieval from Unstructured AEC Documents. Journal of Management in Engineering, 31(1).

Farghaly, K., Soman, R. and Whyte, J. (2024) cSite ontology for production control of construction sites. Automation in Construction, 158.

FIDIC (2017) FIDIC. (2017). Conditions of contract for construction. Lausanne, Switzerland: FIDIC.

Francis, M., Ramachandra, T. and Perera, S. (2022) Disputes in Construction Projects: A Perspective of Project Characteristics. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 14(2).

Gibbs, D.-J., Lord, W., Emmitt, S. and Ruikar, K. (2017) Interactive Exhibit to Assist with Understanding Project Delays. Journal of Legal Affairs andDispute Resolution in Engineering and Constru ction, 9(1).

Guévremont, M. and Hammad, A. (2021) Ontology for Linking Delay Claims with 4D Simulation to Analyze Effects-Causes and Responsibilities. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 13(4).

Guévremont, M. and Hammad, A. (2018) Visualization of Delay Claim Analysis Using 4D Simulation. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 10(3) 05018002.

Hamledari, H., McCabe, B., Davari, S. and Shahi, A. (2017) Automated Schedule and Progress Updating of

IFC-Based 4D BIMs. Journal of Computing in Civil Engineering, 31(4) 04017012.

Hegazy, T. and Kamarah, E. (2022) Schedule optimization for scattered repetitive projects. Automation in Construction, 133.

Hetemi, E., Jerbrant, A. and Mere, J.O. (2020) Exploring the emergence of lock-in in large-scale projects: A process view. International Journal of Project Management, 38(1) 47–63.

HLF (2023) https://hyperledger-fabric.readthedocs.io/en/release-2.5/ Available from https://hyperledger-fabric.readthedocs.io/en/release-2.5/ [accessed 25 December 2023].

Hong, Y., Xie, H., Bhumbra, G. and Brilakis, I. (2022) Improving the accuracy of schedule information communication between humans and data. Advanced Engineering Informatics, 53.

Ibbs, W., Berry, M. and Sun, X. (2017) Visualizing Skipped and Out-of-Sequence Work. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 9(4).

Jelodar, M.B., Yiu, T.W. and Wilkinson, S. (2016) Dispute Manifestation and Relationship Quality in Practice. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 8(1).

Koc, K. and Gurgun, A.P. (2022) Ambiguity factors in construction contracts entailing conflicts. Engineering, Construction and Architectural Management, 29(5) 1946–1964.

Lu, W., Li, X., Xue, F., Zhao, R., Wu, L. and Yeh, A.G.O. (2021) Exploring smart construction objects as blockchain oracles in construction supply chain management. Automation in Construction, 129.

Lu, W., Wu, L., Xu, J. and Lou, J. (2022) Construction E-Inspection 2.0 in the COVID-19 Pandemic Era: A Blockchain-Based Technical Solution. Journal of Management in Engineering, 38(4).

Mason, J. (2021) Innovating construction law: Towards the digital age. . Routledge.

Maylor, H., Geraldi, J., Budzier, A., Turner, N. and Johnson, M. (2023) Mind the gap: Towards performance measurement beyond a plan-execute logic. International Journal of Project Management, 41(4).

Msawil, M., Greenwood, D. and Kassem, M. (2022) A Systematic evaluation of blockchain-enabled contract administration in construction projects. Automation in Construction, 143.

el Nemr, W. (2021) Change Orders after the Contract Completion Date and Contractual Defects in the 'Longest Path' Theory. Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 13(2).

Palaneeswaran, E. and Kumaraswamy, M.M. (2008) An integrated decision support system for dealing with time extension entitlements. Automation in Construction, 17(4) 425–438.

al Qady, M. and Kandil, A. (2015) Automatic Classification of Project Documents on the Basis of Text Content. Journal of Computing in Civil Engineering, 29(3) 04014043.

Seneviratne, K. and Michael, G.V. (2020) Disputes in time bar provisions for contractors' claims in standard form of contracts. International Journal of Construction Management, 20(4) 335–346.

Seo, W. and Kang, Y. (2020) Performance Indicators for the Claim Management of General Contractors. Journal of Management in Engineering, 36(6).

Seo, W., Kwak, Y.H. and Kang, Y. (2021) Relationship between Consistency and Performance in the Claim Management Process for Construction Projects. Journal of Management in Engineering, 37(6).

Sheng, D., Ding, L., Zhong, B., Love, P.E.D., Luo, H. and Chen, J. (2020) Construction quality information management with blockchains. Automation in Construction, 120.

Soman, R.K. and Molina-Solana, M. (2022) Automating look-ahead schedule generation for construction using linked-data based constraint checking and reinforcement learning. Automation in Construction, 134.

Tao, X., Das, M., Liu, Y. and Cheng, J.C.P. (2021) Distributed common data environment using blockchain and Interplanetary File System for secure BIM-based collaborative design. Automation in Construction, 130.

Wang, Z., Wang, T., Hu, H., Gong, J., Ren, X. and Xiao, Q. (2020) Blockchain-based framework for improving supply chain traceability and information sharing in precast construction. Automation in construction, 111 103063.

Wu, L., Lu, W. and Xu, J. (2022) Blockchain-based smart contract for smart payment in construction: A focus on the payment freezing and disbursement cycle. Frontiers of Engineering Management, 9(2) 177–195.

Zhang, L., Ding, T. and Fu, Y. (2023) Do measurement methods matter? A meta-analysis of antecedents and outcomes of contractual governance in interorganisational relationships. Production Planning and Control,.

Zhong, B., Guo, J., Zhang, L., Wu, H., Li, H. and Wang, Y. (2022) A blockchain-based framework for on-site construction environmental monitoring: Proof of concept. Building and Environment, 217.

BLOCKCHAIN SUPPORTED CLOSED LOOP IN CIRCULAR ECONOMY

Marijana Srećković[1], Dominik Hartmann[1], Thomas Preindl[2], Martin Kjäer[2], Andrea Kessler[3], and Peter Kneidinger[3]
[1]Institute of Building and Industrial Construction, TU Wien, Vienna, Austria
[2]Institute of Computer Engineering, TU Wien, Vienna, Austria
[3]materialnomaden, Vienna, Austria

Abstract

Existing challenges in policy, practice, and digital innovation, and their integration into Circular Economy (CE) result in a complex undertaking. Currently, not many frameworks for the digital optimization of a building's End-of-Life (EoL) and its related processes exist. This paper explores the digitalization of the EoL process with blockchain technology, by analysing an applied use case of a closed loop reuse product in Austria. A blockchain-based solution designed to enhance the verifiability of the reuse process is proposed, by introducing a privacy-preserving zero-knowledge proof application which can increase, among other things, transparency, and trust to promote CE.

Introduction

The transition towards a Circular Economy (CE) represents a paradigm shift from traditional, linear economic models towards ones that are restorative and regenerative by design (Bocken et al. 2016; Kirchherr et al. 2023). This shift is underpinned by the integration of digital technologies, which play a pivotal role in enhancing the sustainability and efficiency of various sectors, from waste management to construction (Kamble et al. 2018). The End-of-Life (EoL) process of products and materials is a critical aspect of the CE and gets increasingly important due to resource shortages and CO_2 emissions benchmarks (Gervasio & Dimova 2018) for certification of buildings (e.g. DGNB), also required for compliance with the EU Taxonomy (2024) for buildings. With the gradual digitalization of EoL processes (Dodampegama et al. 2024) a fundamental database can be created to promote recycling, reuse, and reduction of construction demolition waste (Papamichael et al. 2023), contributing to the conservation of resources and the mitigation of environmental pollution.

Studies like Pagoropoulos et al. (2017) emphasize the role of digital technologies, including Big Data and IoT, in advancing CE. Additionally, in the construction sector, digital platforms for Circular Economy are aimed at integrating digital technologies and data management to optimize resource use, addressing the industry's challenges (Kovacic et al. 2020). Findings from Çetin et al. (2022), highlight the potential of digital twins and scanning technologies in supporting CE strategies despite existing challenges. These insights underscore the digitalization's crucial role in advancing CE, calling for continued exploration and application of digital solutions.

This paper explores the digitalization of the EoL process with blockchain technology by analysing a use case of a closed-loop reuse-product, applied in Austria. This blockchain-based solution is designed to enhance the digital perspective of the reuse process, by introducing a verifiable certification concept. This concept ensures the correctness of provided claims without breaking privacy for the involved stakeholders, with zero-knowledge proofs. In this way the solution increases, among other things, transparency and trust to promote CE implementation.

Literature Background

Circular Economy

Kircher et al. (2017) present a comprehensive definition of CE derived from extensive literature review. They define CE as an economic system that shifts away from the EoL concept by prioritizing reduction, reuse, recycling, and recovery of materials throughout the production, distribution, and consumption processes. This operates at various levels including micro (products, companies, consumers), meso (eco-industrial parks), and macro (city, region, nation, and beyond), facilitated by innovative business models and responsible consumer behaviour. Consequently, there are diverse interpretations of CE and circularity (Kirchherr et al. 2017; Alhawari et al. 2021), all aiming at incorporating strategies for smarter product use (refuse, rethink, reduce), longer product lifetimes (reuse, quality, repair), and material circulation at the EoL (recycling, cascading), necessitating new business models for implementation (Brändström & Eriksson 2022).

In Architecture, Engineering, and Construction (AEC), the building life cycle comprises four stages (EN 15978:2011): production, construction, use, End-of-Life including deconstruction/demolition, transport, waste processing, and disposal. Currently, the EoL phase of the buildings' life cycle remains the least sustainable (Charef 2022). Defining End-of-Life within the CE context in AEC poses challenges due to the lack of clear standardization and regulation across different countries, and essential prerequisites for digitalization including the use of Building Information Modeling (BIM). The disconnection between BIM tools and EoL tools hampers comprehensive considerations of this phase (Akbarieh et al. 2020). The built environment represents a significant, yet poorly documented stock of material resources (Heisel & Rau-Oberhuber 2020). Absence of guidelines for material and component databases impedes conclusions drawn from EoL considerations, which should ideally be integrated into the design phase. Adopting circular design strategies (e.g., design for deconstruction, disassembly, adaptability, and flexibility) alone does not reduce environmental impact unless integrated and considered during the design phase to manifest in whole buildings' life cycle. Moreover, EoL asset management necessitates extensive data, requiring new workflows, processes, and business models (Charef 2022).

The digital potential within the reuse process, highlights how advancements in Artificial Intelligence and BIM are unlocking new possibilities for effective reuse strategies in the construction sector (Khosrowshahi, 2017). Initiatives like the EU-funded RRReMaker project, employ Artificial Intelligence for design for deconstruction, aimed at resolving current problems when planning with reuse components and their associated uncertainties such as quality etc. Simpler versions of digitalization are online platforms for reuse (e.g. Harvest Map; Concular), which are increasing in number and significance, and facilitating material reuse, especially with the goal of shifting the paradigm of this topic in society. The success of these platforms depends on the provision of comprehensive and detailed information about available materials, their provenance in combination with material condition (Veit et al. 2014).

The legal framework for CE in Austria is mainly regulated in the Austrian Waste Management Act (AWG 2002), which introduces a waste hierarchy. In addition, the Austrian standard *Dismantling of buildings as standard method for demolition* (ÖNORM B3151:2022) is applied on a voluntary basis, and gives a guideline on how to proceed in the building's EoL: demolition typically follows the reverse sequence of construction, focusing on maximizing reuse, preparation for reuse, or recycling of materials resulting from the process.

The digital potential within the reuse process, highlights how advancements in Artificial Intelligence and BIM are unlocking new possibilities for effective reuse strategies in the construction sector (Khosrowshahi, 2017). Initiatives like the EU-funded RRReMaker project, employ Artificial Intelligence for design for deconstruction, aimed at resolving current problems when planning with reuse components and their associated uncertainties such as quality etc. Simpler versions of digitalization are online platforms for reuse (e.g. Harvest Map; Concular), which are increasing in number and significance, and facilitating material reuse, especially with the goal of shifting the paradigm of this topic in society. The success of these platforms depends on the provision of comprehensive and detailed information about available materials, their provenance in combination with material condition (Veit et al. 2014).

Closed Loop in CE

The closed-loop system is an integral part of the CE, and it stands in contrast to the traditional linear economy, which follows a *take, make, dispose* (EMF 2013) model of production. By maintaining the utility and value of products through multiple life cycles, closed loop systems can significantly reduce environmental impact by lowering the demand for resource extraction and minimizing waste. Moreover, closed loop practices can lead to economic benefits by reducing the raw material procurement and waste disposal. Socially, they can contribute to job creation in the refurbishment and maintenance sectors (Zhang et al. 2024).

Braungart et al. (2008) differentiate between *cradle-to-grave* flows of materials and cyclical, *cradle-to-cradle* flows, whereby Stahel (1982; 1994, p.179; 2010) refers to *closed loop systems* instead of cyclical systems, introducing herewith two essentially different types of loops within a closed loop system (1) reuse of goods and (2) recycling of materials. Referring to the aforementioned authors including McDonough & Braungart (2002), Bocken et al. (2016, p. 309) define two cardinal strategies for the cycling of resources: 1) *slowing resource loops* - By crafting durable goods and implementing strategies like product-life extension, such as service loops for repairs and remanufacturing, the lifespan of products is prolonged, leading to a reduction in the rate of resource consumption; 2) *closing resource loops* - Recycling closes the loop between post-use and production, establishing a circular flow of resources.

In a closed loop system, the End-of-Life concept is replaced with restoration/refurbishment; products must be designed with their next use in mind. This means that once the initial use of a product is completed, it can be disassembled and the materials can be used again in their original form instead of being downcycled into lower-quality applications (Braungart et al. 2007).

However, the transition to a closed loop system is not without challenges. It requires a paradigm shift in design philosophy, a move towards business models that can support product life extension, and the development of markets for secondary materials. It also depends on the establishment of effective collection and processing systems that can support the return of used products for remanufacturing or recycling (Korhonen et al. 2018; Figge et al. 2023).

Expanded from Bocken et al. (2016, p. 309), for the purpose of the presented research, Figure 1 shows a circular life extension of a good/product in CE where the left section (black) demonstrates a system focused on waste recycling. The complete picture however, which

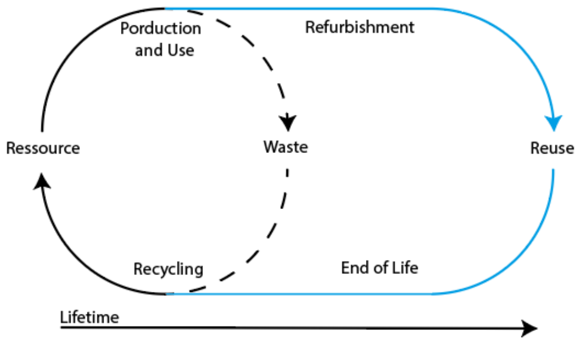


Figure 1: Closed Loop in CE

includes the right side (blue), reveals a fully realized CE that emphasizes keeping goods/products in use for as long as possible (extending the life span) through refurbishment and therefore in a closed loop system. The concept begins with resource extraction and product use, followed by product recycling at the end of its life. The expansion into refurbishment aims at slowing down the resource loop, hence promoting reuse and reducing the dependence on new resources.

Blockchain in EoL

Blockchain technology, synonymous with Distributed Ledger Technology (DLT) and initially introduced through Bitcoin (Nakamoto, 2008), has gained momentum across various sectors. Although the term was originally coined to describe Bitcoin's data structure, it now encompasses the underlying technology that permits decentralized state synchronization. This encompasses not only account balances, as with Bitcoin, but also a range of data types where the current state is critical for all participants. DLT's utility is further enhanced by Smart Contracts (Szabo 1997), with Ethereum blockchain (Buterin et al. 2013) emerging as the standard-bearer for Smart Contracts (SC) execution.

Blockchains, particularly in their original iterations like Bitcoin and Ethereum, operate as open networks with public data visibility. In contrast, private blockchains, which are generally permissioned and limit data access to known participants, are considered more appropriate for corporate use (Wüst et al. 2017). However, the efficacy of private blockchains has been questionable, potentially due to the absence of a "network effect." A hybrid approach that utilizes SCs on public networks alongside private databases interfacing through standardized protocols may offer a resolution.

Several studies illustrate blockchains potential to boost the life cycle approach within the AEC sector (Scott et al. 2021). Blockchain can ensure comprehensive traceability of materials and energy, aiding in the prediction of recycling and reuse outcomes for materials utilized in construction (Shojaei et al. 2021). The technology can also enhance the transparency and reliability of data in Environmental Product Declarations - EPDs (Rangelov et al. 2021). It facilitates end-to-end lifecycle assessment by documenting the metadata of raw materials, thereby verifying their provenance from source to construction site (Shojaei 2019). Copeland and Bilec (2020) have integrated RFID, BIM, and blockchain to specify components with a verifiable data trail throughout their lifecycle, and explored the integration with a crypto-economic incentive model for asset recycling. Additionally, blockchains can be employed for post-occupancy evaluations, with BIM serving as the repository and blockchain as the validator for built environment assets (Di Giuda et al. 2020).

Methodology

The presented research follows an integrated research approach, which, in a first step, was based on a comprehensive literature review, secondly on conducted and evaluated interviews with key stakeholders operating in AEC and EoL. These findings laid the groundwork for the following use case analysis of a reuse product in a CE closed loop, as outlined in the literature background. Furthermore, an important element of this ongoing research within the realm of the DiCYCLE project, is the integration of a blockchain into the framework of the presented use case called re:parkett. To create the presented use case, a workshop was conducted, where the details of the individual process steps were documented, and the current challenges in CE were discussed. The latter were the starting point for the applicability of our framework. The focus was on ensuring that the “executed” process was true to the “planned” process enabling herewith transparency and authenticity verification of the reuse product. In addition, a proposal for the creation of CO_2 certificates for the reuse product is laid out, adding also economic value, based on the partially required CO_2-audits for buildings. The aim is therefore to enable the creation of an unfalsifiable proof that the parquet flooring has been reused and is a closed loop product as shown in the case of re:parkett.

Use Case re:parkett

The starting point for re:parkett was the dismantling of significant areas of parquet flooring at the former Vienna Directorate and subsequently the requirements of a new project that needed installation of approximately 1200 square meters of parquet flooring for a building contractor in St. Pölten, Austria. This resulted in the novel idea of mechanically processing and further refurbishing individual parquet strips for reuse in new building projects.

In collaboration with a floor manufacturer, this concept transitioned into an implementation phase focused on product development. The objective was set to install 500 square meters of the reuse product by the end of 2022. A key aspect of this activity was the partnership with a floor laying company, which played a crucial role in both the dismantling and installation processes, thereby closing the loop of the product cycle. To ensure a systematic approach to the dismantling process, all three involved companies (the “circular design” company responsible for the assessment and collection of reusable parquet flooring, the manufacturer of flooring and the parquet floor layer) aimed at developing a manual that would integrate a standardized harvesting process into their business models. By April 2022, the first sample of the processed parquet was showcased publicly, signalling the readiness for market introduction. Efforts to expand distribution channels across various regions are ongoing, reflecting a scalable model for the adoption of sustainable practices in construction and renovation projects. This use case exemplifies the potential of collaborative innovation in promoting CE principles within the AEC industry.

Use Case Analysis & Discussion

In the research project DiCYCLE, a systematic approach to reprocessing harvested parquet floors has been developed, encompassing several key steps, as shown in Figure 2. The process is divided into four levels, with transport, deconstruction, manufacturing and certification. It initiates upon receiving an order, where builders, flooring layers, or flooring owners seek out the services of a company skilled to perform dismantling, recovery, or reprocessing of used/old parquet flooring. The second step involves a detailed and professional inventory of an existing or potential repository, with an assessment of the current condition of parquet flooring. Hence, an expert determines data regarding quantity, quality, and other features of the parquet, which are documented in a data sheet. Additionally, external parties can submit their own data sheets of potential repositories of reusable parquet flooring to the online platform/database (e.g. Harvest Map) of the “circular design” company which is then further checked and processed.

A crucial decision is made based on the collected information: whether to renovate the parquet flooring on-site, dismantle it for reuse, or dispose of it. In scenarios, where on-site renovation is deemed feasible, sometimes there might be a shift into dismantling and processing of flooring into a reuse product, due to previously false classification or the parquet owner’s change of plans. However, in some cases this leads to improper dismantling by untrained individuals due to time or budget constraints, potentially reducing the amount of material available for reuse.

In the “circular design” company’s database (Harvest Map) all relevant data on the dismantled parquet flooring regarding, quality and quantity is collected. However, the following step, which should entail an automatic digital data transfer between the “source database” and the manufacturer responsible for the refurbishment of the reusable parquet flooring, is currently missing due to the absence of a suitable implementable solution.

Going further into the process the manufacturer after receiving data regarding quality and availability of the flooring area, integrates this information into its own Enterprise Resource Planning (ERP) system, which is then utilized to determine the market demand for the product, as well as production plans and associated sales activities.

To summarize, the final decision on demolition or delivery to the manufacturer, is based on all previous steps in the laid-out process by the “circular design” company. For existing buildings, demolition is initiated, or the inventory in stock is delivered to the manufacturer for further processing. The dismantled product is then handled in the manufacturer's own production facility deemed at restoring/refurbishing the product, eventually ending up either being stored in the manufacturer's own warehouse or sold directly to the customers on demand, depending on the sales planning and production capacities.

In addition to all these steps, there is currently an ongoing effort to create a CO_2 certificate in the form of an EPD for the finished reuse product. This initiative aims to quantify the environmental impacts across all process steps, potentially offering the certificate with the floor delivery, or positioning the closed loop product on the market with its environmental background information. As mentioned, the schematic process is illustrated in Figure 2, providing an overview of the approach adopted in the DiCYCLE project.

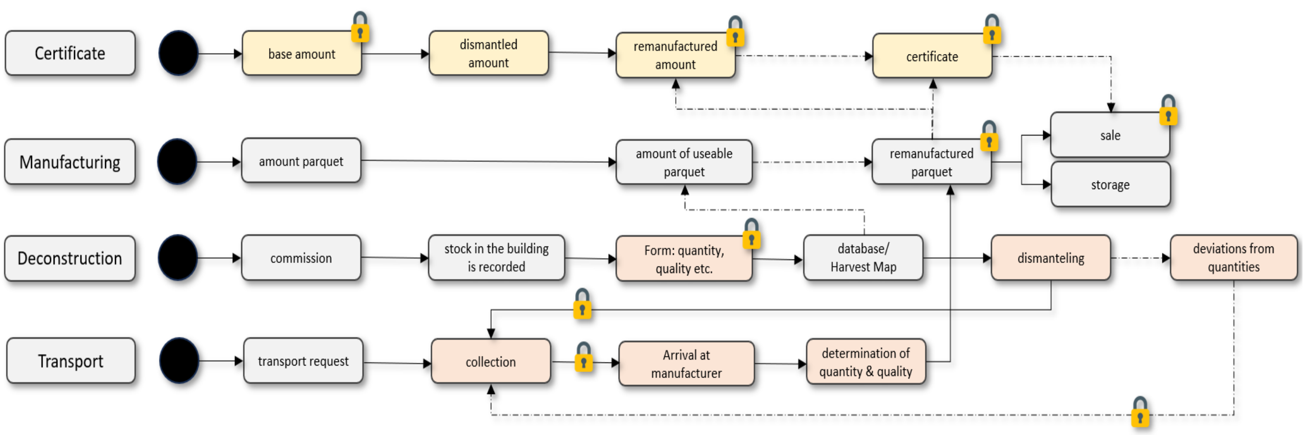


Figure 2: re:parkett process analysis

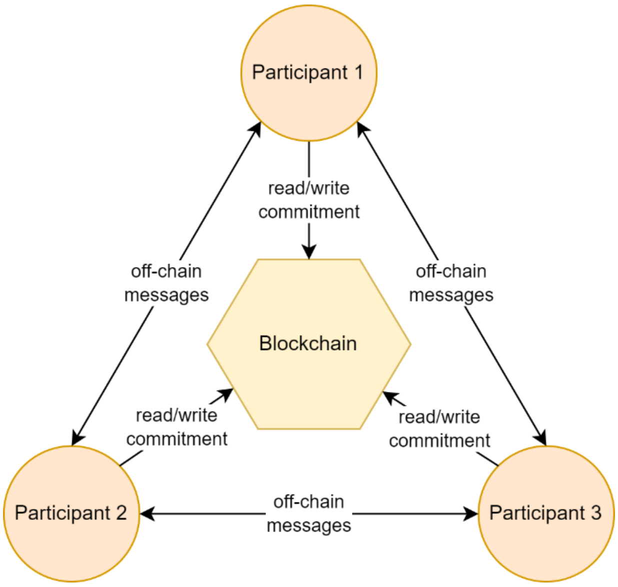


Figure 3: Overview of interactions on the blockchain

Blockchain application

Based on findings from the projects BIM.design & FMChain and the ongoing research presented here, we developed a system architecture for the use of blockchain to address challenges in the EoL process. The proposed framework allows for sustainable documentation of essential steps and data in the EoL value chain and ensures verifiable compliance to process models.

The first use of blockchain technology in the field of business process management mapped processes models entirely on blockchains by translating models to Smart Contracts (Weber et. al. 2016, García-Bañuelos et. al. 2017, López-Pintado et. al. 2019). This initial approach of mapping all data (including actors, gateways, etc.) on the blockchain is problematic for several reasons: Generally, when implementing processes, which involve multiple stakeholders, two fundamental obstacles occur. Firstly, basic privacy requirements of companies are violated: companies generally want to avoid having their own data associated with their business processes (e.g. supply chain data, production process data, etc.) to be openly visible to third parties. This was also reaffirmed by our interview partners in our requirements analysis for our use case. Furthermore, storing data on a blockchain is also unattractive due to high transaction costs. This is especially true for so called public blockchains, like Bitcoin or Ethereum. These are, as the name suggests, visible and open to be used by everyone. However, this is also a disadvantage, since the limited possible number of transactions per second can result in high transaction costs for participants. Private blockchains, i.e., blockchains that cannot be viewed by everyone and are operated by a small number of participants, have this limitation to a much lesser extent, but have other disadvantages, such as missing network effects; and also lack the capability for public verification (Carminati et. al. 2018, Marangone et. al. 2022).

In summary, as little data as possible should be stored on the blockchain but enough to fully exploit the capabilities of this technology. The data on the blockchain should also meet the privacy requirements.

With that in mind, Figure 3 shows our developed framework and how participants interact with the blockchain and each other. The main communication between the individual participants takes place off-chain, i.e. using peer-to-peer communication channels. The interaction with the blockchain is limited to commitments that represent irrevocable statements. These need to be transparent and verifiable for everyone and are therefore stored on the blockchain. These statements include only a fraction of the data that is held by the individual participants. A zero-knowledge proof (ZKP) ensures the relation between the commitments and the private data as well as the correct execution of state updates. The private data hence becomes immutable and tamper-proof.

To harness these privacy and security features in our framework we model use cases as a choreography of message exchanges. Figure 4 displays the message exchanges of an End-of-Life value chain, based on our use case re:parkett in a simplified form as a BPMN choreography diagram. Each box represents a message exchange task and displays the two stakeholders in terms of sender and recipient, as well as the conducted action. This modelling method results in a clear definition of the order in which stakeholders are involved and the information being shared between them. The interaction with the blockchain is limited to essential commitments

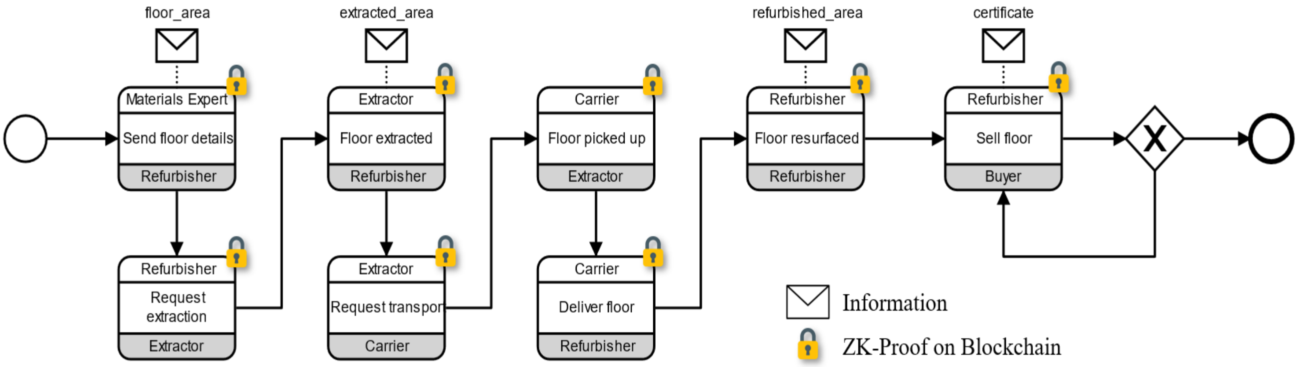


Figure 4: BPNM choreography diagram for re:parkett

and associated ZKPs (illustrated as padlocks in Fig. 2 and Fig. 4), verified with a smart contract on the blockchain. This facilitates immutable and public-verifiable assignment of data, without revealing any additional or non-disclosable private information.

A Zero-Knowledge Proof (ZKP) is set up as an interaction between a prover and a verifier. The prover's goal is to convince the verifier of the correctness of a computation without revealing the complete set of inputs (Goldreich and Oren 1994; Tomaz et al. 2020). ZKPs therefore provide the fundamentals to secure authentication and privacy protection in blockchain transactions, by creating a platform for stakeholders to interact without disclosing sensitive data (Meralli, 2020; Čapko et al. 2022).

While ZKPs were envisioned as multi-round interactive protocols, current variants are non-interactive i.e., only a single message is exchanged between the prover and the verifier. This property allows ZKPs to be included in blockchain transactions. Furthermore, zk-SNARKs, a highly efficient non-interactive variant with small proof sizes, enable storage and verification on the blockchain (Partala et. al. 2020).

Our framework utilizes this technique to ensure the correctness of private processes on a public blockchain. More specifically, participants signal the completion of a choreography tasks by sending off-chain messages to each other. Furthermore, they commit to those messages by sending a transaction consisting of a reference to the message and the process state together with a ZKP to the blockchain. A Smart Contract verifies the correctness of the ZKP and only updates the process state in case of success. The ZKP ensures that only transitions defined in the model can be executed, that only the permissioned participants can trigger a transition, and that the passed information satisfies some predefined rules. For example, in the case of the re:parkett use case, only the *Extractor* may complete the *Floor extracted* task by sending a message to the *Refurbisher*. Furthermore, although not shown in the figure, the value of *extracted_area* needs to be smaller or equal to the *floor_area* value that was provided by the *Materials Expert* in the first task.

Applied to the whole re:parkett use case the framework allows to create a certificate of provenance. A reference to the state in the SC proofs the validity of the certificate, as only conforming process updates can be submitted to the blockchain and hence the claims of the certificates must be correct. In case of re:parkett the certificate ensures that the amount of floor certified as recycled cannot exceed the amount identified by the expert.

The example demonstrates that our framework is highly suitable for many supply-chain use cases in AEC. By adopting our approach of tracing goods, it becomes feasible to comply with privacy requirements and legislation, while increasing the level of trust and security for digital processes in CE.

Conclusion

In this paper, we have explored the integration of blockchain technology into EoL processes within a CE framework, using a specific case study of a closed-loop reuse-product in Austria. The application of a blockchain-based solution has demonstrated substantial potential to enhance challenges regarding transparency, authenticity and trust through verifiable, privacy-preserving mechanisms, specifically zero-knowledge proofs.

Our research confirms that digital technologies, like blockchain, are vital for advancing CE by ensuring transparent and reliable data throughout product and material lifecycles. Implementing our blockchain framework in parquet flooring reuse demonstrates effective digital strategies for real-world CE challenges. This enhances environmental sustainability and creates economic value by potentially generating CO2 certificates (e.g. possibly published on a token, referring to CO2 emissions of a product or part of a product), herewith aligning with EU Taxonomy compliance. The findings support stronger integration of digital tools in CE processes, especially in the building sector, for resource efficiency improvements.

Ultimately, this paper calls for ongoing research and collaboration across sectors to refine and expand the use of digital solutions in promoting sustainable practices within the circular economy. This will not only enhance environmental outcomes but also foster economic resilience and social well-being.

Acknowledgments

DiCYCLE – Reconsidering digital deconstruction, reuse and recycle processes using BIM and Blockchain (Grant no. 886960), is funded by the Austrian Research Promotion Agency (FFG), Program “Stadt der Zukunft”. BIMd.sign - BIM digitally signed with blockchain and smart contracts, is realized in cooperation with the project FMChain - Automated payment and contract management in construction with Blockchain technology and BIM 5D are funded by FFG - Program ICT of the Future - and BMVIT, Grants No. 873842 and 873827.

References

Alhawari, O., Awan, U., Bhutta, M. K. S. & Ülkü, M. A. (2021). Insights from Circular Economy Literature: A Re-view of Extant Definitions and Unravelling Paths to Fu-ture Research. Sustainability, 13(2): 859.

Akbarieh, A., Jayasinghe, L. B., Waldmann, D. & Teferle, F. N. (2020). BIM-Based End-of-Lifecycle Decision Making and Digital Deconstruction: Literature Review. Sustainability, 12(7): 2670.

AWG (2002): Abfallwirtschaftsgesetz 2002. https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=20002086

Braungart, M., McDonough, W. & Bollinger, A. (2007). Cradle-to-cradle design: creating healthy emissions – a strategy for eco-effective product and system design, Journal of Cleaner Production, Volume 15, Issues 13–14, Pages 1337-1348, ISSN 0959-6526, 10.1016/j.jclepro.2006.08.003.

Braungart, M., P. Bondesen, A. Kälin & B. Gabler, (2008). Specific Public Goods for Economic Development: With a Focus on Environment, British Standards Institution (eds), Public Goods for Economic Development. Compendium of Background papers, United Nations Industrial Development Organisation, Vienna).

Brändström, J. & Eriksson, O. (2022). How circular is a value chain? Proposing a Material Efficiency Metric to evaluate business models. Journal of Cleaner Production, 342. 130973.

Bocken, N. M. P., de Pauw, I., Bakker, C. & van der Grinten, B. (2016). Product design and business model strategies for a circular economy. Journal of Industrial and Production Engineering, 33(5): 308-320. 10.1080/21681015.2016.1172124

Buterin, V. (2013). Ethereum White Paper

Čapko, D., Vukmirović, S. & Nedić, N. (2022). State of the Art of Zero-Knowledge Proofs in Blockchain. 2022 30th Telecommunications Forum (TELFOR). pp. 1-4

Carminati, B., Rondanini, C. & Ferrari, E. (2018). Confidential Business Process Execution on Blockchain. In: 2018 IEEE International Conference on Web Services (ICWS). pp. 58–65. 10.1109/ICWS.2018.00015

Çetin S., Gruis V. & Straub Ad. (2022). Digitalization for a circular economy in the building industry: Multiple-case study of Dutch social housing organizations, Resources, Conservation & Recycling Advances, Volume 15, 2022, 200110, ISSN 2667-3789, 10.1016/j.rcradv.2022.200110.

Charef, R. (2022). Supporting construction stakeholders with the circular economy: A trans-scaler framework to under-stand the holistic approach. Cleaner Engineering and Technology, 8. 100454

Copeland, S. & Bilec, M. (2020). Buildings as material banks using RFID and building information modeling in a circular economy. Procedia CIRP, 90, 143-147. 10.1016/j.procir.2020.02.122

Concular. https://concular.de [accsess date 16.01.2024]

DGNB. https://www.dgnb.de/en [access 13.4.2024]

Di Giuda, G. M., Pattini, G., Seghezzi, E., Schievano, M., & Paleari, F. (2020). The Construction Contract Execution Through the Integration of Blockchain Technology. In: Digital Transformation of the Design, Construction and Management Processes of the Built Environment (pp. 27-36). Cham: Springer International Publishing.

Dodampegama, S., Hou, L., Asadi, E., Zhang, G. & Setunge, S. (2024). Revolutionizing construction and demolition waste sorting: Insights from artificial intelligence and robotic applications. *Resources, Conservation and Recycling,* 202: 107375. https://doi.org/10.1016/j.resconrec.2023.107375

EN 15978:2011 Sustainability of construction works – Assessment of environment performance of buildings Calculation method.

EMF (2013).Towards the Circular Economy 1: Economic and business rationale for an accelerated transition. Retrieved, Ellen MacArthur Foundation

EU Taxonomy (2024) https://ec.europa.eu/sustainable-finance-taxonomy/home [accessed 14.4.2024]

Figge, F., Thorpe, A. & Gutberlet, M. Definitions of the Circular Economy - Circularity Matters (2023). Ecological Economics, Vol. 208, 2023, Available at SSRN: https://ssrn.com/abstract=4398717

García-Bañuelos, L., Ponomarev, A., Dumas, M. & Weber, I. (2017). Optimized Execution of Business Processes on Blockchain. In: Business Process Management. pp. 130–146. Lecture Notes in Computer Science, Springer International Publishing. 10.1007/978-3-319-65000-5_8

Gervasio Dos Santos, H. & Dimova, S. (2018). Environmental benchmarks for buildings. *EFIResources: Resource Efficient Construction towards Sustainable Design.* Ispra: European Commission. doi:10.2760/073513

Goldreich, O., & Oren, Y. (1994). Definitions and properties of zero-knowledge proof systems. Journal of Cryptology, 7(1), 1-32.

Heisl, F. & Rau-Oberhuber, S. (2020). Calculation and evaluation of circularity indicators for the built environment using the case studies of UMAR and Madaster. Journal of Cleaner Production, 243. 118482.

Kamble, S. S., Gunasekaran, A. & Gawankar, S. A. (2018). Sustainable Industry 4.0 framework: A systematic literature review identifying the current trends and future perspectives. *Process Safety and Environmental Protection,* 117: 408-425. 10.1016/j.psep.2018.05.009

Kirchherr, J., Reike, D. & Hekkert, M. (2017). Conceptualizing the circular economy: An analysis of 114 definitions. Re-sources, Conservation and Recycling, 127. 221-232.

Kirchherr, J., Yang, N.-H. N., Schulze-Spüntrup, F., Heerink, M. J. & Hartley, K. (2023). Conceptualizing the Circular Economy (Revisited): An Analysis of 221 Definitions. *Resources, Conservation and Recycling,*194:107001.10.1016/j.resconrec.2023.107001

Khosrowshahi F. (2017). Building Infromation Modelling (BIM) a Paradigm Shift in Construction, Building information modelling building performance, design and smart construction. Springer, 47 ff

Korhonen, J., Honkasalo, A. & Seppälä, J. (2018). Circular Economy: The Concept and its Limitations, Ecological Economics, Volume 143, Pages 37-46, ISSN 0921-8009.10.1016/j.ecolecon.2017.06.041.

Kovacic, I., Honic, M. & Sreckovic, M. (2020). Digital Platform for Circular Economy in AEC Industry, Engineering Project Organization Journal (October 2020) Volume 9

López-Pintado, O., García-Bañuelos, L., Dumas, M., Weber, I., & Ponomarev, A. (2019) Caterpillar: A business process execution engine on the Ethereum blockchain 49(7), 1162–1193. 10.1002/spe.2702

Marangone, E., Di Ciccio, C. & Weber, I. (2022). Fine-Grained Data Access Control for Collaborative Process Execution on Blockchain. In: Business Process Management: Blockchain, Robotic Process Automation, and Central and Eastern Europe Forum. BPM 2022. Lecture Notes in Business Information Processing, vol 459. 10.1007/978-3-031-16168-1_4

McDonough, W. & Braungart M. (2002). Cradle to Cradle: Remaking the Way We Make Things, North Point Press, New York, NY

Meralli, S. (2020). Privacy-preserving analytics for the securitization market: a zero-knowledge distributed ledger technology application, Financial Innovation 6 (7). 10.1186/s40854-020-0172-y

Nakamoto, S. (2008). Bitcoin: a peer-to-peer electronic cash system

ÖNORM B 3151:2022. Dismantling of buildings as a standard method for demolition. Wien: Austrian Standards International

Pagoropoulos, A., Pigosso, D. C. A. & McAloone, T. C. (2017). The Emergent Role of Digital Technologies in the Circular Economy: A Review. *Procedia CIRP,* 64: 19-24. 10.1016/j.procir.2017.02.047

Papamichael, I., Voukkali, I., Loizia, P. & Zorpas, A. A. (2023). Construction and demolition waste framework of circular economy: A mini review. *Waste Management & Research,* 41(12): 1728-1740. 10.1177/0734242x231190804

Partala, J., Nguyen, T. H. & Pirttikangas, S. (2020). Non-Interactive Zero-Knowledge for Blockchain: A Survey, IEEE Access, vol. 8, pp. 227945–227961. 10.1109/ACCESS.2020.3046025

Rangelov, M., Dylla, H., Mukherjee, A. & Sivaneswaran, N. (2021). Use of environmental product declarations (EPDs) of pavement materials in the United States of America (U.S.A.) to ensure environmental impact reductions. Journal of Cleaner Production, 283, 124619. 10.1016/j.jclepro.2020.124619

RRReMaker https://www.rrremaker.com/ [Accessed 14.01.2024]

Scott, D. J., Broyd, T. & Ma, L. (2021). Exploratory literature review of blockchain in the construction industry. Automation in Construction, 132, 103914. 10.1016/j.autcon.2021.103914

Shojaei, A. (2019). Exploring applications of blockchain technology in the construction industry, in: 10th International Structural Engineering and Construction Conference (ISEC), ISEC Press, University of Illinois at Chicago. 10.14455/isec.res.2019.78.

Shojaei, A., Ketabi, R., Razkenari, M., Hakim, H. & Wang, J. (2021). Enabling a circular economy in the built environment sector through blockchain technology. Journal of Cleaner Production, 294, 126352. 10.1016/j.jclepro.2021.126352

Stahel, W. R. (1982). The product life factor. An Inquiry into the Nature of Sustainable Societies: The Role of the Private Sector (Series: 1982 Mitchell Prize Papers), NARC: 74-96.

Stahel, W. R. (1994). The utilization focused service economy: Resource efficiency," in B. R. Allenby and D. J. Richards (eds), The Greening of Industrial Ecosystems, National Academy Press, Washington, DC, 178–190

Stahel, W. R. (2010). The Performance Economy, Palgrave Macmillan Hampshire, Hampshire UK

Szabo, N. (1997). Formalizing and securing relationships on public networks, First Monday

Tomaz, A. E. B., Nascimento, J. C. D., Hafid, A. S. & Souza, J. N. D. (2020). Preserving Privacy in Mobile Health Systems Using Non-Interactive Zero-Knowledge Proof and Blockchain. IEEE Access, 8: 204441-204458. 10.1109/ACCESS.2020.3036811

Veit, D., Clemons, E., Benlian, A., Buxmann, P., Hess, T., Kundisch, D., Leimeister, J. M., Loos, P. & Spann, M. (2014). Geschäftsmodelle - Eine Forschungsagenda für die Wirtschaftsinformatik. In: WIRTSCHAFTSINFORMATIK - Research Notes, Number: 1, p.55-64

Weber, I., Xu, X., Riveret, R., Governatori, G., Ponomarev, A. & Mendling, J. (2016). Untrusted Business Process Monitoring and Execution Using Blockchain. In: Business Process Management, vol. 9850, pp. 329–347. Springer International Publishing. 10.1007/978-3-319-45348-4_19

Wüst, K. & Gervais, A. (2018). Do you need a blockchain?, 2018 Crypto Valley Conference on Blockchain Technology (CVCBT).

Zhang, J., Bhuiyan, M., Zhang, G. & Sandanayake, M. (2024). Life cycle assessment of kerbside waste material for an open-looped and closed-loop production– towards circular economy designs, Journal of Cleaner Production, Volume 434, 139991, ISSN 0959-6526. 10.1016/j.jclepro.2023.139991.

Data Analysis, Simulation and Resilience

A COMPARATIVE ANALYSIS OF MULTI-TARGET FEATURE SELECTION METHODS IN DATA-DRIVEN MODELS FOR BUILDING ENERGY AND THERMAL PERFORMANCE PREDICTION

Esther Borkowski, Xiaojun Fan and Arno Schlueter
Architecture and Building Systems, Institute of Technology in Architecture, ETH Zurich, Zurich, Switzerland

Abstract

Building energy management increasingly utilises Machine Learning (ML) to use data from sensor-rich environments. A significant challenge in this context is managing high-dimensional data, which can affect model performance. This study addresses this by applying multi-target feature selection, an underexplored method that reduces dimensionality by analysing inter-feature relationships. From 182 features, two were key for developing three ML models predicting the energy and thermal performance of the HiLo living lab. These models achieved a robust fit with an average Root Mean Squared Error (RMSE) of 0.18 kW and 1.03 °C, demonstrating multi-target feature selection's effectiveness in enhancing building performance predictions.

Introduction

Buildings, accounting for 36% of global energy consumption, play a key role in the transition towards a more sustainable energy grid (United Nations Environment Programme, 2021). With the projected increase in the electrification of many end-use sectors and the broader integration of renewable energy sources, the focus shifts from supply-side to demand-side flexibility in grid operations. This shift is crucial for buildings to contribute significantly to decarbonisation targets by managing their energy consumption and generation dynamically, a concept known as energy flexibility (Li et al., 2022).

In this evolving landscape, buildings are expected to provide nearly half of the flexible demand capacity by 2030. This transformation is accelerated by the 'Internet of Things', which has increased the use of sensors in Heating, Ventilation and Air-Conditioning (HVAC) systems in buildings, resulting in the generation of vast amounts of data (Kathirgamanathan et al., 2019). However, these data are often high-dimensional, which creates challenges for Machine Learning (ML) algorithms due to the presence of irrelevant, redundant or noisy features that can slow down the learning process and degrade performance. Feature selection is recognised as a crucial method for reducing dimensionality. It involves choosing a subset of relevant features based on certain criteria, thereby simplifying model interpretation and decreasing memory usage, computational costs and learning time (Hashemi et al., 2021). Furthermore, should these models be applied in real buildings, the careful selection of sensors through feature selection would necessitate only essential sensors. Such efficient sensor usage would enhance economic viability and conserve critical raw materials, thereby supporting resource efficiency and circularity in buildings by reducing sensor network requirements.

In feature selection, a distinction is made between single-target and multi-target data. While single-target data involves predicting one output per sample, multi-target data consists of samples with multiple outputs, where the correlation between features and these outputs determines feature relevancy. This distinction directly influences the choice and application of feature selection methods. According to Hashemi et al. (2021) and as shown in Figure 1, feature selection methods can be broadly categorised into search methods and supervision methods, with search methods being often designed with single-target problems in mind.

On the other hand, supervision methods can address both single-target and multi-target feature selection. Within supervised ML algorithms, multi-target feature selection methods can be further divided into binary transformation and algorithm adaptation. Binary transformation initially converts multi-target data into several independent single-target data and then evaluates features for each target separately. However, this method may overlook the correlations between different targets (Masmoudi et al., 2020). In contrast, algorithm adaptation is specifically designed for multi-target data and aims to preserve and utilise these inter-target relationships.

Recent years have seen the development of various single-target feature selection methods, with numerous studies evaluating their application in ML models related to building controls (e.g. Olu-Ajayi et al., 2023; Kathirgamanathan et al., 2019; Zhang and Wen, 2019). Yet, real-world complexities often require addressing multi-target problems, especially in buildings where simultaneous prediction of multiple targets is common. Despite their significance, multi-target feature selection methods have received limited attention in research, both generally and within the built environment. While some studies have explored binary transformation for multi-target issues, to the best of the authors' knowledge, research on algorithm adaptation in this area is notably lacking. This study set out to fill this gap by comparing multi-target feature selection methods, with a particular focus on algorithm adaptation in data-

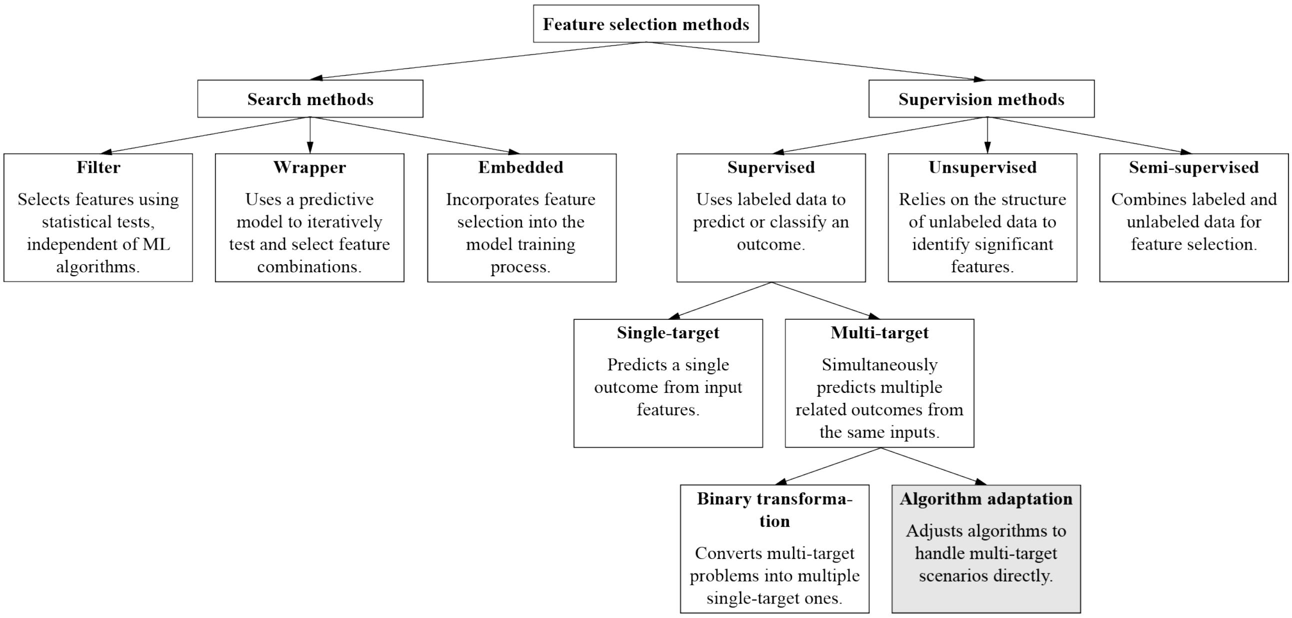


Figure 1: Categorisation of feature selection methods, highlighting the method used in this study. Based on categorisation by Hashemi et al. (2021).

driven models for accurately predicting building energy and thermal performance.

To achieve this aim, investigations have been conducted on data from the HiLo living lab in Dübendorf, Switzerland, as can be seen from Figure 2 (Block et al., 2017). Designed with the guiding principle of 'high performance, low emissions', HiLo features innovative building systems, such as Thermally Activated Building Systems (TABS) embedded into lightweight concrete structures. These TABS use a hydronic pipe network integrated into HiLo's structural components to transform the internal ceilings of two of its offices into radiant surfaces for heating and cooling (Lydon et al., 2019). Utilising concrete's thermal properties for heat storage and thermal inertia, TABS facilitate a time shift of peak heat load and a time lag of sensible heat transfer from the concrete to the interior air. These characteristics make TABS particularly effective for enhancing building energy flexibility (Arteconi et al., 2014).

Figure 2: Exterior view of HiLo from the southwest. Photo: Roman Keller.

However, controlling TABS in heavyweight concrete structures, which typically exhibit extended time constants, presents significant challenges as the regulation of heat output's timing and amount often becomes complex (van der Heijde et al., 2017). In response, this research project investigates the use of TABS in lightweight concrete structures, as exemplified in two of HiLo's offices. To model one of these similarly designed offices, i.e. Office 1, a data-driven (black-box) approach was selected due to its computational efficiency and accuracy against a more resource-intensive physics-based (white-box) model.

This study's central focus is applying multi-target feature selection during data pre-processing. This step was instrumental in accurately predicting two key aspects: (1) the thermal power of the TABS to ensure optimal control over the timing and amount of heat output and (2) the indoor air temperature in the office to maintain it within comfortable ranges. By comparing various multi-target feature selection methods, the study aims to identify the most influential sensors within HiLo's extensive sensor network. The aim is to develop a model that predicts TABS thermal power and indoor air temperature and aligns closely with the measured data. Achieving this aim is expected to provide valuable insights into improving building energy efficiency and occupant comfort and enhancing the methodological tools available for built environment research.

Methods

This section details the study conducted at the HiLo living lab, which was used to assess the effectiveness of multi-target feature selection methods in predicting the thermal power of the TABS and the indoor air temperature in

HiLo's Office 1.

Dataset

The study utilised a comprehensive dataset from HiLo, which is equipped with over 1,500 data points. For this study, a subset of 184 sensors was selected based on their relevance to predicting the cooling performance of the TABS in HiLo's Office 1. These sensors recorded data at ten-minute intervals for two summers (i.e. from 1 June to 31 August 2022 and from 1 June to 21 August 2023), covering a range of parameters from environmental conditions to HVAC system performance metrics. This dataset was then used to study the effect of 182 input variables on two key output variables: the TABS's thermal power and the office's indoor air temperature. Data extraction and analysis were conducted using Python, with libraries such as scikit-learn, pandas and matplotlib.

Data pre-processing

Data pre-processing is crucial for ML algorithm performance, as it addresses raw data imperfections and irregularities, such as high dimensionality and missing data (Olu-Ajayi et al., 2023). Therefore, data were pre-processed in this study to ensure dataset quality and avoid difficulties during model development. The process involved seven main steps:

1. **Data splitting:** Initially, the dataset was divided into features and targets, followed by splitting these into training and testing subsets for model evaluation.
2. **Biased reduction:** To prevent model overfitting and ensure accuracy, features strongly correlated with targets were removed, reducing the feature count to 162.
3. **Constant feature removal:** Non-variable features were eliminated for lacking predictive value, narrowing down the dataset to 106 features.
4. **Redundancy elimination:** Highly correlated features with a Pearson correlation above 0.9 were identified, and duplicates were removed to prevent multicollinearity, leaving 73 features.
5. **Missing data imputation:** Missing values, often a result of communication errors, were imputed using the K-Nearest Neighbours (KNN) algorithm to maintain dataset integrity.
6. **Cyclical feature encoding:** Cyclical aspects like time and solar position were encoded using sine and cosine transformations, leading to 97 features in the dataset.
7. **Supervised learning transformation:** To effectively model the TABS response, 12 lagged observations spanning a two-hour window were included to capture immediate past effects, expanding the feature set to 1,261.

Feature selection methods

The feature selection process leveraged the algorithm adaptation method, employing three ML models to determine the most predictive features:

- **Random Forest (RF):** This model combines multiple decision trees to form a more robust and accurate prediction. It is particularly adept at handling large datasets with complex relationships between features.
- **Gradient Boosting Machine (GBM):** Employing a stage-wise additive model, this technique progressively builds a model to optimise a loss function, which makes it effective for both bias reduction and variance control.
- **Support Vector Regression (SVR):** SVR applies the principles of support vector machines for regression purposes, hence providing a flexible approach to capture both linear and non-linear relationships within the data.

These models were selected due to their ability to handle diverse data characteristics and provide comprehensive insights into feature relevance.

Given that each of these models has its inherent method of assigning feature weights, which can differ significantly in scale and interpretation, feature weights were normalised to ensure comparability. This normalisation process adjusted the feature weights from each model to a common scale, thereby allowing for a direct and meaningful comparison of feature importance across different methods.

The effectiveness of each model in feature selection was then assessed based on these normalised weights, along with the impact on model performance and the computation time. The comparison aimed at identifying the most effective feature selection method for the HiLo dataset.

Model development

This study used the scikit-learn library to implement the three ML models listed above: RF, GBM and SVR. These models were chosen because of their performance in various prediction tasks, especially in scenarios involving complex and high-dimensional data like the real-world data in this study. The development process focused on optimising the predictive performance through strategic feature selection and iterative model tuning. Standard hyperparameters from scikit-learn were used due to the study's exploratory nature.

The process to determine the optimal feature set for each model involved systematic experimentation. Starting with the most predictive feature for the thermal power of the TABS and the most predictive feature for the indoor air temperature of Office 1, features were added incrementally – first in pairs, then in larger groups of ten and eventually by hundreds. This approach allowed testing all features up to the entire dataset of 1,261 features. Duplicate features were carefully removed to ensure a balanced emphasis on feature importance.

Model evaluation

The model's performance was assessed primarily using the Root Mean Squared Error (RMSE) metric, where a lower RMSE value indicates better model performance with fewer errors in prediction (Naser and Alavi, 2023).

For predicting the thermal power of the TABS (0-1 kW range) and the indoor air temperature of Office 1 (22-28 °C range), RMSE values up to 0.3 kW and 1 °C, respectively, were considered indicative of satisfactory model accuracy. This evaluation ensured that the model was not only accurate but also reliable and applicable in practical settings.

Results analysis and discussion

The study conducted with data from the HiLo living lab revealed that only a selected group of features significantly influenced the prediction of the thermal power of the TABS and the indoor air temperature in Office 1. The feature selection methods used – RF, GBM and SVR – differed in identifying these influential features. Notably, RF and GBM models assigned high importance to a few features, whereas the SVR model distributed importance more evenly across a broader range of features, as shown in Figures 3 and 4. Descriptions of the symbols for the top predictive features are provided in Table 1.

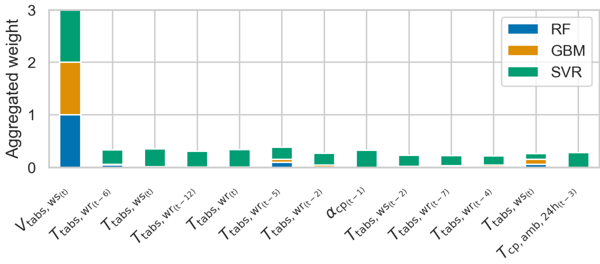


Figure 3: Aggregated weights of top six features for predicting TABS thermal power across RF, GBM and SVR models.

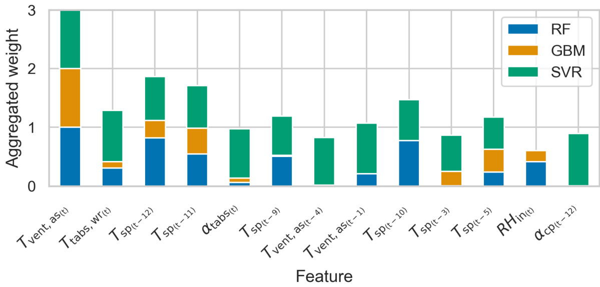


Figure 4: Aggregated weights of top six features for predicting indoor air temperature across RF, GBM and SVR models.

Despite methodological differences, all models consistently recognised two features as most important for their predictions: the supply water flow rate of the TABS at the current time step, $V_{tabs,s_{(t)}}$, for predicting the TABS's thermal power, and the air supply temperature of the ventilation system at the current time step, $T_{vent,s_{(t)}}$, for predicting the indoor air temperature. Beyond these top two features, the divergence in feature selection became more apparent. For instance, when predicting the indoor air temperature, RF and GBM models identified the zone regulation's setpoint temperature at various preceding time steps as the next two most important features. RF focused on the measurements twelve and ten steps prior, $T_{sp_{(t-12)}}$ and $T_{sp_{(t-10)}}$, whereas GBM focused on the measurements eleven and five steps prior, $T_{sp_{(t-11)}}$ and $T_{sp_{(t-5)}}$. In contrast, the SVR model yielded a different pair of features as most influential: the circulation pump's valve opening percentage twelve steps before the current time, $\alpha_{cp_{(t-12)}}$, and the water return temperature of the TABS at the current time, $T_{tabs,r_{(t)}}$.

Table 1: Top feature symbols and descriptions without time-lag notation.

Symbol	Description
$V_{tabs,s}$	Supply water flow rate of TABS
$T_{tabs,r}$	Water return temperature of TABS
$T_{tabs,s}$	Water supply temperature of TABS
α_{cp}	Valve opening of circulation pump
$T_{amb,24h}$	24h average of ambient temperature
$T_{vent,s}$	Air supply temperature of ventilation
T_{sp}	Setpoint temperature of zone regulation
RH_{in}	Indoor relative humidity

This variance in feature selection could be attributed to the inherent characteristics of each model. RF and GBM, both being ensemble tree-based models, tended to focus on similar aspects of feature interaction and non-linear relationships. On the other hand, SVR's unique approach, particularly its sensitivity to data characteristics, led to a distinct interpretation of feature importance.

These differences also slightly influenced model performance. As Figures 5 and 6 illustrate, RMSE values remained relatively stable across varying feature counts and models, which may suggest that the models' reliability was not heavily dependent on the feature count or selection methodology. Overall, all models demonstrated a strong fit to the data, with average RMSE values of 0.18 kW for thermal power and 1.03 °C for indoor air temperature across various feature counts and selection methods.

A closer examination of the RMSE values for TABS's thermal power prediction revealed that they remained relatively constant, regardless of the number of features added, as detailed in Table 2. Including just the top two features yielded RMSE values of 0.15 kW for RF, 0.13 kW for GBM and 0.14 kW for SVR, closely aligning with the overall average. This suggests that more features did not necessarily improve RMSE. Interestingly, expanding the feature set to four or six initially degraded prediction ac-

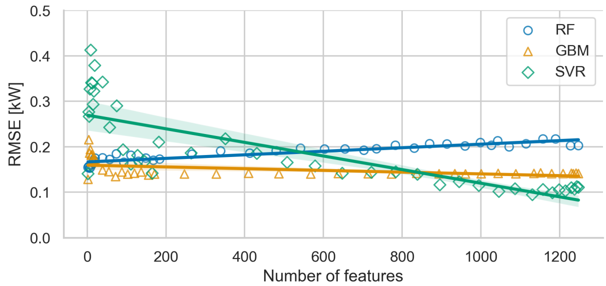


Figure 5: RMSE comparison for TABS's thermal power prediction across RF, GBM and SVR models and varying feature counts.

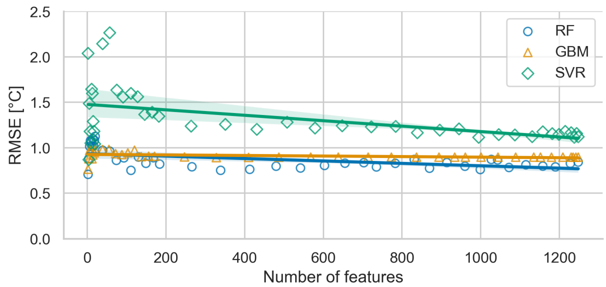


Figure 6: RMSE comparison for indoor air temperature prediction across RF, GBM and SVR models and varying feature counts.

curacy, followed by a slight improvement. However, even with the full feature set, RMSE values were 0.20 kW for RF, 0.14 kW for GBM and 0.11 kW for SVR, which indicated minimal gains from adding more features. This pattern suggests that optimising predictions for TABS's thermal power and indoor air temperature in HiLo's Office 1 might not require more than two important features, which demonstrates effectiveness in feature utilisation from an RMSE standpoint.

Table 2: RMSE comparison for TABS's thermal power prediction [kW] across RF, GBM and SVR models with top two, four and six features and overall feature average μ.

Model	Feature count			
	2	4	6	μ
RF	0.15	0.16	0.15	0.19
GBM	0.13	0.22	0.19	0.15
SVR	0.14	0.28	0.27	0.19

RF and GBM models showed a similar trend for indoor air temperature predictions to those observed for the TABS's thermal power predictions. Starting with the top two features provided better-than-average results, with an RMSE of 0.87 °C for RF and 0.91 °C for GBM. This performance diminished with the inclusion of four or six top features, then slightly improved again as more features were considered. This pattern, however, diverged for the SVR model. With only the top two features, the SVR's RMSE was 2.04 °C, which improved to 0.87 °C with the four top features but worsened again with six. As detailed in Table 3, the fluctuations observed with the SVR model underscore the distinct feature selection strategies inherent to each model. With its unique handling of features, the SVR model underscores the critical influence of feature selection methodology on model performance. This variability further accentuates the need for model-specific feature selection to enhance predictive accuracy and reliability.

Table 3: RMSE comparison for indoor air temperature prediction [°C] across RF, GBM and SVR models with top two, four and six features and overall feature average μ.

Model	Feature count			
	2	4	6	μ
RF	0.71	0.88	1.01	0.87
GBM	0.76	0.88	0.97	0.91
SVR	2.04	0.87	1.49	1.32

The runtime analysis complements these findings. As depicted in Figure 7, the runtime increased significantly with the addition of more features, underscoring the need for a careful balance between computational efficiency and model accuracy. The comparative analysis of RMSE and runtime suggests that a model with two features provided an optimal balance between accuracy and efficiency. Using two features, the runtime was 0.14 seconds for RF, 0.56 seconds for GBM and 0.63 seconds for SVR. This indicates that SVR was 77.77% slower than RF and 11.23% slower than GBM, a significant efficiency loss that could be disadvantageous in more complex modelling scenarios. This balance was consistent across different feature selection methods, offering a practical and robust solution for predicting TABS thermal power and air temperature in the given context.

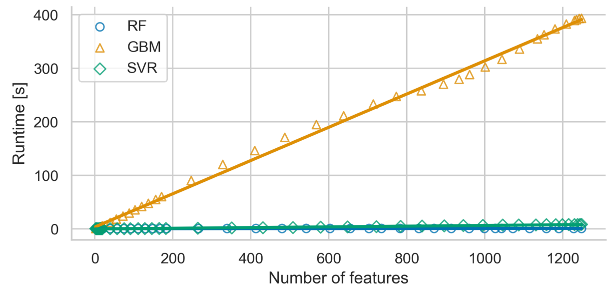


Figure 7: Runtime comparison across RF, GBM and SVR models and varying feature counts.

The overall performance metrics of the models, characterised by average RMSE values of 0.18 kW for thermal power and 1.03 °C for indoor air temperature across various feature counts and selection methods, indicate a robust fit to the dataset. This alignment is further evidenced by Figures 8 and 9, which compare the model's predicted outcomes against the measured data for TABS thermal power and indoor air temperature in HiLo's Office 1. These comparisons span a one-week period in mid-August 2023 within the test set, with each model iteration using the top two features identified by the respective feature selection method.

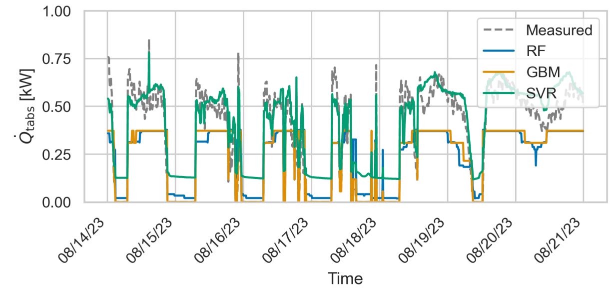


Figure 8: Comparison of predicted and measured TABS thermal power across RF, GBM and SVR models for one week in the test set.

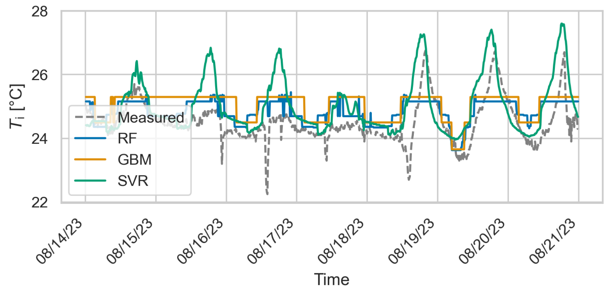


Figure 9: Comparison of predicted and measured indoor air temperature in Office 1 across RF, GBM and SVR models for one week in the test set.

The figures highlight the differences in the algorithms of RF, GBM and SVR models, despite utilising the same features. RF and GBM models yielded more constant outcomes, while the SVR model, characterised by greater variability, performed better at capturing peak values. This was particularly noticeable on the weekend, as indicated in the last three days of the graphs. A plausible explanation for this divergence could be SVR's sensitivity to extreme values (peaks) that arose from its objective to minimise the error within a certain threshold. This inherent sensitivity allowed SVR to better capture and predict extreme values or peaks in the data, leading to its observed variability but effective handling of peaks.

The analysis underlines the importance of model selection based on specific use cases, which factor in considerations such as prediction accuracy, computational efficiency and the ability to manage peak values. While RF and GBM showed higher accuracy and faster execution with the top two features, they did not match SVR's capability in considering peak values. This capability to accurately predict peak values might be crucial for certain applications and emphasises the need to choose models that align with particular performance criteria and application-specific requirements.

Conclusions

This study set out to evaluate how effectively multi-target feature selection methods, through algorithm adaptation, identify influential sensors to enhance building performance predictions. Employing a data-driven approach, the research focused on modelling and predicting the thermal power of the TABS and the indoor air temperature, using a subset of features from the extensive sensor network in the HiLo living lab.

The key findings highlighted that while RF and GBM models prioritised a similar set of features, the SVR model demonstrated a preference for a different subset. Despite these variations, the comparative analysis of the RMSE metric and the computational runtime suggested that models incorporating just two features achieved an optimal balance between accuracy and effectiveness. This balance is particularly crucial in real-time building control applications where both precision and computational speed are essential.

However, the study faced several limitations. First, while the focus was on balancing model complexity and effectiveness, further refinement through hyperparameter tuning or exploring other ML algorithms could enhance performance. Additionally, the inherent differences in how each feature selection method evaluates and ranks feature importance might have impacted the model's ability to generalise across different scenarios.

Despite these limitations, this research underscores the advantages of algorithm adaptation in multi-target feature selection, particularly its ability to leverage inter-feature relationships. This is crucial for complex building systems like TABS, where the interactions between different features significantly impact the system's performance.

Looking ahead, this study opens avenues for future research. A direct extension could utilise the findings to explore the energy flexibility potential of the TABS within HiLo's lightweight concrete structures. This would contribute to a more nuanced understanding of energy flexibility in building systems and its implications for sustainable building operations.

On a broader scale, future research should delve into other multi-target feature selection methods, exploring their strengths and weaknesses in different contexts. Such studies would enrich the toolbox of ML techniques available for building energy management, paving the way for more intelligent control strategies. This would be a significant step toward advancing smart building technologies, which align with global efforts to reduce energy consumption and carbon emissions in the built environment.

Acknowledgments

The data used in this study were collected and made available by the Swiss Federal Laboratories for Materials Science and Technology (Empa) with the support of the Swiss Federal Office of Energy (SFOE) and the Swiss National Science Foundation. It was partly financed by Mitsubishi Electric R&D Centre Europe B.V. [contract numbers EMPA NEST2019 / TT 2019-124, ETH ID No 16234, ETH ID No 20500, ETH ID-No 23035] and the SFOE [contract number SI/502623-01].

References

Arteconi, A., Costola, D., Hoes, P., and Hensen, J. L. M. (2014). Analysis of control strategies for thermally activated building systems under demand side management mechanisms. Energy and Buildings, 80:384–393.

Block, P., Schlueter, A., Veenendaal, D., Bakker, J., Begle, M., Hischier, I., Hofer, J., Jayathissa, P., Maxwell, I., Echenagucia, T. M., Nagy, Z., Pigram, D., Svetozarevic, B., Torsing, R., Verbeek, J., Willmann, A., and Lydon, G. (2017). NEST HiLo: Investigating lightweight construction and adaptive energy systems. Journal of Building Engineering, 12:332–341.

Hashemi, A., Dowlatshahi, M. B., and Nezamabadi-pour, H. (2021). VMFS: A VIKOR-based multi-target feature selection. Expert Systems with Applications, 182:115224.

Kathirgamanathan, A., De Rosa, M., Mangina, E., and Finn, D. P. (2019). Feature Assessment in Data-driven Models for Unlocking Building Energy Flexibility. pages 366–373, Rome, Italy.

Li, R., Satchwell, A. J., Finn, D., Christensen, T. H., Kummert, M., Le Dréau, J., Lopes, R. A., Madsen, H., Salom, J., Henze, G., and Wittchen, K. (2022). Ten questions concerning energy flexibility in buildings. Building and Environment, 223:109461.

Lydon, G. P., Caranovic, S., Hischier, I., and Schlueter, A. (2019). Coupled simulation of thermally active building systems to support a digital twin. Energy and Buildings, 202:109298.

Masmoudi, S., Elghazel, H., Taieb, D., Yazar, O., and Kallel, A. (2020). A machine-learning framework for predicting multiple air pollutants' concentrations via multi-target regression and feature selection. Science of The Total Environment, 715:136991.

Naser, M. Z. and Alavi, A. H. (2023). Error Metrics and Performance Fitness Indicators for Artificial Intelligence and Machine Learning in Engineering and Sciences. Architecture, Structures and Construction, 3(4):499–517.

Olu-Ajayi, R., Alaka, H., Sulaimon, I., Balogun, H., Wusu, G., Yusuf, W., and Adegoke, M. (2023). Building energy performance prediction: A reliability analysis and evaluation of feature selection methods. Expert Systems with Applications, 225:120109.

United Nations Environment Programme (2021). 2021 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector. Technical report, Nairobi, Kenya.

van der Heijde, B., Sourbron, M., Vega Arance, F. J., Salenbien, R., and Helsen, L. (2017). Unlocking flexibility by exploiting the thermal capacity of concrete core activation. Energy Procedia, 135:92–104.

Zhang, L. and Wen, J. (2019). A systematic feature selection procedure for short-term data-driven building energy forecasting model development. Energy and Buildings, 183:428–442.

DETECTION AND CLASSIFICATION OF CONSTRUCTION OBJECTS BY USE OF MACHINE VISION AND DEEP LEARNING

Nicolas Nicolaou, and Symeon E. Christodoulou
Department of Civil and Environmental Engineering, University of Cyprus, Nicosia, Cyprus

Abstract

This research paper focuses on the detection and classification of objects at construction sites and analyzes the utility and potential of such detection and classification activities in the construction industry. Object detection and classification are performed by applying technologies such as machine vision (MV) and deep learning (DL) to image processing and/or in combination with object segmentation and labeling. These activities have varied applications at construction sites, including but not limited to: (1) the monitoring of workers and machinery for productivity measurement and for the prevention of accidents and collisions; and (2) the monitoring and classification of procured and installed construction materials for the evaluation of work progress. This serves as a valuable, low-cost measurement tool in the context of the management and monitoring of construction projects.

Introduction

The construction sector, even in modern times, faces numerous perennial problems and challenges related to the effective management and monitoring of construction projects.

Some of these challenges are, for example, associated with safety and health on construction sites. Despite significant progress achieved through the introduction of regulatory frameworks and legislation, a considerable number of occupational accidents are still recorded today. These accidents are not exclusively personal but are often linked to the reckless use of mechanical equipment and vehicles, as well as insufficient coordination, organization, and monitoring at construction sites. Therefore, a system using machine vision (MV) and deep learning (DL) technologies could, in this case, detect a worker in a restricted zone of the construction site, or determine whether the worker is wearing appropriate safety equipment or, even better, continuously appraise the ergonomic risks to workers (Lambrides and Christodoulou, 2023). Additionally, the system could be used to ensure the proper operation of mechanical equipment in designated areas and the maintenance of safe distances by the working personnel at the construction site.

Another critical issue commonly encountered at construction sites is resource management and construction progress monitoring. Delays, particularly in large construction projects, often occur due to material shortages and insufficient logistics in material delivery. Conflicts and construction delays are also frequent occurrences. Therefore, a mechanism based on the aforementioned technologies could be employed to monitor the transport of construction material to sites and visually document the construction process. This visual documentation can be valuable for reference and analysis.

In light of the aforementioned and other challenges faced by the modern construction industry, the use of machine vision and deep learning technologies is imperative. These technologies enable the automation of numerous technical processes on the construction site while facilitating the monitoring and resolution of various problems within it.

Literature Review

In recent years, a significant number of research studies has been conducted on the use of machine vision (MV) and deep learning (DL) technologies for detecting and classifying construction elements at construction sites. This trend began in the early 2010s, with works on the automated generation of parametric BIMs (Brilakis et al., 2010), and despite the significant improvements in relevant applications, there is still room for further development.

The article by Czerniawski and Leite (2020) introduces the automation of digital modeling of existing buildings through reality capture devices and computer vision algorithms. The goal is to facilitate the use of digital building representation technologies, promoting new forms of simulation, automation, and information provision. The research results suggest that achieving a more complete semantic coverage of building infrastructures will require a revision and intensification of relevant efforts.

Nath and Behzadan (2020) propose the validation of a genetic adversarial network (GAN) based on a deep convolutional neural network (CNN). The research involves photos taken, trained, and tested at the construction site from two internal datasets to increase image resolution when generating missing pixel information. Results demonstrate that using GAN-enhanced images can further improve the average accuracy of pre-trained models for object detection while maintaining overall processing time for real-time object detection.

In a subsequent work, Paneru and Jeelani (2021) provided an up-to-date and categorized overview of computer vision applications in construction by examining recent developments in the construction sector and the challenges that future research must address to maximize the benefits of computer vision. The authors focus on specific areas considered most likely to benefit significantly from computer vision, such as safety management on construction sites, progress and productivity monitoring, and work quality control.

One year later, Duan et al. (2022), focused on developing a large-scale image dataset specifically collected and processed for construction sites, named SODA (Site Object Detection Dataset). This dataset includes 15 types of objects categorized into mechanical means, materials, and labor personnel. Specifically, 20,000 images were collected from various construction sites, considering different construction site conditions, weather conditions, construction phases, and shooting angles. After careful examination and processing, 19,846 images were selected, containing 286,201 objects accompanied by corresponding labels from predefined categories.

An analysis conducted indicated that the developed dataset is advantageous in terms of diversity and volume. Further evaluation using two widely accepted object detection algorithms based on deep learning (*YOLO v3 / YOLO v4*) demonstrated the dataset's effectiveness in visualizing typical construction scenarios, achieving a maximum mean Average Precision (mAP) of 81.47%. This research contributes a large-scale dataset for the development of deep learning applications in object detection within the construction industry. It serves as a reference point for the further evaluation of corresponding algorithms in this field.

In their work, Wang et al. (2022) proposed a new semantic method aiming to extract information by integrating deep learning object detection and image captioning. This method explores important information from construction images or videos. In the proposed approach, object detection serves as an encoder to extract features of construction objects and the holistic image. By adopting this method, semantic information from construction images can be presented to project managers as a valuable tool for making crucial decisions on the construction site.

In the research work of Hou et al. (2022), a multi-object detection method based on the improved *YOLOv4* model is proposed to overcome the problem of low detection accuracy. Research results indicate that the average accuracy (mAP) of the improved *YOLOv4* model for many objects can reach 97.03%, which is 2.16% higher than that of the original *YOLOv4* detection network. At the same time, the detection speed reached 31.11 fps, a decrease of 0.59 fps, a result quite satisfactory for real-time detection data.

Zhou et al. (2022) propose an object detection method based on an improved *YOLOv5* model with high sorting accuracy of construction waste. It involves creating a dataset from images of construction waste taken in situ at construction sites. This improved model was trained, validated, and tested based on the collected images and compared with other conventional models such as *Faster-RCNN*, *YOLOv3*, *YOLOv4*, and *YOLOv7*. The *YOLOv5* model recorded an average accuracy (mAP) on the test dataset of 0.9480, indicating better performance than other conventional models in object detection.

In a recent research paper by Jog et al. (2022), full-scale validation experiments of a multi-object location tracking method for its application to resource tracking in large-scale, congested, outdoor construction sites are presented. The validation stage involved testing under harsh conditions on various large project sites. This research paper describes the process of data collection and testing, as well as the measurements and results obtained. The validation showed that the new vision tracking provides a good solution for tracking different entities in large and congested construction sites.

Research Methodology

The research work discussed herein focuses on the automated detection and classification of construction objects, and the applied research methodology was based on utilizing the Python programming language along with machine vision and deep learning technologies. In the realm of these technologies, several terms are often used interchangeably, yet they entail distinct tasks and methodologies. Classification entails assigning predefined labels or categories to input data based on their inherent features or attributes. Its primary objective is to categorize input instances into one of several predetermined classes, such as determining whether an image depicts a cat or a dog. On the other hand, the term "prediction" encompasses various interpretations, but within the domain of object classification, it involves assigning a probability score to each class to indicate the model's confidence level in its classification decision. Detection, meanwhile, pertains to the identification and localization of specific objects or phenomena within an input scene or data stream. It focuses on discerning the presence and position of objects of interest within images, videos, or sensor data, often using bounding box annotations. However, recognition, frequently conflated with detection, refers to the process of identifying and comprehending objects or patterns within an image or scene. Unlike detection, recognition entails a more profound analysis of visual content, which may include grasping the context, identifying specific object features or traits, and drawing higher-level associations or inferences based on observed patterns.

The goal was to create software, or leverage existing tools, capable of learning a series of construction objects present at a construction site. Subsequently, the software should successfully detect and classify these objects using either images from a dataset or random images. To achieve this objective, *ImageAI* (v.3.0.3) was employed. *ImageAI* (Moses, 2018) is an open-source Python library that simplifies machine vision and deep learning tasks. It is built on other libraries such as *TensorFlow* and *Keras*. From the array of tasks offered by *ImageAI*, specific codes related to image classification and object detection were utilized - activities directly aligned with the focus of this research. For each of the two tasks, a code was used for custom model training process based on the custom classes, resulting in the creation of a model. Additional codes were employed for result extraction, verification of the resulting accuracy-performance, and broader evaluation of the respective trained models, primarily through the utilization of unseen data.

Furthermore, a dataset was created for each task, incorporating photos of all the examined objects. These data resulted from a combination of my own photos from construction sites, ready-made datasets from Kaggle, which is a platform for data science and machine learning competitions, and generally photos obtained by Google Images search service. In the context of this research, the decision was made to initially explore two distinct classes to clarify the operational mode and compatibility of *ImageAI* with the research goals. These objects were the 'column' and the 'excavator'. However, at a later stage, seven more classes (totaling 9) were added to the detection, as follows: 'beam', 'masonry', 'slab', 'window', 'person', 'safety helmet', and 'reflective jacket'. The choice of some of these object classes relates to the intent of using developed algorithms and trained models for use in health & safety applications at construction sites.

Image Classification Framework

For this task, a set of 6000+ images of the object classes to be examined was collected. Initially, a general folder was created, which contained two additional folders named 'train' and 'test,' respectively. Within each folder, a subfolder was created for each prediction object. The training photos, used to train the classification model, and the corresponding test photos, used to evaluate it, were placed in these subfolders.

In the 'train' folder/dataset, 500 photos were included for each object, while in the 'test' folder/dataset, 200 photos were included. This dataset was then utilized in the training code, where various tasks were performed, including the selection of the algorithm. *ImageAI* offers the option to use four different algorithms for training custom image classification models (*MobileNetV2*, *ResNet50*, *InceptionV3*, and *DenseNet*), each with different speed and prediction accuracy characteristics.

Additionally, other parameters such as 'batch_size' (the number of images the network will process simultaneously) and 'num_experiments' (the number of network training iterations on all training images) were set in this code. For the purposes of this work, the MobileNetV2 algorithm was chosen due to its fastest prediction speed in compare with other algorithms.

Upon each execution of training code, the model with the highest accuracy was generated and stored in the dataset folder. Additionally, the other parameters mentioned above were systematically varied during each run to elucidate their impact on the accuracy of the respective model. This measure was undertaken to facilitate the incremental enhancement of the model, which became evident with each successive iteration. In this context, accuracy represents the percentage probability that a detected object belongs to a specific class. The accuracy is calculated using the following formula:

$$\text{Accuracy} = \frac{\text{Number of Correctly Classified Images}}{\text{Total Number of Images}} * 100$$

This percentage reflects the model's confidence in the correctness of its prediction. Higher percentage probabilities generally indicate that the model is more confident in recognizing a particular class of object in the image.

At a later stage, this model was employed in another code, where its effectiveness in predicting the examined and subsequently trained objects was evaluated using both trained and random photos. A schematic overview of this methodology is depicted in Figure 1.

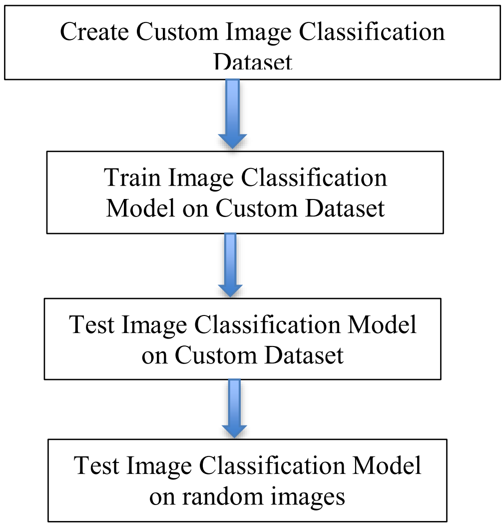


Figure 1: Image Classification methodology flowchart

Object Detection Framework

For this task, a set of 2700+ images was collected for the examined classes. Initially, a general folder was created, which included two subfolders named 'train' and 'validation,' respectively. Within each of these folders, two additional subfolders were created. The first, named 'images,' contained photos of the examined objects without separating them based on the object they depict. The second, named 'annotations,' contained the corresponding assignments for these classes, in txt format.

To create the assignments, an open-source graphic annotation tool for images, Labelimg, was employed. The process involved creating bounding boxes and labels in each photo and assigning them to each of the examined objects for the purpose of learning object detection. As part of this activity, 300+ photos were collected for each object, with 70-80% stored in the 'train' folder for training the detection model and the remainder in the 'validation' folder for evaluating the model's performance during training.

This dataset was then input into the training code, where, among other tasks, algorithm selection was performed. *ImageAI* provides the option to use two different algorithms to train custom image object detection models, namely *YOLOv3* and *TinyYOLOv3*, each with varying speed and accuracy characteristics for prediction. In this code, additional parameters such as 'batch_size' and 'num_experiments' were set, as previously explained.

During the training process for object detection, the initially used model did not include specific objects such as those found at construction sites. The model training with additional, construction site objects enriches the utilized pre-trained model and facilitates its use on construction-related image detection applications. Additionally, the option for training using a pre-trained *YOLOv3* model was specified. For the purposes of this work, both algorithms were employed. Future work shall aim the incorporation of newer releases of YOLO models (e.g., *YOLOv8*) and training datasets (e.g., *SODA*).

Each time the code was executed, the model with the highest accuracy in terms of mAP50 (mean Average Precision at 50%) was generated and stored in the dataset folder. Additionally, the other parameters mentioned above were systematically varied - in conjunction with the practical application of non-maximum suppression (NMS) - during each run to elucidate their impact on the accuracy of the respective model. During the training of each model, in addition to mAP50, additional metrics such as precision, recall, and mAP50-95 were obtained. However, these metrics were not automatically saved. Precision is a measure of the accuracy of a model's positive predictions and is derived from the following relationship:

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives}$$

On the other hand, recall, also known as sensitivity or true positive rate, is a term used to evaluate the performance of a classification or object detection model and is calculated as follows:

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives}$$

By using these two terms, it is possible to calculate the F1 Score, another widely used metric for evaluating classification models. The formula for the F1 score is:

$$F1\ Score = 2 * \frac{Precision * Recall}{Precision + Recall}$$

The term 'mAP' (mean Average Precision) is a metric that assesses the precision-recall tradeoff of a model. It evaluates how well a model performs at different confidence levels in its predictions. Specifically, 'mAP50' evaluates the model's precision and recall at a specific 50% Intersection over Union (IoU) threshold. Higher mAP50 values indicate better performance, with a maximum value of 1.0 representing perfect precision and recall at the specified IoU threshold. IoU is a metric that measures the overlap between the predicted bounding box and the actual location of the object. A 50% IoU means there is at least a 50% overlap between the predicted and actual contexts. This evaluation system is commonly used in assessing object detection models, including those trained for custom object detection tasks. At a later stage, this model was employed in another code, where its effectiveness in detecting the examined and subsequently trained objects was evaluated using random photos. A summary flowchart of this methodology is presented in Figure 2.

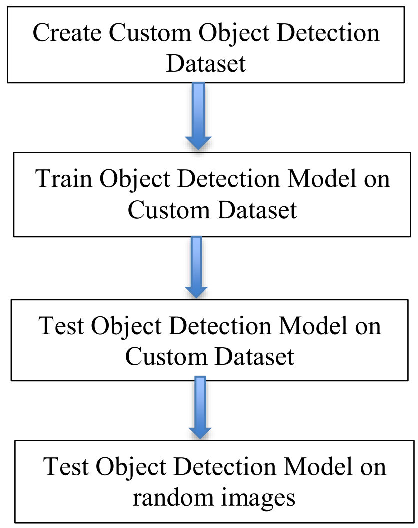


Figure 2: Object detection methodology flowchart

Additionally, the capability was provided by creating a Python code in combination with trained detection models to identify health and safety issues in photographs from construction sites, specifically printing a relevant warning message for the absence of part or all of the necessary safety equipment (protective helmet and reflective jacket) in case a person is detected in those areas.

Case Studies, Findings and Discussion of Results

In the pursuit of fulfilling this work's objectives, a series of tests were conducted through the execution of custom training Python programming language codes, as described. Throughout these tests, specific parameters were systematically varied in each training code, including the dataset itself, to generate two models - one for each task - with the highest accuracy and optimal performance. These models aimed to best fulfill the intended purpose for which they were created.

The final analysis results for the nine classes described in the previous stage are as follows. For image classification, considering the case of nine classes, a *MobileNetV2* model achieved an accuracy of 81.06%. This relatively high accuracy indicates the near certainty of the model in the correctness of its predictions, specifically in successfully predicting the nine trained classes in any given photo. This result was further validated by the model's performance on various photos, consistently yielding generally high probabilities for correctly predicting the depicted object. The model was tested on both trained and random images, and during the conducted tests, no significant change in performance was observed between these two categories of images. It is thus evident that the model's performance was proportional to its accuracy rate. Some example results are provided below.

The figure below presents the outcomes of a random image illustrating various objects at a construction site, including columns, beams and slabs. Utilizing the aforementioned image classification model, a probability of 52.24% was assigned to the depiction of a beam in the photo, 26.02% for column and 21.38% for slab. Predictions for the remaining classes were notably low, aligning with expectations given that only these three classes were prominently featured in the image under examination. The difference between the three main predicted classes is based mainly on the extent of viewing each class from the angle of the photo. This result affirms the high predictive capability of the trained model.

```
beam : 52.2404
column : 26.02
slab : 21.3805
masonry : 0.1741
person : 0.1149
Safety_Helmet : 0.0305
excavator : 0.0174
Reflective_Jacket : 0.0142
window : 0.0081
```

Figure 3: Result of random image in image classification (example 1)

Figure 4 displays corresponding outcomes from a random photo depicting masonry and window, which are equally positive. In this photo the prediction percentage is relatively high and almost equal between the two mainly predicted classes and very low for the rest classes that are not represented. Similar results were obtained for the rest of the classes among random photographs with the characteristic of combining several classes in the same examined image with the results being generally satisfactory as can be seen in Figures 5 & 6.

```
window : 50.7926
masonry : 44.4139
Safety_Helmet : 1.6045
Reflective_Jacket : 1.513
person : 0.6469
beam : 0.5646
column : 0.315
slab : 0.1471
excavator : 0.0023
```

Figure 4: Result of random image in image classification (example 2)

The image in Figure 5 depicts a female worker at a construction site, with appropriate personal construction protection gear. The results of the applied model to this case confirm the successful prediction of the three main classes that appear in the photo in question, with the corresponding prediction percentages showing significant

```
Reflective_Jacket : 49.1006
person : 38.3086
Safety_Helmet : 12.3392
excavator : 0.2281
beam : 0.0159
masonry : 0.0066
slab : 0.0006
window : 0.0003
column : 0.0001
```

Figure 5: Result of random image in image classification (example 3)

fluctuations. Specifically, a higher percentage was given to the reflective jacket class (49.10%), with the person class following (38.31%) and the safety helmet class registering a significantly lower percentage (12.34%) among these three classes.

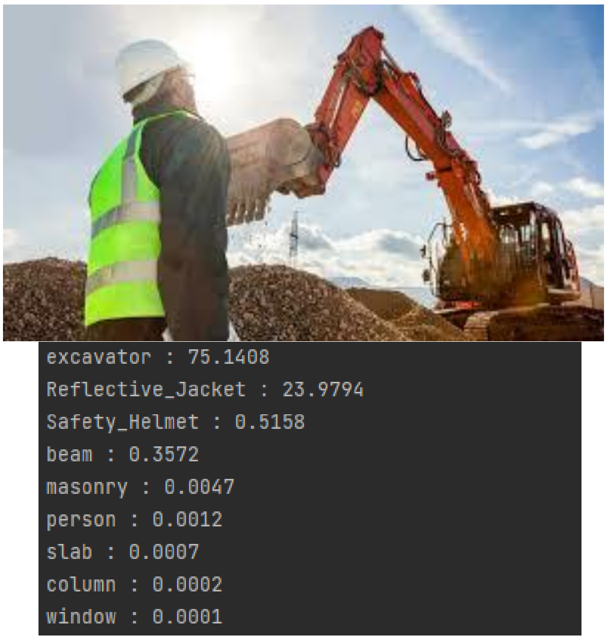


Figure 6: Result of random image in image classification (example 4)

Similarly, Figure 6 shows the results for an image showing a worker with appropriate personal protection gear near an excavator. The trained model correctly identified the existence of the excavator and the safety vest but not the worker and the safety helmet. One possible explanation for the very low prediction rates of the mentioned existing classes in the respective photograph is the intense lighting precisely at the point where the safety helmet is located, along with the posture of the human figure.

Frequently, models such as the one developed for image prediction encounter challenges related to their real-world effectiveness when presented with random images or images combining the examined objects. For instance, a model may perform well on the trained dataset but struggle to generalize its effectiveness to new, random data. However, as previously mentioned, no such phenomena were observed for this particular model.

Accordingly, for object detection, a *YOLOv3* model with an average accuracy (mAP) of 67.41 % was achieved. This figure indicates the relatively average to good accuracy of the specific model in terms of detecting and successfully classifying the objects under study in examined photographs, however efforts are being made to enhance the performance of this model to achieve even higher success rates. This result was further validated by the model's performance on various photos, where several satisfactory results were observed in terms of the true positive detection and classification of objects. The *YOLOv3* model was tested on both trained and random images, and during the conducted tests, no significant change was observed in terms of the model's performance between these two categories of images. Therefore, in this case as well, it is evident that the performance of the model is proportional to its accuracy rate. Some example results are provided below.

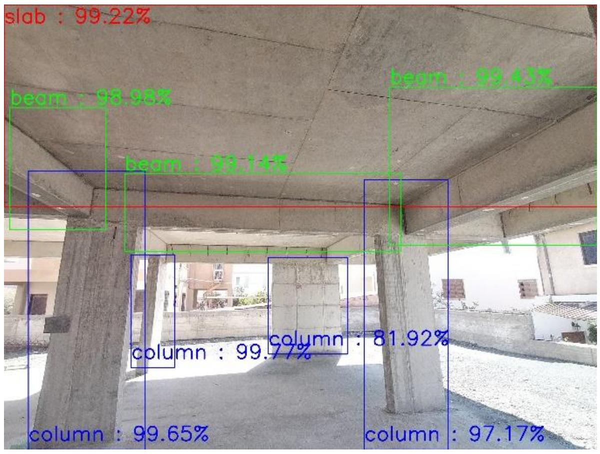


Figure 7: Result of rabdom image in object detection (example 1)

The figure above displays the outcomes of a random image depicting various objects on a construction site, including columns, beams and slabs. Utilizing the aforementioned object detection model, several load-bearing elements were detected with a probability of successful classification exceeding 81% affirming the high performance of the trained model. However, some other objects that would be expected to be detected by the model were not detected. These objects may have been influenced by the phenomenon of occlusion due to the relative angle of the photograph's capture. Nevertheless, the results largely coincide with those obtained using the classification model, which is deemed satisfactory. Additionally, Figure 8 demonstrates corresponding outcomes from a random photo of a worker on a construction site with appropriate personal protection measures. From the results of the model for this case, it emerged the successful detection of the three main classes that appear in the photo in question, with the corresponding confident scores showing very high (>96%) as in the case of the classification model for the same photo, the results of which were previously presented.

Figure 8: Result of random image in object detection (example 2)

Equally satisfactory results were obtained from tests conducted on other random photographs depicting the examined classes, as evident from Figures 9 & 10. The photo in question, presented in Figure 9, shows that using the model, masonry and a window were successfully detected with a confidence score exceeding 94%. Therefore, the behavior of the model under these conditions is considered satisfactory.

Figure 9: Result of random image in object detection (example 3)

Furthermore, in Figure 10, an example is provided depicting multiple objects (of those under examination), with the detection model yielding satisfactory results for most of them and moderate results for a single class (person). However, a better performance of the detection model is observed in the said image compared to its classification counterpart (Figure 6), as noticeably higher prediction rates are evident.

Figure 10: Result of random image in object detection (example 4)

Using the same detection model, results were obtained from other construction site photographs, with an additional capability introduced: the detection and notification of safety and health issues concerning the necessary and recommended personal protective measures on the construction site. Specifically, if a person was detected in these photographs without either or both of the reflective jacket and safety helmet, a corresponding warning was issued, as demonstrated in Figure 11.

Figure 11: Result of random image in object detection (health & safety example)

In conclusion, Figure 12 presents the confusion matrix of the detection model examined in the study. In a confusion matrix, each row represents the actual labels depicted in the validation set images, while the columns represent the corresponding labels predicted by the detection model. From the presented matrix, it is observed that some classes are positively evaluated due to a high number of true positive detections (e.g., excavator, window, etc.), while others are characterized as moderate to negative. The results of the matrix are to some extent expected, as the examined detection model did not achieve particularly high levels of accuracy. However, an NMS implementation with value equal to 0.4 led to an increase in the overall accuracy of the model in terms of F1-Score by approximately 30%. Essentially, the matrix provides insights into which classes the model struggles to predict accurately and can guide further improvements in the model, such as fine-tuning class-specific features or collecting more diverse training data for those classes. Therefore, there is room for significant future improvements.

Complete Confusion Matrix:

	column	excavator	beam	masonry	slab	window	person	Helmet	Jacket
column	349	0	111	10	67	3	3	0	0
excavator	0	76	0	0	0	0	3	0	0
beam	144	0	223	6	51	7	3	1	0
masonry	5	0	6	130	5	4	6	1	0
slab	66	0	41	3	78	1	2	0	1
window	13	0	5	7	2	277	26	0	0
person	7	5	4	6	4	15	298	58	74
Helmet	0	2	1	1	0	0	95	107	62
Jacket	0	0	0	1	0	0	85	70	65

Figure 12: Results of confusion matrix based on object detection model

Conclusions

The utilization of artificial intelligence, particularly technologies such as Machine Vision (MV) and Deep Learning (DL), in the construction industry is deemed imperative. The applications and benefits that can arise from these technologies are crucial, especially during the transition to a new era fraught with challenges. The real-time application of MV and DL can enhance the monitoring of safety and health issues on construction sites, extending to the broader and more essential oversight of labor management, mechanical equipment, vehicles, and materials, all while considering the relatively low costs resulting from the use of these technologies.

The present study focused on the automated detection and classification of construction elements at construction sites using the *ImageAI* library, built on the foundation of Python's *TensorFlow* and *Keras* libraries. The entire process was based on the integration of Machine Vision and Deep Learning technologies, combined with a dataset collected for the objects under consideration. The extracted results were analyzed in relation to the accuracy of the corresponding models from which they were derived. As part of future work, the following actions are to be taken to enhance the performance and accuracy of the relevant models based on *ImageAI*:

- Improvement of the annotation functions for bounded frames and labels, using advanced rendering practices in conjunction with the practical application of non-maximum suppression (NMS).
- Use of a balanced dataset with respect to all examined objects to prevent overfitting and the memorization of specific objects by the trained model for each activity.
- Exploration and testing of other custom activities offered by the *ImageAI* library for accuracy and usefulness, particularly by using video streams of related content.

References

Ahmadzada, A. (2020). People Image Dataset, many pictures of people performing different activities. https://www.kaggle.com/datasets/ahmadahmadzada/images2000/data

B Naik, N. (2023). Safety Helmet and Reflective Jacket, images of Individuals Wearing Safety Helmets and Reflective Jackets. https://www.kaggle.com/datasets/niravnaik/safety-helmet-and-reflective-jacket

Brilakis, I., Lourakis, M., Sacks, R., Savarese, S., Christodoulou, S., Teizer, J. and Makhmalbaf, A. (2010). Toward automated generation of parametric BIMs based on hybrid video and laser scanning data. Advanced Engineering Informatics, 24(4), pp.456-465.

Czerniawski, T. & Leite, F. (2020). Automated digital modeling of existing buildings: A review of visual object recognition methods. Automation in Construction, 113, p.103131.

Deshmukh, R., Wenguang, M. & Wei, M. (2020). Window Detection in Street Scenes, selected images from Paris Street-View Dataset with Window Annotations. https://www.kaggle.com/datasets/rude009/window-detection-in-street-scenes

Duan, R., Deng, H., Tian, M., Deng, Y. & Lin, J. (2022). SODA: site object detection dataset for deep learning in construction. arXiv preprint arXiv:2202.09554.

Hou, L., Chen, C., Wang, S., Wu, Y. & Chen, X. (2022). Multi-object detection method in construction machinery swarm operations based on the improved YOLOv4 model. Sensors, 22(19), p.7294.

Jog, G.M., Brilakis, I.K. & Angelides, D.C. (2011). Testing in harsh conditions: Tracking resources on construction sites with machine vision. Automation in construction, 20(4), pp.328-337.

Lambrides, E., & Christodoulou, S.E. (2023). Human action detection and ergonomic risk assessment at construction sites, by use of machine vision and deep learning. In: EC3 Conference 2023 (Vol. 4). European Council on Computing in Construction, Crete, Greece.

Moses, O. (2018). ImageAI, an open source python library built to empower developers to build applications and systems with self-contained computer vision capabilities. https://github.com/OlafenwaMoses/ImageAI.

Nath, N. & Behzadan, A.H. (2020). Deep generative adversarial network to enhance image quality for fast object detection in construction sites. In: 2020 Winter Simulation Conference (WSC) (pp. 2447-2459). IEEE.

Paneru, S. & Jeelani, I. (2021). Computer vision applications in construction: Current state, opportunities & challenges. Automation in Construction, 132, p.103940.

Tzutalin (2015). LabelImg, a graphical image annotation tool. https://github.com/HumanSignal/labelImg

Umer Yasin, M. (2022). Bricks Under Construction or Old Building / Houses, an image dataset that contains pictures of buildings and houses under construction. https://www.kaggle.com/datasets/mumeryasin/bricks-under-construction-or-old-building-houses/data

Wang, Y., Xiao, B., Bouferguene, A., Al-Hussein, M. & Li, H. (2022). Vision-based method for semantic information extraction in construction by integrating deep learning object detection and image captioning. Advanced Engineering Informatics, 53, p.101699.

Zhou, Q., Liu, H., Qiu, Y. & Zheng, W. (2022). Object Detection for Construction Waste Based on an Improved YOLOv5 Model. Sustainability, 15(1), p.681.

2024 EC³

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

A COMPARATIVE STUDY OF DEEP LEARNING MODELS FOR GRANULOMETRY IMAGE BASED ESTIMATION OF CONCRETE AGGREGATE

Benjamin Pasquier and Houda Chabbi Drissi
iCoSys - Institute of Artificial Intelligence and Complex Systems
HEIA-FR, HES-SO Haute école spécialisée de Suisse occidentale

Abstract

Obtaining the granulometry is the starting point of our pipeline for automating the calculation of concrete properties using images. For this reason, we focus on developing the best deep learning model that can compute aggregate gradation and can generalize to images obtained from different aggregate producers. We investigate two established approaches: Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs). Our analysis includes a dedicated CNN model trained from scratch, alongside pre-trained CNN and ViT models adapted through transfer learning.

To evaluate the performances and the generalization ability of the models, we use three different datasets: two publicly available and one of our own. Our analysis shows that transfer learning followed by fine-tuning on ViT_16 outperforms the other models, on both classification and regression tasks, with smaller errors and greater generalization capabilities.

Introduction

Deep learning applied to images is successfully used for prediction in many fields. Our aim is to use images of aggregates and a deep learning model to predict the properties of concrete. Our assumption is that such a model, trained on a dataset of aggregate images and corresponding concrete property data, would extract meaningful features from the visual representations of aggregates. These features can then be used to predict concrete properties such as compressive strength and workability.

The most influential factor on the properties of concrete mixtures is the type of used aggregates and their granulometry. Therefore, we propose to build a system that uses aggregate images as the foundation for the predictive final system, which will then use additional inputs to predict concrete properties based on the aggregate size distribution, as shown in Figure 1. Integrating this model with a camera system observing the aggregate conveyor belt during concrete production allows for real-time granulometry determination through image analysis. Leveraging this real-time information, the entire system can then continuously estimate key concrete properties enabling real-time monitoring of the mix design and ensuring consistent concrete quality.

We are interested in concrete mixes in which recycled or

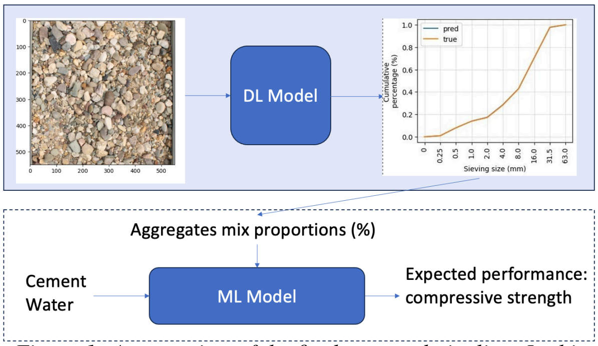


Figure 1: An overview of the final expected pipeline. In this paper, we are interested in finding the best DL model for the granulometry task.

natural aggregates may be present in varying proportions. Therefore, we seek for a model that can accurately and automatically extract granulometry distribution from aggregates images. In addition, this model should have a strong generalization capability, adapting to new images and new particle size distributions of natural or recycled aggregates, without the need to explicitly specify the aggregate type.

To carry out our study, we compare the performances of different deep learning architectures for granulometry estimation, first as a classification task, then as a regression task. To obtain a deep learning model, two main approaches can be considered: either build an own network and train it, or take advantage of generalist pre-trained models to perform the desired downstream task using transfer learning. The latter leverages knowledge from large datasets to improve efficiency and performance on new tasks. In our evaluation, we take AggNet, a specialized Convolutional Neural Network (CNN) model (Coenen et al., 2022) on the one hand, and perform transfer learning on pre-trained CNN models from different families on the other: MobileNetV2 (Sandler et al., 2018), ResNet50 (He et al., 2015) and finally the Vision Transformer (ViT) model ViT_16 (Dosovitskiy et al., 2020), a recent architecture using self-attention mechanisms (Vaswani et al., 2017).

In addition to leveraging transfer learning to enhance the performance of pre-trained models, we use hyperparameter optimization techniques to refine the predictions of these models. ResNet50 and MobileNetV2 are two popular deep learning CNN models that has been shown to be effective for a wide range of tasks and are relatively easy to train and fine-tune. ViT_16 is a vision transformer (ViT) that also shown to be state-of-the-art for image classification. ViT architecture models are interesting in this study because they are expected to allow a better accuracy and generalize better than CNNs (Maurício et al., 2023).

The paper is organized as follow: the next section reviews related works. Then, the two following sections introduce our methodology and our experimental setup to carry on the study for classification and regression tasks. Before concluding, we present the results of our experiments.

Related Work

Classification and granulometry tasks to determine the distribution of particle is very important not only for estimating concrete properties but for ore field in general. To avoid the need for costly manual techniques such as sieving, work has been carried out to automate the process by analyzing images of aggregates or ores. Since the emergence of deep learning, research has embraced CNN as a primary approach. A comprehensive survey on ore image processing using deep learning can be found in WANG Wei and Hao (2023), which highlights similarities to approaches employed for aggregates. The application of deep learning for aggregates can be divided into two main categories: those that start from scratch by building their own models, as seen in Lau Hiu Hoong et al. (2020), Qin et al. (2023) or Coenen et al. (2022), and those that leverage existing pre-trained models and employ transfer learning techniques to adapt them to specific tasks, such as Olivier et al. (2020). In addition to this distinction, there is another categorization based on the image processing strategy employed. The first one is to classify individual aggregate images as described in Sun et al. (2022), while the second way is to regress from images of aggregate mixtures to granulometry distributions.

In the first category, the authors of Lau Hiu Hoong et al. (2020) propose a customized Residual Network (ResNet) model and a dataset of 36'000 images of individual grains. They achieve a classification accuracy of 97% (brick, ceramic, stone, etc.). They also proposed a segmentation method to predict the nature of each grain in an image of a multi-grain sample. In Qin et al. (2023), the study is based on the concept of instance segmentation, using a specialized neural network model (AS Mask RCNN) to detect and classify individual aggregates within mixed aggregate images. The results of their study indicate that the AS Mask RCNN model achieved an accuracy of over 89.13%. The approach requires a dataset made up of images and their segmentation masks for each aggregate to be provided for each image. All these papers demonstrate the relevance of using deep learning models when calculating granulometry based on the segmentation of individual aggregates.

In the second category, where classifying and regressing are used without segmentation, deep learning has also been successfully used. In Olivier et al. (2020), the authors use the CNN architecture VGG16 Simonyan and Zisserman (2015) with transfer learning to predict the ten size fractions considered for an ore. The obtained results show the effectiveness of a CNN in predicting the size distribution of ore, with a mean model error of -0.012 and a standard deviation of 0.107. Coenen et al. (2022) presents a deep learning model, AggNet, for real-time determination of concrete aggregate grading curves. They propose a dedicated CNN network model with multi-scale feature extraction to handle diverse particle sizes and showed good results on a classification task with an accuracy of 95.5%, which is the best according to our knowledge.

We are interested in the second category of approaches because of their simplicity and industrial applicability. Once trained, these models dispense with the need for labor-intensive data preparation, allowing for direct estimation of aggregate granulometry from images. The main difficulty lies in preparing a dataset of varied aggregates images with their size distribution. Authors of AggNet published their dataset Coenen (2022) which we rely on as it meets our needs: each aggregate image is associated with its particle size distribution. We use this model as a reference for our analysis to study how it generalizes to our own dataset and to compare it to our approach which is based on adapting pre-trained models. In Coenen et al. (2023), the authors propose to use vision transformers and developed again their own model based on this architecture. They demonstrate the technical feasibility and interest of this approach. However, we believe that we can leverage the feature extraction capabilities of the pre-trained models (CNN or ViT), acquired through training on vast image datasets, and tailor them to our downstream tasks of classification and regression.

Methodology

As said in the previous sections, we compare four neural network architectures for estimating the granulometry of aggregates from images. We separate this task into two sub-tasks: the first one aim to classify the aggregates images toward the corresponding DIN 1045-2 Deutsches Institut für Normung (2008) standard granulometry class and the second one aim to directly estimate the mass percentage for each bin size considered. To perform this comparative study, we use three different datasets, two are publicly available and one is own-made.

Data

We use two publicly available datasets that contain image samples of natural aggregates: the Visual Granulometry dataset (Coenen, 2022) and the Deep Granulometry dataset (Coenen, 2023). The first one is designed for a classification task and the latter for a regression task. The Visual Granulometry dataset contains 900 images of aggregates along with their corresponding DIN 1045-2 standard granulometry class. There are nine classes in the standard (see Figure 2), each representing a grading curve, i.e. the size distribution of the aggregates. For each class, two samples of 5 kg of aggregates were produced and mixed to obtain a total of 50 images per sample, and 100 images per class.

The Deep Granulometry dataset contains 1650 images of coarse aggregate samples with different particles sizes ranging from 0.1 mm to 32 mm. Each image is accompanied by the mass percentage of each particle size bin considered, following 33 different granulometries (11 per largest grain size).

We then use a custom dataset with our own data that we use only for evaluation purposes, in order to measure the generalization of the trained models. This dataset contains 174 images of both recycled or natural aggregates from seven different sources, i.e. seven different granulometries, in an unbalanced fashion. For the classification task, we assign the DIN 1045-2 class to each of these granulometries by minimizing the mean squared error between them and the granulometry of each class. As shown in Figure 2, our grading curves can be far from the standard classes, therefore the assignation is not exact but still allows us to evaluate the model on our own data for a classification task. As two granulometries fall in the same class, we have only 6 of the 9 DIN 1045-2 classes that are represented. For the regression task, we simply report the mass-percentage for each size bin considered in the Deep Granulometry dataset.

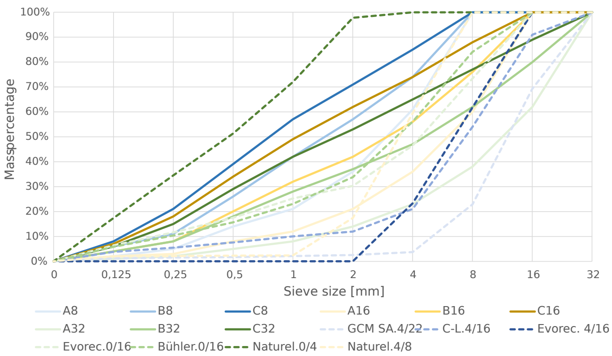


Figure 2: Grading curves of the DIN 1045-2 standard classes (solid lines) and those of our sources (dashed lines).

Since the two publicly available datasets contains images rectified by homography that were taken with a ground sampling distance (GSD) of 0.125mm, we also rectified our images in a similar way. However, as the setup was not exactly the same for all taken images, 80 of images have been cropped manually and can therefore present some deformations due to the perspective. In addition, the GSD of our images is not exactly equal to 0.125mm. These variations in our own dataset will serve to assess the generalization capabilities of the models we test. The Table 1 summarizes the size and characteristics of the two datasets used.

Neural networks

Classification, i.e. classifying images of aggregates towards the right DIN1045-2 standard class, is the first task we consider. We assume that this task is simpler than the regression one which consists of predicting the real percentage for each size bin considered. Therefore, we evaluate three CNN-based models, namely ResNet, MobileNetV2 and AggNet, as well as one ViT model. The AggNet model is a dedicated CNN model for granulometry estimation and the source code of its architecture has been made available by the authors (Coenen, 2022). The remaining three models are pre-trained models that we adapt to the granulometry estimation task using transfer learning. We then adapt and evaluate the two best performing classification models on the regression task, i.e. estimating the mass percentage for each size bin considered.

Transfer learning

As the state-of-the-art computer vision models are composed of millions of parameters, they need to be trained on large datasets. In order to adapt these models on a new task, it is often recommended to use the weights of a pre-trained model instead of training the model from scratch and risking to overfit the data. This can be done by freezing the weights of the pre-trained model and adding a new fully connected layer on top of it, which will be trained on the new task. This process is called transfer learning. In this study, we freeze the weights of the pre-trained model feature extraction layers, using it as a feature extractor, and train a new fully connected layer on top of them, as shown in Figure 3.

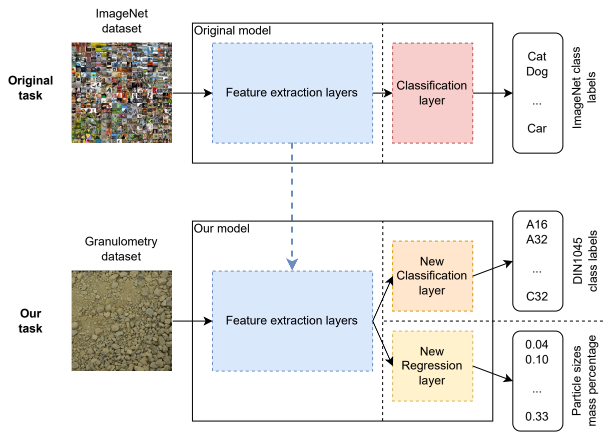


Figure 3: An overview of the transfer learning process used in this study. Feature extraction layers of the pre-trained model are frozen and new trainable classification or regression layers are added on top of them. Top image from (Deng et al., 2009).

To perform our comparative study, we use transfer learning on three pre-trained models, namely ResNet50, MobileNetV2 and ViT_16, all originally trained on the ImageNet (Deng et al., 2009) dataset.

Training

The network parameters θ are learned by optimizing a loss function $\mathscr{L}(\theta)$, which differs depending on the network task. The loss function is computed for each batch of data and the network weights are updated according to the gradient of the loss function with respect to the weights. The compared classification networks aim to classify images of aggregates towards M classes, where M is the number of DIN 1045-2 standard classes. To do so, they are

Table 1: *Summary of the datasets used.*

Dataset	Task	Size	Images size	Particles size	Nb. of classes / size bins
Visual Granulometry (Coenen, 2022)	Classification	900	2200x3000px	0-32mm	9
Deep Granulometry (Coenen, 2023)	Regression	1666	2200x3000px	0-32mm	33
Own	Both	174	1072x1472px	0-32mm	6 / 7

trained by minimizing the well-known cross-entropy loss function, defined as

$$\mathscr{L}_{CE}(\theta) = -\sum_{i=1}^{M} y_i \log(\hat{y}_i) \tag{1}$$

where y_i is the ground truth label for the i^{th} class (either 0 or 1) and $\hat{y}_i$ is the predicted probability for the i^{th} class.

For the regression task, we aim to predict the mass percentage of the M size bins considered. Therefore, the networks are trained by minimizing the Kullback-Leibler divergence, which is the same loss used in (Coenen et al., 2022). It is defined as

$$\mathscr{L}_{KL}(\theta) = \sum_{i=1}^{M} y_i \log(\frac{y_i}{\hat{y}_i}) \tag{2}$$

where y_i is the ground truth mass percentage for the i^{th} size bin and $\hat{y}_i$ is the predicted mass percentage for the i^{th} size bin.

For training, we systematically perform early stopping to avoid overfitting and use the Adam optimizer (Kingma and Ba, 2017) to optimize model weights. In order to obtain the best performances, we then perform different optimizations and evaluate their impact on a validation set before selecting the best model.

– **Data augmentation :** As the dataset used for both classification and regression are relatively small (900 to 1666 images), we perform data augmentation to allow models to generalize better and avoid overfitting on training data. This data augmentation is performed during the training, on the fly, so that it is highly unlikely for the model to see the same image twice. We test two kind of data augmentation:

1. A tuned augmentation by evaluating the model many times on different combinations of geometric transformations, such as rotation, shift, zoom or shear.
2. The augmentation proposed in paper (Coenen, 2022), performing geometric and radiometric (e.g. hue shift) transformations.

We compare the best tuned augmentation to the one proposed in (Coenen, 2022) and keep the one that performs the best on the validation set.

– **Hyperparameter tuning :** We then tune the hyperparameters of the models in order to optimize their performances. This is done by evaluating many times the model on different combinations of hyperparameters, which are the batch size, the neurons number in the fully connected layers, the dropout rate and the learning rate. We keep the combination of hyperparameters that gives the best performances on the validation set.

– **Fine-tuning :** Finally, we perform a final fine-tuning by unfreezing the weights of the pre-trained model feature extraction layers and continuing the training for a few epochs. We then evaluate if this fine-tuning improves the performances of the model on the validation set.

Experimental Setup

In order to evaluate the performances of the different models, we use an identical experimental setup for all of them. We first split the dataset into a training and test set with a ratio of 80% and 20% respectively, the latter being only used for the final evaluation of each model. Before training, the training set is further split into a training and validation set with the same ratio, allowing us to select the best model according to different configurations and to ensure its good generalization. The Table 2 summarizes the number of images used for training and evaluation for each task.

Table 2: *Number of images used for training and evaluation for each task. Evaluation is both performed on the Deep or Visual Granulometry (V/DG) datasets and on our own dataset.*

Set		Classification	Regression
Training		720	1326
Evaluation	V/DG	179	340
	Own	174	174

As the architecture of the three pre-trained models we consider are designed to take images of 224x224 pixels as input, all the images are cropped to be squared and then resized to this size. For the AggNet model, images are simply down-sampled to 550x750 pixels, keeping their original size ratio.

Figure 4 summarizes the training and evaluation procedures. More details for each task we consider are given in the next sections.

Classification task

For the classification task, we train and evaluate respectively four models : ResNet50, MobileNetV2, ViT_16 and AggNet. Hyperparameter tuning, data augmentation tuning and fine-tuning are only performed on ResNet50 and MobileNetV2, as ViT_16 showed already good performances without doing so (see section *Results and dis-*

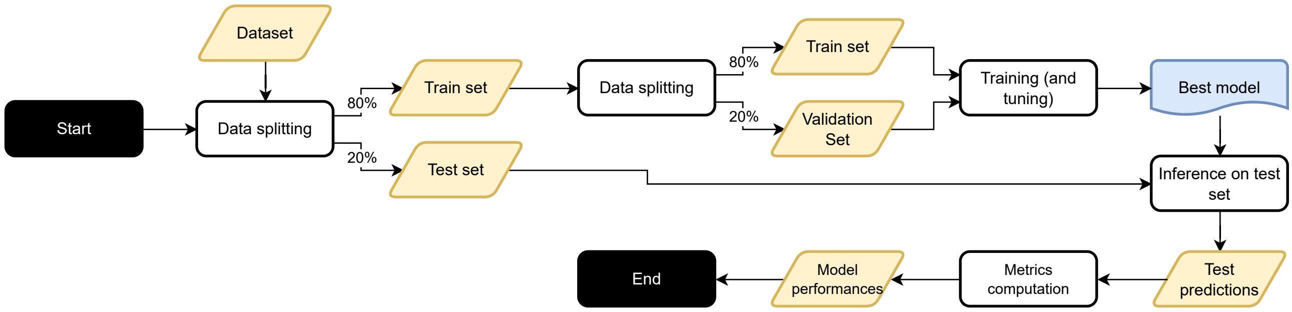


Figure 4: Training and evaluation procedure for each model architecture we compare, on both classification and regression tasks.

cussion) and as the authors of the AggNet model already performed these optimizations.
To evaluate their performances, we use the accuracy metric, defined as

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \quad (3)$$

We also compute the confusion matrix, that allows us to see which classes the evaluated model confuses the most. It also allows us to quickly calculate other metrics per class, such as the precision, recall and F1-score.

Regression task

For the regression task, we only train the two best performing models on the classification task, namely ViT_16 and AggNet (see section *Results and discussion*). This time, we also perform hyperparameter tuning, data augmentation tuning and fine-tuning on ViT_16 in order to obtain the best possible model and see if it can outperform AgGNet. Training configuration for AggNet model is once again taken from (Coenen, 2022), while the best configuration we find for ViT is the following :

- **Tuned hyperparameters** : batch size of 64, dropout rate of 0, hidden size of 512 and learning rate of 0.007
- **Tuned data augmentation** : horizontal and vertical flip, shift range of 0.1, zoom range of 0.3 and fill mode on reflect.
- **Fine-tuning** : learning rate of 0.0001.

To evaluate the performances of each model, we use the mean absolute error (MAE) metric and the root mean squared error (RMSE) metric, defined as

$$MAE = \frac{1}{N}\sum_{i=1}^{N}|y_i - \hat{y}_i| \quad (4)$$

and

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2} \quad (5)$$

where N is the number of predictions, y_i is the ground truth mass percentage for the i^{th} image and $\hat{y}_i$ is the predicted mass percentage for the i^{th} image. These metrics are calculated for each size bin and then averaged to obtain a single measure of model performances.
We compute the RMSE along with the MAE because the RMSE penalizes more the large errors than the MAE, and thus gives us an other important information about the model performances. We use the RMSE instead of the MSE because it is more interpretable as it is in the same unit as the ground truth mass percentage vector.

Results and discussion

We first discuss the results obtained by the different classification models, before focusing on the results obtained by the two best performing ones on the regression task.

Classification

Table 3 summarizes the performances in terms of accuracy obtained by each compared models on the two test datasets. While they all performs significantly better on the Visual Granulometry dataset, the ViT_16 outperforms the other models with respective accuracies of **97%** and **34%** on both datasets. As neither hyperparameter tuning nor data augmentation tuning were performed on ViT_16, we can assume that even better performances could be obtained by doing so. This shows how powerful transformers can be on various tasks, including computer vision tasks. The AggNet model is also performing very well on the Visual Granulometry dataset Coenen (2022), which coincide with the results reported in Coenen et al. (2022). All the models show a low accuracy on our own dataset, with accuracies ranging from 26% (AggNet) and 34% (ViT_16), showing a poor generalization of the model on the classification task.

Table 3: Performances of the different classification models on the two test sets, i.e. the Visual Granulometry (VG) and our own dataset.

Model	**Accuracy**	
	VG data	**Our data**
ResNet-50	0.85	0.30
MobileNetV2	0.87	0.29
ViT_16	**0.97**	**0.34**
AggNet	0.94	0.26

Figure 5 shows the confusion matrix obtained by the ViT_16 model on our own dataset and help to understand why models are under performing on it. As it shows high

Table 4: *Regression results on the Deep Granulometry dataset with both models. Errors are computed for each grain size bins considered and then averaged.*

Grain size bins [mm]		**0.25**	**0.5**	**1**	**2**	**4**	**8**	**16**	**31.5**	**63**	**Avg.**
AggNet	MAE [%]	0.22	1.23	1.36	0.67	1.65	1.61	1.73	1.69	0.21	1.15
	RMSE [%]	0.28	1.56	1.78	0.86	2.11	2.18	2.61	2.72	0.47	1.62
ViT	MAE [%]	0.13	0.88	0.86	0.46	0.51	0.98	0.94	0.49	0.04	**0.59**
	RMSE [%]	0.18	1.20	1.29	0.66	0.73	1.59	1.61	1.02	0.14	**0.93**

accuracy on Visual Granulometry data, we do not present here the confusion matrix obtained on this data. On our own data, ViT_16 frequently misclassifies B16 as A32 or A16, and A32 as A16, likely due to the fact that our data samples do not exactly follow the particle size distribution of the classes defined by DIN 1045. For better performances, we should add more classes instead of simply assign our sample to one of the standard classes. Analyzing ViT's regression performances on our data might be more revealing, as it predicts continuous values of the real granulometry instead of inferred discrete classes.

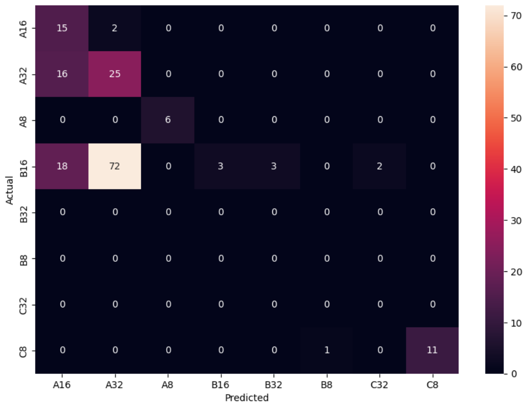


Figure 5: *Confusion matrix obtained on our own dataset with the ViT_16 model.*

Regression

Since the best models on the classification task on Visual Granulometry data are the ViT_16 and AggNet models, we only train and evaluate these for a regression task.

Table 4 first shows the results obtained by both models on the Deep Granulometry dataset (Coenen, 2023). The ViT model therefore fares better than the AggNet model, with an average MAE of **0.59%** versus 1.15%, i.e. half as much. The obtained RMSE with both models also confirmed this observation. As far as errors by size are concerned, the ViT model seems to have more difficulty in predicting proportions for sizes from 0.5mm to 1mm and from 4 to 31.5mm, as does the AggNet model. It shows that either some sizes are more difficult to differentiate than others, either the data is much more varied in certain bin sizes than others.

The worst respectively the best predicted grading curves by ViT on this first dataset are shown in Figure 6, along with the ground truth. We see that the model achieve to predict the perfect grading curve in some case, and to predict a grading curve that is still very close from the ground truth in the worst case.

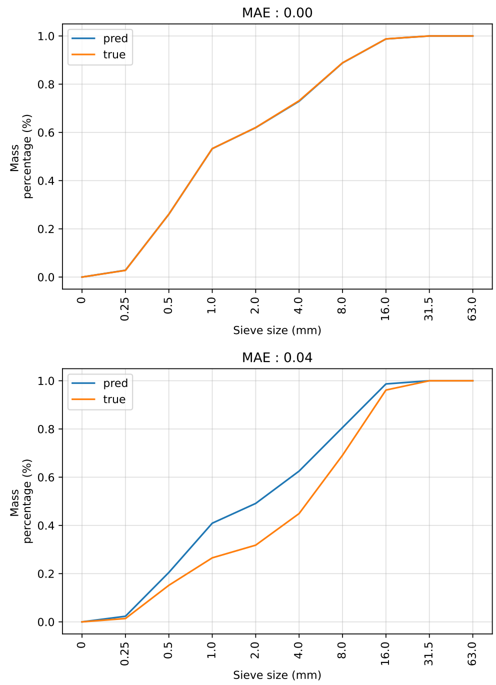


Figure 6: *Best (top) and worst (bottom) predicted grading curves (in blue) by ViT on the Deep Granulometry dataset, along with the ground truth (in orange)*

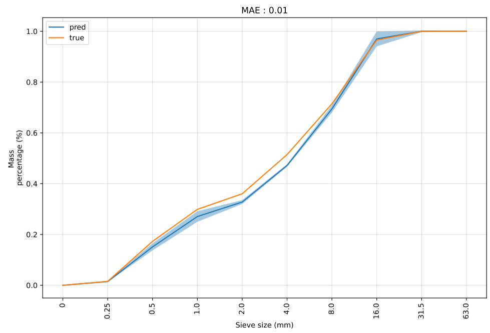


Figure 7: *ViT worst resulting grading curve (in blue) obtained by averaging predictions of a same aggregates mixture on the Deep Granulometry dataset, along with the ground truth.*

If we average predictions over samples following the same grading curve, we obtain a new prediction that is much closer to the ground truth even in the worst case, as shown in Figure 7. This shows that in a real setup where many images of the same mixture of aggregates are taken, we can predict a much more precise granulometry by averaging the predictions made by the model.

Table 5 then shows the results obtained by both models on our own dataset. The results are significantly worse than those obtained on the Deep Granulometry dataset, with an

Table 5: Regression results on our own dataset with both models. Errors are computed for each grain size bins considered and then averaged.

Grain size bins [mm]		0.25	0.5	1	2	4	8	16	31.5	63	Avg.
AggNet	MAE [%]	6.06	9.0	7.22	4.64	9.46	9.87	9.04	13.33	0.41	7.67
	RMSE [%]	10.01	9.45	7.88	5.44	10.15	12.07	11.66	15.63	0.71	9.22
ViT	MAE [%]	6.35	4.57	4.14	5.00	8.47	11.87	9.53	10.82	0.18	**6.77**
	RMSE [%]	10.22	5.28	4.98	6.21	10.76	13.82	11.44	12.97	0.29	**8.44**

average MAE of **6.77%** for the ViT model and 7.67% for the AggNet model. While these errors are similar, ViT is still able to generalize better that the AggNet model. This overall increase in errors can be explained by the difference in particle size distributions between the two datasets. Indeed, the granulometries of our dataset are significantly different from those of the Deep Granulometry dataset. As the models are trained on the latter, it is very likely that they will have difficulty generalizing to other data. In addition, some images in our dataset have not been rectified by homography, this difference in images may therefore also have an influence on model performances.

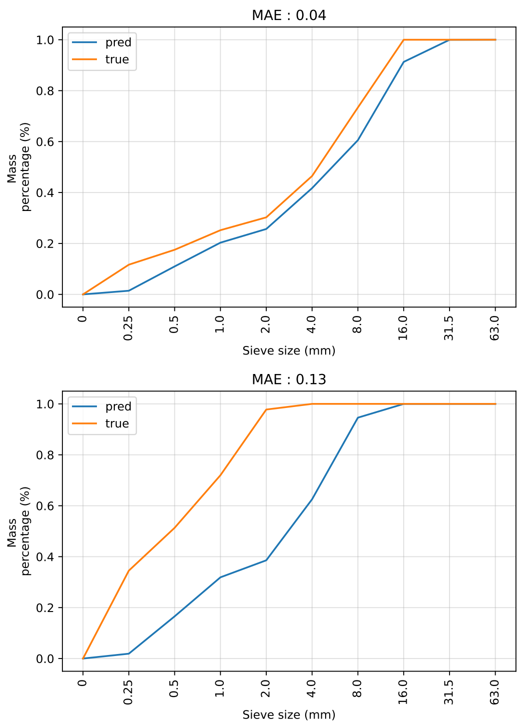


Figure 8: Best (top) and worst (bottom) predicted grading curves (in blue) by ViT on the Deep Granulometry dataset, along with the ground truth (in orange)

The worst respectively the best predicted grading curves by ViT on our dataset are shown in Figure 8, along with the ground truth. This time, we see that the model may struggled to predict a grading curve close from the ground truth, especially when aggregates follow a granulometry far from the ones the model was trained on. These results indicate that we may need aggregates training data that follows a wider range of granulometries in order to increase the generalization of the model. We can still note that for different but close granulometries, the model is able to make prediction with few errors, which is encouraging. If we average predictions over samples following the same grading curve, we achieve to reduce the MAE but still remains relatively high (11%) in the worst case, as shown in Figure 9. The need for more varied training data therefore remains.

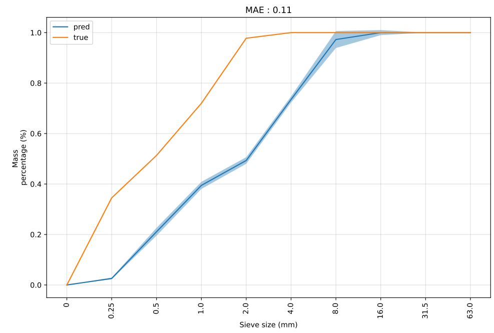


Figure 9: Worst resulting grading curve (in blue) obtained by averaging predictions of a same aggregates mixture made by ViT on our own dataset, along with the ground truth (in orange).

Conclusion

The result of our study shows that using the dedicated CNN AggNet model of Coenen et al. (2022) for the task of classification and calculation of the distribution of aggregates from their images is a better approach than performing transfer learning on the most popular pre-trained CNNs that we used: ResNet50 and MobileNetV2. On the other hand, our tests show that applying transfer learning on a pre-trained model based on transformers (ViT_16) allows to achieve better results on the two considered tasks. Iman et al. (2023) positions transfer learning as a valuable technique to unlock the full potential of deep learning. In our case, where the images are very specific and different from datasets of the pre-trained models, the smaller AggNet model, using multi-scale feature extraction layers, effectively handles the diverse aggregate sizes compared to the generalist pre-trained CNNs. This specialization likely contributes to its better performance. Similarly, ViT_16 excels in this task due to the inherent ability of transformers to capture both local and global interactions within the image, potentially explaining the advantage of ViT over AggNet.

To evaluate the generalization ability of each model, we employed our own dataset for testing. This dataset differs from the publicly available one used to train and evaluate the models in two key ways. Unlike the public dataset, it includes both natural and recycled aggregates and it presents different aggregate size distributions. Again, we obtained better results with ViT_16, which reinforces the idea of deepening this approach to improve its adaptability to datasets that do not perfectly follow the granulometry of the training data. While the current results of this generalization are not optimal, we believe incorporating a subset of our own data into the training set has the potential to significantly improve model performance.

Therefore, our next future task is to enlarge our own dataset in order to cover a greater variety of grading curves. Besides, our dataset is currently imbalanced, and we aim to achieve a balanced distribution with at least 100 images per class. While this imbalance was not a disadvantage for the present study (used for testing only), it needs to be addressed to determine the best strategy for ViT_16 generalization. We will explore two options: fine-tuning the model on a subset of our own data or reapplying transfer learning with hyperparameter tuning incorporating this subset into the training dataset. Our goal is also to identify the optimal dataset size that maximizes the generalization performance of the ViT model for granulometry estimation.

Acknowledgments

We are thankful to Daia Zwicky, Julien Ston and Sonia Anselmina from iTeC institute of HEIA-FR (Institut des Technologies de l'Environnement Construit) for the preparation of our own dataset.

This research was partially supported by a grant of the Programme de recherche HEIA-FR / Smart Living Lab (AGP:119149).

References

Coenen, M. (2022). Dataset: Visual granulometry: Image-based granulometry of concrete aggregate.

Coenen, M. (2023). Dataset: Deep granulometry.

Coenen, M., Beyer, D., and Haist, M. (2023). Granulometry transformer: image-based granulometry of concrete aggregate for an automated concrete production control.

Coenen, M., Beyer, D., Heipke, C., and Haist, M. (2022). Learning to sieve: Prediction of grading curves from images of concrete aggregate. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, V-2-2022:227–235.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248–255.

Deutsches Institut für Normung, B. (2008). Din 1045-2: Concrete, reinforced and prestressed concrete structures.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. CoRR, abs/2010.11929.

He, K., Zhang, X., Ren, S., and Sun, J. (2015). Deep residual learning for image recognition. CoRR, abs/1512.03385.

Iman, M., Arabnia, H. R., and Rasheed, K. (2023). A review of deep transfer learning and recent advancements. Technologies, 11(2).

Kingma, D. P. and Ba, J. (2017). Adam: A method for stochastic optimization.

Lau Hiu Hoong, J. D., Lux, J., Mahieux, P.-Y., Turcry, P., and Aït-Mokhtar, A. (2020). Determination of the composition of recycled aggregates using a deep learning-based image analysis. Automation in Construction, 116:103204.

Maurício, J., Domingues, I., and Bernardino, J. (2023). Comparing vision transformers and convolutional neural networks for image classification: A literature review. Applied Sciences, 13(9).

Olivier, L. E., Maritz, M. G., and Craig, I. K. (2020). Estimating ore particle size distribution using a deep convolutional neural network this work is based on research supported in part by the national research foundation of south africa (grant number 111741). IFAC-PapersOnLine, 53(2):12038–12043. 21st IFAC World Congress.

Qin, J., Wang, J., Lei, T., Sun, G., Yue, J., Wang, W., Chen, J., and Qian, G. (2023). Deep learning-based software and hardware framework for a noncontact inspection platform for aggregate grading. Measurement, 211:112634.

Sandler, M., Howard, A. G., Zhu, M., Zhmoginov, A., and Chen, L. (2018). Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation. CoRR, abs/1801.04381.

Simonyan, K. and Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition.

Sun, Z., Li, Y., Pei, L., Li, W., and Hao, X. (2022). Classification of coarse aggregate particle size based on deep residual network. Symmetry, 14(2).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L. u., and Polosukhin, I. (2017). Attention is all you need. In Guyon, I., Luxburg, U. V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., and Garnett, R., editors, Advances in Neural Information Processing Systems, volume 30. Curran Associates, Inc.

WANG Wei, LI Qing, Z. D.-z. L. H. and Hao, W. (2023). A survey of ore image processing based on deep learning.

MACHINE LEARNING MODEL PREDICTION OF PROJECT SUCCESS

K. Moravej, S. L. Laroche, and S. M. Musick
Canadian Nuclear Laboratories, Chalk River, Canada

Abstract

Managing projects can be challenging with institutionalism and enterprise environments. Artificial Intelligence can be used to augment decision-making by enhancing the understanding of "drivers" enabling project performance. Research was performed to determine if a supervised machine learning (ML) model could predict the project's performance at fiscal year end, utilizing mid-year information. Enterprise Data, from annual performance data was used to train and evaluate several ML models. The results are promising, and indicate the potential of identifying projects that require monitoring at mid-year. Organizations could use a similar methodology to support decision-making and focus on the highest-value projects with predictable outcomes.

Introduction

The tendency for an organization to manage projects with the determination to drive successful outcomes can be challenging with institutional dynamics and enterprise environmental factors. Decisions to initiate, continue, or stop projects are frequently based on short-sighted assumptions, such as flawless implementation of processes and "as planned" project execution. These assumptions often result in risks that are realized where projects require longer durations to complete and at higher costs, a study found up to a 38% extension in schedule duration and an increase in the original budget of up to 45% (Chandrasekaran, et al., 2021).

To help mitigate these risks, new and evolving tools, approaches, and methodologies are being applied to enable decisions earlier. One such tool is ML. When augmented with an organization's Enterprise Data, ML will enhance the project teams' awareness and understanding of the "drivers" enabling project performance.

An agile organizational change management approach is critical to the success of implementing any ML model. Primarily, this approach is directly associated with the tendency of an organization to have complex data environments that do not readily lend themselves to direct integration between systems with the ever-changing enterprise environment, (Bou Hatoum, Nassereddine, Musick, & El Jazzar, 2023). However, it is essential that the exhaustive amount of data being generated daily throughout the project lifecycle is captured using a holistic framework that enables collaborative access by decision makers (Hatoum, Piskernik, & Nassereddine, 2023).

The development of new tools for project management is evident as industry continues to invest in projects with only 35% of these projects being considered successful (Nieto-Rodriguez & Viana Vargas, 2023). Projects management activities continue to rely on traditional tools, such as spreadsheets and slides. Artificial Intelligence (AI) can be used to leverage the existing information in enterprise systems to reveal trends earlier and enable better selection and prioritization of projects.

A Gartner Survey reported that 37% of organizations had implemented AI in some form by 2019, increasing 27% from the previous four years (Costello, 2019). This increasing trend to use AI has continued with the availability of commercial off the shelf items that integrate AI into their service offerings.

In this project, the application of ML in project management was investigated. A research initiative was undertaken to determine whether a supervised ML model could accurately predict the probability of project success. This research aimed to illustrate that organizations could enhance their project management methodologies by integrating their own model designs with existing approaches. The initiative utilized Enterprise Data, incorporating annual project performance data from almost 700 projects over a five-year period. These projects were labelled with binomial project identifiers, such as "On Track" or "Monitor Performance", determined from prescribed cost and schedule performance indices success criteria ranges. The outcomes from the labelling effort were then used to train the ML model algorithms.

This research sought to determine how the trained model classifies each project, by evaluating the explain-ability of the ML model. These results furthered the understanding of the behavior of ML models and revealed that after six months, cost and schedule performance indices were lead indicators in predicting project success.

The ML model, designed under this research initiative, provided advanced insights into project performance ahead of project completion. This additional insight provided management with additional information to determine if the project should proceed, and a firm foundational technique that other organizations can use to leverage ML. Once implemented, organizations would be able to weigh project benefits while supporting decision-making process through the gating and sanctioning process, focusing on the highest value projects with predictable outcomes.

Objective

This project aims to explore the feasibility of designing a supervised ML model capable of predicting the project's status at the end of the fiscal year, utilizing the information accessible during the mid-year point. To achieve this objective, the following steps were taken:

- **Data Source:** Identify and extract relevant data sources from the company's database.
- **Data cleaning and Processing:** Evaluation, cleaning, and preprocessing of the extracted data to prepare it for ML model training.
- **Model Training and Evaluation:** Training, tuning, and evaluating ML models using the prepared dataset.

Data Sources

The identification of suitable enterprise data is a critical step when starting to develop the ML model. A complete understanding of the data is important in determining what is possible or not based on the data available and its quality.

A review of historical project data identified three types of data sources to serve as project information with similar characteristics. These data sources were Project Data (including Earned Value Management System (EVMS) Data), Milestones, and Deliverables. The Milestones and Deliverables datasets are crucial tools for project management as they contain information about each project's ability to achieve the desired outcomes. The enterprise data was consolidated into a single location, or a dataverse, as shown in Figure 1. This process required the careful integration of data from all sources, rigorous data cleaning, and preprocessing steps to remove irrelevant data points. Data consistency was ensured across all sources by standardizing the data format and aligning the data fields.

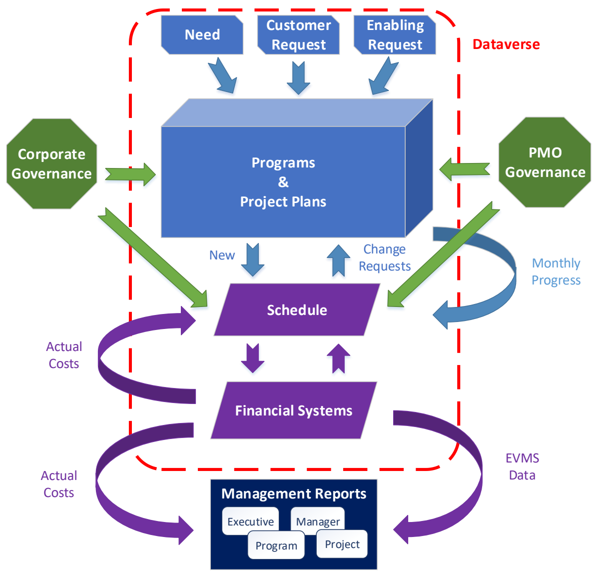


Figure 1: Dataverse contents

The extracted dataset consisted of nearly 700 projects spanning a five-year timespan, with multi-year projects included for every year of their duration. Figure 2 illustrates the number of projects for each fiscal year.

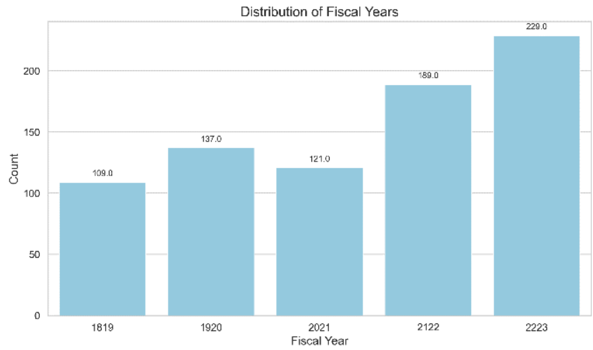


Figure 2. The number of projects for each fiscal year. The horizontal axis represents the fiscal year, where, for example, "1819" corresponds to the fiscal year 2018-19, and "2223" corresponds to the fiscal year 2022-23.

Table 1 summarizes the types of information available for each project and included free text, categorical, and numerical. This project information served as input to the ML model after appropriate pre-processing and is referred to as project input or indicators in this paper.

Table 1: Information available for each project.

Free Text	Categorical	Numerical
WorkPackage_Title Project Description Project Objective Background Tasks details Milestone details Deliverable details Etc.	fiscalyear Program, Program Manager Theme_Name, SubTheme_Name Project Lead Principal Investigator planned employees actual employees Etc.	BurdenedCost April-March EVBurdened April-March PVBurdened April-March ACBurdened Etc.

Data cleaning and Processing

The resulting structured dataset was designed for ML by ensuring the data was well-formatted, consistent, and complete. All dataset fields were retained during this phase to determine which fields should be used as inputs for the ML model. This decision ensured that all potentially useful information was available for analysis and optimization of the ML model's performance. Several appropriate feature engineering and pre-processing techniques, such as normalization, scaling, encoding of categorical data (Müller & Guido, 2016), text preprocessing, and topic modeling (Tong & Zhang, 2016) using natural language processing (NLP) were implemented.

Project Success Indicators

Labelled data is required to train a supervised ML model. To classify the projects, a success metric was established and each project tagged with the appropriate label. The project data could be classified as either "On Track" or "Monitor Performance".

The criteria for labelling the dataset were determined based on Project Management Office governance and annual Cost and Schedule Performance Indices, SPI, and CPI. These indices were used to label the historical project performance data which could then be used to train the supervised ML model. Table 2 shows the metric used for labeling the projects. For example, if the project's CPI at the end of the fiscal year is 0.98, that project is labeled as "On Track," and if the CPI at the end of the year is 1.25, the project is labelled as "Monitor Performance." The same idea has been applied to label each project based on the SPI metric. This approach provided two labels for each project: one based on the CPI and the second one based on the SPI performance.

Table 2: The CPI and SPI thresholds for labelling projects.

Metric	Threshold	Label
CPI & SPI	0.95 ≤ metric < 1.2	On Track
CPI & SPI	metric < 0.95, metric ≥ 1.2	Monitor Performance

Model Training and Evaluation

Given that CPI and SPI are distinct variables driven by dynamically different factors, two separate models were created. The first model was formulated to predict the CPI label at the end of the year, while the second model concentrated on predicting the SPI label. The input parameters for both models remained consistent with only the target variable being different between the two models. For the SPI model, the target variable is the SPI label, while for the CPI model, the target variable is the CPI label.

To ensure a fair evaluation of our model's performance post-training, the data was divided for each model into a training dataset (75%) and a test dataset (25%). This dataset split was organized to preserve a similar ratio of monitor performance and on-track instances in both the training and validation sets. Subsequently, the ML model underwent training using the training set and performance evaluation using the test datasets—essentially, unseen data—during the training phase. This approach serves as a reliable indicator of the ML model's effectiveness and performance.

Multiple ML models have been trained to investigate their performances on our dataset. In this paper, the Decision Tree (Charbuty & Abdulazeez, 2021), Random Forest (Parmar, Katariya, & Patel, 2019) and XGBoost (Chen & Guestrin, 2016) models were used to evaluate the suitability.

The hyperparameters of each model were tuned using the K-Fold cross-validation technique (Gupta, Gupta, Kumar, & Sardana, 2021). K-Fold cross-validation is very important in ML development and hyperparameter tuning, especially when dealing with a small dataset. This method improves model performance and robustness while optimizing hyperparameters for better generalization. We adjusted the hyperparameters of our models using the F1 score which is defined as:

$$F1 = 2 * \frac{(Precision * Recall)}{(Precision + Recall)} \quad (1)$$

where, precision measures the proportion of true positive predictions out of all positives and recall measures the proportion of actual positives correctly predicted by the model. Precision and recall are defined as:

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

where,

- *True Positives (TP)* are the number instances that are correctly predicted as positive by the model.
- *False Positives (FP)* are the number of instances that are incorrectly predicted as positive by the model when they are actually negative.
- *False Negatives (FN)* are the number of instances that are actually positive but are incorrectly predicted as negative by the model.

These metrics are determined by comparing the actual and predicted labels generated by the trained model.

Model evaluation and comparison

Recall was prioritized for the Monitor Performance class when evaluating the performance of each model. This prioritization is because the primary goal is to identify and flag all projects belonging to this class, while being less concerned if some flagged instances designated as Monitor Performance, turn out to be On Track. Table 3 summarizes the performance of different models on our dataset for CPI label prediction, and Table 4 summarizes the performance of different models for SPI prediction.

Table 3: The performance of different models on test data set for CPI label prediction.

	Monitor Performance			On Track		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Decision Tree	0.69	0.69	0.69	0.56	0.57	0.56
Random Forest	0.74	0.85	0.79	0.72	0.57	0.64
XGBoost	0.74	0.77	0.76	0.66	0.62	0.64

Table 4: The performance of different models on test data set for SPI label prediction.

	Monitor Performance			On Track		
	Precision	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Decision Tree	0.72	0.58	0.64	0.51	0.66	0.57
Random Forest	0.80	0.79	0.79	0.69	0.70	0.70
XGBoost	0.76	0.77	0.76	0.65	0.63	0.64

Table 3 and Table 4 show that for both CPI and SPI, the Random Forest model outperforms other models, reasonable in both Recall and Precision metrics for the 'Monitor Performance' class. XGBoost also demonstrates reasonable performance, with its results closely aligning with those of the Random Forest Model.

The models were also compared in terms of computational costs and the results are summarized in Table 5. The models were trained and tested on a Windows laptop with an 11th Gen Intel(R) Core(TM) i5-1145G7 processor running at 2.60GHz and 16.0 GB of RAM. Table 5 shows that XGBoost is the most expensive one during the training phase and Random Forest is the most expensive one during the test phase.

Table 5: The computational cost of different models.

	Train time (s)	Test Time (s)
Decision Tree	0.003	0.0004
Random Forest	0.266	0.022
XGBoost	0.495	0.008

Performance Factor Importance

This initiative aimed to investigate how the trained Random Forest model classifies each project and determine which project input/performance indicator holds the highest contribution or importance in the prediction of the Random Forest model. This approach is known as feature importance (Menze, et al., 2009) in the ML community. Figure 3 shows the top ten project indicators in terms of importance in the prediction of Random Forest Model.

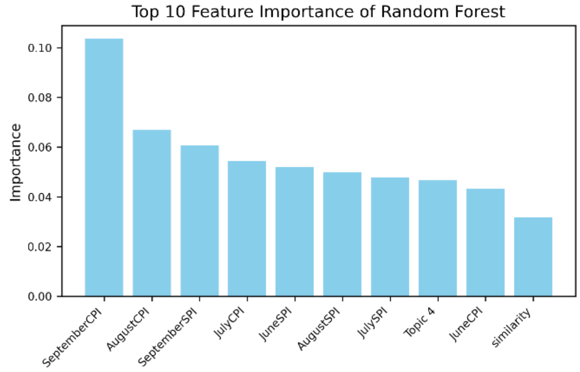


Figure 3: Importance of Performance Indicators for CPI model.

Figure 4 displays the top 10 project inputs in terms of their importance and impact on the Random Forest ML model for predicting the CPI label. This assessment revealed that mid-year cost performance holds the highest significance. Schedule performance had a comparatively lower significance, which can be attributed to the flexibility in the delivery approach, allowing for adjustments to recover the schedule.

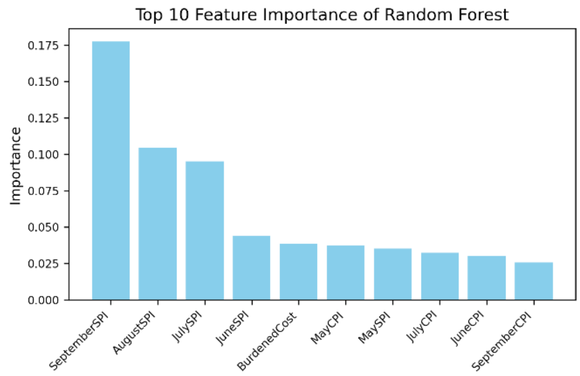


Figure 4: Importance of Performance Indicators for SPI model.

Figure 4 illustrates the top 10 project inputs in terms of their importance and impact on the Random Forest ML model for predicting the SPI label. Results showed that mid-year SPI is more important than other project inputs. Following September, August, and July SPI, Burdened Cost and mid-year CPI emerge with the highest importance.

Methodology for Implementation

Two separate models were designed and optimized, one for CPI and one for SPI, to predict the project status at the end of the fiscal year, utilizing information available at mid-year. Both the CPI and SPI models independently labeled each project, resulting in two labels for each project: one based on outcomes from the Cost model and another from the Schedule model (refer to Table 6). Subsequently, the results were consolidated into a single outcome per project for review by the project team and decision-makers. The classification methodology used is outlined in Table 6.

Table 6: Labelling methodology based on the outcome of CPI and SPI models.

CPI Model Label	SPI Model Label	Resultant
On Track	On Track	Project on Track
Monitor	On Track	Monitor Cost Performance
On Track	Monitor	Monitor Schedule Performance
Monitor	Monitor	Project at Risk

Conclusions and Discussion

In this project, multiple ML models were developed to predict the project's status at the end of the fiscal year based on the information available at mid-year. The results are promising, and testing the models indicates that we can potentially identify projects that need monitoring at mid-year. However, it is important to acknowledge that, due to the complexity of factors affecting project management, achieving perfect precision and recall with such an application is inherently impossible. This highlights the need for a balanced approach in interpreting and applying the model's predictions in practical scenarios.

The ML model designed under this research initiative, provided advanced insights into project performance ahead of project completion. This additional insight provided management with additional information to determine how to proceed with the project, and a firm foundational technique that other organizations can use to leverage ML. Once implemented, organizations would be able to weigh project benefits while supporting the decision-making process throughout the gating and sanctioning processes while focusing on the highest value projects with predictable outcomes. Using a ML model as part of enabling a more powerful framework for monitoring project performance, the following prospective outcomes can be inferred:

- Stabilizing a central dataverse to house all project performance data is critical to enabling a ML model for projects of any scale or complexity.
- Ability for a more robust determination on which projects should proceed during the gating and sanctioning process from a fulsome understanding of the true as-is condition based on current performance and the potential for a project's future success.
- Project teams will be able to more readily see and understand where and how their optimism bias is impacting the overall performance of the project.
- Actionable forecasted intelligence on where the project team's efforts should be focused to correct the predicted outcome.
- Business decisions can be made with greater confidence while projects are evaluated more closely earlier on.

These findings from this research furthered the idea that organizations can use ML to support project management activities with project outcomes and can be predicted after capturing six months worth of project performance data. Further, cost and schedule performance indices were viable leading indicators in predicting project success when utilizing the earned value management system.

Despite the simplicity of this model, it is anticipated and expected that the subsequent steps and implementations for leveraging ML within an organization would be dramatic and exponential. The foundational needs and designs required to establish this simplistic model will enable many future avenues to explore to further enhance predictable project execution.

Acknowledgments

The authors would like to acknowledge the support during the execution building, testing and documenting the results from the ML model from Ashlea Colton, Catalina Anghel, Larkin Mosscrop, Jesse Rhodes and Bryan Simpson.

References

Bou Hatoum, M., Nassereddine, H., Musick, S., & El Jazzar, M. (2023). Investigation of PESTEL Factors Driving Change in Capital Project Organizations. *Frontiers in Built Environment*, 1207564.

Chandrasekaran, S., Kalidas, M. S., Nel, G., Parbhoo, P., Rao, A., & Srikanthan, S. (2021). Capital Projects 5.0: Reimagining capital project delivery. *McKinsey and Company, Oprations Practice*.

Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of Applied Science and Technology Trends*, 20-28.

Chen, T., & Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *roceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, (pp. 785-794).

Costello, K. (2019). Gartner survey shows 37 percent of organizations have implemented AI in some form. *https://www. gartner. com/en/newsroom/press-releases/2019-01-21- gartner-survey-shows-37-percentof-organizations-have, last accessed on*, 2022.

Gupta, V. K., Gupta, A., Kumar, D., & Sardana, A. (2021). Prediction of COVID-19 confirmed, death, and cured cases in India using random forest model. *Big Data Mining and Analytics*, 116-123.

Hatoum, M. B., Piskernik, M., & Nassereddine, H. (2023). A holistic framework for the implementation of big data throughout a construction project lifecycle. *Proceedings of the 37th International Symposium on Automation and Robotics in Construction (ISARC)*, (pp. 1299-1306). Kitakyshu, Japan.

Menze, B. H., Kelm, B. M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., & Hamprecht, F. A. (2009). A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC bioinformatics*, 1-16.

Müller, A. C., & Guido, S. (2016). *Introduction to machine learning with Python: a guide for data scientists.* O'Reilly Media, Inc.

Nieto-Rodriguez, A., & Viana Vargas, R. (2023). How AI will transform project management. *Harvard Business Review*.

Parmar, A., Katariya, R., & Patel, V. (2019). A review on random forest: An ensemble classifier. *International conference on intelligent data communication technologies and internet of things (ICICI) 2018* (pp. 758-763). Springer.

Tong, Z., & Zhang, H. (2016). A text mining research based on LDA topic modelling. *International conference on computer science, engineering and information technology*, (pp. 201-210).

SCALABILITY IN BUILDING COMPONENT DATA ANNOTATION: ENHANCING FAÇADE MATERIAL CLASSIFICATION WITH SYNTHETIC DATA

Josie Harrison[1], Alexander Hollberg[1], and Yinan Yu[1]
[1]Chalmers University of Technology, Gothenburg, Sweden

Abstract

Computer vision models trained on Google Street View images can create material cadastres. However, current approaches need manually annotated datasets that are difficult to obtain and often have class imbalance. To address these challenges, this paper fine-tuned a Swin Transformer model on a synthetic dataset generated with OpenAI's DALL-E and compared the performance to a similar manually annotated dataset. Although manual annotation remains the gold standard, the synthetic dataset performance demonstrates a reasonable alternative. The findings will ease annotation needed to develop material cadastres, offering architects insights into opportunities for material reuse, thus contributing to the reduction of demolition waste.

Introduction

Demolition waste from construction and renovation activity is a growing problem internationally, making up 25%-30% of all waste generated in the EU (Anastasiou et al., 2014). There is a general consensus that current practices have room for improvement for diverting and recovering materials from demolition waste (Kabirifar et al., 2020).

Recovering demolition waste is not a common activity in many countries because there remain limitations that practitioners currently face. Not having insight into the availability of which specific materials will be available and at what time is often identified as a top barrier (Akanbi et al., 2020). The planning of projects that could use recovered materials typically begin several years before the actual construction begins. A material cadastre that maps detailed building material information for every building at the scale of a country with an estimate of when the materials could become available would significantly alleviate the current time window bottleneck.

Because of the significant potential of this solution, several researchers have turned their attention to developing city-scale material cadastres with computer vision models trained on images of building exteriors (Raghu et al., 2023; Arbabi et al., 2022). A top pain-point in developing these models is acquiring the training data necessary for machine learning (ML). For deep learning models, some practitioners say that the size of training data should be 10x the number of weights in a network (Baum and Haussler, 1988)—often leading to six figure digits and more.

Additionally, because the type of material is not known before the GPS coordinates are requested to collect training data, this process can lead to some materials (i.e. classes) having higher counts than other classes—also known as class imbalance. ML practitioners try to avoid class imbalance because a ML model will place importance on a class that has a higher chance of occurring in the training data. For example, if brick is the most common material in the training data, then the model is likely to predict brick when deployed in a real-world situation. This effect can be negligible if brick actually is the most common material in a city, but it is preferable to assume equal class counts since the true distribution of materials in a city is unknown.

The scale achieved by previous studies has been at the scale of a city; however, the blue sky vision for a tool like this would be at the scale of a country. This would mimic current practices of sourcing materials while staying within the bounds of maximum distance for a recovered material to travel before its sustainability benefits become negligible (Ginga et al., 2020). There is some truth to the logic that if a model works at the scale of a city then it should have similar performance at the scale of a country; however, previous studies have identified that several problems can arise when scaling an image classification task (Maggiori et al., 2017; Perronnin et al., 2010; Hendrycks et al., 2022). For example, wood siding can take on different colours and textures in different cities since it is often a regional material, but an image classification model may struggle to perform well if it hasn't been trained on wood with a particular colour. Because it's impossible to foresee all potential quirks that may arise with a given use case, it's important to move beyond proof-of-concepts at the scale of a city and demonstrate that the model maintains reasonable performance at the scale of a country.

Furthermore, it would be desirable to test an image classification model's performance on residential and office interiors since this layer of a building typically has a high rate of change, resulting in a potentially high yield of reusable renovation waste. However, obtaining large and diverse datasets of building interiors can be challenging because of privacy concerns, data access limitations, and the labor-intensive process of data collection and annotation. As a result, public datasets for this task are limited in size, diversity, and representativeness, which may not fully capture the complexity and variability of real-world interior environments. Addressing the scarcity of this type of dataset could unlock an important layer in the quest for creating country-wide material cadastres.

Therefore, the aim of this paper is to investigate the potential of using synthetically generated images to augment and extend previous research done to classify façade materials in Google Street View (GSV) images. More specifically, the research questions that this paper addresses are:

1. What are the impacts on an image classifier's performance when correcting the class imbalance of a manually annotated dataset with synthetic data?
2. Is the error distribution of an image classifier trained exclusively on synthetic data comparable to a classifier trained on manually annotated data when both are evaluated on a manually annotated test set?
3. Does a model trained and tested only with synthetic data have a similar error distribution to the manually annotated dataset?

The results from these research questions make several contributions to the field of urban mining for material cadastres. To the authors' best knowledge, no previous studies have utilized synthetic images that mimic GSV images for augmenting and extending an image classification model to detect façade material. As a result, we present a novel approach to reduce the amount of time needed to develop a training dataset for the classification of façade materials, as well as provide an indication of the potential to use synthetic images to create datasets that would otherwise be difficult to obtain. Additionally, we present two methods to decrease errors in the current approaches: utilizing a higher resolution model by reducing the problem space and augmenting a manually annotated dataset to correct class imbalance. The workflow developed in this research can be extended to detect façade materials that have no manually annotated training data and the workflow can be used to improve performance of the current state-of-the-art for this specific use case.

Background

Material stock datasets

In 2023, researchers from ETH Zurich published the first urban-scale ground truth dataset—hereon referred to as the Urban Resource Cadastre (URC) dataset—to detect façade materials in Tokyo, New York City, and Zurich (Raghu et al., 2023). The researchers collected this data by first identifying GPS coordinates of interest, requesting the corresponding GSV image for these GPS coordinates, and manually annotating all resulting images. This dataset contains 971 annotated 400x600 pixel images with the labels 'brick', 'stucco', 'rustication', 'metal', 'siding', 'wood', 'null' (for images with no façade), and 'other' (for façade material that did not belong to any of the other labels). However, the authors noted that collecting this type of dataset can easily become a time-consuming process.

Generating images for training data

In recent years, there were significant developments with synthetic images generated for the purpose of training downstream models. Research in this area is often fueled by use cases where acquiring more data is time/cost prohibitive, it is necessary to mask personally identifiable information for privacy, or it is not possible to collect data at all (Man and Chahl, 2022). There exists a wide variety of generative text-to-image models; however, it has recently been accepted that diffusion models outperform previous text-to-image generative models. Inspired by non-equilibrium thermodynamics, diffusion models replicate the diffusion process observed in physics by teaching a deep learning model to add and reverse noise in images (Sohl-Dickstein et al., 2015).

The lack of diversity within a synthetic dataset is a known limitation associated with generating images for training models. To overcome this, He et al. (2023) used "language enhancement" to add variety to the prompts given to the generative model while also using a filter to discard images that didn't resemble the target image. The "language enhancement" involved using a simple keyword-to-sentence Natural Language Processing (NLP) model based on the T5 model (He et al., 2023) to make sentences from keywords. This method to diversify the generated output is prone to straying far from the desired image, however. To counteract this effect, He et al. (2023) proposed a "CLIP Filter" that leveraged CLIP zero-shot classification confidence. However, the authors aimed to create a fully automated process, which could be excessive for use cases where there is enough value in just generating synthetic images alone—which is this paper's use case.

Façade materials with high reusability potential

At this point, it becomes apparent that we should justify the selection of specific façade materials for the synthetic generation of images. For example, there exist buildings built with ETFE plastic bubbles, but since these projects are rare, their inclusion in a training dataset would confuse a ML model. While research in the area of the reusability potential of specific building materials is rather sparse, Icibaci (2019) and Iacovidou and Purnell (2016) identified stone with lime-based mortar, curtain walls, and concrete panels as materials with high reusability potential—hereon referred to as the High Reusability Potential (HRP) labels.

Image classification

Originally developed for NLP tasks (Vaswani et al., 2017), vision transformers (ViTs) are deep learning models that take an image separated into patches (with its positional information saved), encode these patches into a memory, and compare that memory with a target value (Dosovitskiy et al., 2020). The memory is created with 'attention', which maps all the patches to each other to establish relative importance between them all (Vaswani et al., 2017). The Swin Transformer is a variant of the ViT methodology, which has shown to outperform many other models on image classification tasks while also being conscious of computational efficiency (Liu et al., 2021).

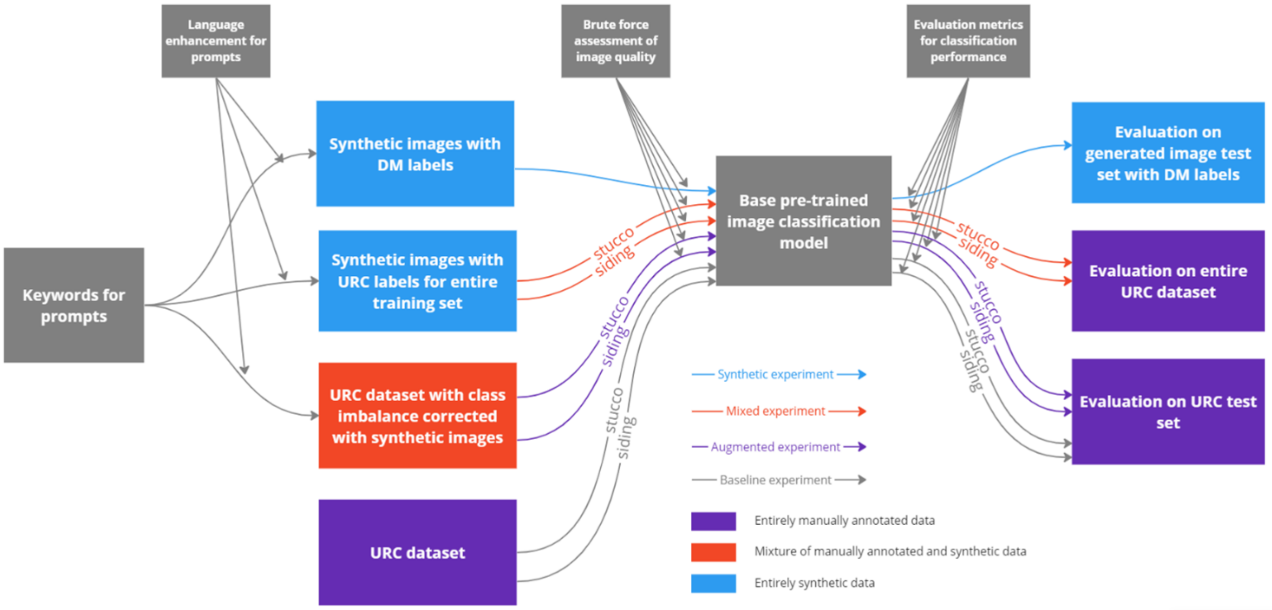


Figure 1: Workflow showing all four experiments: baseline, augmented, mixed, synthetic with sub-experiments for stucco and siding.

Experiment

Workflow

The workflow shown in Figure 1 assembled all materials and methods together into four experiments. There are three main experiments compared to a baseline experiment: augmented, mixed, and synthetic. For each experiment, a pre-trained image classification model was fine-tuned with the respective dataset, with synthetic images first passing through a brute-force image quality evaluation. After model training, the predictions were evaluated on the experiment's test dataset. The synthetic experiment's test set is entirely synthetic while the other three experiments use a manually annotated test set.

Materials

Drawing on previous research, four datasets (baseline, augmented, mixed, and synthetic) were collected for the purposes of training an image classification model to replicate previous work and to also train a model on unseen labels. The baseline, augmented, and mixed datasets used the URC labels, while the synthetic dataset used the HRP labels. For the augmented, mixed, and synthetic experiments (shown in Figure 2), varying amounts of 512x512 pixel synthetic images were generated. This resolution was suitable for image classification models that can be trained on consumer-grade hardware.

For the baseline, augmented, and mixed experiments, the number of labels was reduced to three to create sub-experiments: null, other, and one label of interest in order to reduce the parameter space of the image classification model. 'Stucco' and 'siding' were isolated as two labels of interest because 'stucco' had a similar class distribution as the 'null' and 'other' labels, and because 'siding' had a high class imbalance in the original dataset. Therefore, there were two sub-experiments for the baseline, augmented, and mixed eyxperiments using two sets of labels:"null, other, stucco" and "null, other, siding". As mentioned earlier, the URC dataset came with class imbalance, so the augmented experiment corrected this class imbalance with synthetic images. The mixed experiment used only synthetic images for training while leaving the entire URC dataset as the test dataset. Finally, the synthetic experiment trained and tested entirely on synthetic images using the HRP labels: 'stone', 'curtain wall', and 'concrete panels'. The synthetic experiment dataset size was made to match the mixed stucco experiment's dataset set to maintain comparability.

Methods

There are several open source versions of diffusion models that are capable of creating synthetic images; however, the effort required to get them into an operational state can often be time-prohibitive while also requiring intensive hardware. For that reason, OpenAI's DALL-E 2 model was chosen to generate images because it is the most effective to leverage with just a simple API request. A portmanteau of the famous surrealist artist Salvador Dalí and the Pixar character WALL-E, DALL-E takes in a user defined text prompt to create highly realistic images. The DALL-E 2 text-to-image generation model was trained on 650 million images (Ramesh et al., 2022); however, OpenAI does not disclose the source of training data for any of their models (Jiang et al., 2023). To reduce this dependency on an external organization, it would be ideal to develop an open source model in-house, however.

A limitation of generating synthetic GSV images was the inability for DALL-E to create believable images with more than one façade material. Other researchers found similar difficulties when requesting more than three components with descriptions about shape or colours (Marcus et al., 2022). Additionally, it was possible to utilize

Figure 2: Class distributions for the different datasets.

a higher resolution image classification model by reducing the label set to one material of interest combined with the ‘null’ and ‘other’ labels. For these reasons, it was decided to generate images with one label. Since this paper’s use case would not require predicting several labels in real-time, it is proposed that the implementation would require several models with each model focusing on one label (as well as the non-buildings and other materials).

As mentioned earlier, there was enough value in creating the synthetic images alone to not utilize a fully automated filtering process. Therefore, a brute force version of this process was suitable. This brute force method shown in Figure 3 involved generating a collection of prompts from keywords, generating a few hundred images by randomly sampling a prompt from the collection, manually removing irrelevant images from the final dataset, and taking note of the few prompts that were likely to produce the best images. The remaining dataset was generated using this final collection of ‘high batting average’ prompts.

To achieve a classification model that can perform at the scale of a country, it is desirable to encourage diversity in the synthetic images by injecting additional keywords into the prompt. For this experiment, the facade material label was given possible synonyms, a time period when the façade material was likely to be common, and cities where the facade material can be found. However, these keywords could be extended to any aspect of a building, such as type, size, or neighbourhood density. For example, ‘siding’ was assigned the synonyms ’shiplap’, ’feather edge’, ’fiber cement’ and the time period ’20th century’. For all labels in the augmented and mixed experiments, the original URC dataset cities (New York City, Zurich, and Tokyo) were used as keywords. For the synthetic experiments, the cities ’Vancouver’, ’San Francisco’, and ’Amsterdam’ were chosen arbitrarily but would benefit from additional cities in future experiments.

For the image classification task, it was decided to use a Swin Transformer model pre-trained on the ImageNet-21k dataset to maximize comparability to previous studies and to leverage a well-known image classification model. A pre-trained image classification model is trained on a large generic dataset with the intention to later ’fine-tune’ the ’head’ of the model on tailored datasets. This method generally improves accuracy and reduces the need for a large tailored dataset. The ImageNet-21k dataset is commonly used for pre-training because it is 14 million images with 21,000 labels and contains a variety of natural images. Raghu et al. (2023) achieved a macro-averaged F1 score of 0.93 by applying the pre-trained Swin Transformer v2 model to the URC dataset. However, the authors used the model version with a resolution of 192 pixels, which was not adequate for the synthetic image dataset. It was seen that the synthetic images struggled with creating the same texture detail as real images, and this data loss was compounded when used with a low-resolution model. This is akin to taking a photocopy of a photocopy. Therefore, the experiments in this study use the Swin Transformer v2 model with 384 pixel resolution.

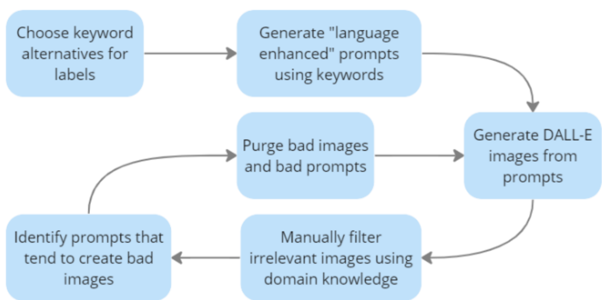


Figure 3: Brute force method to filter irrelevant synthetic images from accepted images.

To provide a quantitative and objective measurement of the image classification performance, it was decided to calculate the weighted F1 score, precision, recall, as well as create a receiver operating characteristic (ROC) graph and confusion matrices for each experiment. The confusion matrix is a grid of all target labels with the y-axis representing the true test values and the x-axis representing the predicted test values. In this way, it’s possible to index how many samples were predicted for a specific label: the first cell in the grid shows the number of samples predicted as 'null' were correctly classified. It is not possible to decide which dataset performed the best by looking at these confusion matrices, however, so quantitative metrics were calculated: precision and recall.

$$precision = \frac{tp}{tp + fp} \tag{1}$$

$$recall = \frac{tp}{tp + fn} \tag{2}$$

Precision (1) refers to the likelihood that a prediction will be correct, while recall (2) refers to the likelihood that all samples will be correctly classified. In its basic binary form, the precision is the true positives (tp) divided by all samples marked as positive ($tp + fp$) (2). Similarly, the recall formula (2) divides the true positives (tp) by the sum of true positives and false negatives ($tp + fn$). The F1 score (3) provides a harmonic mean of both precision and recall.

$$F = \frac{2 * precision * recall}{precision + recall} \tag{3}$$

When working with multi-class problems, it is necessary to choose an averaging mechanism for the F1 score, precision, and recall metrics. The 'weighted' average is recommended for imbalanced classes because it assigns different weights to each class based on their prevalence in the dataset. This prevents the dominant class from strongly influencing these metrics.

Lastly, the ROC curve graph was created to visually compare the performance of all four experiments. In this study's context, the ROC curve can be used to visualize the characteristics of the performance; at what threshold between the true positive rate and false positive rate the model becomes uncertain about its predictions.

Discussion and result analysis

Image generation

As with many NLP models, the generated prompts contained a fair amount of hallucination; however, these nonsensical prompts surprisingly created some of the best images. For example, a successful prompt for the 'stone' label was, "The construction of building made of stone is in the medium of the Artistic Gymnastics." It was valuable to refrain from editorial oversight of the initial prompts to allow for spontaneity in the generated images. Because of this, the resulting datasets had more diversity than if one prompt created all the synthetic images.

Some labels had a higher number of generated images flagged as irrelevant than other labels, with an average of 24% of all images being flagged as irrelevant. In most cases, the irrelevant images featured materials different from the target label. For example, the 'stucco' label often created images showing metal, siding, or rustication, which resulted in 45% of all 'stucco' images being flagged as irrelevant. In some cases, the generative image model created images with no building shown at all.

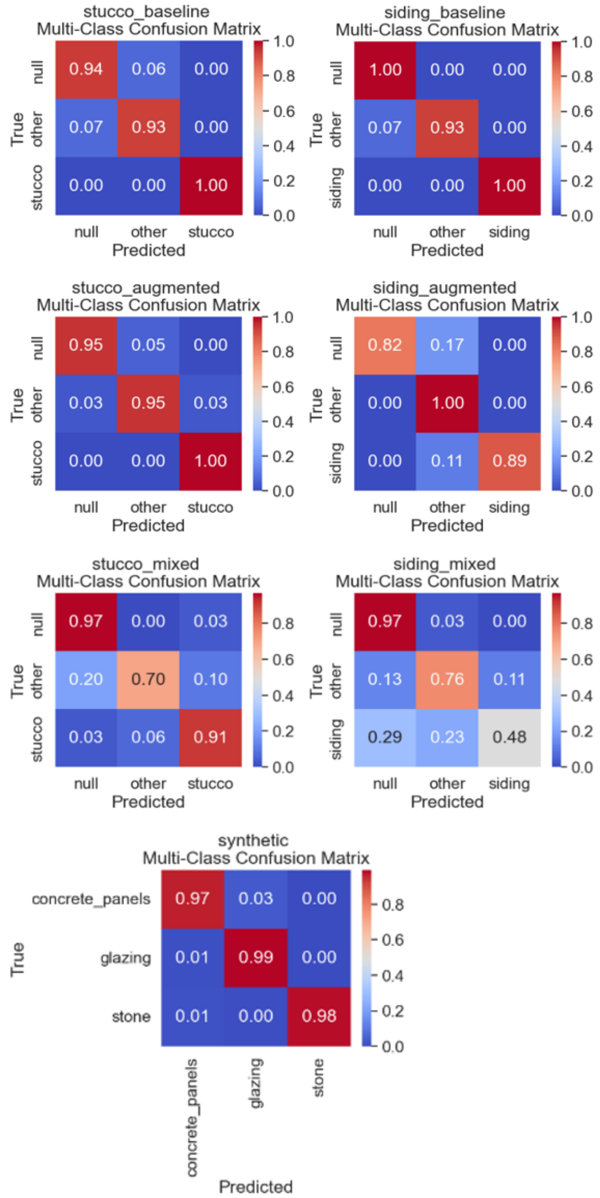


Figure 4: Confusion matrices for all experiments.

Image classification

By reviewing the confusion matrices, ROC curve graph, and bar graph comparing F1 scores, precision, and recall, we are able to achieve a nuanced and holistic evaluation of the image classification results in the context of the research questions.

The confusion matrix compares the number of true labels versus predicted labels. To aid in comparability and interpretability, the numbers of each row in the matrix were normalized to a percentage scale of 0% to 100%. The set of confusion matrices (Figure 4) show which labels perform the worst and best for each experiment and whether any degradation of performance is acceptable. Comparing the augmented stucco experiment to the baseline, it can be seen that the augmented dataset offered a slight advantage for the 'null' and 'other' labels, while maintaining perfect

performance for the 'stucco' label. When using the mixed dataset, there was decrease in performance (100% of true 'stucco' samples predicted as 'stucco' versus 91% of true 'stucco' samples predicted as 'stucco'). Moving to the siding experiment set, the augmented siding experiment had a decrease in performance (as opposed to the augmented stucco experiment). This is likely because the 'siding' augmented class had a higher proportion of synthetic to manually annotated images, which triggers the question that there might be a threshold where the ratio of synthetic to manually annotated images hinders performance. The mixed siding experiment showed further degradation with only 48% of predicted 'siding' samples matching the true 'siding' label. This may have been caused by the size of the mixed siding dataset (248 training samples and 62 testing samples for each class), which was half the size of the mixed stucco dataset (596 training samples and 149 testing samples for each class). Lastly, it is no surprise that the synthetic experiment achieved near perfect results since the model had no way to test its performance on 'in the wild' images. Comparing the synthetic experiment results to the mixed experiment offers insight into the possible ways a model trained and tested on purely synthetic data might degrade when brought into a real data scenario.
The ROC curve graph is helpful in comparing each experiment's discrimination capability between different classes. A model with a curve that hugs closely toward the top left corner suggests high performance in sensitivity and specificity, while the straight, black, dotted line represents the performance of a random classifier. A model with a ROC curve too close to this random classifier line indicates that the model is essentially predicting a class at random.

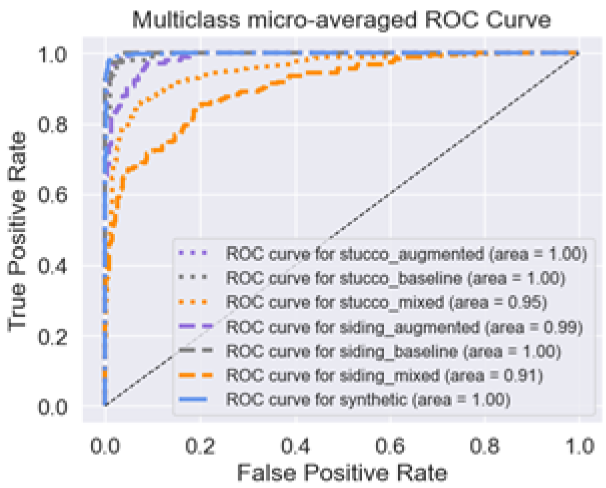


Figure 5: ROC curve graphs comparing the stucco and siding sub-experiments.

From the ROC curves (Figure 5), it can be seen that there are a few experiments that achieved both high sensitivity and specificity across all thresholds, indicated with an Area Under the Curve (AUC) value of 1.00: both baseline experiments, the augmented stucco experiment, and both synthetic experiments. The two mixed experiments achieved AUC values of 0.95 and 0.91 for stucco and siding, respectively, which is generally considered quite a good score. Since those ROC curves are positioned well above the random classifier line, it can be interpreted that these mixed models achieved a meaningful separation between true positives and false positives.
Evaluating the weighted precision, recall, and F1 score of the four experiments provides a global view of the overall performance of each experiment. A higher precision indicates a lower rate of false positives, a higher recall indicates a lower rate of false negatives, while a higher F1 score indicates a balance between precision and recall.
The bar graphs showing the weighted F1 score, precision, and recall (Figure 6) further supports the observations that models trained with some augmentation can increase performance and models trained with synthetic images maintain reasonable results on a manually annotated test dataset. The highest performing experiments were the synthetic experiment and the baseline siding experiment, achieving a value of 0.98 for all three metrics. Although the confusion matrices differed slightly, the baseline stucco and augmented stucco experiments achieved the same score of 0.96 for all three metrics. The augmented siding experiment performed slightly worse than the baseline siding experiment with a value of 0.91 for F1 score and recall (and a value of 0.93 for precision). This further supports the finding that too much augmentation can degrade performance. Although the mixed experiments had a decrease in metrics, the mixed stucco experiment maintained a reasonable performance with the values 0.86, 0.87, 0.86 for F1 score, precision, and recall. The mixed siding experiment showed further degradation with values 0.72, 0.75, 0.74 for F1 score, precision, and recall. However, the training and testing sets were half the size of the mixed stucco experiment, which could have influenced performance.

There were several key findings from this study; some that offer an improvement over the state-of-the-art, and some that provide a nuanced analysis into the trade-offs made when training a model on synthetic images. In terms of overall improvement, this study achieved a higher F1 score, precision, and recall than the state-of-the-art by utilizing a higher resolution model (Swin Transformer v2 at 384 pixel resolution) and reducing the problem space to three labels (null, other, and a label of interest). This study achieved a 0.98 weighted F1 score, 0.98 recall, and 0.98 precision on the baseline siding experiment while the state-of-the-art achieved 0.93 macro F1 score, 0.91 recall, and 0.96 precision (Raghu et al., 2023). For marginal improvements, it was found that augmenting a dataset with 27% synthetic data (for the 'stucco' class) offered a slight improvement over the baseline. This had the opposite effect when using a dataset with 80% synthetic data (for the siding class). As for the possibility to train a model on entirely synthetic data, the performance was generally worse than the baseline, but still a reasonable performance with the mixed stucco experiment capturing 91% of true posi-

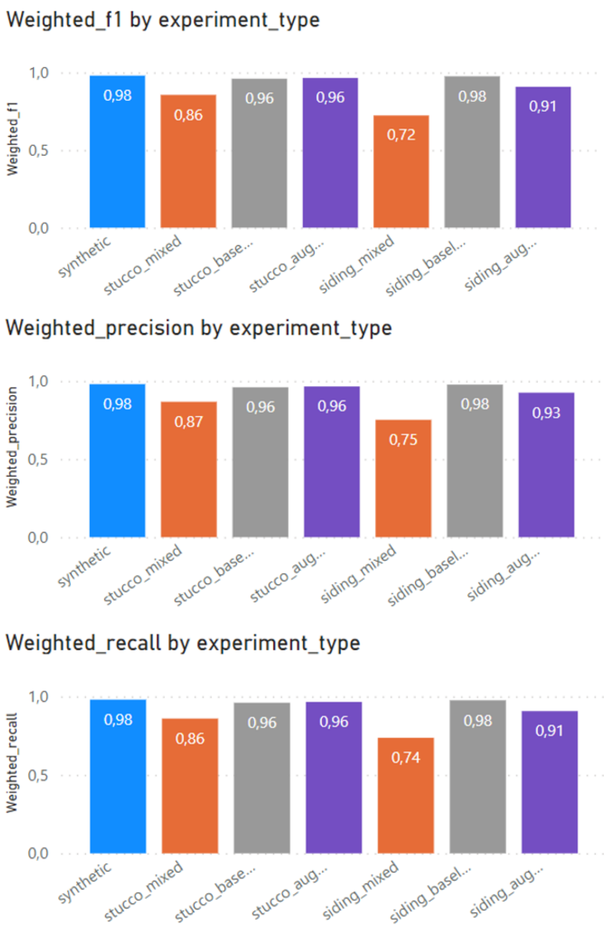


Figure 6: Weighted F1, precision, and recall for all experiments.

tives in the confusion matrix, achieving an AUC value of 0.95, and scoring 0.87 for precision and 0.86 for the F1 score and recall. The results from the mixed siding experiment may have been adversely affected by the size of the dataset (62 samples for test and 248 samples for training); potentially signaling a need to have a larger dataset when training on synthetic images. The near perfect results of the synthetic experiment show that it is possible to train a model on synthetic images for a new label set, but demonstrates the value of testing a model on manually annotated images.

Conclusions

Recent advancements in text-to-image generation have made it feasible to train image classification models on synthetic images, addressing previous challenges of data scarcity and enabling scalability to a national or global level for façade material identification. In this study, we explored the potential benefits and trade-offs that could be made when using varying amounts of synthetic images for training a classification model, as well as proposed an improvement in overall performance with a higher resolution model. The research contributes a novel approach in urban mining for material cadastres by utilizing synthetic images to augment and extend an image classification model for detecting façade materials, offering a potential solution to reduce training dataset development time and address data scarcity challenges while improving current approaches' performance. This, of course, comes with caveats: the text-to-image prompts should trigger diversity in the generated images, the model should be at least 384 pixel resolution, the problem space should be reduced to three labels (null, other, and label of interest), the synthetic training data should be large enough, and it is worthwhile to still manually annotate a test set to develop an intuition for the types of errors that will happen. The results from this study are limited to replicating GSV images for the study's specific set of labels and would not necessarily guarantee reasonable performance for synthetic images of building interiors or niche facade materials (such as ceramic tile cladding). Access to a high-performance GPU (an nVidia GeForce RTX 3070 8GB GPU) enabled utilization of the pre-trained Swin Transformer v2 384 pixel resolution model, which may not be possible to train on a CPU alone. Additionally, there could be privacy concerns associated with remotely compiling building material inventories, particularly regarding historical (materials that should not be moved from their location) or high-security façade materials (materials that may reveal structural vulnerabilities).

These limitations indicate future research that could progress the field of urban mining for material cadastres further. While this experiment achieved good results mimicking GSV images, it would not be safe yet to conclude that any type of building image could be replicated with synthetic images. Therefore, it would be valuable for future work to explore applying the workflow to interior office and residential synthetic images. For the resolution of synthetic images, there was a considerable difference in performance when moving from the 192 pixel model to the 384 pixel model resolution, which begs the question whether generating higher resolution images and training with an even higher resolution model may increase performance further. When generating the synthetic images, a brute-force method was used to separate irrelevant images from the training set because it was suitable for this use case; however, an automatic filtering method would greatly reduce the time needed to prune the output of the generative model. Further studies using synthetic images could benefit material cadastre research by unlocking building interior datasets, by increasing performance of image classification models even more, and by reducing the time needed to filter synthetic images. Developing a robust image classification model that has proven reliable performance in detecting building materials in a wide variety of contexts, scales, and building layers would be an important milestone in the goal to develop country-wide material cadastres for reusable demolition and renovation waste.

Acknowledgments

This project was financially supported by the Swedish Foundation for Strategic Research.

References

Akanbi, L. A., Oyedele, A. O., Oyedele, L. O., and Salami, R. O. (2020). Deep learning model for demolition waste prediction in a circular economy. Journal of Cleaner Production, 274:122843.

Anastasiou, E., Georgiadis Filikas, K., and Stefanidou, M. (2014). Utilization of fine recycled aggregates in concrete with fly ash and steel slag. Construction and Building Materials, 50:154–161.

Arbabi, H., Lanau, M., Li, X., Meyers, G., Dai, M., Mayfield, M., and Densley Tingley, D. (2022). A scalable data collection, characterization, and accounting framework for urban material stocks. Journal of Industrial Ecology, 26(1):58–71.

Baum, E. and Haussler, D. (1988). What size net gives valid generalization? In Advances in Neural Information Processing Systems, volume 1. Morgan-Kaufmann.

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., and Houlsby, N. (2020). An image is worth 16x16 words: Transformers for image recognition at scale. ICLR.

Ginga, C. P., Ongpeng, J. M. C., and Daly, M. K. M. (2020). Circular economy on construction and demolition waste: A literature review on material recovery and production. Materials, 13(13):2970.

He, R., Sun, S., Yu, X., Xue, C., Zhang, W., Torr, P., Bai, S., and Qi, X. (2023). Is synthetic data from generative models ready for image recognition? arXiv preprint arXiv:2210.07574.

Hendrycks, D., Basart, S., Mazeika, M., Zou, A., Kwon, J., Mostajabi, M., Steinhardt, J., and Song, D. (2022). Scaling out-of-distribution detection for real-world settings. ICML, abs/1911.11132.

Iacovidou, E. and Purnell, P. (2016). Mining the physical infrastructure: Opportunities, barriers and interventions in promoting structural components reuse. Science of The Total Environment, 557–558:791–807.

Icibaci, L. (2019). Re-use of building products in the netherlands: The development of a metabolism based assessment approach. A+BE | Architecture and the Built Environment, page 1–422.

Jiang, H. H., Brown, L., Cheng, J., Khan, M., Gupta, A., Workman, D., Hanna, A., Flowers, J., and Gebru, T. (2023). Ai art and its impact on artists. In Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society, AIES '23, page 363–374, New York, NY, USA. Association for Computing Machinery.

Kabirifar, K., Mojtahedi, M., Wang, C., and Tam, V. W. Y. (2020). Construction and demolition waste management contributing factors coupled with reduce, reuse, and recycle strategies for effective waste management: A review. Journal of Cleaner Production, 263:121265.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., and Guo, B. (2021). Swin transformer: Hierarchical vision transformer using shifted windows. Proceedings of the IEEE/CVF international conference on computer vision, abs/2103.14030.

Maggiori, E., Tarabalka, Y., Charpiat, G., and Alliez, P. (2017). Convolutional neural networks for large-scale remote-sensing image classification. IEEE Transactions on Geoscience and Remote Sensing, 55(2):645–657.

Man, K. and Chahl, J. (2022). A review of synthetic image data and its use in computer vision. Journal of Imaging, 8(1111):310.

Marcus, G., Davis, E., and Aaronson, S. (2022). A very preliminary analysis of dall-e 2. arXiv preprint arXiv:2204.13807.

Perronnin, F., Sánchez, J., and Mensink, T. (2010). Improving the fisher kernel for large-scale image classification. In Daniilidis, K., Maragos, P., and Paragios, N., editors, ECCV 2010, page 143–156. Springer.

Raghu, D., Bucher, M. J. J., and De Wolf, C. (2023). Towards a 'resource cadastre' for a circular economy – urban-scale building material detection using street view imagery and computer vision. Resources, Conservation and Recycling, 198:107140.

Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M. (2022). Hierarchical text-conditional image generation with clip latents. arXiv e-prints. `http://arxiv.org/abs/2204.06125`.

Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., and Ganguli, S. (2015). Deep unsupervised learning using nonequilibrium thermodynamics. International conference on machine learning, pages 2256–2265.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems 30 (NIPS 2017), volume 30.

DEVELOPING A BUILDING SIMULATION IDENTITY CARD FOR ENHANCED SAFETY AND COLLABORATION IN EMERGENCY EVACUATION SIMULATION

Muhammad Farhan Jahangir[1], Aliakbar Kamari[1], and Carl Peter Leslie Schultz[2]
[1] Department of Civil and Architectural Engineering, Aarhus University, Aarhus, Denmark
[2] Department of Electrical and Computer Engineering, Aarhus University, Aarhus, Denmark

Abstract

Ensuring that a building can be safely evacuated in emergencies necessitates using various specialized simulators for prediction, including earthquake propagation, safe evacuation route planning, etc. However, there is no common vocabulary nor standardized approach for integrating diverse building simulation models, hindering their direct integration in an unified collaborative simulation. This is a crucial limitation because no single simulator alone can accurately predict evacuation scenarios in the building industry. We propose the concept of Building Simulation Identity Cards (BSIC) that characterizes and integrates properties of specialized simulation models and demonstrates its role in collaborative simulations for emergency evacuation.

Keywords: Collaborative Simulations, Building Simulation Identity Cards, BIM, Emergency Evacuation

Introduction

Disasters, whether resulting from human activities or natural phenomena, pose a significant risk to life and property, even when they can be anticipated. Emergency evacuation, as a primary safety measure, involves the organized movement of people from disaster areas to safe zones. This practice has been extensively studied and implemented to mitigate the devastating consequences of such events (Yueming & Deyun, 2008).

Emergency evacuation simulations in buildings involve technological innovation, safety engineering, and collaborative planning. As our urban landscapes continue to expand and undergo transformation, safety and resilience in built environments become increasingly important. Research in the field of effective emergency evacuation gained momentum in the 1980s after nuclear incidents at Three Mile Island and Chernobyl (Battegazzorre et al., 2021; Johnson Jr & Zeigler, 1986). These incidents served as evidence of the inability of many emergency authorities to handle such catastrophic events. The subsequent development of specialized simulators has become indispensable for forecasting evacuation scenarios, encapsulating essential information such as geometry, behavior, occupant characteristics, environmental conditions, assessment of structural integrity under seismic conditions, identification of potential challenges through physical simulations, and estimation of evacuation time, among others.

However, despite the availability of powerful, specialized simulators, there is not a single simulator that can account for the highly diverse range of scenarios that users may need to make predictions about. For example, consider a prediction scenario for an earthquake event. An evacuation planning simulator alone often does not consider elements such as structural failure resulting from earthquake loading or considerations of lighting, electricity, weather, and time of day.

In response to these challenges, our study selectively focuses on specific simulations, including: (a) earthquake propagation within the building, (b) development of secure evacuation routes, (c) geometric modeling of the building structure, (d) assessment of structural integrity of the building under seismic conditions, (e) identification of potential challenges in evacuation through physical simulations, (f) estimation of evacuation time, and (g) analysis of occupant movement pattern to ensure the effectiveness of results.

Addressing the aforementioned challenge involves integrating specialized simulators effectively within a collaborative simulation, or so-called co-sim (Gomes, Thule, Broman, et al., 2018; Hansen et al., 2024; Thule et al., 2019) framework to ensure the validity, reliability, and formal verifiability of co-simulation results.

Toward achieving a comprehensive solution, our study introduces a novel concept, the Building Simulation Identity Cards (BSIC) which is aimed at addressing inconsistencies during the integration of simulation models, ensuring that different modules or components can interact correctly, proposed as a comprehensive research agenda and vision for the community. In this paper, we will introduce and outline the development of the BSIC concept, drawing inspiration from key technological initiatives in other industries such as manufacturing, automobile, and aviation, where similar approaches have proven successful in enhancing operational efficiency and interoperability between simulation models.

The process for the development of BSIC involves the creation of standardized ontologies and meta-models to characterize building simulators, which will incorporate available information regarding the simulation model across various aspect (Aßmann et al., 2006; Mahdavi et al., 2023). It involves creating standardized interfaces, frameworks, and algorithms to facilitate their integration into collaborative simulations (Hansen et al., 2022) . An example of this approach is the Model Identity Card (MIC) that has been developed for characterizing simulators in the automotive industry (Sirin et al., 2015).

Attached to the emergency evacuation planning domain, The BSIC concept will incorporate Functional Mockup Interface (FMI) (Schwan et al., 2017) and co-simulation theory and frameworks. This ensures the standardized storage and sharing of simulation model information, including interfaces, input/output parameters, methods, and usage data. This integration will enhance collaboration and knowledge transfer between simulation models, promoting interoperability and mitigating existing barriers.

BSIC will serve as a novel BIM-based solution for harmonizing the properties and attributes of various simulation models employed in the context of emergency evacuation. BSICs act as standardized profiles for simulation models integrated with BIM standards (Eastman, 2011). In this framework, the present study will address the following research question:

RQ: How can the integration of two independent simulation models be systematically achieved to facilitate the co-simulation process in the building sector for emergency evacuation planning in case of earthquake?

By leveraging BSIC, we have a potential avenue for improved collaboration and communication among stakeholders, researchers, and practitioners involved in emergency evacuation simulations. In this paper, we introduce the concept of BSIC as a solution to the challenge of the integration of specialized simulators including earthquake analysis and emergency evacuation simulator for predictive simulations in case of emergency evacuation planning from buildings.

Related Work

There are several research studies focused on emergency evacuation simulation within buildings. (Gelenbe & Desmet, 2013) used graph theory to identify critical positions for emergency evacuations in buildings. In another study. (Forssberg et al., 2019) aimed to improve the reliability of evacuation design assessments by analyzing the variability in pre-movement times during building evacuations. (Mirahadi & McCabe, 2021) proposed the utilization of Dijkstra's algorithm to identify the most secure path-planning strategy for evacuations, particularly in the context of fire emergencies. The model establishes a risk factor for each compartment, taking into account factors such as fire location and potential blockages. By employing a modified version of Dijkstra's algorithm, the model calculates the path with the lowest risk. (Kirby et al., 2015) in their study, they used AnyLogic software to create an evacuation simulation model, demonstrating the importance of data on optimal staff, materials, space, and time resources required for evacuation planning. Additionally, (Chu et al., 2019) in their study investigated the emergency evacuation simulation and management optimization in urban residential communities. A framework was developed for data acquisition, scenario development, evacuation simulation, and emergency management analysis.

In parallel, collaborative simulations (co-sim) have emerged as a topic of increasing interest in the building industry. (Alfalouji et al., 2023) investigated the integration of multiple tools for co-simulation of buildings and smart energy systems. This method has significant potential in advancing building envelopes due to its ability to accommodate complex control strategies and sequences.

Framework for co-simulations for building

By pursuing our aim in this paper, the Building Simulation Identity Card (BSIC) concept is developed to overcome challenges in integrating and orchestrating specialized simulators in a collaborative simulation framework. In this light, we propose a roadmap for BSIC development inspired by major technological initiatives, i.e MIC, to characterize simulators in the automotive industry.

In order to address occupant safety challenges in buildings, a design and engineering team develops a comprehensive digital model of the building, very often incorporating BIM-based programs such as Revit. Subsequently, the structural engineer analyzes the building for potential failure and risk under earthquake loading (Smith, 2016) using the BIM model in a structural integrity simulator. The earthquake loading is sourced from another plateform, e.g., PEER Ground motion data base (PEER-Center, 2013). The data from this simulator needs to be integrated into the evacuation planning simulator. The data may include information about the occupants, output from the structural integrity simulator, and available evacuation routes. Taken together, the purpose of conducting these simulations is to plan the evacuation of people from the building in the event of an earthquake. Yet, neither of the two aforementioned simulators can independently execute such a scenario. Therefore, running a co-simulation is imperative to obtain the desired results. However, the lack of a common vocabulary for integrating different simulators poses a high risk of simulation model integration failure. The framework for co-simulations for emergency evacuation planning from buildings is represented in Figure 1. The evacuation planning engineer uses BSIC to gather the structural integrity simulator output, which serves as input for the emergency evacuation simulator. This process involves manual implementation, where a structural engineer formulates the BSIC for the structural integrity simulator. Subsequently, an evacuation planning engineer utilizes this BSIC as input for the emergency evacuation simulator.

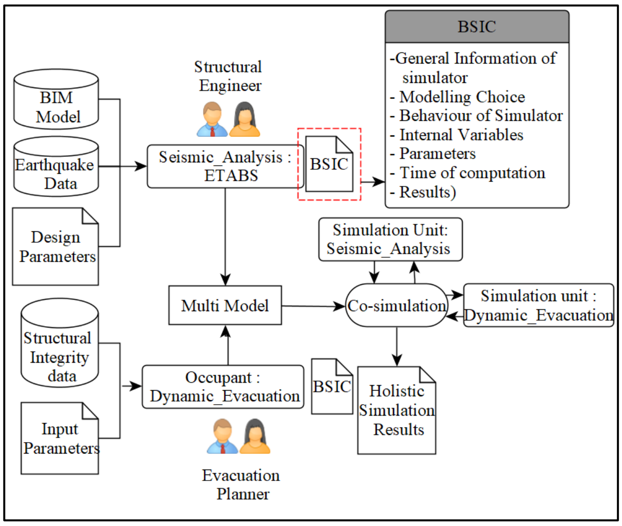


Figure 1: Framework for co-simulations for emergency evacuation from buildings.

In co-simulation, the process begins by defining the connections between simulation unit instances at the system modeling level. These simulation units interact through designated interfaces, known as ports. For instance, in the seismic analysis and dynamic evacuation simulation unit, there is a float output port. The framework also incorporates the Functional mockup interface (FMI) (Palmieri et al., 2020), a common interface standard crucial for co-simulation. FMI provides a standardized approach for encapsulating simulation functionalities, enabling seamless transfer of inputs between simulation units. In assosiation with Functional mockup units (FMUs), which are standardized models for exchanging simulation data, multi-model facilitates the transfer of input to each simulation unit at the current time step. Subsequently, within each simulation unit, the time step for the next major iteration is calculated and used in the multi-model (see next sectoin for multi-model). Each simulation unit advances in a time step until it reaches the point where the maximum shear and moment causing structural failure of the beam are obtained. A collaborative simulation using this multi-model is finally executed by a Co-sim Orchestration Engine (COE), which will act as the central control unit, coordinating interactions among simulation instances. This process is done manually as a part of proof of concept.

Building Simulation Identity Card (BSIC)

We develop BSIC that represents key attributes of simulation models, including general information about the simulator, modeling choices, its behaviour, internal variables, parameters, and simulation results, which serve as the basis for input for other simulators. As presented in Table 1, these attributes are part of the proof-of-concept study, and a more detailed formalization of the BSIC will be undertaken in future research.

Table 1: Classification of attributes used to formulate BSIC.

Attributes	Description
Information	Name , Purpose, description, run time & version of simulation model
Modelling Choice	Modelling field, model dimension and time scale
Required Simulation tool	Name , version and alternative simulator tool
Behaviour	Behaviour specification, analysis type
Interface	Nature, domain , sub-domain
Internal Variables	Type, description, value, unit
Parameters	Classification, quantity, scale
Simulation Results	Time step, accuracy, outputs

Multi-model

FMUs and their interdependencies establish the framework known as a multi-model system. The concept of multi-modeling involves the integration of diverse simulation models to analyze complex systems (Fishwick et al., 1994). This enables communication among different simulators and provides information that cannot be obtained through the use of a single model. This approach allows for a more comprehensive analysis of the system's overall dependability and performance under various scenarios, such as emergency evacuation planning from buildings. For instance, in the context of emergency evacuation planning scenarios, multi-modelling allows the integration of structural integrity and emergency evacuation models. Each model represents a different aspect of the case, facilitating a more holistic understanding of the system as a whole.

In this study, we develop a multi-model that uses BSIC to aid in the solution of potential inconsistencies in the creation and combinations of specialized simulation models. Figure 2 shows the multi-model for co-simulation of emergency evacuation planning incorporating structural integrity.

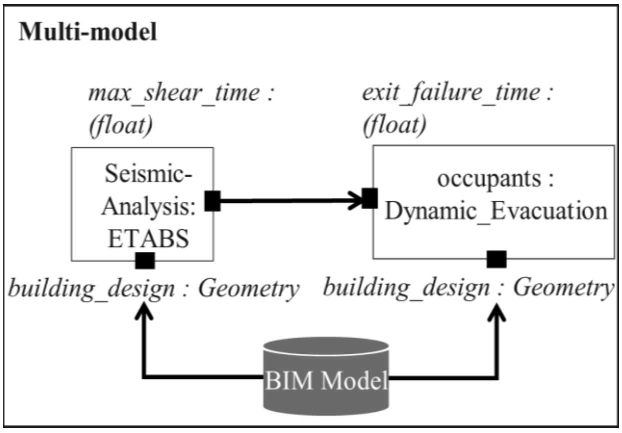


Figure 2: Emergency evacuation planning multi-model.

Proof of concept demonstration

In this section we present a proof-of-concept demonstration of an evacuation scenario that integrates an earthquake simulator and occupant egress simulator. The structure under examination is intended to be simple in terms of geometry, a single-story building featuring two rooms measuring 20 by 30 feet each. The building has been modelled using Autodesk Revit software. The structure has one normal exit, referred to as Evacuation Point (E.P)1, and an emergency exit labelled E.P2. The primary evacuation route is E.P1, which is used by most occupants during normal times. In case E.P1 is inaccessible, individuals will use E.P2 as an emergency exit. The total area of the building encompasses 1200 square feet, accommodating 20 individuals. Figure 3 represents the floor plan of the building.

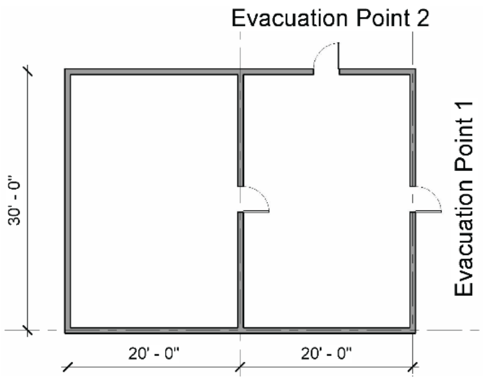


Figure 3: Building floor plan.

A comprehensive structural analysis of the building was conducted, revealing a fundamental time period (Çelebi, 2007) of 0.84 seconds. ETABS software was used for a comprehensive analysis of structural stability and earthquake impact (Guleria, 2014). The time history of the 1989 Loma Prieta earthquake (PEER-Center, 2013) served as a reference for this analysis. The structural analysis revealed the maximum stresses experienced by structural elements at t=5 sec, with corresponding shear and moment recorded. During this critical instance, the shear and moment data suggest a potential beam failure above E.P1. This is because the design shear and moment capacity of the beam is less than the shear and moment caused by the applied loads. Subsequently, these findings were transcribed into the BSIC (see the next section), forming the basis for input in an emergency evacuation analysis. To facilitate evacuation analysis, a dedicated agent-based simulator was employed using the Python language.

Results

The ETABS analysis provides critical information regarding the structure's seismic response. This is represented through a comprehensive time history plot, illustrating the dynamic movement of the structure's floors. The plot represents base shear over time, describing the structure's response to seismic forces. Figure 4 represents a time history plot illustrating the seismic response of the building under earthquake loading, indicating a peak base shear of -78 kips occurring at t = 5 seconds in X direction. This plot is utilized to analyze the behavior of the building under seismic loading (Rathod & Gupta, 2020).

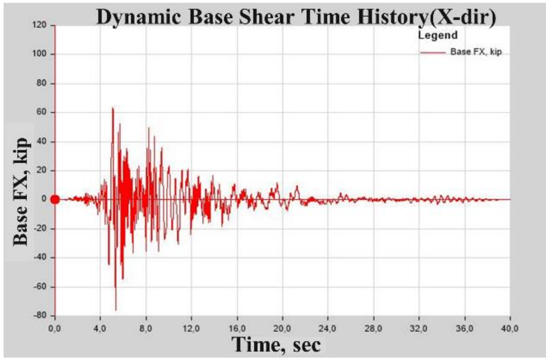


Figure 4: Dynamic base shear time history plot (X-dir) for the building under earthquake loading.

The analysis results indicate that at t = 5 seconds, the structural beam above E.P1 experience maximum stress. At this point, the stress level has reached a critical level, which indicates a potential vulnerability in the structure. It's important to note that the beam above E.P1 may fail due to a critical stress condition, leading to its blockage.

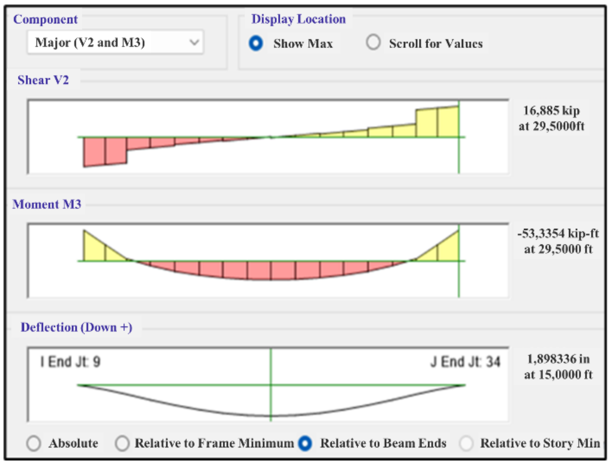


Figure 5: Screenshot of the ETABS tool showing the maximum moment and shear diagram for beam above E.P1 at t=5 sec

The seismic forces during the earthquake event have created conditions where the moment and shear forces acting on the beam above E.P1 exceed its design structural capacity. This ultimately leads to beam failure. Figure 5 illustrates the moment and shear diagram for the beam under consideration subjected to earthquake loading. Figure 6 illustrates the lateral displacement of the structure in the X-dir, providing a visual representation of structural deformation under earthquake loading. The displacement in the X-dir is more than in the Y-dir. Notably, the lateral displacement of the structure at t=5 seconds exceeds the allowable limit of H/500 (Abd Samat et al., 2017), reaching 0.6 inch. This will cause structural failure starting from the beam above E.P1.

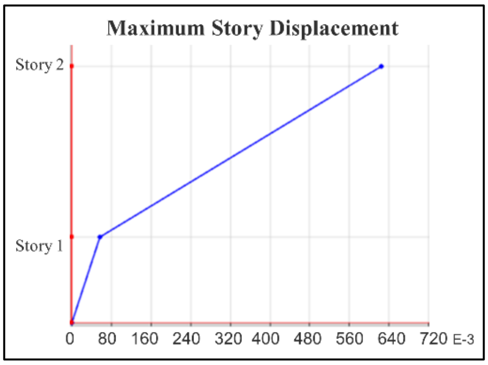


Figure 6: Lateral story displacement in X-direction.

The integration of outcomes derived from the seismic analysis simulator into the evacuation planning simulator has been executed manually, serving as a proof of concept. This process is represented in the sequential diagram illustrated in Figure 7. A master algorithm is utilized by a Co-sim Orchestration Engine (COE), such as the Jacobi algorithm, which performs a simulation step by initiating the advancement of all FMUs in time, retrieving the necessary outputs, and setting the necessary inputs (Gomes, Thule, Larsen, et al., 2018). The simulation follows a series of iterative steps until a termination condition is met, typically determined by the maximum timestep specified by the user. The procedural sequence can be described as follows: The COE initiates the simulation cycle by dispatching inputs to all FMUs. Subsequently, the COE retrieves the next time step from each FMU. Advancing the simulation, the COE progresses the state of each FMU's simulation to the minimum (earliest) time reported by the respective units. Finally, the COE consolidates the output data generated by all FMUs.

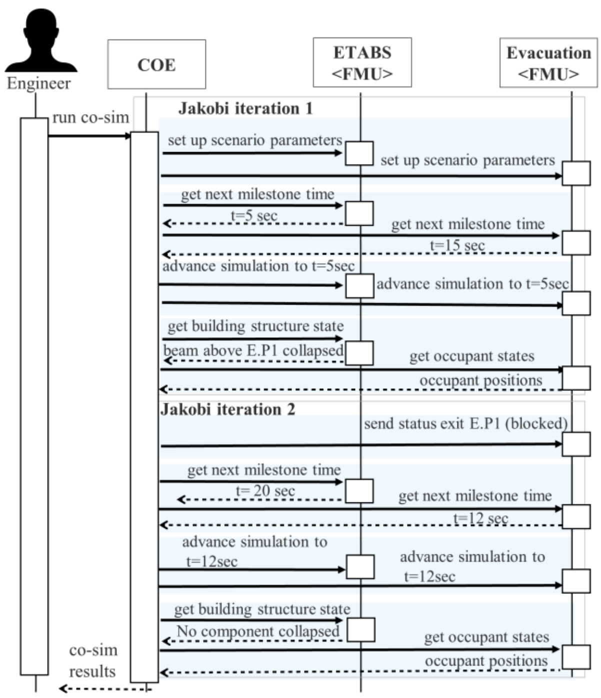


Figure 7: Sequence diagram for co-simulation in Emergency evacuation planning

BSIC used in co-sim process for structural integrity simulation is represented in Figure 8. It reflects a understanding of simulation model characteristics. It provides general information about the simulation model, such as its name, purpose, and the simulator employed for execution. Additionally, it covers details related to the modeling and behavior of the simulation model. The BSIC also encompasses variables and parameters that are necessary for the design of the simulation model. Finally, it presents the results obtained from the simulation model.

Subsequently, the BSIC information is then integrated into simulation model for evacuation analysis. The latter is executed by developing Python code that establishes a comprehensive and integrated approach. This approach will accommodate the seismic activity and emergency evacuation scenarios within our model.

SeismicAnalysis_ETABS	
Attributes	**Description**
General Information	
Name of Simulation Model	SeismicAnalysis ETABS
Simulation purpose	Earthquake analysis
Description	This simulation model serves the purpose of providing information on the maximum stresses occurring at a specified time and location
Simulation run time	Varies based on the complexity of the model
Software/language used	ETABS (Extended Three-Dimensional Analysis of Building Systems)
File format name	.e2k, .edb, .f2k, .exr, .DXF/.DWG, CIS/2, IFC, .igs, .stl, .obj, .xlsx, .docx, .accdb
Modelling Choice	
Modelling Field	Structural Engineering
Model dimension	2 rooms (20x30x11)
Applied Earthquake	Loma Prietan1989
Time scale	Time-dependent analysis capturing seismic events
Behaviour	
Behaviour specification	Material nonlinearity, geometric nonlinearity, and boundary condition nonlinearity considered
Analysis type	Dynamic Time-History Analysis for seismic response
Internal Variables	
Time History	X and Y-direction
Base Shear X-direction	56Kips
Base Shear Y-direction	78Kips
Damping	0.5%
Parameters	
Spectral Acceleration (X-dir)	0.78g
Spectral Acceleration (Y-dir)	0.75g
Concrete compressive strength	4000psi
Poission´s ratio	0.2
Results	
Time of max response	5 sec
Fundamental time period	0.82sec
Design moment	-52kips-ft(top), 27kip-ft(bottom)
Shear capacity	11.95kip
Moment at beam due to E.Q loading	-53-3kip-ft
Moment at beam due to E.Q loading	16.9 kip
Lateral displacement (allowable H/500)	0.26 in
Lateral displacement at E.Q loading (X-dir)	0.6inch
Lateral displacement at E.Q loading (Y-dir)	0.57inch

Figure 8: BSIC for structural integrity simulator

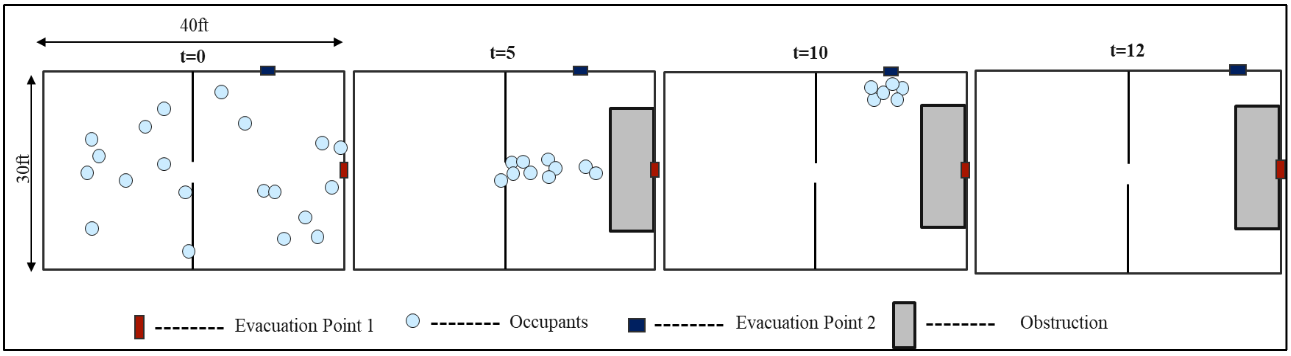


Figure 9: Emergency evacuation scenario w.r.t time.

Agent-based simulation is employed to predict the evacuation of people within the building during a seismic event. Figure 9 illustrates thse flow of people towards the primary exit, also known as E.P1, at times t = 0 and t = 5 sec. However, in the event of a structural failure occurring at evacuation point 1 at time t = 5 sec, the primary exit becomes obstructed by rubble. As a result, people will move towards E.P2, identified as the secondary exit. The figure illustrates that at t = 12 sec, all the occupants have evacuated the building safely.

The dynamics of the evacuation rate, as derived from the data collected during the evacuation simulation, are depicted in Figure 10. Notably, the rate of evacuation increases significantly when a large number of people are at an evacuation point at the same time. In the observed time frame from t = 0 to t = 5 seconds, there is a noticeable upward trend in the cumulative departure of occupants from the building.

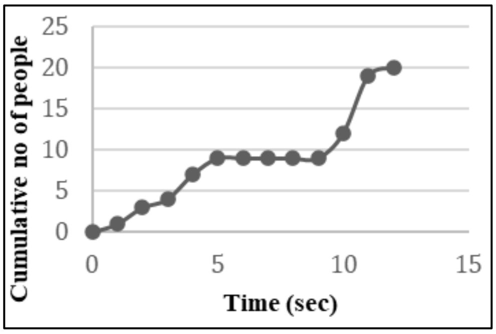


Figure 10: Cumulative number of people leaving the building.

Conversely, between t = 5 seconds and t = 9 seconds, a consistent evacuation rate is observed. This stabilization corresponds to occupants redirecting their movement towards E.P2 following the closure of E.P1. At t=12 seconds, where the graph reaches its peak, all 20 occupants have noticeably vacated the room.

Figure 11 represents the concise BSIC version of the dynamic evacuation simulation model, summarizing crucial information about the simulator used, its inputs, and outputs. It provides a comprehensive understanding of the simulation process, including different attributes of simulation model. This BSIC will prove valuable in the iterative development of the structural integrity simulator and in designing additional simulation models for co-simulation purposes.

BSIC: Dynamic_Evacuation	
Attributes	**Description**
General Information	
Name of Simulation Model	Dynamic_Evacuation
Simulation purpose	Emergency evacuation
Description	The simulation models the movement of people towards evacuation points considering room structures and obstacles.
Simulation run time	20 seconds
Software/language used	Python
Modelling Choice	
Modelling Field	Evacuation Planning
Model dimension	2 rooms (20x30)
Time scale	Discrete
Behaviour	
Behaviour specification	Individuals exhibit dynamic movement, responding to room structures and obstacles, with a focus on reaching evacuation points.
Analysis type	Real-time Visualization
Evacuation Strategy	Dynamic pathfinding based on real-time structural information.
Internal Variables	
Description	Instances representing individuals spped and obstacles in the simulation
No of structural component failure	1
Speed of occupants	3ft/s
Parameters	
Name	Room width, room length, no people, speed, total time
Description	Parameters defining the simulation environment, evacuation points, number of people, and speed
No. of occupants	20
No. of evacuation points	2
Model dimension	2 rooms (20x30x11)
Results	
Time step	1 sec
Time of structural failure	5 sec
Evacuation Success Rate	100%
Total evacuation time	12 sec

Figure 11: BSIC for emergency evacuation simulator

Discussion

To address the issues related to the combination of a specialized simulation model for information storage and sharing, we demonstrated the first proof of concept of BSIC. For conducting emergency evacuation simulations from a building in case of an earthquake, the provided proof of concept serves as a foundational framework for co-simulations. To improve emergency evacuation planning, integrating data from diverse simulations is imperative, e.g., understanding earthquake propagation, managing dynamic crowd flows during egress, and creating building structure models are crucial components. Furthermore, emergency evacuation planning should consider assessing structural integrity during earthquakes, estimating evacuation time to prevent injuries and analyzing occupant movement patterns.

In this study, we used a seismic analysis simulator to determine if the structure would fail during an evacuation due to earthquake loading. Then the data from the seismic analysis simulation model is integrated with an evacuation planning simulation model to understand the significance and usability of BSIC in combining diverse simulation models.

The integration process between simulation models is executed manually, creating the challenge of automating the co-simulation process, which needs consideration. This includes the implementation of ETABS and the dynamic evacuation simulator as proper FMUs integrated into a COE, and the processing of inputs and outputs for seamless communication between the two simulators. An intricate issue in automating the process concerns the sequential aspects of simulation data. The ETABS simulator focuses on the structural time scale, whereas the dynamic evacuation simulator focuses on crowd behavior dynamics. Future studies will focus on running this co-simulation process automatically.

Furthermore, the integration between specialized simulation models not only highlights the interaction between seismic analysis and the evacuation planning simulation model but also emphasizes the practical application of BSIC in enhancing the understanding of the emergency evacuation planning process.

BSIC outlined in Figure 8 and Figure 11 represents a limited understanding of Seismic analysis and evacuation planning simulation model properties. BSIC is used for classifying analysis modelling, including input/output parameters and quality expectations. A comprehensive BSIC at full scale will be developed in a future study.

Conclusion

Our study, focused on improving collaboration and integration among simulators for predictive simulation during emergency evacuation planning from buildings in earthquake scenarios, has yielded significant insights into co-simulation complexities. The efforts were put in place to answer the research question projected in the introduction.

We first introduced and configured the concept of BSIC to facilitate the co-simulation process for independent simulation models. Then, we utilized BSIC for integrating specialized simulators into emergency evacuation prediction scenarios. Our study achieved significant progress by surpassing the limitations of individual simulators through the successful implementation of a standardized and shared vocabulary.

We implemented BSIC in emergency evacuation predictions by standardizing model representation, ensuring an accurate representation of various factors from both seismic analysis and emergency evacuation simulation models.

This addresses challenges related to a lack of common terminology and interface inconsistencies during the integration phase of a simulation model, whether performed manually or facilitated by software. Through standardized interfaces and frameworks, BSIC facilitates informed decision-making for design teams, enabling effective collaboration in the integration of seismic analysis and evacuation planning simulators.

This research lays the foundation for future investigations, where we aim to further consolidate data from various simulation models, building upon the capabilities of BSIC to achieve a comprehensive understanding of the emergency evacuation planning process. In this regard, BIM will play a crucial role in emergency evacuation planning by facilitating realistic scenario modeling, aiding in the planning of evacuation routes, assessing potential damage, and effectively allocating resources (Stančík et al., 2018). BIM's ability to create a digital replica of real-world structures and environments, complete with intricate details, provides a strong foundation for disaster preparedness. In the next stage, our project focuses on how and in which way integration of BIM with specialized building simulators should be developed to use extensive datasets, offering insights into complex details related to structural integrity and evacuation planning.

References

Abd Samat, R., Chua, F. T., Mustakim, N. A. H. M., Anuar, F. I., Saad, S., & Bakar, S. A. (2017). Lateral Displacement And Shear Lag Effect Of High-Rise Buildings With Diagrid SystemThat Is Constructed Above A Frame. MATEC Web of Conferences,

Alfalouji, Q., Schranz, T., Falay, B., Wilfling, S., Exenberger, J., Mattausch, T., Gomes, C., & Schweiger, G. (2023). Co-simulation for buildings and smart energy systems — A taxonomic review. Simulation Modelling Practice and Theory, 126, 102770. https://doi.org/https://doi.org/10.1016/j.simpat.2023.102770

Aßmann, U., Zschaler, S., & Wagner, G. (2006). Ontologies, Meta-models, and the Model-Driven Paradigm. In C. Calero, F. Ruiz, & M. Piattini (Eds.), Ontologies for Software Engineering and Software Technology (pp. 249-273). Springer Berlin Heidelberg. https://doi.org/10.1007/3-540-34518-3_9

Battegazzorre, E., Bottino, A., Domaneschi, M., & Cimellaro, G. P. (2021). IdealCity: A hybrid approach to seismic evacuation modeling. Advances in Engineering Software, 153, 102956.

Çelebi, M. (2007). On the variation of fundamental frequency (period) of an undamaged building–a continuing discussion. International Conference on

Experimental Vibration Analysis for Civil Engineering Structures. Porto,

Chu, H., Yu, J., Wen, J., Yi, M., & Chen, Y. (2019). Emergency Evacuation Simulation and Management Optimization in Urban Residential Communities. Sustainability, 11(3), 795. https://www.mdpi.com/2071-1050/11/3/795

Eastman, C. M. (2011). BIM handbook: A guide to building information modeling for owners, managers, designers, engineers and contractors. John Wiley & Sons.

Fishwick, P., Narayanan, N., Sticklen, J., & Bonarini, A. (1994). A Multi-Model Approach to Reasoning and Simulation. Systems, Man and Cybernetics, IEEE Transactions on, 24, 1433-1449. https://doi.org/10.1109/21.310527

Forssberg, M., Kjellström, J., Frantzich, H., Mossberg, A., & Nilsson, D. (2019). The Variation of Pre-movement Time in Building Evacuation. Fire Technology, 55(6), 2491-2513. https://doi.org/10.1007/s10694-019-00881-1

Gelenbe, E., & Desmet, A. (2013). Graph and Analytical Models for Emergency Evacuation. Future Internet, 5. https://doi.org/10.3390/fi5010046

Gomes, C., Thule, C., Broman, D., Larsen, P. G., & Vangheluwe, H. (2018). Co-simulation: a survey. ACM Computing Surveys (CSUR), 51(3), 1-33.

Gomes, C., Thule, C., Larsen, P. G., Denil, J., & Vangheluwe, H. (2018). Co-simulation of continuous systems: a tutorial. arXiv preprint arXiv:1809.08463.

Guleria, A. (2014). Structural analysis of a multi-storeyed building using ETABS for different plan configurations. Int. J. Eng. Res. Technol, 3(5), 1481-1485.

Hansen, S. T., Thule, C., Gomes, C., Lausdahl, K. G., Madsen, F. P., Abbiati, G., & Larsen, P. G. (2024). Co-simulation at different levels of expertise with Maestro2. Journal of Systems and Software, 209, 111905. https://doi.org/https://doi.org/10.1016/j.jss.2023.111905

Hansen, S. T., Thule, C., Gomes, C., van de Pol, J., Palmieri, M., Inci, E. O., Madsen, F., Alfonso, J., Castellanos, J. Á., & Rodriguez, J. M. (2022). Verification and synthesis of co-simulation algorithms subject to algebraic loops and adaptive steps. International Journal on Software Tools for Technology Transfer, 24(6), 999-1024. https://doi.org/10.1007/s10009-022-00686-8

Johnson Jr, J., & Zeigler, D. J. (1986). Modelling evacuation behavior during the Three Mile Island reactor crisis. Socio-Economic Planning Sciences, 20(3), 165-171.

Kirby, A. M., Dietz, J. E., Matson, E. T., Pekny, J. F., & Wojtalewicz, C. (2015). Major city evacuation planning using simulation modeling. International Journal of Disaster Resilience in the Built Environment, 6(4), 397-408.

Mahdavi, A., Wolosiuk, D., & Berger, C. (2023). Toward a theory-driven ontological framework for the representation of inhabitants in building performance computing. Journal of Building Engineering, 73, 106804. https://doi.org/https://doi.org/10.1016/j.jobe.2023.106804

Mirahadi, F., & McCabe, B. Y. (2021). EvacuSafe: A real-time model for building evacuation based on Dijkstra's algorithm. Journal of Building Engineering, 34, 101687. https://doi.org/https://doi.org/10.1016/j.jobe.2020.101687

Palmieri, M., Bernardeschi, C., & Masci, P. (2020). A framework for FMI-based co-simulation of human–machine interfaces. Software and Systems Modeling, 19(3), 601-623.

PEER-Center. (2013). PEER Ground Motion Database. https://ngawest2.berkeley.edu/

Rathod, K. V., & Gupta, S. (2020). A nonlinear time history analysis of ten storey RCC building. International Research Journal of Engineering and Technology, 7, 7153-7160.

Schwan, T., Unger, R., & Pipiorke, J. (2017). Aspects of FMI in Building Simulation. Modelica,

Sirin, G., Paredis, C. J., Yannou, B., Coatanéa, E., & Landel, E. (2015). A model identity card to support simulation model development process in a collaborative multidisciplinary design environment. IEEE Systems Journal, 9(4), 1151-1162.

Smith, P. (2016). Structural design of buildings. John Wiley & Sons.

Stančík, A., Macháček, R., & Horák, J. (2018). Using BIM model for Fire Emergency Evacuation Plan. MATEC Web of Conferences, 146, 01012. https://doi.org/10.1051/matecconf/201814601012

Thule, C., Lausdahl, K., Gomes, C., Meisl, G., & Larsen, P. G. (2019). Maestro: the INTO-CPS co-simulation framework. Simulation Modelling Practice and Theory, 92, 45-61.

Yueming, C., & Deyun, X. (2008). Emergency evacuation model and algorithms. Journal of Transportation systems engineering and information technology, 8(6), 96-100.

CAN MACHINE LEARNING AUTOMATE CARBON CLASSIFICATION OF MATERIALS WITHIN A BUILDING INFORMATION MODEL?

Abdulkadir Adeola Abdulrahman[1], Rogage Kay[2]
[1,2]Northumbria University, Newcastle, United Kingdom.

Abstract

Buildings account for an estimated 35% of the UK's total greenhouse gas emissions. Assigning carbon data to building designs can aid contractors in sustainable material selection. Authoring inconsistencies in building data mean addition of environmental data is difficult and expensive. Machine learning enables classification of materials at a scale manual approaches cannot match. A machine learning approach is documented for classifying building products that enables automatic augmentation of environmental data from carbon databases. Our findings provide foundational research on automating data authoring, that can reduce costs and simplify processes associated with adding environmental assessment data to building designs.

Introduction

In June 2019, UK parliament made a legally binding commitment to bring all Greenhouse Gas (GHG) emissions in the UK to Net Zero by 2050 (Department for Business, Energy & Industrial Strategy, 2019). The construction sector is a major contributor to the UK's carbon footprint and is responsible for 50% of all extracted materials and 35% of GHG emissions (Rogage et al., 2019). Construction activity accounts for around 50m tonnes of CO2 emissions, over half of this is linked to construction products and materials (Government Commercial Function, 2022). Modern approaches to collecting information about buildings across the supply chain offer unprecedented data resources for analysing and improving processes. Building Information Modelling (BIM) is an established term within academia and across the sector that is defined as a process for collaborative processing and management of design and construction information (Tallet et al., 2021). Whilst in its essence BIM is an information management approach, it is more specifically associated with standardised schemas and classification systems that enable interoperability and automated data processing. Cavalliere et al. (2019) stressed carbon management in construction is becoming an increasing priority as buildings make up such a high percentage of GHG emissions. Frameworks such as the Leadership in Energy and Environmental Design (LEED) (2024) and Building Research Establishment Environmental Assessment Method (BREEAM) (2024) certifications, enable contractors to calculate and certify embodied carbon impact when selecting building materials. Whilst calculating embodied carbon of buildings is not the aim of our research, our findings could link to existing frameworks such as LEED and BREEAM as part of the calculation method for the material selection process. Tools are needed that accurately classify materials within BIM and match them up to associated suitable carbon database items in order to optimise carbon management strategies (Cavalliere et al., 2019).

Whilst several classification systems exist, often product material data is unclassified or classified using non-standard approaches, making it difficult to automate assigning carbon information to products. This research explores a Machine Learning (ML) approach to address the problem of assigning carbon information to products. We seek to understand the potential for ML to accurately identify material descriptions in BIM data, so information from carbon databases can be automatically augmented to building products. Our developed approach uses text and sentiment analysis, and classification techniques in order to match BIM material descriptions with those found within carbon databases. A comprehensive literature review of classification algorithms using BIM data and carbon management provides the foundation for developing the ML approach. We then select and evaluate a number of algorithms for classifying BIM data against carbon impact categories in the Inventory of Carbon Energy (ICE) (Circular Ecology, 2020) database. We propose a novel application of ML techniques for predicting and classifying incomplete or inaccurate BIM data.

Related Work

BIM has rapidly become an indispensable approach in AEC industries for creating digital representations of building designs (Basbagill, 2013). A BIM consists of any elements that represent assets within a building such as doors, windows, walls etc. Consistent and accurate classification of BIM elements including materials, remains an ongoing challenge for AEC professionals. Honic et al. (2019) found that inconsistent data and a lack of cooperation among different stakeholders, created a barrier to automation of data related to recycling of building materials. Manual classification of BIM elements can be time consuming, ineffective and create inconsistencies within data. An automated classification approach utilising ML could significantly enhance accuracy and efficiency when applied to element classification of BIM data.

Studies have highlighted ML’s utility for accurately classifying heritage buildings (Bassier et al., 2017), maintenance issues (McArthur et al., 2018), wall and door elements (Koo et al., 2021) and quantifying environmental impact assessments (Cavalliere et al., 2019; Xu et al., 2022). Deep Learning (DL) is a subcategory of ML that uses historical experience as the basis for future predictions. DL specifically has shown promise as an accurate way of classifying complex BIM elements based on geometry (Koo et al., 2021; Rogage and Doukari, 2024). One area in which ML algorithms have been applied to BIM is in assessing embodied environmental impacts of building designs. Basbagill et al. (2013) and Cavalliere et al. (2019) offer examples of this application using Life Cycle Assessment (LCA), continuous BIM-based assessment and an automated LCA approach. Xu et al. (2022) proposed an innovative BIM-integrated LCA for prefabricated buildings as an automated way of performing the embodied carbon assessment process. All three studies demonstrate the potential of ML/BIM combination in mitigating environmental impacts while optimising building design performance. ML algorithms have also been successfully utilised with BIM for classification of maintenance issues and enhanced data collection, according to an approach devised by McCarthy et al. (2018) utilising ML visualisation techniques in BIM to classify issues while increasing data collection rates. Zabin et al. (2022) conducted an in-depth literature review pertaining to applications of ML to BIM projects and highlighted its capacity for automated quality control and error detection. Koo et al. (2021) introduced an automated approach for accurate classification of BIM elements using 3D geometric Deep Neural Networks (DNN) as walls and door elements. Rogage and Doukari (2024), further demonstrate use of 3D geometric DNN for classifying a further range of products. These studies demonstrate the promise of ML with BIM for improving accuracy and efficiency of classification of elements within BIM. However, consideration must be made regarding possible challenges or restrictions associated with its usage; including issues surrounding data quality, privacy concerns and interpretability (Zabin et al., 2022).

Recurrent Neural Networks

Here we consider ML models for classifying BIM data. Recurrent Neural Networks (RNN) are a type of artificial neural network that use layers to process data. RNN with Long Short-Term Memory (LSTM) architecture comprises three layers: input; recurrent; and output (Nosouhian et al., 2021). RNN input layers (shown on the left of figure 1) accept input sequences such as time series data or natural language text as vectors; then feed these vectors one at a time into their networks for processing. A type of RNN called LSTM units can store long-term memory, which enables them to comprehend text sequences better by capturing context and semantics (Kaur and Mohta, 2019). An RNN's recurrent layer serves to maintain its memory of previous inputs and generates outputs based on each cell taking an input and producing an output, along with updating their hidden state, each time step based on current and previous hidden states - thus giving rise to RNNs' ability to recall past data inputs (Kaur and Mohta, 2019). Within each layer are nodes or neurons, represented by circles in figure 1, which perform simple operations and pass signals to other neurons. Connections between neurons are represented by lines, the strength of connections determined by weights. Signals travel through the network from left to right via these weighted connections. As data passes through each layer, the network learns patterns and extracts meaningful features. An RNN's output layer (shown on the right of figure 1) takes the final hidden state from its recurrent layer and generates its output; it could take the form of either a simple feedforward layer, or it could employ complex features like softmax layers for classification tasks (Nosouhian et al., 2021). This output produces the network's predictions or classifications based on what it has learned. In many applications, neural networks contain multiple hidden layers to solve increasingly complex problems. The number and arrangement of layers and neurons can be modified to suit different tasks. One of the primary strengths of RNNs lies in their capacity to handle variable-length data sequences efficiently, due to using identical weights across time steps of sequence processing allowing it to process sequences of various lengths without disrupting operation of the network.

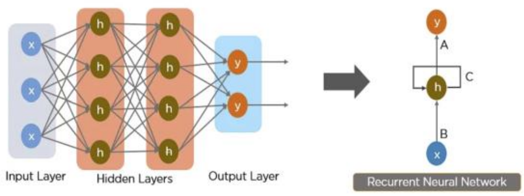


Figure 1: Deep Learning architecture (Badhan et al, 2024)

Figure 2 demonstrates how RNNs process sequential data. RNNs are commonly used for natural language processing tasks where the order of inputs is important, such as text classification. We can think of this grid as representing a sequence of inputs an RNN would read, from left to right and top to bottom. Each cell contains a token or word from the sequence. As the RNN processes each step, it considers both current input and information from the previous step, represented by the cell to the left. This looping, recurrent structure allows RNNs to connect previous information to later inputs in a sequence. The network develops an internal state capturing features of the entire sequence processed so far which is passed from each cell to the next. By reading the sequence from this grid in order, an RNN would build up understanding of word patterns and relationships over multiple time steps. This visualisation conveys how RNNs can process variable-length sequential data by sharing parameters across every step. The 2D array representation provides a simple way to conceptualise how RNNs flow information through a network from start to end of a sequence, capturing context at each point to help classify or generate sequential outputs.

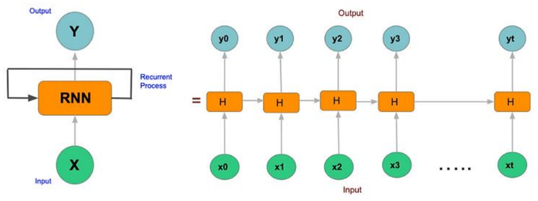


Figure 2: RNN architecture (Nagesh, 2020)

Naive Bayes

Naive Bayes methods are a set of supervised learning algorithms based on applying Bayes' theorem with the "naive" assumption of conditional independence between every pair of features given the value of the class variable (Qiang, 2010). Naive Bayes is a probabilistic classification algorithm that calculates the probability of an instance belonging to a class based on the probabilities of its features. It uses Bayes' theorem:

$$P(y|x) = \frac{(P(x|y)\ P(y))}{P(x)} \qquad (1)$$

Where:

P(y|x) is the probability of class y given the features x.
P(x|y) is the probability of features x given class y.
P(y) is the prior probability of class y.
P(x) is the probability of features x.

Naive Bayes estimates probabilities P(x|y) and P(y) from a training dataset. During prediction, it calculates the probability of each class label for a new instance and assigns it to the class with the highest probability. The model assumes features are independent, meaning the presence or absence of one feature does not affect the presence or absence of others. This assumption simplifies calculations and allows the model to handle high-dimensional feature spaces efficiently. Despite its simplicity, naive Bayes performs well in text classification tasks, such as spam detection and sentiment analysis. It is computationally efficient and can handle large datasets. However, it may not perform well when the independence assumption is violated or when there is a lack of sufficient training data.

Random Forest

Random Forest is an ensemble learning method that combines multiple decision trees to make predictions (Wang and Wang, 2021). It randomly selects subsets of training data and features to train each tree. The final prediction is made by aggregating predictions of all trees, either through majority voting for classification or averaging for regression to improve prediction performance. The algorithm is illustrated by creating multiple decision trees, each trained on different subsets of data and features, and then combining their predictions. Random Forest is widely used for its effectiveness and ability to handle high-dimensional datasets.

Figure 3 shows three decision trees each making independent classifications. Branches of Decision Tree-1 and Decision Tree-2 represent two individual decision trees classifying an input sample and arrive at a "Result". These predictions are combined through majority voting. Majority voting is the process of having each tree "vote" for a class and selecting the class receiving the most votes overall. This aggregation helps to reduce variance and prevent overfitting compared to a single decision tree. By growing many decision trees on randomly sampled subsets of training data and combining their results, Random Forests are able to capture relationships any individual tree may miss. This ensemble approach typically yields better generalisation performance than a single estimator. This ensemble technique is effective for both classification and regression tasks.

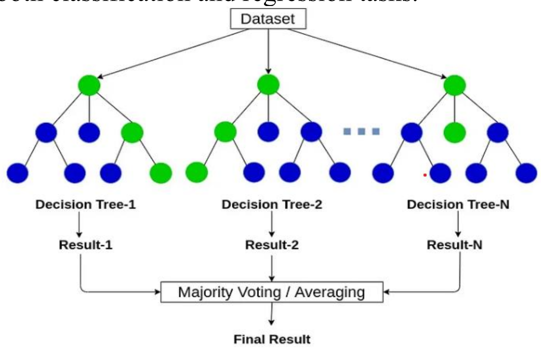


Figure 3: Random Forest structure (Zhu and Spachos, 2021)

Simulation and Experiment

Methodology

The data science pipeline and methodology is shown in figure 4. This approach involves selecting data, then cleansing and preparing data using feature extraction with consideration to the model to be employed. Data exploration precedes model construction to comprehend the significance of various features, the relationships among features, data distribution patterns and hypotheses formulated about them. A model is then built with the aim of producing inferences or future event predictions from it or discovering a "root cause" behind an already observed event; its results are then assessed and presented upon evaluation.

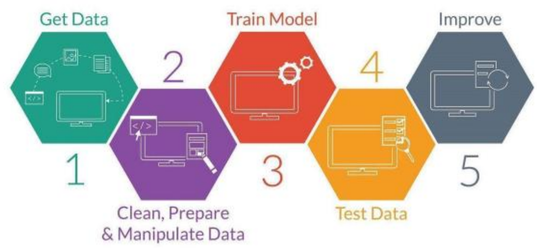


Figure 4: Machine Learning workflow (Muhammad, 2022)

Data Sources

The first data source contains approximately 10,000 material descriptions extracted from the ICE database (Circular Ecology, 2020). These include detailed information on embodied carbon and energy of various construction materials, reflecting their significant influence in the construction industry. The second data source encompasses BIM data, including labels and classifications for different building products. The ICE database was selected for its comprehensive coverage and credibility in documenting embodied carbon and energy of construction products. This dataset is split into two parts: one containing detailed material descriptions, and the other consisting of material names or labels. Material descriptions provide extensive information about each building material, including its embodied carbon, which is crucial for the analysis.

Exploring the Data

Data exploration began with data cleaning and preparation. The BIM dataset comprised 11,488 rows with four columns covering Family, IFC Type, Material Types and categories such as "ICE Category" (this is the category the model will predict to automate the mapping of BIM data directly into the ICE database). Table 1 provides a dataset sample. After a preliminary study of the data, it became apparent two columns: Family and IFC Type; were not relevant because they are both classified in generic terms which are applicable to all material that falls under the general family representation respectively, so they were removed from the dataset entirely leaving only the Material and ICE Category columns.

Table 1: Sample data from the BIM dataset

IFC Type	Material	ICE category	ICE Material
Air Terminal	Air Terminal - Diffuser Body	NaN	NaN
Air Terminal	Air Terminal - Perforated Plate	NaN	NaN
Air Terminal	Aluminium - Koolair	Aluminium	NaN
Air Terminal	Aluminium	Aluminium	NaN
Air Terminal	Anodised Aluminium - Koolair	Aluminium	NaN

Further visualisation of the data (figure 5) revealed an uneven distribution of building materials. More specifically, certain samples had significantly more entries than others indicating an uneven representation in the data and potentially leading to biased predictions and reduced accuracy when training ML models. To address this challenge, data balancing was employed; an approach in which samples from each category are adjusted to create an even distribution for more accurate predictions through ML models trained on representative samples of data. There are various strategies available for balancing data, including oversampling and undersampling (Bonatti and Kirrane, 2019). Oversampling involves increasing the sample population from minority classes while undersampling involves decreasing it; both strategies aim to achieve an even distribution across categories. This study used oversampling, specifically the Synthetic Minority Over-sampling Technique (SMOTE), to balance data (Chu et al., 2016). SMOTE involves creating artificial samples in minority classes by interpolation among existing ones; thus ensuring each minority class receives enough samples, leading to more balanced distribution of data. By using SMOTE to balance data, the model was then trained on a more balanced distribution of materials within the dataset for more accurate predictions and vital insights for sustainable building practices.

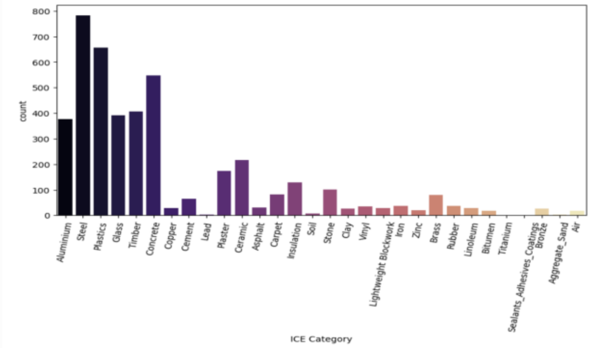


Figure 5: Dataset Distribution

A Bag of Words function was introduced to search and verify the presence of specific materials in the database, offering insights into their carbon footprints. This method represents each document as a vector of word frequencies, without regard to order or structure of words (Anello, 2021). This serves as indicators for classification of building materials, assisting in the ML process for accurate categorization and analysis. Use of these datasets allows for a detailed examination and understanding of the relationship between material properties and their environmental impact. This concluded the data preparation and manipulation stage ready for the final stage of training the model.

Model Development

The ML model was developed using three algorithms: RNN with LSTM, Random Forest, and naive Bayes. The combination of these algorithms offers a comprehensive approach to the classification of building materials, leveraging strengths of each to enhance overall effectiveness and reliability of the model in classifying building materials based on their carbon impact, as per the ICE recommendations.

Model Training

In this project, a common technique for training network parameters and performing classification was employed: splitting available data sets 80/20 on a train/test split. This strategy is known as train-test split and commonly utilised within ML to assess model performance. At this phase of training, the model was trained on its training set by using an optimization algorithm to adjust network parameters and minimise its loss function. Alpaydin (2020) defines loss function as measuring the difference between predicted output of the model and actual output and an optimization algorithm can adjust these network parameters so as to decrease this difference as much as possible. Once the model is trained and evaluated on its test set, the next stage is to improve the model. Cross-validation is used for measuring generalisation ability, testing the model on new data allows for a more accurate representation of its performance on unseen sets of information and provides an opportunity to improve the model. Hyperparameter tuning was conducted to improve the Random Forest classifier model. The number of estimators was set to 400, and the criterion for measuring the quality of a split was chosen as "entropy". The random state was also fixed to ensure reproducibility of results. This careful selection of hyperparameters aims to optimise the model's performance by adjusting its complexity and learning capacity. For the Bidirectional LSTM model, specific architecture choices such as embedding dimension, number of LSTM units (16), and use of dropout for regularization indicate a tuning process aimed at balancing model capacity with the need to avoid overfitting. While the description doesn't detail an iterative search over a range of values, these choices reflect considerations typical in model tuning, focusing on optimising the network's structure for the task at hand. The metrics used to evaluate the accuracy and completeness of the model's predictions are Precision, Recall, and F1 Score.

Precision is the ratio of true positives (correct predictions) to the total number of positive predictions. It measures how precise the model is in identifying relevant cases. A high precision means that the model has a low rate of false positives (incorrect predictions).

Recall is the ratio of true positives to the total number of actual positive cases. It measures how well the model can recall or retrieve all relevant cases. A high recall means the model has a low rate of false negatives (missed predictions).

F1 Score is the harmonic mean of precision and recall. It balances both metrics and gives a single score that reflects overall performance of the model. A high F1 score means the model has both high precision and high recall. From table 2, Random Forest model has the highest scores across all metrics at 99%. This means the Random Forest model is the most accurate and complete in its predictions, compared to the other two models. Both naive Bayes and RNN have a Precision of 98%, Recall of 97%, and an F1 Score of 97%. These scores are slightly lower than the Random Forest model, but still very high and impressive.

Discussion and result analysis

Multiple experiments using the RNN (LSTM), naive Bayes, and Random Forest models to classify building materials into various ICE categories were conducted. Training the model involved using datasets of building material data before testing its performance on test sets. A subset of building material data were manually classified with corresponding ICE categories before comparing accuracy between each classification method. Results of the experiment demonstrated high levels of accuracy on their test sets, achieving 97%, 98.38%, and 99.59% (See table 2) respectively. During data preprocessing, the dataset was simplified by removing columns ('IFC Type', 'Family', 'ICE Material') not directly relevant to the primary goal of predicting the 'ICE Category' from the 'Material' description. This decision represents a targeted form of feature selection, focusing the model's learning on the most relevant data available. The strategy adopted for imputation was straightforward yet effective: rows with missing 'ICE Category' values, which are critical for supervised learning, were dropped. This approach ensures the models are trained on complete records, enhancing the reliability of their predictions. However, it's important to note that this method of handling missing values might not be suitable for datasets where such omissions could lead to significant loss of valuable information. In those cases, more sophisticated imputation techniques might be necessary. It is crucial to note that these results, while promising, require further validation to ensure their reliability and applicability in real-world scenarios. The comparison with manual classification methods, although indicative of potential efficiency gains, should not overshadow the importance of accurate and safe classification of building materials. Table 2 shows how each model performed using different evaluation metrics.

Table 2: Model comparison Between Random Forest, naive Bayes and RNN results

	Precision	Recall	F1 Score
Random Forest	99%	99%	99%
Naive Bayes	98%	97%	98%
RNN	98%	97%	97%

Table 3 shows the first 5 rows of materials predicted using the RNN model. This means that the RNN model has successfully classified the building materials in the ICE Category based on what the model understands from the training dataset. This was the first model developed, while it successfully predicted the materials, the accuracy of its prediction could not be confirmed as this model does not suppose the confidence level measure.

Table 3: RNN Prediction

Material	ICE Category
Air Terminal - Diffuser Body	Air
Air Terminal - Perforated Plate	Air
ConnectorSolidMaterial	Concrete
Duct_Afkast	Steel
Duct_Udsugning	Brass

Table 4 shows the first 5 rows of predictions with their confidence level by the naive Bayes model. The results show the model has classified some materials such as “Air Terminal - Diffuser Body” and “Air Terminal - Perforated Plate” under the “Air” category correctly, with high confidence levels of 18.05 and 17.53 respectively. However, the model has made errors, such as classifying “Connector/Solid/Material” and “Duct_Lining/lining” as “Linoleum”, with low confidence levels of 3.43 each. These materials are likely made of metal or plastic, not linoleum, which is a type of flooring material. Similarly, “Duct_Airseal” is categorised as “Steel” with a confidence level of 9.05, which may not be accurate if the duct is made of another material.

In ML, confidence level is a measure of how likely a prediction is to be correct. It is usually expressed as a percentage between 0 and 100, where higher values indicate higher confidence. In the case of the naive Bayes prediction in table 4, where the model predicts that a building material belongs to the “Air” category with a confidence level of 18.05, it means that the model got 18.05% vote after splitting the total available percentage to the possible materials, materials with the highest vote gets predicted and the higher the percentage the better chance of the prediction being accurate (Srivastava et al., 2018).

Table 4: Naive Bayes prediction

Material	ICE Category	Confidence Level
Air Terminal - Diffuser Body	Air	18.05%
Air Terminal - Perforated Plate	Air	17.53%
ConnectorSolidMaterial	Linoleum	3.43%
Duct_Afkast	Steel	9.05%
Duct_Udsugning	Linoleum	3.43

Table 5 below also depicts the first 5 rows of material predictions and their confidence level by the Random Forest model. This model performs best although with 1% margin, and higher confidence level out of the models explored. The confidence level is calculated by taking the proportion of votes from the decision trees that agree on the same label. For example, if 100 trees are in the forest, and 96 of them vote for the “Air” category for the “Air Terminal - Perforated Plate” material, then the confidence level is 96/100 = 0.96 or 96%.

Table 5: Random Forest prediction

Material	ICE Category	Confidence Level
Air Terminal - Diffuser Body	Air	99.25%
Air Terminal - Perforated Plate	Air	96.00%
ConnectorSolidMaterial	Concrete	69.75%
Duct_Afkast	Steel	89.00%
Duct_Udsugning	Steel	77.75%

The models demonstrated impressive performance because of the focused feature set. By concentrating on the 'Material' descriptions as the primary input feature and carefully preprocessing this text, the models could learn effectively from the most relevant data. Also, by addressing class imbalance through resampling techniques such as SMOTE or direct resampling helped mitigate biases and improve model performance across less represented categories. Finally, optimised model configurations including deliberate setting of hyperparameters and model architecture choices, such as number of estimators in the Random Forest and structure of the LSTM network, contributed to the fine-tuning of each model's ability to capture the nuances of the dataset.

Conclusions

The contribution of this project lies in its exploration of ML techniques for automating classification of building data with carbon information. Whilst LCA tools and frameworks enable the calculation of embodied carbon, they rely on parametric modelling for environmental assessment, however, its application within BIM has been relatively underutilised and, more importantly, does not exist within commercial software packages used in BIM (Alwan et al., 2021). In the absence of this data, we offer an automated technique for classifying BIM products so that environmental assessment data within carbon inventory databases can be augmented to models, thus removing the need for authors to manually generate this data at design stages.

Our approach offers a foundation upon which further research can develop model accuracy and further validate the results. It demonstrates the potential of ML in automating classification tasks, which could lead to significant productivity gains in the construction industry. However, we acknowledge that our results are a preliminary step in this direction and that further research is needed to refine these techniques, improve accuracy, and fully realise their potential. Future research in this area should focus on enhancing the validation processes for ML models in construction, ensuring that they not only achieve high accuracy in controlled tests but also maintain this accuracy in practical applications. Further work is also required to investigate the integration of additional variables and data sources to enrich the models and capture a more comprehensive range of factors affecting building material classification. By advancing these aspects, subsequent studies can build upon our findings, contributing further to the knowledge and application of ML in construction and ensuring that the transition towards automated processes prioritises safety, accuracy, and reliability.

Limitations and Further Recommendations

Although this project demonstrated the efficiency of using ML to classify building materials into different ICE categories, some limitations must be acknowledged within its research design. Performing the study on a larger dataset would improve generalisability and make networks more resilient against fluctuations. Additionally the quality of data used for training and testing the ML models could be improved. Accuracy depends heavily on input quality; any errors or inconsistencies could compromise network performance significantly.

This study concentrated on classifying building materials based on their carbon impacts according to ICE. While this approach is instrumental in understanding and mitigating the environmental footprint of construction materials, it's important to acknowledge that this focus on carbon impacts alone presents certain limitations. For instance, the classification does not account for other environmental aspects such as the life cycle impacts of materials beyond carbon emissions, which can also be significant in sustainable construction practices. Future studies might expand the scope to include a more holistic environmental assessment of building materials, encompassing a broader range of environmental metrics beyond just carbon impacts.

It is also recognised that our proof of concept prototype solution utilised the UK based ICE database for linking product data to carbon calculations via materials. The embodied carbon coefficients of building materials, vary from country to country. To address this, further validation is required to assess the generalisability of the solution on databases focusing on other countries such as Öekobaudat (Federal Institute for Building, Urban and Spatial Research, 2024), Gabi (Sphera, 2024) and Bauteilkatalog (2024).

Future Work

Future work could include expanding the study to cover other properties of building materials such as durability, strength and cost. Several other ML techniques can also be evaluated in order to determine which algorithm best predicts the materials. By employing ML techniques to predict these characteristics of materials used in construction projects, more informed decisions could be made regarding selection and usage for more sustainable and cost-effective outcomes. More research could also focus on creating ML models specifically tailored for sustainability in construction industry applications, by including sustainability criteria into ML model development processes.

Acknowledgements

The datasets used within this study were provided by xbim Ltd who provided access to the ICE data alongside industry BIM data. The proprietary aspect of the ICE dataset means that it cannot be made available for further analysis by researchers outside the research team. Additionally, whilst the BIM data provided by xbim Ltd was anonymised, we were not authorised to make this dataset public. We would like to thank xbim Ltd for providing access to both datasets, and time to support the researchers in better understanding the data, without their support this project would not have been possible.

References

Alpaydin, E. (2020). Introduction to machine learning. MIT Press, pp. 003-012.

Alwan, Z., Nawarathna, A., Ayman, R., Zhu, M., & ElGhazi, Y. (2021). Framework for parametric assessment of operational and embodied energy impacts utilising BIM. Journal of Building Engineering, 42, 102768.

Anello, E. (2021). A friendly guide to NLP: Bag-of-Words with Python example. Available at:https://www.analyticsvidhya.com/blog/2021/ 08/a-friendly-guide-to-nlp-bag-of-words-with- python-example/ (Accessed: 31 October 2023). Badhan, A.K., Bhattacherjee, A. and Roy, R., 2024. Deep Learning Techniques in Big Data Analytics. In Data Analytics and Machine Learning: Navigating the Big Data Landscape (pp. 171-193). Singapore: Springer Nature Singapore.

Bassier, M., Vergauwen, M. and Van Genechten, B., 2017. Automated classification of heritage buildings for as-built BIM using machine learning techniques. ISPRS Annals of the photogrammetry, remote sensing and spatial information sciences, 4(2W2), pp.25-30.

Basbagill, J., Flager, F., Lepech, M. and Fischer, M., 2013. Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts. Building and Environment, 60, pp.81-92.

Bonatti, P.A. and Kirrane, S., 2019, July. Big Data and Analytics in the Age of the GDPR. In 2019 IEEE International Congress on Big Data (BigDataCongress) (pp. 7-16). IEEE.

Building Research Establishment Group (2024). BREEAM. Available at: https://bregroup.com/products/breeam/ (Accessed: 21 March 24).

Chu, X., Ilyas, I.F., Krishnan, S. and Wang, J., 2016, June. Data cleaning: Overview and emerging challenges. In Proceedings of the 2016 international conference on management of data (pp. 2201-2206).

Circular Ecology (2020) Embodied Carbon Footprint Database. Available-at: https://circularecology.com/embodied-carbon-footprint-database.html (Accessed: 22 November 2023).

Department for Business, Energy & Industrial Strategy (2019) UK becomes first major economy to pass net zero emissions law. Available at: https://www.gov.uk/government/news/uk- becomes-first-major-economy-to-pass-net-zero-emissions-law (Accessed: 22 November 2023).

Federal Institute for Building, Urban and Spatial Research. (2024). Oekobaudat database. Available-at:https://www.oekobaudat.de/en.htm (Accessed: 21 March 2024).

Government Commercial Function, 2022, Promoting Net Zero Carbon and Sustainability in Construction. Available-at:https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1102389/20220901-Carbon-Net-Zero- Guidance-Note.pdf (Accessed: 22 November 2023).

Holliger Consult. (2002). bauteilkatalog.ch component catalogue. Available at: https://www.bauteilkatalog.ch/Home. (Accessed: 21 March 2024).

Honic, M., Kovacic, I., Sibenik, G. and Rechberger, H., 2019. Data-and stakeholder management framework for the implementation of BIM-based Material Passports. Journal of building engineering, 23, pp.341-350.

Kaur, M. and Mohta, A., 2019, November. A review of deep learning with recurrent neural networks. In 2019 International Conference on Smart Systems and Inventive Technology (ICCSIT) (pp. 460-465). IEEE.

Koo, B., Jung, R. and Yu, Y., 2021. Automatic classification of wall and door BIM element subtypes using 3D geometric deep neural networks. Advanced Engineering Informatics, 47, p.101200.

McArthur, J.J., Shahbazi, N., Fok, R., Raghubar, C., Bortoluzzi, B. and An, A., 2018. Machine learning and BIM visualisation for maintenance issue classification and enhanced data collection. Advanced Engineering Informatics, 38, pp.101-112.

Muhammad Mudasser Afzal., 2022. 5core steps to understand machine learning work flow–a guide for beginners. Medium, Page 1. https://medium.com/@mudasserch1/5-core- steps-to-understand-machine-learning-workflow-a-guide-for-beginners-737040850d9b (Accessed: 21st March 24).

Nosouhian, S., Nosouhian, F. and Khoshouei, A.K., 2021. A review of recurrent neural network architecture for sequence learning: Comparison between LSTM and GRU.

Qiang, G., 2010, May. An effective algorithm for improving the performance of Naïve Bayes for text classification. In 2010 Second international conference on computer research and development (pp. 699-701). IEEE.

Rogage, K., Clear, A., Alwan, Z., Lawrence, T. and Kelly, G., 2019. Assessing building performance in residential buildings using BIM and sensor data. International Journal of Building Pathology and Adaptation, 38(1), pp.176-191.

Rogage, K., & Doukari, O. (2024). 3D object recognition using deep learning for automatically generating semantic BIM data. Automation in Construction, 162, 105366.

Srivastava, S., Shukla, A. and Tiwari, R., 2018. Machine translation: from statistical to modern deep-learning practices. arXiv preprint arXiv:1812.04238.

Tallet, E., Gledson, B., Rogage, K., Thompson, A. and Wiggett, D., 2021. Digitally-Enabled Design Management. In Handbook of Research on Driving Transformational Change in the Digital Built Environment (pp. 63-89). IGI Global.

Nagesh. SC., 2020. Introduction to RNN and LSTM. The AI dream. Available at: https://www.theaidream.com/post/introduction-to-rnn-and-lstm. (p1) (Accessed: 21st March 24).

United States Green Building Council (2024) LEED rating system. Available at: https://www.usgbc.org/leed (Accessed: 21st March 24).

Wang, H. and Wang, G., 2021. Improving random forest algorithm by Lasso method. Journal of Statistical Computation and Simulation, 91(2), pp.353-367.

Zabin, A., González, V.A., Zou, Y. and Amor, R., 2022. Applications of machine learning to BIM: A systematic literature review. Advanced Engineering Informatics, 51, p.101474.

Zhu, L. and Spachos, P., 2021. Support vector machine and YOLO for a mobile food grading system. *Internet of Things*, *13*, p.100359.

ENHANCING RFI ANALYSIS IN CONSTRUCTION PROJECTS: A COMPARATIVE STUDY OF TEXT CLUSTERING METHODS AND VISUALIZATION TECHNIQUES

Neziha Yilmaz[1] and Esin Ergen[1]
[1]Istanbul Technical University, Istanbul, Turkey

Abstract

The Request for Information (RFI) is a vital communication tool in construction projects, aiding teams in addressing queries and navigating challenges. Unstructured RFIs hinder manual analysis for extracting hidden knowledge. Prior research in RFI analysis employed NLP and text clustering and often relied on a single clustering method. This paper performs a comparative analysis using diverse clustering methods (LDA, NMF, and K-means) and visualization techniques to determine the most suitable methods. The study offers project managers and quality engineers an effective tool for extracting hidden knowledge in RFI.

Introduction

The construction industry heavily relies on efficient information management throughout the project lifecycle, generating a substantial volume of information in various formats (Al Qady and Kandil, 2014). Approximately 80% of corporate information is embedded in project documents, and challenges arise as professionals spend considerable time searching for and reading documents (Abbaszadegan and Grau, 2015). Insufficient information or delays in information exchange can lead to significant issues at construction sites (Abdirad et al., 2022).

Manual knowledge extraction from text documents is time-consuming and error-prone (Khuzaimah and Hassan, 2012), prompting researchers to adopt advanced Natural Language Processing (NLP) and text mining techniques, including sentiment analysis, semantic analysis, and clustering methods (Afzal et al., 2023; Jallan et al., 2019; Gheeseewan and Pudaruth, 2020).

In compliance and quality assurance, the Request for Information (RFI) document is a crucial text document in construction. It serves as a performance indicator and is analyzed conventionally through manual methods or text mining to extract insights and lessons learned (Herrera et al., 2019; Kim et al., 2022). Extracting insights and lessons learned from RFIs facilitates the early identification of potential issues, project uncertainties, and risks (Lee and Yi, 2017), allowing for timely corrective actions to be implemented in the early stages. Moreover, uncovering hidden information within RFIs enables the identification of the root cause of nonconformance. In the literature, Koc et al. (2024) have investigated the impact of non-conformities on project costs using machine learning methodologies. Similarly, the envisioned approach in this study is to extract insights from RFIs and utilize them to predict potential non-conformities for corrective and preventive action development. This approach fosters a culture of quality and compliance while mitigating reworks over time. Clustering the RFI texts to determine topics is the first step for utilizing RFI to predict unconformity and uncover hidden information within RFI texts. Exploring the effectiveness of various text topic modeling algorithms is crucial due to the unique characteristics of RFI texts. For instance, RFIs and responses, typically one to seven pages long and containing 20 to 5000 words, may be stored in unstructured format within a CDE. These texts often contain a mix of uppercase and lowercase letters, spelling errors, abbreviations, and inconsistent terminology (such as "frame" versus "beam" or "drawing" versus "CAD"), necessitating preprocessing techniques. Given their intricate details, RFI texts require micro-level clustering focus.

While some publications focus on text clustering and visualization techniques for RFIs (Lee and Yi, 2017; Afzal et al., 2023), they have utilized a single clustering method, lacking a comparison with other methods and determining the preferred visualization technique for experts. To address these gaps, this paper performs a comparative term analysis of RFI text using various clustering methods, namely Latent Dirichlet Allocation (LDA), Non-negative Matrix Factorization (NMF), and K-means, and visualization techniques (i.e., pyLDAvis, scatter plots, word cloud, and bar charts). The study involves expert interviews to determine the most suitable clustering and visualization methods for RFI clustering analysis. The research contributes to the limited comparative studies on text clustering and visualization of RFIs, offering a valuable tool for project managers to grasp RFI status.

Literature Review

Text-based data clustering methods are important for diverse applications in the construction industry. Moreover, text mining plays a crucial role in the quality management of construction engineering, involving tasks such as text cleaning, text segmentation, and semantic network analysis on unstructured data within text collections (Wang et al., 2018). The construction sector benefits from the interactive utilization of user-centric dictionaries, sentiment analysis based on word clusters, and the automated extraction and clustering of relevant terms for dictionary generation (Kohita et al., 2020; Alshari et al., 2020). These approaches ensure the construction industry's access to comprehensive and industry-specific term dictionaries, facilitating effective text analytics and knowledge management.

RFI texts, rich in information about project history and design functionality, offer detailed insights into potential quality issues in the construction site, allowing for a comprehensive focus on field operations. Traditionally, extracting insights from RFIs involved manual data mining and content analysis (Bhat et al., 2017; Afzal et al., 2023). Several studies in the literature focus on classifying RFIs. Dantas Filho et al. (2016) revised constructability concerns for evaluating RFIs in residential projects, introducing categories such as Design Correction, Divergence of Information, Design Change, Design Failure, Validation Information, and Design Verification. Morales et al. (2022) analyzed 2690 RFIs from 17 projects using statistical classification methods to associate RFIs with BIM usage. The RFIs were classified, such as alternative design solutions, approvals, clarifications of information, and other categories. Subclassifications include conflict RFIs, incorrect RFIs, insufficient RFIs, and questionable RFIs. Ham and Yuh (2023) used 872 BIM RFIs to classify them based on their purpose across four completed projects and 12 construction sites.

However, existing studies generally focus on the manual classification of RFIs, and a limited number of studies have concentrated on extracting important terms from RFIs via text clustering methods (Lee and Yi, 2017; Afzal et al., 2023). Lee and Yi (2017) used topic modeling to predict uncertainty in RFIs, focusing on the pre-bidding stage. Yet, this study did not consider the construction phase, although numerous RFIs are typically created as projects progress. Afzal et al.'s (2023) study proposes an innovative approach using the LDA algorithm to extract patterns from RFIs, offering valuable insights for the construction industry but lacking a critical comparison of clustering and visualization techniques.

LDA is an unsupervised generative probabilistic model commonly used for text topic modeling for uncovering hidden topics within a collection of text documents (Buntine et al., 2004; Lee and Yi, 2017; Afzal et al., 2023). LDA was used for text clustering in annual reports of construction companies (Jagannathan et al., 2022), focusing on macro-level clustering to identify key concerns related to understanding corporate and industry strategies. Lai and Konstanta (2019) also investigated job descriptions, which are relatively short, for building permits using LDA. In another study, LDA was used for clustering on-site inspection text, and this data is entered within constrained timeframes of on-site inspections, thus including typos and variations in upper/lower case (Lin et al., 2020). Lastly, construction-defect litigation cases were clustered using LDA, and in this type of text, many words are used interchangeably, for example, grading versus slope (Jallan et al., 2019).

On the other hand, NMF is a linear algebraic model. Unlike other topic modeling approaches, NMF utilizes matrix factorization and multivariate analysis to generate coefficients rather than probabilities for each word, associating them with specific topics (Naik, 2016; Jagannathan et al., 2022). NMF effectively reduces the dimensionality of large datasets by finding a low-rank approximation of the original data matrix, capturing essential features while discarding noise and irrelevant information (Lee et al., 2012). However, NMF works better with shorter texts, such as tweets or titles, and smaller datasets (Egger and Yu, 2022). Previous studies in non-construction contexts, particularly analyzing large text streams and review data, have demonstrated varying performance between LDA and NMF, with one algorithm often outperforming the other (George and Vasudevan, 2020).

K-Means is another unsupervised learning algorithm that groups data into n clusters with equal variance by minimizing the inertia or sum-of-squares within each cluster (Wu, 2012). K-means efficiently process large datasets by grouping data points into clusters based on similarities (Farhang, 2017). However, the initial selection of cluster centers in advance can impact the outcome (Rana et al., 2011). Additionally, K-means are sensitive to noise and outliers, potentially compromising clustering accuracy (Kaur & Aggarwal, 2017). In the literature, Liu et al. (2021) investigated the reasons for pipeline incidents and contributory factors using the K-means clustering method.

Application of text mining and visualization techniques, including sentiment analysis, semantic analysis, content analysis, and clustering methods, is crucial for unlocking insights and analyzing construction issues in RFI documents. These techniques extract meaningful knowledge from unstructured data, providing valuable insights for addressing construction challenges. However, previous studies have not determined which of these methods is the most suitable for analyzing RFI texts.

Methodology

Data Collection and Pre-Processing

The project's unique characteristics and phases (i.e., pre-bid, design, or under construction) can significantly alter the terms and topics within RFI texts. Using a common RFI dataset from various project types might lead to overlooking project-specific hidden information. Therefore, this study analyzed the RFI documents of a single project, which is an airport project. The RFIs were retrieved from a common data environment (CDE), which has streamlined the RFI handling process by enabling online communication and ensuring tracked information (Afzal et al., 2023).

During the data preparation phase, if an RFI conveyed multiple topics, these were treated as separate data instances (i.e., RFI texts); for example, a clarification topic might be observed at multiple points in a structure but communicated through a single RFI. Each clarification point is treated as an individual RFI text in this case. Additionally, an RFI should be evaluated in conjunction with its responses, as submitting an RFI triggers opinions and additional inquiries from multiple parties, and the responses should provide information about the RFI topic. A subject stored as 4-5 distinct RFI documents in CDE was considered a single RFI text to maintain the RFI context. Consequently, 288 RFI texts were extracted, representing different disciplines and locations within the structure. A similar number of RFI documents (243 RFI data) were analyzed in another study

(Lee and Yi, 2017). To ensure the extraction of pertinent and meaningful information from the RFI dataset while minimizing irrelevant or confounding data, the following pre-processing steps were employed: (1) lowering case, (2) removing punctuation, (3) stop words removal, (4) filtering out alphanumerical patterns, (5) removing currency symbols, (6) copyright symbols removal, (7) tokenization, and (8) lemmatization.

Text Topic Clustering

Despite preprocessing efforts, not all obtained words (terms) hold equal significance on the RFI topic. Therefore, to prioritize terms based on their weights, the "Term Frequency-Inverse Document Frequency" (TF-IDF) stands out as the most representative weighting method for word prioritization (McArthur et al., 2018). TF-IDF considers the frequency of words across different documents and their frequency within each document. If a word occurs frequently in a document but is rare in the general collection, its importance within the document increases (Christopher et al., 2008). This situation forms the basis for using TF-IDF to reduce common words, focus on significant terms that differentiate documents from others, and extract key terms based on TF-IDF scores. Then, RFI texts were transformed into a TF-IDF matrix. Subsequently, this TF-IDF matrix is processed using four different clustering algorithms, and distinct key terms are identified.

Evaluation of Topic Clusters and Visualization Techniques

Identifying themes for each topic involves subjective human interpretations, necessitating domain knowledge from the building and construction industry (Lai and Kontokosta, 2019). To highlight the importance of domain experts, Lai and Kontokosta (2019) stated that they plan to gather a panel of experts, including architects and contractors, to further validate the topic model. Therefore, the efficacy of clustering methods and visualization techniques for complex RFI data analysis was assessed through expert evaluation by three quality engineers with over five years of experience. The following question was posed to the experts, "How successful are text clustering algorithms' outputs in understanding an RFI text's content (topic) and determining the keywords that reveal its content?". Similar questions are included in measurement tools, such as The Usability, Satisfaction, and Ease of Use (USE) and The System Usability Scale (SUS), which are commonly used in assessment of prototype, process, and model evaluation studies (Lund, 2001; Brooke, 1996). Results from each algorithm were evaluated on a 5-point Likert scale, ranging from most to least useful, practical, and applicable (i.e., successful).

Additionally, engineers were presented with visualization techniques tailored to each algorithm's output, rated on a 5-point Likert scale. This comprehensive approach identified the most advantageous clustering methods and visualization techniques for analyzing RFI texts.

Findings

All clustering algorithms were applied to the same dataset, forming 10 clusters with 20 distinct key terms. The researchers systematically named the topics to enhance clarity and facilitate interpretation. Tables 1-3 briefly list these keywords into topic clusters.

Within the framework of these three distinct clustering methods, certain pivotal keywords have undergone encryption due to privacy issues. This encryption is necessary because it includes project names, places, people involved, and contract details. For reference, this encryption scheme adheres to the following mapping: A*: (5 characters), B*: (5 characters), C*: (5 characters), D*: (6 characters), E*: (7 characters), F*: (8 characters), G*: (5 characters). Additionally, words with two or three characters have been included in the analysis. These short words hold significance for experts involved in the project and dealing with RFIs, serving as a distinctive feature of a known phenomenon within the project.

Evaluation of LDA Clusters

LDA is a probabilistic model where each word in a specific document can be associated with multiple topics. LDA considers each document a mixture of different topics and assumes that each word has a probability distribution among the topics in the document. Therefore, a word can be associated with multiple topics. The list of the 10 topic sets was obtained from the LDA model in Table 1 and the top 20 keywords with the highest weights that characterize each topic.

Table 1: Partial List of RFI Topics in LDA Clustering Method

Topic	Keywords
0	detail, attach, beam, see, column, connection, level, information, lift, steel
1	subrfi, rfi, A*, date, doc, response, damper, fire, room, rficir
2	ceiling, type, bracket, shutter, roller, camera, height, case, width, custom
3	lock, door, cut, provision, button, require, room, water, out, contact
4	door, schedule, use, cylinder, requirement, fire, confirm, attach, lock, handle
5	fire, report, foundation, contractor, stair, document, structural, provide, wall, include
6	change, rfi, frame, relate, B*, instruct, prab, draw, acm, core
7	lamella, sensor, dual, bliptrack, wifisensor, gate, strip, lift, type, installation
8	comment, mount, request, draw, area, equipment, revise, elevation, update, change
9	prab, show, acm, connection, sounder, design, accord, heating, floor, draw

The explanation of each topic is provided as follows:

Topic 0_Structural Details: Discusses details, attachments, beams, columns, and steel lift information.

Topic 1_RFI and Document Handling: Involves submissions, responses, dates, documents, and fire related topics.

Topic 2_Ceiling and Custom Design: Focuses on ceiling types, brackets, shutters, roller cameras, and custom designs.

Topic 3_Door and Water Requirements: Covers door related discussions, schedules, cylinders, and water requirements.
Topic 4_Door Scheduling and Confirmation: Focusses door scheduling, usage requirements, and confirmation processes.
Topic 5_Fire Safety and Foundation: Relates to fire reports, foundations, contractors, stairs, and structural documents.
Topic 6_Change Requests: Centers on changes, RFIs, frames, instructions, and core design.
Topic 7_Sensor Installations: Discusses lamellas, sensors, dual systems, WiFi sensors, gates, and lift installations.
Topic 8_Comments and Revisions: Involves comments, mounting requests, drawing areas, equipment revisions, and elevation updates.
Topic 9_Prab Design and Heating: Focuses on designs, connections, sounder designs, heating, and floor drawings.

Evaluation of NMF Clusters

NMF decomposes the given data matrix into the product of two lower-dimensional non-negative matrices. In text analysis, one matrix represents the distribution of topics across documents, and the other represents the distribution of words across topics. Due to this decomposition, a word in NMF can possess non-zero weights in multiple topics, resulting in the occurrence of the same keyword across different topics.

The list of the 10 topic sets was obtained from the NMF model in Table 2 and the top 20 keywords with the highest weights that characterize each topic.

Table 2: Partial List of RFI Topics in NMF Clustering Method

Topic	Keywords
0	subrfi, A*, doc, date, response, rfi, C*, mep, pdf, sprinkler
1	door, lock, leaf, cylinder, hardware, bk, closer, handle, schedule, D*
2	rficir, A*, shaft, doc, fcu, date, group, air, return, near
3	drawing, prab, td, acm, wall, structural, str, floor, foundation, design
4	opening, beam, existing, fit, slope, routing, pipe, designer, existing, solution
5	subrfi, A*, damper, control, supply, power, sys, date, doc, td
6	bracket, color, ral, camera, drawing, custom, shop, colour, commented, schedule
7	beam, level, connection, lift, steel, attached, E*, load, column, imd
8	total, gh, lcp, dcl, belongs, area, toilet, principle, room, follow
9	ceiling, shutter, roller, casing, height, mm, large, lamella, type, leeuw

The explanation of each topic is provided as follows:
Topic 0_RFI and Sprinkler Systems: Addresses RFIs, documents, responses, MEP, PDFs, and sprinkler.
Topic 1_Door Hardware and Schedules: Focuses on doors, locks, leaves, cylinders, hardware, and schedules.
Topic 2_RFICIR and Air Return: Discusses RFICIR (indicated the RFI which is in circulation among different stakeholders), shafts, documents, FCUs, dates, groups, air, and near returns.
Topic 3_Drawing and Structural Design: Involves drawings, walls, structural elements, floors, foundations, and designs.
Topic 4_Opening Solutions: Addresses openings, beams, existing structures, fits, slopes, routing, pipes, and design.
Topic 5_RFI and Damper Control: Covers RFIs, dampers, control, supplies, power systems, dates, and documents.
Topic 6_Brackets and Color Design: Focuses on brackets, colors, cameras, drawings, custom designs, shops, and schedules.
Topic 7_Beam Connections and Steel: Discusses beams, levels, connections, lifts, steel attachments, and columns.
Topic 8_Total Area and Ceiling Design: Addresses total areas, belongings, toilets, principles, rooms, and designs.
Topic 9_Design and Casing: Discusses specific design elements, ceiling shutters, roller casings, large lamellas, and leeuw types.

Evaluation of K-means Clusters

In addition to LDA and NMF, in the K-means model, each cluster is defined by the similarity of documents in the feature space. If a particular keyword is relevant to documents in multiple clusters, it can appear in the top keywords for each cluster.

The list of the 10 topic sets was obtained from the K-means model in Table 3 and the top 20 keywords with the highest weights that characterize each topic.

Table 3: Partial List of RFI Topics in K-means Clustering Method

Topic	Keywords
0	door, schedule, switch, bar, push, frame, ral, alarm, need, F*
1	door, lock, cylinder, leaf, bk, hardware, closer, D*, provision, handle
2	drawing, wall, attached, prab, td, acm, document, information, design, rfi
3	fit, opening, routing, existing, slope, pipe, changed, possible, beam, order
4	convector, size, card, reader, cad, lift, send, signal, riser, inspection
5	subrfi, A*, doc, date, rfi, damper, control, response, td, supply
6	beam, opening, existing, solution, ipe, steel, level, connection, attached, E*
7	ceiling, total, shutter, mep, roller, height, lighting, final, casing, type
8	rficir, A*, shaft, doc, fcu, date, group, rfi, air, near
9	lamella, mm, panel, cps, plate, new, G*, circuit, solution, damper

The explanation of each topic is provided as follows:
Topic 0_Door Schedules and Bar: Involves door schedules, switches, bars, pushes, frames, and alarms.
Topic 1_Door and Hardware: Focuses on doors, locks, cylinders, leaves, hardware, closers, and provisions.
Topic 2_Drawing and Design: Discusses drawings, attached documents, information, and design-related topics.

Topic 3_Fit and Routing Solutions: Addresses fits, openings, routings, existing structures, slopes, pipes, changes, and possible orders.

Topic 4_Convector Size and Lifts: Involves convector sizes, cards, readers, CAD, lifts, signals, and risers.

Topic 5_RFI and Supply Control: Covers RFIs, documents, dates, responses, and damper control.

Topic 6_Beam Solutions and Steel: Discusses beams, openings, existing structures, solutions, IPE, steel, levels, and connections.

Topic 7_Ceiling and Lighting Design: Focuses on ceilings, total areas, shutters, MEP, rollers, heights, lighting, final casings, and types.

Topic 8_RFICIR and Lamella Design: Addresses RFICIR, shafts, documents, FCUs, dates, groups, air, near returns, and lamella designs.

Topic 9_Circuit and New Solutions: Discusses specific elements like circuits, new panels, plates, and innovative solutions.

Comparison of All Clusters

Upon reviewing the results, the observed behavior can be attributed to the inherent differences in the assumptions and structures of the three clustering algorithms (LDA, NMF, and K-means). Each algorithm assigns unique topics based on its approach to modeling the connections between RFI documents and words, albeit with some word overlaps. For instance:

- LDA Topic_0, NMF Topic_7, and K-means Topic_6 share five common keywords (beam, level, connection, steel, attach (LDA) / attached (NMF&K-means)).
- LDA Topic_1, NMF Topic_0, and K-means Topic_5, with different weights, share 6 common keywords (subrfi, doc, date, response, rfi).
- LDA Topic_2, NMF Topic_9, and K-means Topic_7, with different weights, share 6 common keywords (ceiling, shutter, roller, case (LDA) / casing (NMF&K-means), height, type).
- LDA Topic_4, NMF Topic_1, and K-means Topic_1, with different weights, share 4 common keywords (door, lock, cylinder, handle). Additionally, leaf, hardware, bk, and closer are common only in NMF Topic_1 and K-means Topic_1.
- LDA Topic_9, NMF Topic_3, and K-means Topic_2, with different weights, share 4 common keywords (draw (LDA) / drawing (NMF&K-means), prab, acm, design).
- Although LDA exhibits similar topics, as seen in the examples above, it diverges with some topics (e.g., LDA Topic_8).

Table 4: Expert Evaluation of RFI Clustering Methods

	Expert #1	Expert #2	Expert #3	Mean
LDA	5	4	5	4,67
NMF	3	5	4	4,00
K-means	3	4	4	3,67
Mean	3,67	4,33	4,33	

The scale for understanding an RFI text's content (topic) and determining the keywords was defined as 1: very unsuccessful and 5: very successful (Table 4).

From the experts' perspective, LDA is more successful than the other two clustering algorithms in clustering RFIs. LDA can create more distinctive topics and provides a focus on micro-level clustering rather than generating general topics. When considering the characteristic features of RFIs compared to other text data used in the construction sector, such as their lengthiness, noisiness, inclusion of typos, combinations of upper- and lowercase letters, and utilization of variable terms for a single concept, the obtained results are consistent with the literature.

Ratings assigned to NMF and K-means models, except for Expert #2, are identical. This topic similarity has led to experts giving similar ratings, as these models create quite identical topics. None of the experts gave a rating below three because clustering algorithms are generally not used in current RFI analyses. Therefore, a clustering study of this kind is perceived to positively impact on understanding the RFI text's content (topic) and determining the keywords that reveal its content. In summary, even the model with the lowest score (K-means) demonstrates average success in the clustering RFI text data.

Comparison of Visualization Techniques

Multiple visualization techniques are employed in clustering problems. The pyLDAvis, specific to LDA models (See Figure 1(a)), is a Python package that empowers users to visualize LDA models effectively. This interactive tool provides features like hovering and clicking for in-depth data exploration. Essentially, pyLDAvis is a critical tool for gaining insights into the topics generated by LDA models and their distribution across documents.

The second method involves scatter plot visualization. For this method, the t-SNE (t-Distributed Stochastic Neighbor Embedding) algorithm was applied first (See Figure 1(b)). t-SNE is a dimensionality reduction technique designed to visualize high-dimensional data compactly by preserving pairwise distances between data points as much as possible in a lower-dimensional space. In essence, t-SNE rearranges data points from a high-dimensional space to a lower-dimensional one, emphasizing the preservation of relationships among similar points.

The third method visualizes topics with word clouds (See Figure 1(c)). Word clouds are highly preferred for identifying and highlighting key themes and subjects within a textual corpus.

They facilitate a quick and effortless analysis of crucial keywords in a dataset by visually displaying frequently occurring words, where the font size is directly proportional to their frequency of occurrence. These word clouds represent a set of words visually, with the most frequent words appearing larger in size and color. They are typically arranged in a circular or elliptical pattern.

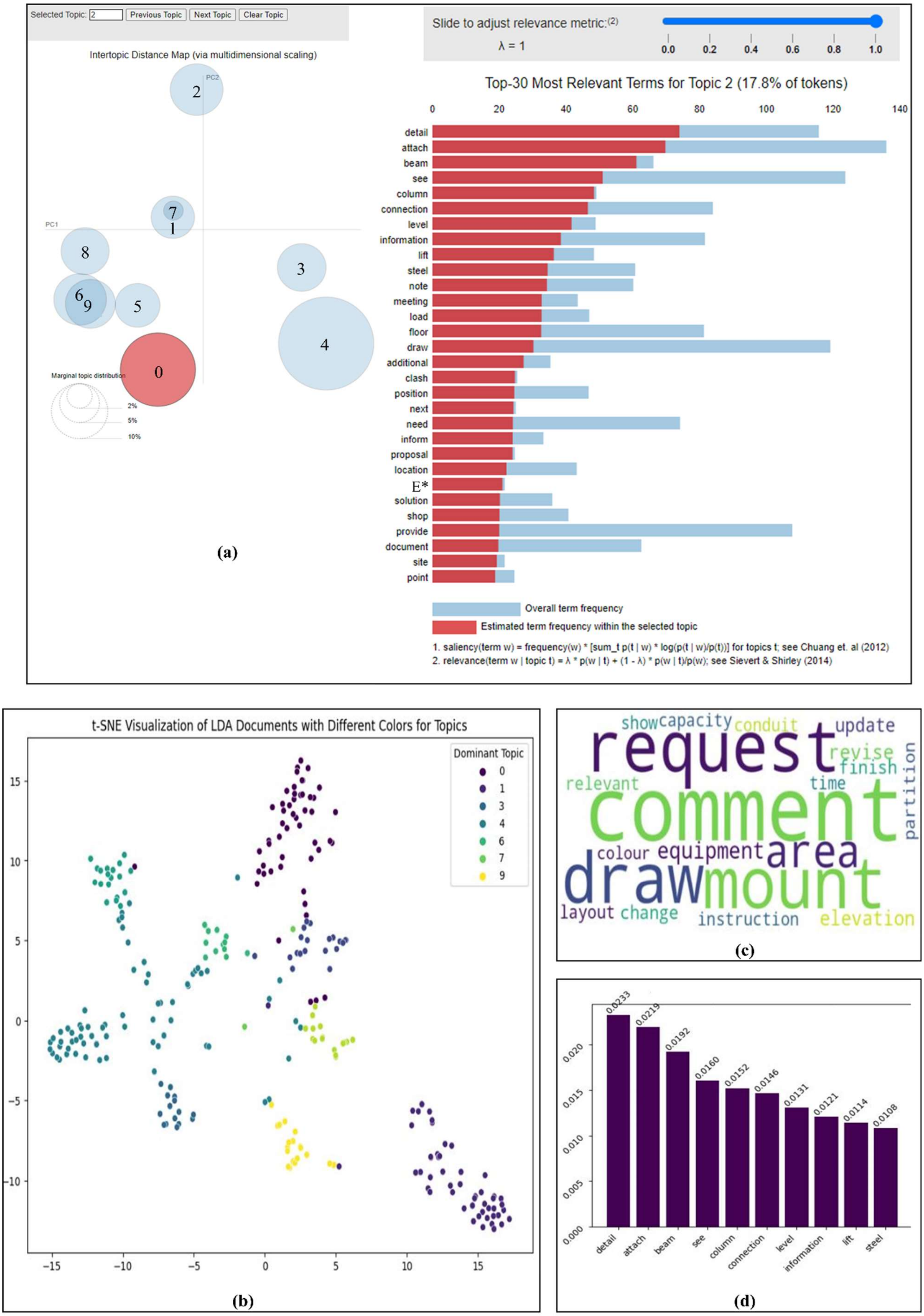


Figure 1: RFI Clustering Visualization Methods (a) pyLDAvis; (b) scatter plot; (c) word cloud; (d) bar charts

The final method utilizes bar charts (See Figure 1(d)). Bar charts are employed in statistical studies for simplicity, interpretability, distribution insight, representation of topic data, and space efficiency. This study applied visualization techniques, including scatter plots, word clouds, and bar charts, to LDA, NMF, and K-means, while pyLDAvis was exclusively used for LDA.

The experts evaluated the outputs of each model with each visualization method and were asked to rank their preferences for visualization tools used for analyzing topics and keywords. Table 5 shows the experts' preferences for visualization techniques regarding RFI topics and keywords.

Table 5: Expert Evaluation of RFI Clustering Visualization Methods

	pyLDAvis	scatter plots with t-SNE	word clouds	bar charts
Expert #1	1	4	3	2
Expert #2	1	4	2	3
Expert #3	1	4	3	2

The most preferred visualization method was identified as pyLDAvis. The experts deemed its ability to filter topic clusters with their keywords rapidly and provide interactive access to be the most useful, practical, and applicable.

Two experts ranked word clouds and bar charts second, providing data on keyword weights. While word clouds offered insights into keyword weights through font sizes, experts found the provision of numeric data more practical.

The scatter plot was evaluated as the least preferred visualization method because of its perceived lack of user-friendliness for quickly analyzing topics despite providing insights into the proximity or overlap of topic clusters.

Conclusions

In this study, 288 RFI texts retrieved from an airport project were analyzed with three different clustering algorithms (LDA, NMF, and K-means), and the results of these algorithms were compared. The generated clusters and keywords were presented using four visualization methods (pyLDAvis, scatter plots, word cloud, and bar charts). In the final stage of the study, with expert opinions, the most preferred algorithm (LDA) and visualization technique (pyLDAvis) were determined. The results of this study provide crucial insights into determining the most effective clustering algorithm and visualization technique for analyzing RFI texts in the literature. Moreover, the findings of this study can benefit project managers and quality engineers for rapidly and effectively analyzing RFI topics to develop proactive measurements for non-conformities before they arise on the construction site.

In the future, comparative analyses can be conducted by increasing the number of RFIs and including different types of projects. Furthermore, the topics in this study's results can be extended by processing them as metadata in RFI texts and analyzing them with machine learning algorithms. While the results of this study play a role in the prior work for visualizing RFIs on the BIM model, visualization and 3D querying practices within the scope of BIM are left for future studies.

References

Abbaszadegan, A., & Grau, D. (2015). Assessing the influence of automated data analytics on cost and schedule performance. *Procedia Engineering*, *123*, 3-6.

Abdirad, H. (2022). Managing digital integration routines in engineering firms: Cases of disruptive BIM cloud collaboration protocols. *Journal of Management in Engineering*, *38*(1), 05021012.

Afzal, M., Wong, J. K. W., & Fini, A. A. F. (2023, August). Unlocking Insights: Analysing Construction Issues in Request for Information (RFI) Documents with Text Mining and Visualisation. In *2023 IEEE 19th International Conference on Automation Science and Engineering (CASE)* (pp. 1-6). IEEE.

Al Qady, M., & Kandil, A. (2014). Automatic clustering of construction project documents based on textual similarity. Automation in construction, 42, 36-49.

Alshari, E., Azman, A., Doraisamy, S., Mustapha, N., & Alksher, M. (2020). Senti2vec: an effective feature extraction technique for sentiment analysis based on word2vec. Malaysian Journal of Computer Science, 33(3), 240-251.

Bhat, A. S. (2017). *Data visualization of requests for information to support construction decision-making* (Doctoral dissertation, University of British Columbia).

Brooke, J. (1996). SUS-A quick and dirty usability scale. *Usability evaluation in industry*, *189*(194), 4-7.

Buntine, W., Lofstrom, J., Perkio, J., Perttu, S., Poroshin, V., Silander, T., ... & Tuulos, V. (2004, September). A scalable topic-based open source search engine. In *IEEE/WIC/ACM International Conference on Web Intelligence (WI'04)* (pp. 228-234). IEEE.

Christopher, D. M., Prabhakar, R., & Hinrich, S. (2008). Introduction to information retrieval. *An Introduction To Information Retrieval*, *151*(177), 5.

Dantas Filho, J. B. P., Angelim, B. M., Guedes, J. P., & Neto, J. D. P. B. (2016). BIM based Request For Information classification and distribution: two residential tower cases. *PARC Pesquisa Em Arquitetura E Construção*, *7*(2), 75-88.

Egger, R., & Yu, J. (2022). A topic modeling comparison between lda, nmf, top2vec, and bertopic to demystify twitter posts. *Frontiers in sociology*, *7*, 886498.

Farhang, Y. (2017). Face extraction from image based on K-means clustering algorithms. *International Journal of Advanced Computer Science and Applications*, *8*(9).

George, S., & Vasudevan, S. (2020). Comparison of LDA and NMF topic modeling techniques for restaurant reviews. *Indian J. Nat. Sci*, *10*.

Gheeseewan, H., & Pudaruth, S. (2020). Categorisation of Computer Science Research Papers using Supervised Machine Learning Techniques. *International Journal of Computing and Digital Systems*, *9*(6), 1165-1175.

Ham, N. H., & Yuh, O. K. (2023). Performance Analysis and Assessment of BIM-Based Construction Support with Priority Queuing Policy. *Buildings*, *13*(1), 153.

Herrera, R. F., Mourgues, C., Alarcón, L. F., & Pellicer, E. (2019). Assessing design process performance of construction projects. In *Proceedings of the CIB World Building Congress* (pp. 1-10).

Jagannathan, M., Roy, D., & Delhi, V. S. K. (2022). Application of NLP-based topic modeling to analyse unstructured text data in annual reports of construction contracting companies. *CSI Transactions on ICT*, *10*(2), 97-106.

Jallan, Y., Brogan, E., Ashuri, B., & Clevenger, C. M. (2019). Application of natural language processing and text mining to identify patterns in construction-defect litigation cases. *Journal of legal affairs and dispute resolution in engineering and construction*, *11*(4), 04519024.

Kaur, N., & Aggarwal, S. (2017). Comparative analysis of hybrid k-mean algorithms on data clustering. *International Journal of Computer Applications Technology and Research,* 6(8), 384-390.

Khuzaimah, K. H. M., & Hassan, F. (2012). Uncovering tacit knowledge in construction industry: Communities of practice approach. *Procedia-Social and Behavioral Sciences*, *50*, 343-349.

Kim, J. J., Petrov, A. L., Lim, J., & Kim, S. (2022). Comparing cost performance of project delivery methods using quantifiable RFIs: cases in California heavy civil construction projects. *International journal of civil engineering*, *20*(3), 323-335.

Koc, K., Budayan, C., Ekmekcioğlu, Ö., & Tokdemir, O. B. (2024). Predicting Cost Impacts of Nonconformances in Construction Projects Using Interpretable Machine Learning. *Journal of Construction Engineering and Management*, *150*(1), 04023143.

Kohita, R., Yoshida, I., Kitamura, H., & Nasukawa, T. (2020). Interactive construction of user-centric dictionary for text analytics. https://doi.org/10.18653/v1/2020.acl-main.72.

Lai, Y., & Kontokosta, C. E. (2019). Topic modeling to discover the thematic structure and spatial-temporal patterns of building renovation and adaptive reuse in cities. *Computers, Environment and Urban Systems*, *78*, 101383.

Lee, J., & Yi, J. S. (2017). Predicting project's uncertainty risk in the bidding process by integrating unstructured text data and structured numerical data using text mining. *Applied Sciences*, *7*(11), 1141.

Lee, S. Y., Song, H. A., & Amari, S. I. (2012). A new discriminant NMF algorithm and its application to the extraction of subtle emotional differences in speech. *Cognitive neurodynamics*, *6*, 525-535.

Lin, J. R., Hu, Z. Z., Li, J. L., & Chen, L. M. (2020). Understanding on-site inspection of construction projects based on keyword extraction and topic modeling. *IEEE Access*, *8*, 198503-198517.

Liu, G., Boyd, M., Yu, M., Halim, S. Z., & Quddus, N. (2021). Identifying causality and contributory factors of pipeline incidents by employing natural language processing and text mining techniques. *Process Safety and Environmental Protection*, *152*, 37-46.

Lund, A. M. (2001). Measuring usability with the use questionnaire12. *Usability interface*, *8*(2), 3-6.

McArthur, J. J., Shahbazi, N., Fok, R., Raghubar, C., Bortoluzzi, B., & An, A. (2018). Machine learning and BIM visualization for maintenance issue classification and enhanced data collection. *Advanced Engineering Informatics*, 38, 101-112.

Morales, F., Herrera, R. F., Rivera, F. M. L., Atencio, E., & Nuñez, M. (2022). Potential Application of BIM in RFI in Building Projects. *Buildings*, *12*(2), 145.

Morchid, M., Bouallegue, M., Dufour, R., Linarès, G., Matrouf, D., & De Mori, R. (2015). Compact multiview representation of documents based on the total variability space. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *23*(8), 1295-1308.

Naik, G. R. (2016). *Non-negative matrix factorization techniques*. Heidelberg: Springer.

Prihatini, P. M., Putra, I. K. G. D., Giriantari, I. A. D., & Sudarma, M. (2017). Fuzzy-gibbs latent dirichlet allocation model for feature extraction on Indonesian documents. *Contemporary Engineering Sciences*, *10*(9), 403-421.

Rana, S., Jasola, S., & Kumar, R. (2011). A review on particle swarm optimization algorithms and their applications to data clustering. *Artificial Intelligence Review*, *35*, 211-222.

Wang, D., Fan, J., Fu, H., & Zhang, B. (2018). Research on optimization of big data construction engineering quality management based on rnn-lstm. Complexity, 2018, 1-16.

Wu, J. (2012). *Advances in K-means clustering: a data mining thinking*. Springer Science & Business Media.

EXPLORING CAUSAL LOOPS IN RESILIENT MAINTENANCE OPERATIONS OF BUILT ASSETS: A LITERATURE REVIEW

Fereshteh Ahmadi[1], Pablo Martinez[1], Barry Gledson[1], SeyedReza RazaviAlavi[1]
[1] Northumbria University, Newcastle, UK

Abstract

Maintenance operations account for a significant portion of project lifecycle costs and are subject to disruptions. A potential solution to this problem is to develop resilient strategies to cope with disruptions, ensuring safe and reliable operation of assets. This paper intends to develop a causal loop diagram to depict system downtime causes among resilient maintenance operations of built assets and represent the complex nature of relationships between maintenance outcomes and disruptions. As a result, root causes of major disruption events to maintenance operations of built assets are mapped to inform technological advances that support predictive maintenance such as simulations.

Introduction

The operation and maintenance (O&M) phase is the "longest and costliest phase" of an asset's lifecycle (Chen and Tang, 2019), which can account for up to 80% of the total cost of the asset (Cavka et al., 2015). The cost of maintenance itself can constitute 25% of the overall operation and maintenance cost (Navas et al., 2020). Therefore, improving maintenance operations can significantly reduce the total cost of built assets. 'Maintenance operations' refers to what maintenance activities should be carried out on the asset, when and how the demand for this operation is activated (Gits, 1992). In this study, maintenance operations refer to activities to be undertaken to ensure the acceptable level of functionality of an asset. Maintenance operation is a complex process due to the requirement of incorporating a variety of disciplines, knowledge, human resources (e.g., maintenance crew), tools (diagnostics and predictive technologies), information technology systems (hardware and software), financial resources and physical resource (e.g., spare parts) (BS EN 15341:2019). This complex process can be affected by different sources of uncertainty and disruptions consequently leading to significant financial losses, threatening survival of asset and unserviceability of the assets (Osei-Kyei et al., 2021). In this regard, with the increase of disruption events (man-made and natural disasters) (Dianat, et al., 2021) there is an immediate need for a resilient strategy to improve responsiveness of the asset to the unexpected events and disruptions (Ali et al., 2017). As Hosseini et al. (2016) stated, improving resilience of systems significantly raised for researchers and industries. However, literature shows that there is a lack of clarity in the maintenance operation resilient definitions, concepts, and strategies (Ali et al, 2017, Burroughs, 2017). Simulation is one of the effective methods for modelling maintenance operations and analyzing the efficiency of various scenarios for improving resilience in response to disruptions. It can model the complexity and uncertainties of maintenance operations in a risk-free environment, allowing organizations to observe a system's behavior under various circumstances. The necessary first step to developing a decision-making model for this purpose is to look at the dynamic behavior of system and potential root causes of disruptions to analyze their interactions with each other and on the system and how they work to create the phenomenon (Dianat et al., 2021). In addition, there are some causes whose sources are independent of the system itself; however, the impacts of them amplify disruptions in the system. These interactions can be illustrated via causal loops using system dynamics (SD). The SD shows how components interact throughout the system by going beyond events and searching for behavior patterns (Khorshidi et al., 2023). It also captures, simulates, and estimates the effects of policies, parameters, and components that change dynamically over time on the whole system (Khorshidi et al., 2023) and demonstrates them in a causal loop diagram (CLD).

Methodology

This study is informed by a deductive approach where the qualitative data (pre-existing theory) is analyzed to anticipate certain core concepts. Thereby, an extensive literature review is conducted to identify root causes of disruptions to maintenance operations. In order to make full sense of the findings, a start list of priori categories was generated in line with previous research using the research questions (mentioned below). Data was deductively analyzed to develop clusters of disruption sources (internal to the organization, external to the organization but internal to the network, and external to the organization). Thereafter, a system dynamics approach was used to develop a novel causal loop diagram (CLD) for demonstrating the causality (causes and effects) and interrelationships between identified disruptions, their variables, and sources. CLD is one of the SD modeling that represents the complex and nonlinear relationships between components and can be developed through gaining knowledge about a system as well as discussing and brainstorming with experts (Khorshidi et al., 2023). The CLD helps the user communicate the feedback structure (root causes of a problem) and underlying assumptions (Sushil, 1993). According to Sterman (2000) CLD can *"1) quickly capture your hypotheses about the causes of dynamics, 2) eliciting and capturing the mental models of individuals or teams; and 3) communicating the important feedback, you believe are responsible for a problem."*

Research questions:

1) What are the root causes of major disruptions to the maintenance operation of assets?

2) How do the identified events impact the maintenance operations (e.g., what aspects of maintenance operation are affected)?

3) How can the resilience of maintenance operations be improved (e.g., what capabilities are required to be improved to reduce the impact of disruptions)?

Literature review

Maintenance operation resilience

Resilience is the ability of a system to absorb, resist, adapt, and recover from disruptions (Osei-Kyei et al., 2021 and Durán et al., 2020) without interrupting the full performance of the system. If a system is substantially affected, its resilience gives it the ability to fully recover its function in the shortest possible time (Hosseini et al., 2016). In this paper, resilience of maintenance operations refers to the ability of the asset to continue its operations while subjected to disruptions. Bukowski and Werbińska-Wojciechowska (2021) developed a resilient-based maintenance support system with four subsystems including monitoring, responding, learning, and anticipating (as shown in Figure 1) to minimize the consequences of disruptions. The main task of the monitoring phase is to detect disruptive events. The response to disruptive events should be authorized, effective, and analyzed after the disruption. However, as resources (e.g., information, maintenance crew, materials, and tools) in organizations are not infinite, responses can only be prepared for a limited number of disruptive events or situations that occur frequently (Bukowski and Werbińska-Wojciechowska, 2021). The learning subsystem is based on the organization's behavior and actions in specific situations. The primary purpose of learning is to improve the organization's ability to respond, monitor, and anticipate disruption as well as changing values and criteria in the organization if required. The final phase is anticipation. The key purpose of this phase is to think and imagine outside the event horizon, conceive different possibilities, and predict what can happen in the future.

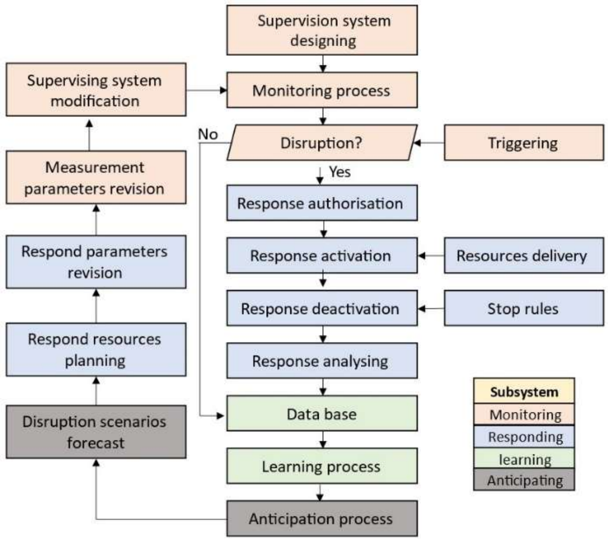


Figure 1: A functional diagram of a maintenance support system (after Bukowski and Werbińska-Wojciechowska. 2021)

Sun et al. (2022) examined another framework which is a resilience-based approach to maintenance asset and operational cost planning comprising three principal capacities: absorption, adaption, and restoration. As shown in Figure 2, 'Absorption' (R0) is the capacity of a system to withstand a disruption, absorb its consequences, and return to its original state. The strength of the absorption capacity is based on the structure of the system and the intensity of the interruption. A system with higher absorption capacity requires less effort to recover from a disruption (Sun et al., 2022). 'Adaption' (R1) is the ability of a system to recover a certain amount of lost performance without the need for external maintenance actions (Abimbola and Khan, 2019). And 'restoration' (R2) is the phase in which the system is restored to a new equilibrium state which can be lower, equal, or greater (green dotted line) than its original state by employing external maintenance actions.

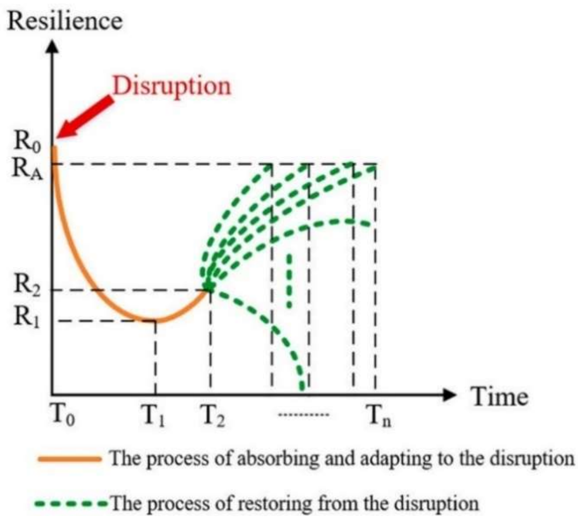


Figure 2: The resilience behaviour of a system subject to disruptions (after Sun et al., 2022)

Shekarian and Mellat Parast (2021) defined four resilience enhancers (flexibility, agility, collaboration, and redundancy) and examined the effect of each enhancer on five types of disruptions to a network (demand, supply, process, control, and environmental disruptions) and realized that adopting appropriate enhancers has great impacts on organizational performance while being affected by disruptions. To improve resilience in an organization, all the affective variables are required to be identified. In this regard, disruptions and their root causes to the maintenance operations categorized are discussed below.

Disruption to maintenance operation

Sources of disruptions to a system can be categorized into three types: 1) internal to the organization, 2) external to the organization but internal to the network, and 3) external to the network (Shekarian and Mellat Parast, 2021, Sun et al., 2022). Shekarian and Mellat Parast (2021) defined five types of disruptions to a network including process and control risks (internal to the firm), demand and supply risks (external to the firm but internal to the network) and environmental risks (external to the network) and classified supply chain disruptions into these categories as shown in Figure 3 below.

The following section discusses the sources of disruptions to the maintenance operation of built assets based on the identified categories discussed.

Maintenance process risks: process risks involve potential deviations from producing the desired quality and quantity at the right time (Shekarian and Mellat Parast, 2021). Root causes of maintenance operations process risks can be referred to lack of maintenance crews or skilled workers (e.g., labor strikes), lack of spare parts (caused by deficiency in inventory management and logistic network), adequate budgetary resources for maintenance costs, lack of diagnostic and preventive technologies tools, and interruptions to information communication technology (ICT). For example, cyber-attacks are one type of ICT disruption. In recent years, the development of digital platforms for the operation and maintenance of assets to exchange and store data has benefited organizations but it has also made them more vulnerable to cyber-attacks (Ghadiminia et al., 2022).

There are some additional sources of process risks depending on the maintenance approach adapted in the organization, such as failure in the data transmission network (failures in sensors and IoT networks), breakdown of external or internal IT infrastructure, insufficient maintenance records, inaccessibility to data in real time, and lack of decision support systems.

Control risks: "control risk or network risk involves the assumptions, rules, systems, and procedures that govern how an organization exerts control over its processes" (Shekarian and Mellat Parast, 2021). Various maintenance approaches, including corrective maintenance (unscheduled and event-driven tasks) and proactive maintenance (time-based or planned preventative maintenance, and condition-based maintenance) have different maintenance standards, regulations, and legislation (Miles et al., 2019). However, many maintenance standards have been developed in a way that they provide a common ground for a harmonized maintenance approach (Miles et al., 2019). These legislation and regulations outline the general requirements for defining all types of maintenance (e.g., EN 13306), present general recommendations for the technical documentation of maintenance (e.g., EN 13460), provide generic descriptions of maintenance process (e.g., EN 17007) among others but they lack local legislation to mitigate health, safety, and environmental (HSE) risks (Miles et al., 2019). On the other hand, compliance with such requirements is typically governed and managed by experienced system experts in the organization; therefore, lack of interorganizational policies to embed these rules is another source of control risk. Other examples of sources of this type of risk are lack of safety policies and asset management policies.

Demand risks: this type of risk involves any possible gap between actual and anticipated demand and any potential disruptions in the flow of material and information within the network or between the focal firms and the market (Shekarian and Mellat Parast, 2021).

Demand forecasting is a challenging task as the demand is intermittent and lumpy. From maintenance operations perspective, unanticipated demands (e.g., skilled workers), high demand services, uncertain maintenance demands, and insufficient information for forecasting demands are examples of demand risks.

Supply risks: these risks entail: 1) failure to supply spare parts in terms of time, quantity, and quality; and 2) disruptions to the flow of products and information within or outside the organization. Turan et al. (2020) stated that unavailability of spare parts accounts for up to 80% of all system downtime in the maintenance operations. Sources of supply risk associated with maintenance operations can be referred to insufficient logistics networks (failures in nodes (facilities) or links), poor inventory planning, and lack of outsourcing and globalization.

Environmental risks: these are risks external to the organization and are beyond the control of organizations. One of the examples of sources of environmental risk is natural disasters. As a result of climate change, natural disasters, such as floods, have occurred more frequently and intensively in the past decades (Song et al., 2016 and Feldmeyer et al., 2020). In fact, a 2% rise in the annual incidence of natural disaster was documented during the past 15 years and built assets are one of the most vulnerable areas to be affected by natural disasters (Bang and Burton, 2021). Other examples include war, epidemics (e.g., COVID-19), and political instability.

In addition to the previously described disruptions, evaluating 'co-occurring' or 'compounding failures' (when two or more sources of disruption occur

Supply chain sources of risks				
Internal to the firm		External to the firm but internal to the supply chain network		External to the network
Process risk	Control risk	Demand risk	Supply risk	Environment risk
Machine failure	Lack of collaborative planning	Volatile demand	Outsourcing and globalisation	Natural disaster
Labor strike	Asset management policy	Market changes	Sudden hike in costs	Terrorism and war
Product quality problem	Safety stock policy	Innovation competitors	Supplier commitment	Political instability
Equipment unreliability	Batch size or order quantity policy	Unanticipated demand	Supplier insolvency	Social and political grievance
Operator unavailability	Transportation management policy	Unusual customer payment delays	Variability of replenishment lead time	Technology changes
Bottleneck or inflexible process		Competition changes	Supplier quality problem	Diseases or epidemics
Breakdown of external or internal IT infrastructure	Asymmetric power relationships	Forecasting errors	Supplier bankruptcy	Economic downturn
			Sudden supplier demise	
Reliability of supporting communication system	Poor visibility along the supply chain	Insufficient information from customer order	Poor logistics performance of suppliers	
			Capacity fluctuations or shortage in supply market	

Figure 3: Supply chain sources of risks and disruptions (after Shekarian and Mellat Parast, 2020)

concurrently) and 'cascading failures' (when a disruption happens after the initial failure (horizontal-correlated cascading failures) or when a disruption can cause failure at the upper layers of the system (vertical-correlated cascading failures) (Moffatt et al., 2021) is crucial because they have high impacts on the maintenance operations and are more difficult to fix.

Due to the variety of sources of disruptions to the maintenance operations, developing strategies to mitigate disruptions by making the maintenance operations more resilient and responsive is crucial. Access to reliable and quality data is critical for achieving this aim and developing a simulation-based decision support system.

Technologies and disruptions

There are two types of disruptions to maintenance operations: 1) 'anticipated' disruptions, like demand risks, which can be predicted based on historical data related to their nature, range, and frequency, and 2) 'unanticipated' disruptions like environmental disruptions, which are not dependable or consistent with historical data (Tsiamas and Rahimifard, 2021).

In recent years, integrating BIM (building information modeling), GIS (geographic information system), IoT (Internet of Things) and computerized maintenance management system (CMMS) such as IBM Maximo, ARCHIBUS, EcoDomus, FM Systems, AssetWorks, or eMaint among others has supported maintenance operations by improving decision-making (Ma et al., 2020). These approaches use asset life-cycle data and maintenance records to detect possible failures (Moradi et al., 2021). For example, BIM can be used to minimize maintenance processes and supply disruptions by providing equipment location data, equipment maintenance data, cost data, and historical maintenance records. Also, it can be adopted to deal with control disruptions by increasing collaboration between participants and identifying potential risks. IoT and machine learning tools can be adapted to minimize supply and maintenance process disruptions, as they can predict when an asset or equipment requires maintenance actions (before breakdown point) and provide an accurate assessment of equipment health conditions based on data collected by sensors (Shamayleh et al., 2020).

Digital twins, through the synergy of several technologies including IoT, artificial intelligence (AI) and BIM integrate multi-source data into one single system to support decision-making for maintenance processes. AI can be adapted to improve maintenance demand planning with identification of demand pattern, market trend, and efficient forecasting. However, in order to take advantage of AI, clean and quality data, sufficient technological infrastructure, resource availability, and area experts are necessary. This will include supplying spare parts.

Results and discussion

The initial contribution to this research is to define the sources of disruptions to the maintenance operations and propose a classification based on these criteria. The classification includes 'Internal disruptions' (process risk, control risks), 'External to the organization but internal to the network' (demand risks, supply risks) and 'External disruptions' (environmental risks) to the maintenance operations. The connections between sources of disruptions and their root causes to improve organizational responses are mapped. To achieve this, SD is adapted due to its capability of capturing the dynamic behavior of complex systems and depicting the dynamic interactions of maintenance operations in different areas such as economic, risks, humanity, environmental, and supply chain. Causal loop diagrams in SD are employed for developing a hypothetical and knowledge based CLD (see Figure 4) to demonstrate maintenance operation system structure, complexity, and feedback processes.

In CLD, each box is a variable and causal relationships between each variable are represented as arrows with polarities (positive and negative signs). In this context, a positive sign means that a change in one variable led to a change in the same direction, and a negative sign led to a change in the opposite direction. The causal loops can be positive (reinforcing) or negative (balancing). "A positive loop is associated with exponential growth. However, a

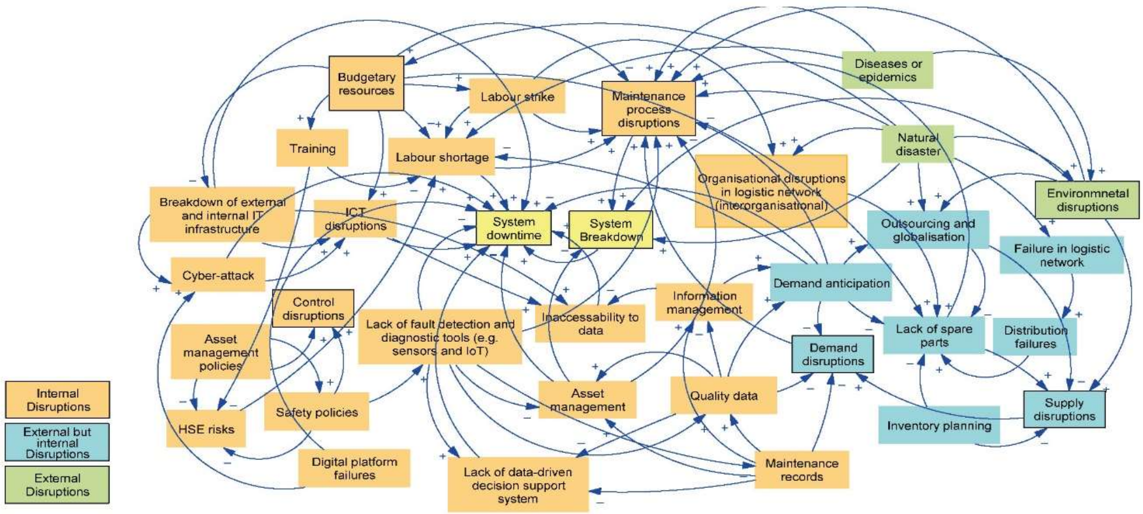


Figure 4: Causal loop diagram - root causes of disruptions to the maintenance operations.

negative loop tends to reach a balance point" (Khorshidi et al., 2023). Color coding is adapted to distinguish different types of disruptions (internal, external, external but internal) and their interconnected factors.

The advantage of CLD is that it can connect non-linear relationships between different sources and varieties and present casualty of system breakdown or system downtime. As it can be seen in Figure 4, the relationship between maintenance process disruptions and training is not a linear relationship. For example, budgetary resources have a positive impact on training; the higher the budget, the more training is conducted. So, training has a negative impact on maintenance process disruptions through budgetary resources, which is a dynamic variable.

The proposed CLD requires it to be tested empirically and validated by experts in the future. Once the CLD is validated, it can be transformed and formulated for the model and run for results. By adopting simulation, different recovery strategies on maintenance operations performance affected by disruptions can be analyzed. To achieve this, different scenarios in terms of the presence or absence disturbances and presence or absence of a mitigation strategy can be examined. The impact of mitigation strategies will be analyzed in regard to time and cost of recovery.

The CLD aims to provide a base for the development of accurate simulation models in the context of maintenance disruptions. In this context, the CLD can be used to ensure that the all the relevant and impactful data on the simulated variable is considered in the simulation. For example, let's envision creating a simulation model that aims to determine the impact of disruptions on 'spare parts supply chain'. Following the CLD across the multiple links that connect 'Supply Disruptions' with the rest of the variables, one can ensure that all the relevant variables are included in the model. On the first instance, 'outsourcing and globalization', 'logistic networks', 'lack of spare parts', 'inventory planning' and 'environmental disruptions' are connected, then the variables linked to those need to be considered and so on. With the CLD in mind, a comprehensive list of variables required to accurately simulate an outcome from a major disruption is given. However, the importance and impact of each variable on the outcome is not provided on the CLD and simplification of simulation models from the initial list of variables is a work in progress.

Conclusion and further study

This study presents a review of the literature to explore potential root causes of disruptions to the maintenance operations of built assets to improve maintenance operations resilience. Due to the dynamic nature of this system, system dynamics theory is used to demonstrate relationships and feedback loops between related factors to enhance understanding of the system complexity and non-linear causality. It also shows the need of multidisciplinary resilience approaches (current research gap). As a result, academics and practitioners are better equipped with the technological advances that support predictive maintenance, such as simulations, for improving resilience of maintenance operations in the practices. It is mentionable that practices required some practical changes for adopting resilience policies that necessitated a rethinking and adaptation of new governance approaches.

The limitations for this study are: 1) there is not much available literature about disruptions to maintenance operations as well as improving resilience in maintenance operations (research gap); and 2) the literature review is secondary data, not primary data: therefore, quality and accuracy of data is limited.

By understanding the current gaps, future research will develop a digital simulation-based decision-support system (based on the proposed CLD) to improve resilience of maintenance operations of built assets. Survey research including expert interviews will be conducted to 1) analyze the intensity and frequency of experienced disruptions in the medium-to-large size organizations, 2) explore organizational mitigation plans to recover from the disruptions, 3) evaluate the efficiency of the mitigation strategies adapted, 4) evaluate resilient capabilities in the organization absorption, adaption, and restoration.

Also, given that most of the research has focused on the pre-disruption stage, future studies are recommended that also investigate the post-disruption stage and its management.

References

Abimbola, M., & Khan, F. (2019). Resilience modeling of engineering systems using dynamic object-oriented Bayesian network approach. Computers & Industrial Engineering, 130, 108-118. https://doi.org/10.1016/j.cie.2019.02.022

Ali, A., Mahfouz, A., & Arisha, A. (2017). Analysing supply chain resilience: Integrating the constructs in a concept mapping framework via a systematic literature review. Supply Chain Management: An International Journal, 22(1), 16-39. https://doi.org/10.1108/scm-06-2016-0197

Bukowski, L., & Werbińska-Wojciechowska, S. (2021). Using fuzzy logic to support maintenance decisions according to resilience-based maintenance concept. Eksploatacja i Niezawodność – Maintenance and Reliability, 23(2), 294-307. https://doi.org/10.17531/ein.2021.2.9

Burroughs, S. (2017). Development of a tool for assessing commercial building resilience. Procedia Engineering, 180, 1034-1043. https://doi.org/10.1016/j.proeng.2017.04.263

Cavka, H., Staub-French, S., & Pottinger, R. (2015). Evaluating the alignment of organizational and project contexts for BIM adoption: A case study of a large owner organization. Buildings, 5(4), 1265-1300. https://doi.org/10.3390/buildings5041265

Chen, C., & Tang, L. (2019). BIM-based integrated management workflow design for schedule and cost planning of building fabric maintenance. Automation in Construction, 107, 102944. https://doi.org/10.1016/j.autcon.2019.102944

Dianat, H., Wilkinson, S., Williams, P., & Khatibi, H. (2021). Planning the resilient city: Investigations into using "causal loop diagram" in combination with "UNISDR scorecard" for making cities more resilient. International Journal of Disaster Risk Reduction, 65, 102561. https://doi.org/10.1016/j.ijdrr.2021.102561

Durán, O., Aguilar, J., Capaldo, A., & Arata, A. (2020). Fleet resilience: Evaluating maintenance strategies in critical equipment. Applied Sciences, 11(1), 38. https://doi.org/10.3390/app11010038

Feldmeyer, D., Wilden, D., Jamshed, A., & Birkmann, J. (2020). Regional climate resilience index: A novel multimethod comparative approach for indicator development, empirical validation and implementation. Ecological Indicators, 119, 106861. https://doi.org/10.1016/j.ecolind.2020.106861

Ghadiminia, N., Mayouf, M., Cox, S., & Krasniewicz, J. (2021). BIM-enabled facilities management (FM): A scrutiny of risks resulting from cyber attacks. Journal of Facilities Management, 20(3), 326-349. https://doi.org/10.1108/jfm-01-2021-0001

Gits, C. (1992). Design of maintenance concepts. International Journal of Production Economics, 24(3), 217-226. https://doi.org/10.1016/0925-5273(92)90133-r

Hosseini, S., Barker, K., & Ramirez-Marquez, J. E. (2016). A review of definitions and measures of system resilience. Reliability Engineering & System Safety, 145, 47-61. https://doi.org/10.1016/j.ress.2015.08.006

Khorshidi, H. A., Marshall, D., Goranitis, I., Schroeder, B., & IJzerman, M. (2023). System dynamics simulation for evaluating implementation strategies of genomic sequencing: Tutorial and conceptual model. Expert Review of Pharmacoeconomics & Outcomes Research, 24(1), 37-47. https://doi.org/10.1080/14737167.2023.2267764

Ma, Z., Ren, Y., Xiang, X., & Turk, Z. (2020). Data-driven decision-making for equipment maintenance. Automation in Construction, 112, 103103. https://doi.org/10.1016/j.autcon.2020.103103

Maintenance process and associated indicators. (n.d.). https://doi.org/10.3403/30340662

Maintenance. Documentation for maintenance. (n.d.). https://doi.org/10.3403/30163968

Maintenance. Maintenance key performance indicators. (n.d.). https://doi.org/10.3403/30140422

Maintenance. Maintenance terminology. (n.d.). https://doi.org/10.3403/30187553

Moffatt, J., Zaitouny, A., Hodkiewicz, M. R., & Small, M. (2021). Detecting asset cascading failures using complex network analysis. IEEE Access, 9, 120624-120637. https://doi.org/10.1109/access.2021.3108427

Moradi, P., Asadi, M. J., Ebrahimzadeh, N., & Yarahmadi, B. (2021). Ilam tunnels inspection, maintenance, and rehabilitation: A case study. Tunnelling and Underground Space Technology, 110, 103814. https://doi.org/10.1016/j.tust.2021.103814

Navas, M. A., Sancho, C., & Carpio, J. (2020). Disruptive maintenance engineering 4.0. International Journal of Quality & Reliability Management, 37(6/7), 853-871. https://doi.org/10.1108/ijqrm-09-2019-0304

Ngenyam Bang, H., & Church Burton, N. (2021). Contemporary flood risk perceptions in England: Implications for flood risk management foresight. Climate Risk Management, 32, 100317. https://doi.org/10.1016/j.crm.2021.100317

Osei-Kyei, R., Tam, V., Ma, M., & Mashiri, F. (2021). Critical review of the threats affecting the building of critical infrastructure resilience. International Journal of Disaster Risk Reduction, 60, 102316. https://doi.org/10.1016/j.ijdrr.2021.102316

Shamayleh, A., Awad, M., & Farhat, J. (2020). IoT based predictive maintenance management of medical equipment. Journal of Medical Systems, 44(4). https://doi.org/10.1007/s10916-020-1534-8

Shekarian, M., & Mellat Parast, M. (2020). An integrative approach to supply chain disruption risk and resilience management: A literature review. International Journal of Logistics Research and Applications, 24(5), 427-455. https://doi.org/10.1080/13675567.2020.1763935

Song, X., Zhang, Q., Sekimoto, Y., Shibasaki, R., Yuan, N. J., & Xie, X. (2016). Prediction and simulation of human mobility following natural disasters. ACM Transactions on Intelligent Systems and Technology, 8(2), 1-23. https://doi.org/10.1145/2970819

Sun, H., Yang, M., & Wang, H. (2022). Resilience-based approach to maintenance asset and operational cost planning. Process Safety and Environmental Protection, 162, 987-997. https://doi.org/10.1016/j.psep.2022.05.002

Tsiamas, K., & Rahimifard, S. (2021). A simulation-based decision support system to improve the resilience of the food supply chain. International Journal of Computer Integrated Manufacturing, 34(9), 996-1010. https://doi.org/10.1080/0951192x.2021.1946859

Turan, H. H., Atmis, M., Kosanoglu, F., Elsawah, S., & Ryan, M. J. (2020). A risk-averse simulation-based approach for a joint optimization of workforce capacity, spare part stocks and scheduling priorities in maintenance planning. Reliability Engineering & System Safety, 204, 107199. https://doi.org/10.1016/j.ress.2020.10719

DEVELOPMENT OF SIMEC-OPT FRAMEWORK FOR OPTIMIZING GLAZING PARAMETERS ENHANCING RESIDENTIAL BUILDING THERMAL COMFORT AND ENERGY PERFORMANCE

Bhavya Gonuguntla, Tripti Singh Rajput, Omprakash Ramalingam Rethnam, and Albert Thomas
Indian Institute of Technology Bombay, Mumbai, India
Corresponding author, e-mail address: albert@iitb.ac.in

Abstract

While glazed buildings have become a symbol of modern construction with their aesthetic appeal, there is a lack of research specifically focusing on the influence of glazing parameters on energy consumption and occupant comfort in residential buildings, especially in tropical climate zones. The study is focused on developing a simulation-based framework for optimizing the glazing parameters of multi-storied residential buildings in Indian climate zones to enhance thermal comfort and energy performance using a baseline model in Mumbai, India. Optimal configurations, identified through the study, suggest potential improvements of up to 30% in energy efficiency and 8% in thermal comfort compared to a baseline design considered in the study. The findings aim to inform architects, builders, and policymakers about tailored glazing configurations in various Indian climates.

Introduction

As global carbon emissions attributed to buildings reach nearly 42 percent, the emphasis on sustainable and energy-efficient building design has become paramount (Architecture 2030, 2023). The key to achieving sustainability in the building sector is the exploration of energy consumption patterns so as to take up energy-conserving measures for the same. An assessment by the Bureau of Energy Efficiency (BEE) shows that in India, lighting and air conditioning use 80 percent of the energy in commercial buildings, whereas fans, lighting, and refrigerators use maximum energy in residential buildings (Kaja, 2015). The energy demands of a building are significantly influenced by its envelope which serves as the primary connection between the external environment and the interior spaces. The architectural world is witnessing a surge in the popularity of providing more glass as a major material in the building envelope (Ayyad, 2011). Glass offers transparency and a sleek, contemporary look that enhances the architectural appeal of structures. Glass allows natural light to enter buildings, positively impacting occupant well-being as exposure to natural light has been linked to improved state of mind, increased alertness, and enhanced productivity (Golmohammadi et al., 2021). As countries such as India experience urbanization and economic growth, there is a desire to shift to contemporary architectural styles that incorporate more glass in the envelopes. However, when it comes to building energy efficiency, the choice and application of glass can significantly influence the outcome. One of the primary challenges with glass in tropical climates is excessive heat gain. Tropical climates typically experience intense and prolonged sunlight which can lead to a greenhouse effect inside buildings, causing them to overheat. This, in turn, requires more energy for air conditioning and cooling, leading to increased energy costs and carbon emissions (Sayed and Fikry, 2019). Hence, the performance of selected glass is of paramount importance in controlling energy consumption.

The process of selecting glass has become increasingly intricate due to the wide array of glass options available, ranging from performance-oriented attributes to aesthetic considerations. Usta and Zengin (2022) performed energy modeling for an office building, considering four distinct types of glazing, to assess the influence of various window glazing properties on the building's energy performance while maintaining consistent indoor thermal and visual comfort conditions. The findings of the study demonstrated that the utilization of appropriate glazing materials could lead to a substantial reduction in energy consumption, specifically about 25% (Usta et al., 2022). Though the study showcased significant improvement in energy consumption, the lack of validation of the model raises concerns about the reliability of the simulated results. Further, the methodology outlined in the study appears to have a limited scope regarding glazing variations as it considers only four types of glazing models. Khalaf et al. (2019) investigated the impact of façade and shading systems on the energy performance of school buildings in Istanbul, emphasizing the need for a balance among lighting energy, heating, and cooling energy. The authors presented a case study comparing traditional clear double-layer windows with various glazing and shading alternatives and revealed a significant energy consumption reduction when using proper glazing and shading systems. However, the authors overlooked the user comfort and the potential effect of glazing and shading type on the productivity of students (Khalaf et al., 2019). Kumar et al., (2018) investigated the thermal performance of commercial and residential buildings constructed with various building materials and window glass types in five different climatic zones of India. The findings demonstrated that, among the eighty building models analyzed, the building featuring a mud brick wall and south-facing bronze-reflective glass windows exhibited the most significant energy savings in terms of minimizing heat gain. While this study made valuable strides in understanding energy efficiency in residential buildings, it beckons further considerations to understand the broader impact of glazing choices. With the increasing prevalence of remote work, residential spaces have become multifunctional and serve as both living and working areas, enhancing the significance of factors such as productivity and comfort in residential buildings. While the above studies make significant contributions to the aspect of energy efficiency, the impact of different glazing options on occupant comfort in residential environments remains underexplored. ASHRAE (American Society of Heating, Refrigerating, and Air-Conditioning Engineers) defines thermal comfort as the conditions in which occupants of a space feel satisfied with their thermal environment (ASHRAE Standard 55). Further, due to the intricate nature of building models, computer simulations emerge as an efficient, comprehensive, and highly accurate method. Naji et al., (2021) used a multi-objective optimization approach using TRNSYS (Transient System Simulation Tool) and EnergyPlus to enhance the envelope components of prefabricated houses across six distinct climate zones in Australia. Optimization with multiple objectives involves the simultaneous consideration of conflicting goals, such as minimizing energy consumption while maximizing occupant comfort. This requires evaluating trade-offs between different design options to identify the most optimal solution that satisfies multiple criteria. The authors in this study considered energy efficiency, thermal comfort, and daylight illuminance as objectives. The results demonstrated significant improvements in various aspects of building performance compared to a baseline design. Compared to a baseline design, the life cycle cost of the optimal solution was reduced by 27-31%, indicating cost savings over the building's lifespan. Further, thermal discomfort hours were reduced by 6-55% (Naji et al., 2021). In a similar study by Ran et al. (2023), DesignBuilder was used, which employs a non-dominated sorting genetic algorithm (NSGA-II), to optimize building energy, comfort, and cost. The optimal solution reduced energy use by 21.1%, and improved thermal comfort by 32.4%, but increased initial investment cost by 35.3% for a reference building in Beijing. However, the above studies focus on various construction aspects, such as building materials, orientation, and insulation, rather than explicitly addressing the direct impact of glazing types. The studies also oversee the importance of validating building models with real-time data to authenticate consumption patterns and occupant behavior. Bridging this gap in literature would provide valuable insights for architectural regulations and building design, especially in the context of Indian climate zones. India is a vast country with different climate patterns and temperature ranges, from the tropical conditions in the south to the colder climates in the north. With residential buildings already accounting for a significant portion of India's energy consumption i.e., 20.4% projections suggest that by 2032, this percentage will increase sevenfold to reach 36.5% emphasizing the need for more sustainable practices in residential buildings (BEE, 2021). Addressing this need, the study is focused on the selection of optimum glazing type and Window-to-Wall ratio (WWR) that help build energy-efficient and thermally comfortable residential buildings.

Methodology

The study employs a simulation-based framework – SIMEC-Opt, illustrated in Figure 1, to find the optimum glazing configuration that reduces heat gains and improves thermal comfort. The building is modeled in DesignBuilder, including detailed information regarding materials, characteristics of HVAC systems, schedules, and occupant behavior. Further, to understand the impact of glazing, an essential aspect involved the identification of glazing properties that will define glazing configurations. Glazing parameters such as U-value and Solar Heat Gain Coefficient (SHGC) are defined from glass types available in the Indian market. Next, optimization is carried out with the objectives of minimizing electricity consumption and maximizing thermal comfort by considering varied glazing options, and WWR to identify optimal combinations for different climate zones. The climate of India has been classified into five climate zones by BEE (BEE, 2021). BEE, in its Handbook of Replicable Residential Designs, has identified five representative cities corresponding to these climate zones. The five climate zones and their associated cities are detailed in Table 1. These 5 cities have been selected as they are a fair representation of the climate characteristics of the zone and have a high residential development potential in the coming years (Handbook of Replicable Residential Designs, 2021). Weather data files are obtained from climate.onebuilding in the EnergyPlus Weather (EPW) format. These EPW files provide comprehensive meteorological data necessary for building energy simulations, including parameters such as temperature, humidity, wind speed, solar radiation, and precipitation (climate.onebuilding).

Table 1: Climate Zones of India and their representative city

Climate Zone	City
Hot and Dry	Aurangabad
Temperate	Bengaluru
Warm and Humid	Bhubaneshwar
Composite	Lucknow
Cold	Srinagar

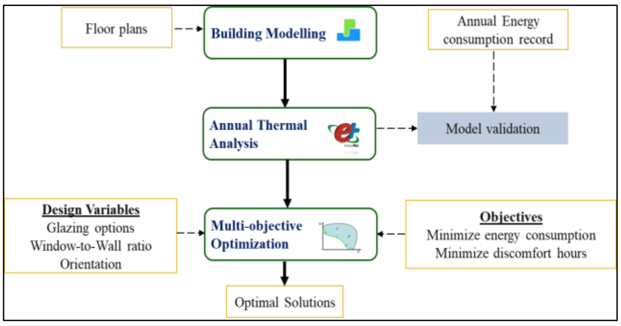


Figure 1: SIMEC-Opt Framework

Energy consumption is quantified by assessing the building's annual lighting, cooling, and heating demands, while thermal comfort is evaluated by calculating annual discomfort hours. Discomfort hours refer to the duration of time during which the indoor environmental conditions within a building fall outside the range considered comfortable for occupants. The criteria are often aligned with industry standards for acceptable indoor temperature and humidity levels. Discomfort hours in DesignBuilder are aligned with ASHRAE Standard 55, which defines the conditions for thermal comfort in indoor environments (DesignBuilder Software Ltd).

Modelling

Model Description

DesignBuilder is chosen as the tool for modeling and analysis in this study as it offers flexible geometry input, extensive material libraries, and load profiles and utilizes EnergyPlus as its simulation engine (DesignBuilder Software Ltd). Using the original drawing files, a residential quarter is replicated in the DesignBuilder software. This building with its actual glazing configuration and WWR, serves as the base model for this study, representing a typical apartment complex in India. In practical applications, the number of windows and building materials differ across geographical locations. Since the present research is designed to assess the impact of various glazing systems, the same building model is employed across all locations, ensuring consistency and comparability in the study. The building was originally located in Mumbai, a city characterized by warm and humid climate conditions, as per the BEE classification (BEE, 2021).

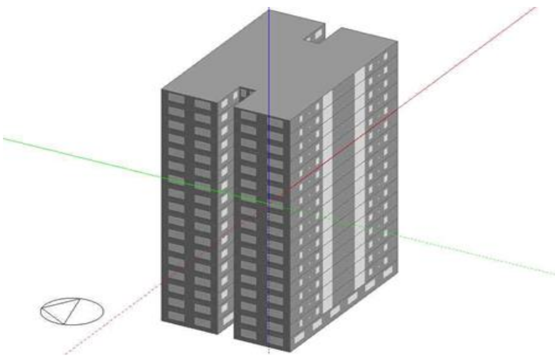


Figure 2: Model of the building

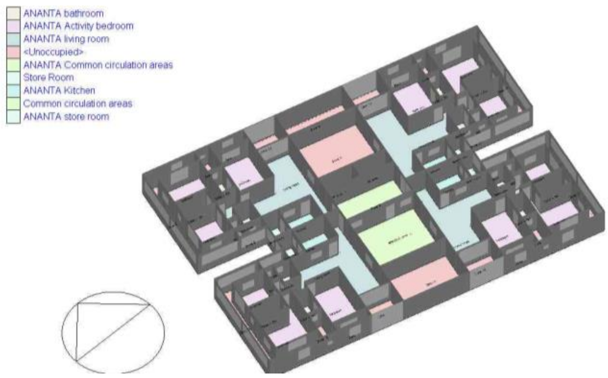


Figure 3: Floor plan of the building

Table 2: Specifications of the building envelope

Component	Layers	Material	U-value (W/m^{2}K)
External Wall	3	1.25 cm cement plaster + 22.5 cm brick + 1.25cm cement plaster	2.192
Internal Wall	3	1.25 cm cement plaster + 20 cm brick + 1.25cm cement plaster	2.335
Roof	2	1 cm RCC + 1 cm lime concrete	2.939
Floor	1	10mm ceramic tiles + 50 mm cement screed + 150 mm concrete floor slab	2.562
Glazing	1	Single clear glass 6 mm	5.88

Table 3: Zone-wise HVAC details of the building

Zones	Ventilation	Lighting	Cooling	Heating
Bedroom 1	Mixed mode	On	On	On
Bedroom 2	Mixed mode	On	On	On
Bedroom 3	Mixed mode	On	Off	Off
Living Room	Mixed mode	On	Off	Off
Kitchen	Mixed mode	On	Off	Off
Storeroom	Mixed mode	On	Off	Off
Bathroom 1	Mechanically ventilated	On	Off	Off
Bathroom 2	Mechanically ventilated	On	Off	Off
Bathroom 3	Mechanically ventilated	On	Off	Off

The building comprises a total of sixteen floors, with each floor accommodating four individual units. These units are equipped with three bedrooms, a kitchen, a store room, three bathrooms, and a living room. Additionally, the ground floor of the building is designated for parking. A visual representation of the building and the floor plan for each level is provided in Figure 2 and Figure 3, respectively. Detailed specifications of the building's envelope and the U-value of components are outlined in Table 2 (ECBC, 2017). The model is equipped with two-pane sliding windows (50% openable area) of aluminum frame and single-glazed clear glass of 6 mm thickness with SHGC of 0.88 and a U-value of 5.8 W/m^{2}K.The HVAC systems in each zone within the apartments are outlined in Table 3. The building follows a specific operational schedule for each zone, considering that occupants are outside and have office working hours starting at 9 am to 6 pm hence occupants are considered active at the house only in hours other than office hours on weekdays and active all day on weekends (Rethnam and Thomas, 2023).

Model Validation

To validate the model, the monthly electricity consumption data for the building is collected from the regional Estate office and then compared with the simulated energy consumption for the Mumbai location. DesignBuilder uses detailed information about the building, including its geometry, materials, HVAC systems, schedules, occupant behavior, and weather data for the required location to simulate annual building performance. The actual annual energy consumption for the building is given to be 223922 kWh, while the simulated value stood at 210065.6 kWh, resulting in a difference of 6.2%, which lies in the acceptable range of error for building energy simulation models (Xu et al., 2017).

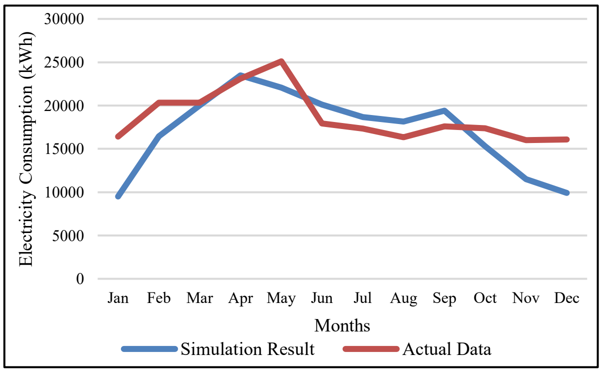


Figure 4: A graphical comparison of monthly total energy consumption between simulated and actual data

Figure 5 below provides a graphical representation of the monthly comparison between simulated and actual energy consumption (kWh). Furthermore, discomfort hours are also calculated by DesignBuilder, in compliance with ASHRAE guidelines, by running detailed energy simulations. The software considers various factors such as indoor temperatures, humidity, airspeed, and other environmental conditions over a year. It uses this data to quantify the duration of time when indoor environmental conditions deviate from the ASHRAE-defined comfort range. Annual discomfort hours of the building are determined to be 1048.46 hours out of 8760 hours annually after the model validation.

Optimization Configuration

Furthermore, the validated model of the building is again simulated for annual energy performance and thermal comfort with initial glazing configuration and WWR for all the representative cities. To find the best design options in each climate zone, the optimization is carried out in the next step considering design variables, glazing configuration, and WWR. Glazing options are provided in Table 4.

Table 4: Glazing types involved in parametric analysis

Glazing options	SHGC	U-value (W/m^{2}K)
Glazing 1	0.24	1.6
Glazing 2	0.53	1.8
Glazing 3	0.34	1.8
Glazing 4	0.3	3.7
Glazing 5	0.14	3.6
Glazing 6	0.41	4.8
Glazing 7	0.47	5.7
Glazing 8	0.69	5.7
Glazing 9	0.49	5.8
Glazing 10	0.58	5.70
Glazing 11	0.42	4.60
Glazing 12	0.34	4.00
Glazing 13	0.45	5.30
Glazing 14	0.33	5.00
Glazing 15	0.44	5.80
Glazing 16	0.38	1.70
Glazing 17	0.59	5.40
Glazing 18	0.29	1.70
Glazing 19	0.21	1.70
Glazing 20	0.38	5.70

For WWR, the range is set from a maximum of 80% to a minimum of 20%, with steps of 5% increments. This means that WWR values are considered at 80%, 85%, 70%, 75%, and so on, down to 20%. The in-built module "Optimization" of DesignBuilder is utilized for conducting multi-objective optimization in this study. Further, objective functions, which quantify the goal as minimizing energy consumption and minimizing discomfort hours, are defined. The optimization process involves an advanced evolutionary algorithm, NSGA- II (Non-dominated Sorting Genetic Algorithm II). As depicted in Table 5, a Simulated Binary Crossover (SBX) with a crossover probability of 0.9 and a mutation probability of 0.1 is used in the optimization process, which includes a population size of 100 individuals. Selection is done using a Binary tournament, and optimization is carried out for a 1000-generation termination criterion to reach optimal design solutions. Finally, NSGA II finds a set of Pareto-optimal solutions in a Pareto front (where no solution is better than the other), which represent trade-offs between conflicting objectives, energy consumption, and discomfort hours (Singh Rajput and Thomas, 2023).

Table 5: NSGA II parameters

Parameter	Value
Population size	100
Crossover Type	Simulated Binary Crossover
Mutation probability	0.1
Crossover probability	0.9
Selection process	Binary Tournament
Termination criterion	1000 generations

Results

Baseline configuration simulation results

Table 6: Baseline configuration simulation details

City	Electricity Consumption (kWh)	Discomfort Hours (hours)
Aurangabad	166638.50	828.55
Bengaluru	142195.60	1181.64
Bhubaneshwar	210194.40	1012.73
Lucknow	184967.80	1033.54
Srinagar	117669.62	1346.13

Table 6 showcases the simulation outcomes for the annual energy performance and thermal comfort, considering the initial glazing configuration and WWR for the representative cities across the five climate zones in India. In the base case simulation results, Bhubaneshwar has electricity and discomfort hours similar to Mumbai's. It could be attributed to the fact that they both belong to warm and humid climates. Bhubaneshwar also has the highest electricity consumption and discomfort hours, followed by Lucknow, which has a composite climate. Srinagar is last, followed by Bengaluru, which has a temperate climate.

Optimization Result

The optimization analysis aimed at minimizing the annual electricity consumption (in kWh) and minimizing the annual discomfort hours (in hours) involving design variables, WWR, and various glazing options, for each representative city, led to a specific number of iterations to find optimal solutions as detailed in Table 7.

Table 7: Number of iterations and optimal solutions of each representative city

City	Iterations	Number of Optimal Solutions
Aurangabad	196	19
Bengaluru	242	62
Bhubaneshwar	200	30
Lucknow	213	32
Srinagar	242	54

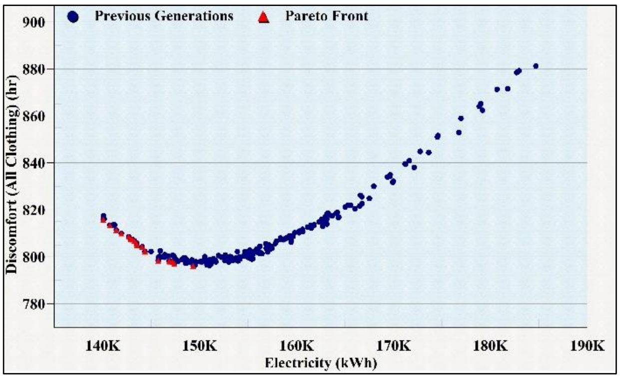


Figure 5: Pareto graph result of Aurangabad

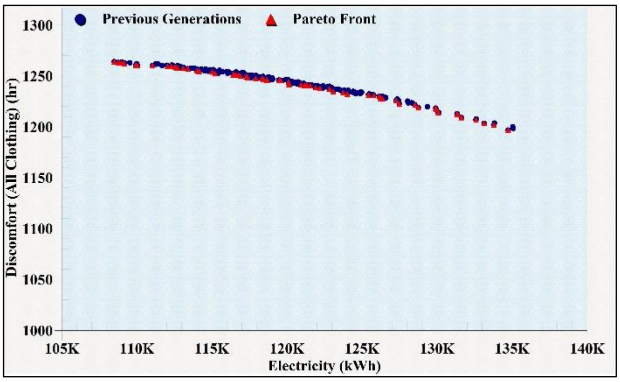


Figure 6: Pareto graph result of Bengaluru

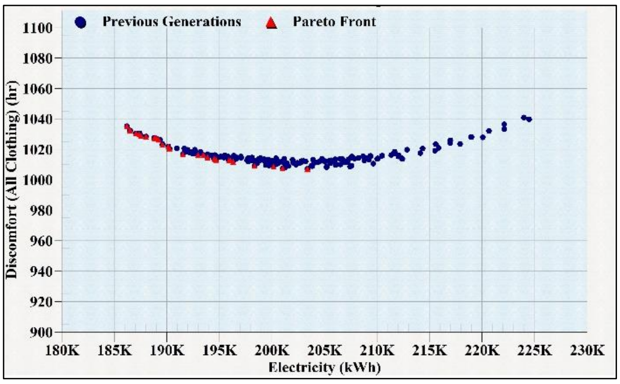


Figure 7: Pareto graph of Bhubaneshwar

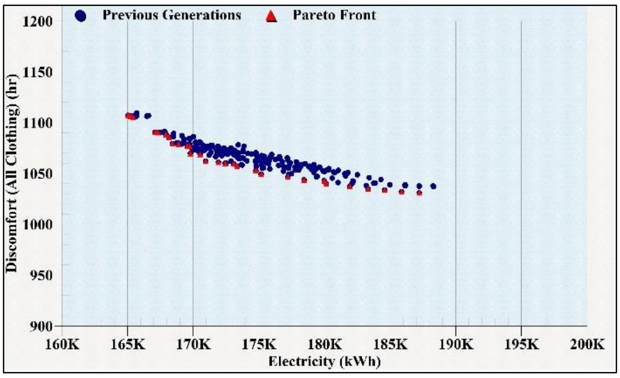


Figure 8: Pareto graph of Lucknow

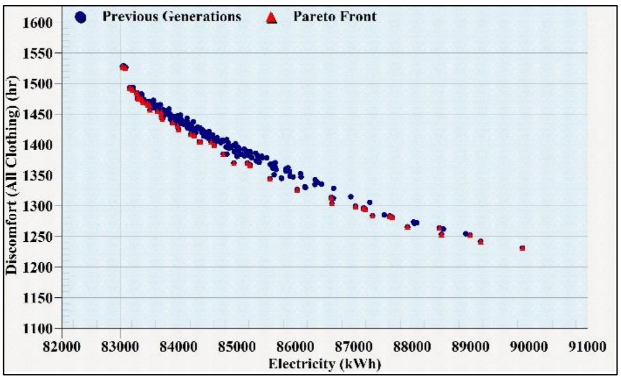


Figure 9: Pareto graph of Srinagar

The Pareto-front graph of the optimization analysis of each city is presented in Figure 5, Figure 6, Figure 7, Figure 8, and Figure 9. The optimization results illustrated a trade-off relationship between electricity consumption and discomfort hours. The electricity consumption and discomfort hours of optimal solutions of each city are compared with the values of the respective city baseline configuration which is equipped single clear glass of 6 mm and 30 % WWR. Aurangabad demonstrated a 10% to 16% reduction in electricity consumption and a 4% reduction in discomfort hours. However, for Bengaluru, this reduction ranged from 5 to 24%, while discomfort hours showed no potential for improvement. In Bhubaneshwar, optimal solutions lead to a 3 to 11% reduction in electricity consumption and a minimal change in discomfort hours. Lucknow, on the other hand, experienced a decrease in electricity consumption by up to 11% and also showed negligible potential for improvement in discomfort hours. Srinagar displayed up to 30% reduction in electricity consumption and up to 8% decrease in discomfort hours. Overall, the results indicated significant improvements in some cities, while others showed minimal enhancements compared to the baseline scenario. Comparable to outcomes of research conducted in Dubai on a two-story building, wherein double glazing resulted in a 9.3% reduction in energy consumption, the results of Aurangabad's (hot and dry climate) energy saving resonate closely (Abdeen et al., 2024).

The optimum solutions with minimum electricity consumption and minimum discomfort hours for each city are illustrated as a decision matrix in Figure 10. Along the top horizontal axis, WWR ranges from 20 to 80 with increments of 5. On the bottom x-axis, cities representing distinct climate zones are listed. The y-axis denotes glazing types numbered from 1 to 20. In this matrix, each block represents possible glazing configurations with different glazing types and WWR. Notably, blocks highlighted in brown signify the optimal configuration of glazing type and WWR for each respective city. These configurations strike a balance between reducing electricity usage and minimizing discomfort hours, tailored to the specific climatic conditions of each location. In Aurangabad, which represents a hot and dry climate, optimal configurations commonly included Glazing 5, 19, and 1. WWR for Glazing 5, which has the lowest SHGC, ranged from 20% to 80%, and for Glazing 1 and Glazing 19 ranged from 20% to 35%. with WWR ranging from 20% to 40%. Bengaluru, representing a temperate climate, tended to favor Glazing 18, Glazing 5, and Glazing 3, with recommended WWR falling between 20% to 80%. Bhubaneshwar, representing a warm and humid climate, showed optimal solutions with Glazing 5 and WWR varying from 20% to 75%. Lucknow, representing a composite climate, tended to benefit from Glazing 2 over all the WWR and from Glazing 1, and Glazing 16, with WWR between 20% to 40%. In Srinagar, representing a cold climate, optimal configurations include Glazing 2 for Glazing 16, and Glazing 17, with WWR ranging from 20% to 80%.

Discussion

The results suggested specific recommendations for different climatic zones regarding the selection of glazing properties. In hot and dry climates, and warm and humid climates, glazing with lower SHGC values and moderate U-values to minimize solar heat gain. Composite climates are recommended to have a mix of U-values and SHGC values when choosing glazing, demonstrating the need for a balanced approach in composite climates. In cold climates, the glazing options with lower U-value and relatively higher SHGC are found to be suitable. This is

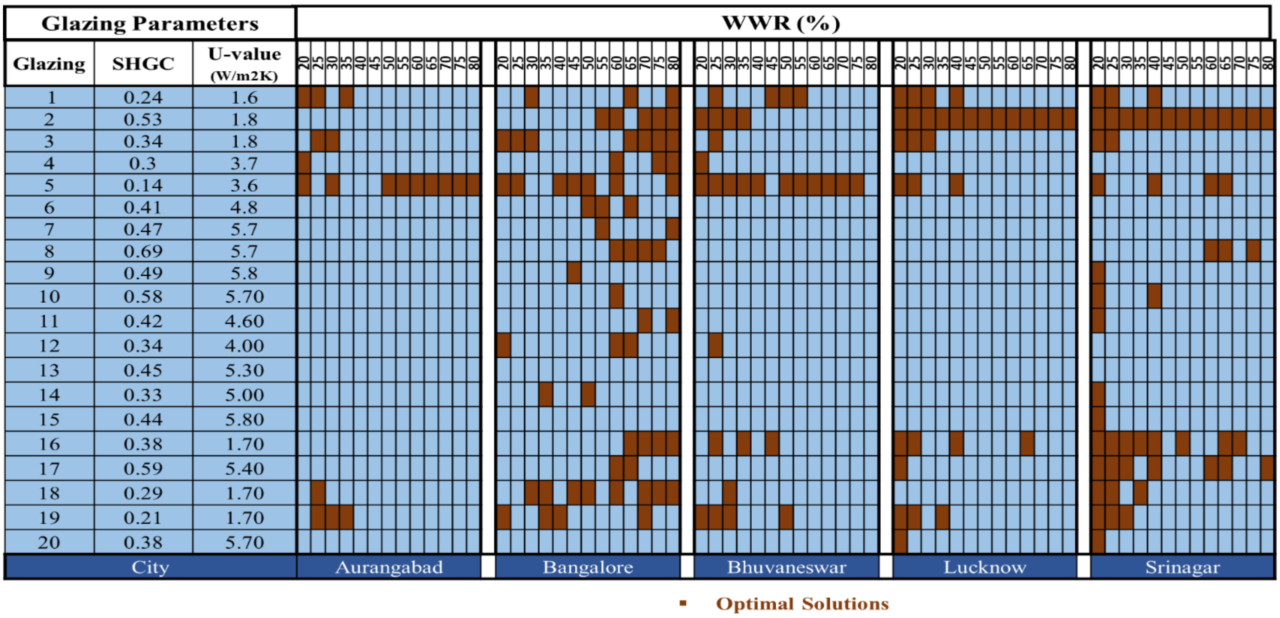


Figure 10: Decision Matrix of glazing solutions for minimum electricity consumption and maximum thermal comfort

because the total heat gain through the glazing of the building is given by relative heat gain (RHG), which combined accounts for both conductive and radiative heat transfer. Conductive heat transfer depends on the temperature difference between the building's interior and the outside environment and is influenced by the U-value. Radiative heat transfer is related to the amount of solar radiation entering the building and is influenced by the SHGC of glass. In tropical climates where temperature differences are minimal, conductive heat transfer plays a minor role, making SHGC more critical than U-value. Conversely, in colder and temperate climates such as Srinagar and Bengaluru, the U-value of glazing options becomes a more critical factor in reducing heat loss. The results also recommend relatively higher SHGC for cold climates so as to facilitate passive heating. The decision matrix derived through the optimization analysis can serve as a valuable guideline for updating building codes, standards, and sustainability regulations. By promoting the adoption of energy-efficient building envelopes tailored to local climatic conditions, policymakers can support the transition towards more sustainable and resilient built environments.

Conclusions

This study aimed to establish a simulation-based framework, SIMEC-Opt for optimizing glazing parameters in Indian residential buildings, employing a validated model of a residential building across diverse climate zones in India. By performing optimization analysis in the study, glazing configuration and WWR as design variables, the best solutions were identified, revealing the potential for significant improvements in both energy efficiency and thermal comfort. Across representative cities in different climate zones, the study demonstrated reductions in electricity consumption ranging from 3% to 30%, coupled with improvements in discomfort hours of up to 8%. Discomfort hours showed minimal variation across a few climate zones, suggesting that factors such as air velocity, ventilation, and occupant preferences may be more influential in analyzing thermal comfort.

The SIMEC-Opt framework represents a significant advancement in the field of building optimization, particularly concerning the selection of glazing parameters to enhance energy efficiency and thermal comfort in residential buildings. Through its simulation-based approach and multi-objective optimization algorithm, SIMEC-Opt offers a systematic and comprehensive methodology for identifying optimal glazing configurations tailored to specific climate zones. By considering a wide range of glazing options and window-to-wall ratios, SIMEC-Opt enables architects, builders, and policymakers to make informed decisions that maximize energy savings while ensuring occupant comfort across diverse environmental conditions. The framework's novel features, including its focus on Indian residential buildings and its consideration of multiple performance metrics, contribute to its effectiveness and relevance in addressing the pressing challenges of sustainable building design and construction.

While the study takes into account the diverse climate zones within India, it's important to note that building characteristics, energy consumption patterns, and occupant behavior also exhibit variability across different regions. The study's findings may be specific to certain typologies of residential buildings reducing the generalizability of results. Factors such as cost, government policies, local building laws, and technological feasibility may also influence the practicality of adopting the recommended solutions. The optimization framework primarily focused on glazing configurations and window-to-wall ratios, neglecting other critical building components such as walls and roofs. Though the narrow focus of the study confined to glazing allowed for a detailed examination of the unique characteristics and effects of glazing on building performance, integrating the wall and roof materials into the analysis can provide a more holistic understanding of building performance. While the study focused on representative cities across different climate zones in India, the underlying principles and optimization approach are transferable to similar climatic regions worldwide. As such, the recommendations can inform building design and retrofitting projects in various geographical contexts, contributing to global efforts to mitigate climate change and reduce energy consumption in the built environment.

Acknowledgments

The authors gratefully acknowledge the Glazing Society of India for supporting this project through the R S Subramanian fellowship (IIT Bombay internal project code- RD/0123-GLAZE18-001).

References

ASHRAE Standard 55, Thermal Environmental Conditions for Human Occupancy. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2021.

Abdeen, A., Emad Mushtaha, Hussien, A., Chaouki Ghenai, Aref Maksoud and Vittorino Belpoliti (2024). Simulation-based multi-objective genetic optimization for promoting energy efficiency and thermal comfort in existing buildings of hot climate. Results in Engineering, 21, pp.101815–101815.

Aboulnaga, M. M. (2006). Towards green buildings: Glass as a building element—the use and misuse in the gulf region.

Ampadu-Asiamah, A. D., and Gyebi-Adjei, E. A. (2014). Sustainable Construction in Ghana - Factors That Influence the Extensive Use of Glass on Facades of Office Buildings in Accra, Ghana.

Architecture 2030 (2023). Why The Built Environment – Architecture 2030. URL: www.architecture2030.org.

Ayyad, T. M. (2011). The Impact of Building Orientation, Opening to Wall Ratio, Aspect Ratio and Envelope Materials on Buildings Energy Consumption in the Tropics.

Bureau of Energy Efficiency, Government of India, Ministry of Power. URL: https://beeindia.gov.in/en.

Climate.onebuilding.org. URL: https://climate.onebuilding.org/.

Designbuilder Software Ltd – Home. URL: https://designbuilder.co.uk/

Eco – Niwas Samhita Energy Conservation Building Code for Residential Buildings (Part I: Building Envelope Design). (2017). URL: https://beeindia.gov.in/en/eco-niwas-samhita-ens

Graiz, E., and Al Azhari, W. (2019). Energy Efficient Glass: A Way to Reduce Energy Consumption in Office Buildings in Amman (October 2018).

Golmohammadi, R., Yousefi, H., Safarpour Khotbesara, N., Nasrolahi, A. and Kurd, N. (2021). Effects of Light on Attention and Reaction Time: A Systematic Review. Journal of Research in Health Sciences, pp.e00529–e00529.

Handbook of replicable designs for energy efficient residential buildings. (2021). URL: https://beeindia.gov.in/en/eco-niwas-samhita-ens

Kaja, N. (2015). An Overview of Energy Sector in India. International Journal of Science and Research, 6, 2319–7064.

Khalaf, M., Ashrafian, T., and Demirci, C. (2019). Energy efficiency evaluation of different glazing and shading systems in a school building.

Kiran Kumar, G., Saboor, S., and Ashok Babu, T. P. (2018). Investigation of various wall and window glass material buildings in different climatic zones of India for energy efficient building construction.

Naji, S., Aye, L., and Noguchi, M. (2021). Multi-objective optimisations of envelope components for a prefabricated house in six climate zones.

Sayed, M. A. A. E. D. A., and Fikry, M. A. (2019). Impact of glass facades on internal environment of buildings in hot arid zone. Alexandria Engineering Journal, 58(3), 1063–1075.

Ran, J., Cui, M., & Liu, J. (n.d.). Multi-objective optimization of building envelope in different climate zones in China based on BP-NSGA-II under the future climate.

Rethnam, O. R., and Thomas, A. (2023). Urban building energy modelling-based framework to analyze the effectiveness of the community-wide implementation of national energy conservation codes. https://doi.org/10.1108/SASBE-09-2022-0210.

Singh Rajput, T., and Thomas, A. (2023). Optimizing passive design strategies for energy efficient buildings using hybrid artificial neural network (ANN) and multi-objective evolutionary algorithm through a case study approach. International Journal of Construction Management, 23(13), 2320–2332.

Tibi, G., and Mokhtar, A. (2014). Glass Selection for High-rise Residential Buildings in the United Arab Emirates Based on Life Cycle Cost Analysis. Energy Procedia, 62, 270–279.

Usta, P., Zengin, B., Usta, P., and Zengin, and. (2022). An evaluation of the glazing type impact on building energy performance through a building simulation. Journal of Energy Systems, 6(1), 1–17.

Xu, L., Pan, Y., Lin, M., and Huang, Z. (2017). Community load prediction: Methodology and a case study. Procedia Engineering, 205, 511–518.

FINE-GRAINED SEGMENTATION OF HIGH-RESOLUTION BRIDGE CRACK IMAGES USING RENDERING TECHNOLOGY

Honghu Chu[1,2], Weiwei Chen[1,3*], Lu Deng[2]
[1] Bartlett School of Sustainable Construction, University College London, London, UK
[2] College of Civil Engineering, Hunan University, Changsha, China
[3]Civil Engineering Division, Engineering Department, University of Cambridge, CB3 0FA, UK

Abstract

Drawing on insights from computer graphics, this study introduces the Crack Boundary Point Rendering Network (CBPRN), an innovative high-resolution (HR) image segmentation framework designed to improve UAV-based bridge crack inspections. We developed an edge-guided branch and an uneven sampling strategy, enhancing detail preservation on crack boundary areas effectively. Through comprehensive ablation experiments, the efficacy of the CBPRN was validated, demonstrating its superior performance with remarkable outcomes: a processing speed of 13.45 FPS and mIoU, mBA, and Dice scores of 87.23%, 93.56%, and 89.59%, respectively, for images beyond 2K resolution. The CBPRN establishes a new standard in HR crack image segmentation.

Introduction

The emergence of surface cracks can effectively reflect the recent load-bearing status of bridges and provide a strong reference for traffic management departments to make reasonable maintenance decisions, thus preventing potential catastrophes (Manjunatha et al., 2023, Park et al., 2019). Therefore, accurate and efficient crack detection is crucial for ensuring the bridge's safety during its service life. In recent years, advancements in crack detection technology, driven by digital image processing algorithms, have seen rapid progress. This development has markedly enhanced efficiency and reduced the high costs traditionally associated with manual inspection methods (Çelik and König, 2022). Presently, extensive research has been undertaken by researchers in the field of crack detection utilizing image processing methods (Munawar et al., 2021, Ren et al., 2020). Within this domain, deep learning (DL)-based crack segmentation algorithms have garnered significant interest over classification and object detection algorithms. This preference is due to the segmentation algorithms' superior ability to accurately delineate the contours and shape characteristics of cracks with pixel-level precision.

A considerable amount of research based on DL for crack segmentation has been conducted, with some advanced algorithms reporting an impressive mIoU of over 93% on certain open-source crack datasets (Yang et al., 2023, Ali et al., 2021). However, it is noteworthy that these studies and their corresponding algorithmic improvements have primarily focused on identifying the main body of cracks while neglecting the recognition of fine-grained representations at crack boundaries. This oversight is significant because the quality of mask boundaries plays a crucial role in image segmentation; precise object segmentation directly benefits various downstream applications, such as damage quantification and assessment (Liu et al., 2023, Li et al., 2017). To comprehend the limitations in achieving fine-grained representations at crack boundaries, a critical analysis of the supervisory principles governing these segmentation algorithms is required, with a particular focus on the Mask Intersection-over-Union (Mask IoU) loss function (Cheng et al., 2021). This loss function, a benchmark in model training, guides models in predicting masks at pixel wise. Under the supervision of Mask IoU loss, models strive to maximize the overlap between the predicted mask and the actual mask. However, this loss evaluates all pixels equally, both internal and boundary pixels, making it less sensitive to the boundary quality of coarser cracks. As the crack size increases, the number of internal pixels grows quadratically and can far exceed the linearly increasing number of boundary pixels. This discrepancy leads to ambiguous segmentation effects in crack boundaries, especially at higher image resolutions where the difference between crack boundary and main body pixels is more pronounced. This trend is problematic for the industry's shift towards HR imaging for crack detection, as it can significantly affect segmentation results and impede accurate structural safety assessments. Therefore, a systematic study of the fine-grained recognition of crack edges is necessary to address this challenging issue.

To this end, this study introduces the Crack Boundary Point Rendering Network (CBPRN), which enhances three key components of traditional point-based rendering architecture, enabling the rendering head to fully utilize its advantages in refined segmentation of cracks. The network architecture is illustrated in Figure 1. Initially, an edge-guided branch based on a super-resolution encoding architecture was designed, ensuring that details of crack boundaries and related tiny crack information are adequately preserved in the deep semantic feature maps used as a source for refined rendering. For the second improvement, an uneven sampling strategy focusing on boundary areas was developed for the rendering-based prediction head, allowing the network to concentrate computational power on challenging areas like crack boundaries. Overall, these two improvements fully leverage the advantages of graphic rendering methods in the fine-grained segmentation of crack images. The CBPRN achieves a good balance between computational resource consumption and practical deployability while providing crack image outputs with precise boundaries, which is significant for accurate structural safety assessments of bridges.

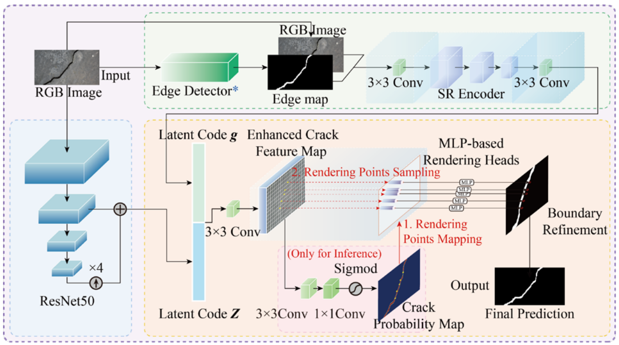


Figure 1: Visualization of boundary region-guided sampling with different dilation coefficients during the training phase

Methodology

The CBPRN proposed in this study consists of three main components: a coarse crack segmentation feature extraction backbone, an edge-guided branch, and a point rendering-based fine-grained prediction head. The coarse crack segmentation feature extraction backbone is built on a ResNet50-based encoding architecture, designed to perform coarse-grained feature extraction from high-resolution crack images. The edge-guided branch comprises a fixed-parameter edge detector and a super-resolution image encoding architecture, which guide the edge details in high-dimensional implicit features for coarse-grained crack features. The fine-grained prediction head based on point rendering primarily aggregates the implicit crack features outputted from the first two components and restores the fine-grained edge details of the cracks through point-by-point refined rendering based on the shared-weight multi-layer perceptron (MLP). Figure 1 visually presents some algorithmic details and computational logic of the proposed CBPRN. The edge-guided branch and the fine-grained prediction head constitute two innovative enhancements to the original PointRend architecture, respectively, and are detailed in the following subsections.

Edge-guided Branch

To ensure that the deep semantic feature maps of cracks sufficiently retain the details of crack boundaries and minute cracks for refined rendering, this study customizes an edge-guided branch in addition to the coarse crack segmentation feature extraction backbone. In fact, previous research has utilized guided image filtering (He et al., 2012) for boundary guidance in natural scene image segmentation, an effective edge-preserving smoothing operator based on guided images. However, edge recognition methods based solely on morphological operations are easily disturbed by environmental noise like cracks and struggle to accurately extract edges of small-sized cracks. To address this, the authors retain the fixed-parameter guided image filtering operator while introducing an encoder designed for super-resolution reconstruction tasks, aiming to eliminate noise interference in the boundary feature map while enhancing the representation of minute crack boundary features. This super-resolution encoding architecture, as shown in Figure 2, is primarily constructed using three residual modules. To improve the extraction performance of tiny crack details, the authors incorporate a transformer module in each residual module (b1, b2, b3), hoping to model long-distance dependencies for crack pixels scattered across the global view through the inherent self-attention mechanism of the transformer.

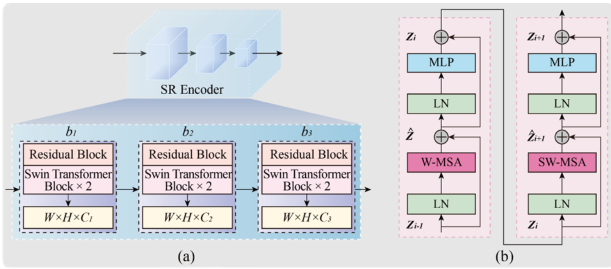


Figure 2: Details of the super-resolution reconstruction encoding architecture and some internal components in the edge-guided branch, (a) implementation details of the super-resolution reconstruction encoding architecture, (b) implementation details of two consecutive Swin Transformer blocks

Fine-Grained Prediction Head Based on Point Rendering

Inspired by refined rendering methods in computer graphics, the authors propose a crack fine-grained feature prediction decoding architecture based on point rendering. It is important to note that the principle of boundary refinement rendering point sampling remains consistent with the original PointRend architecture (Kirillov et al., 2020), ensuring computational efficiency during the training phase while enabling the trained model to perform effective end-to-end inference. Therefore, two different point sampling methods are designed for model training and inference.

For the training phase: due to the availability of precise pixel-level labels for effective supervision of prediction results, selecting boundary points for adequate sampling based on these refined pixel-level labels is a more accurate and computationally efficient method compared to the original approach in PointRend, which involved boundary prediction based on the most uncertain points. Therefore, during the training phase, boundary information is directly extracted from the refined pixel-level labels of crack images to guide the selection of sampling points. Specifically, an edge detection algorithm is used to extract the edge areas of the crack labels, and some of the sampling points, originally uniformly distributed across the background and the main body of the crack, are concentrated in these extracted edge areas. It is important to note that to prevent a decline in model performance due to an imbalance in the ratio of positive to negative samples during training, and to ensure efficient training, the total number of sampling points per image is set to $N = \frac{H\times W}{20}$. These sampling points are randomly distributed in the crack body, crack boundary, and background areas in a ratio of 0.3:0.4:0.3. Figure 3 visually demonstrates the sampling strategy for the rendering points used in training the model. Ultimately, all the sampling points identified on the labels are mapped onto the corresponding enhanced crack feature maps for model training.

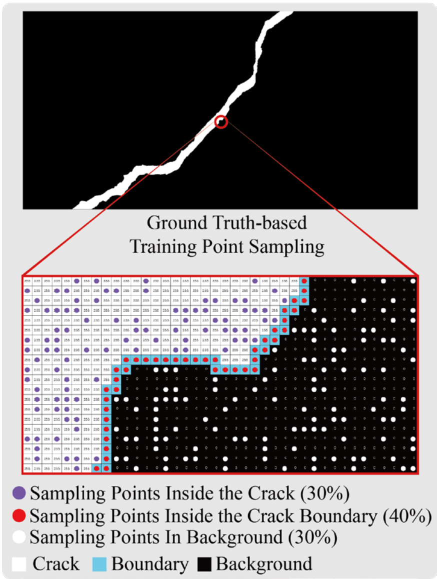


Figure 3: Visualization of the sampling points selection strategy in the training phase for the point rendering-based prediction head

For the inference phase: Refined labels, which are the ultimate prediction results, are not available at the beginning of the inference phase, hence the sampling method used during the training phase cannot be applied. To effectively focus computational resources on accurately predicting minute cracks and crack boundaries, this study proposes a boundary-guided rendering point sampling strategy based on a probabilistic heatmap. Specifically, two convolutional modules are added to the enhanced crack feature map to generate a refined crack pixel probability heatmap. Rendering points are guided by identifying pixels with high uncertainty in predicted values on the probability heatmap. Based on the probability, pixels on the heatmap can be roughly divided into three areas: pixels with probabilities close to 0 and 1 represent background and crack body, which are easily recognized by the network; pixels with probabilities around 0.5 represent areas of minute cracks or boundaries that are difficult for the model to determine with certainty. During the inference phase, for pixel areas on the probability heatmap with probabilities close to 0 and 1, the corresponding background and crack labels are directly used to represent the final prediction values, eliminating the need for further refined rendering. However, for pixel areas with probabilities around 0.5, refined rendering is required to further refine these ambiguous predictions. The refined rendering points are uniformly distributed in these hard-to-identify pixel areas. It is important to note that the probabilistic heatmap is used instead of the coarse segmentation results from the original PointRend architecture because the enhanced crack feature map used to obtain the heatmap is unified in size according to the second block of ResNet. It retains more details of minute cracks with only one downsampling compared to the original coarse segmentation and requires less computational resources than coarse segmentation. To visually represent the refined rendering point sampling method during the inference phase, Figure 4 displays a visualization of a randomly selected probabilistic heatmap example. On the heatmap, probabilities in the background and main body of the crack are concentrated near 0 and 1, respectively, while in the boundary area, due to issues like manual annotation errors and insignificant color differences, the probabilities of pixels fluctuate around 0.5. This study sets the probability range for these hard-to-identify pixels between 0.3 and 0.7, and in the subsequent refined rendering phase, only samples with probabilities between 0.3 and 0.7 undergo refined inference. The parameter settings for sampling points during the training and inference phases will be detailed in subsection 3.3.2.

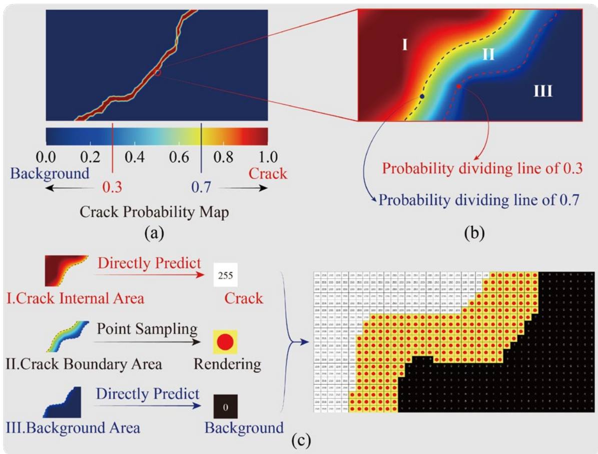


Figure 4: Visualization of the sampling points that require refined rendering during the inference phase

Experimental Setup and Results

Dataset

The CFD (Shi et al., 2016), Crack500 (Yang et al., 2019), and Deepcrack537 (Zhou et al., 2022) were chosen for model training. The crack images in these three datasets almost encompass most crack forms in engineering structures and include crack samples collected under different lighting conditions, which is beneficial for enhancing the model's robustness. Importantly, these three datasets all have finely annotated pixel-level labels, which are helpful for accurately evaluating the advantages of the algorithm proposed in this study in terms of boundary refinement segmentation. It is noteworthy that before training the model, images from the three datasets were uniformly resized to 256 × 256 pixels to facilitate uniform training input and save computational resources required for training. In total, 1200 resized crack images from the three open-source datasets were used, including 900 training samples, 150 validation samples, and 150 test samples.

Furthermore, a total of 60 crack images were collected in the urban area of Changsha, as shown in Figure 5, to test the segmentation performance of CBPRN on HR crack images.

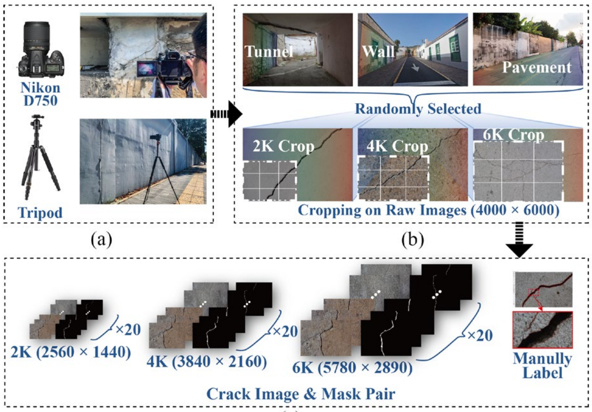


Figure 5: Details of establishing the high-resolution crack image dataset with pixel-level annotations

Implementation Details and Metrics

Training Hyperparameters: The experiments were conducted on a system equipped with an i9-9820X CPU and two NVIDIA RTX 3090 Ti GPUs, running Ubuntu 20.04, and the network model was implemented within the PyTorch framework. The total number of iterations was set to 1000 epochs, with a batch size of 16. The initial learning rate was set at 0.01, using a warming-up strategy for the first 100 epochs followed by a poly learning rate decay strategy with a decay rate of 0.9. Additionally, to ensure the global optimum of the loss function can be obtained during training, Adam optimizer, which combines the advantages of momentum and RMSprop, was used with a momentum of 0.9 and a weight decay of 1×10^{-4}. With these parameters, the preliminary model trained on the low-resolution open-source crack dataset was further fine-tuned using field-collected crack images.

Evaluation Metrics: Two common metrics were chosen for quantitative assessment of the experimental results: mean Intersection over Union (mIoU) and Dice Similarity Coefficient (Dice). Additionally, to highlight the performance of the proposed method in boundary areas, Mean Boundary Accuracy (mBA) was used as an additional evaluation metric. The core concept of mBA involves calculating the IoU between the Ground Truth (GT) and the predicted mask within the boundary area (Cheng et al., 2020).

Ablation Study

Ablation study for the point rendering-based prediction head: To fully illustrate the advantages of performing fine-grained crack mask using the point rendering method, a performance comparison was first made between the proposed decoding architecture and the traditional decoding architecture based on multiple convolutional layers and upsampling operations. The relevant experimental results are listed in Table 1. From the first two rows of Table 1, it can be seen that the point rendering-based decoding architecture has achieved improvements in mIoU, Dice, and mBA compared to the traditional decoding architecture, with the most significant improvement observed in mBA, reaching 86.78%. This is because the MLP in the point rendering-based decoding architecture is position-sensitive, calculating the prediction value for each pixel independently. Therefore, it can flexibly capture details and spatial relationships in crack images. In contrast, the traditional upsampling-based decoding architecture is limited by discrete feature sampling and struggles to capture local crack details.

After demonstrating the superiority of the proposed decoding architecture, parameter performance experiments need to be conducted to obtain the optimal parameters matching the model architecture. Since the training and inference stages use distinct fine-grained rendering point sampling methods, as described in Section 2.2, the following two sets of parameter performance experiments are conducted to obtain relatively optimal point sampling parameters for both the training and inference processes.

Point sampling study in the training phase: It is necessary to emphasize again before conducting parameter experiments that the regions most likely to have erroneous predictions are mainly concentrated at the boundaries and their adjacent areas, because these regions often have colors and contrasts similar to those of nearby cracks in RGB images. However, if the training point sampling is carried out as shown in Figure 3, where only a one-pixel-wide edge area is used for boundary guidance, it may not be possible to avoid the biasing guidance caused by the subjective nature of human annotation, resulting in errors between the true boundary and the labeled boundary. To avoid the negative impact of erroneous guidance on the model during training, it is necessary to expand the guided boundaries. Specifically, in this study, simple morphological operations were used, with the outermost pixels of the crack label as the center of dilation, and dilation operations were performed with the same dilation factor towards the background area and inside the crack. Considering the image size and the average pixel width of cracks in the training dataset, four different edge dilation coefficients were used to expand the boundary region. As shown in Figure 6, the widths of the expanded boundary regions (yellow outlined areas) after dilation are 1, 3, 5, and 7, respectively. The total number of sampling points on each training image is $N = \frac{H \times W}{20}$, distributed randomly in the background, expanded crack edges, and inside the cracks at proportions of 30%, 40%, and 30%, respectively.

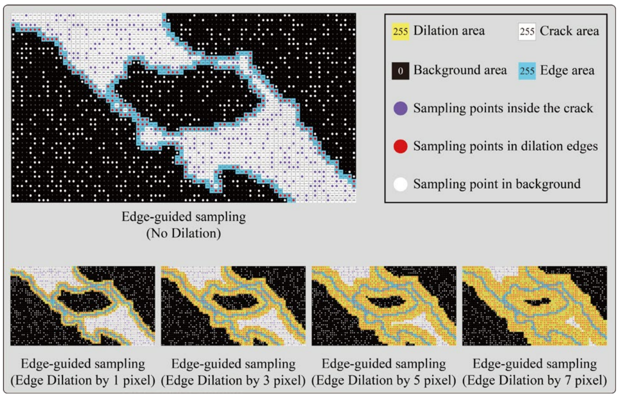


Figure 6: Visualization of boundary region-guided sampling with different dilation coefficients during the training phase

The set No.2 to set No.5 in Table 1 provide statistics on the performance of the model, which was trained in the training phase using four different widths of expanded boundaries for guiding sampling points. It can be observed that the model performs best when the dilation coefficient is 2 (corresponding to a boundary region width of 5), with mIoU, mBA, and Dice reaching 86.49%, 90.52%, and 88.54%, respectively. This best performance occurs due to the fact that the manual labeling bias range in the crack training dataset used in this study falls within this interval. Combining the experimental results in Table 2 with the visual effects of the probability heatmaps in Figure 4, the following conclusions can be drawn: dilation coefficients that are too low (0 or 1) or too high (3) result in improved performance compared to no dilation, but they respectively lead to insignificant improvements due to the inability to fully encompass error boundaries in the sampling region or the dispersion of computational resources beyond the error interval. Specifically, when the dilation coefficient is 0 or 1, the dilated boundary region is not sufficient to encompass the bias generated near the boundary during manual labeling. When the dilation coefficient is 3, the crack's main region and too many background regions without artificial labeling errors are included as ambiguous boundary regions requiring fine sampling. These unnecessary simple sample regions take away computational resources that should belong to the ambiguous boundary regions, thereby reducing the model's learning and representation capabilities for ambiguous boundary regions, resulting in a very limited improvement in recognition accuracy brought by boundary-guided sampling.

Point sampling study in the inference stage: In order to effectively improve the model's fine-grained inference performance in the boundary regions while saving computational efficiency, it is necessary to determine reasonable probability intervals for areas with uncertain prediction results concentrated around 0.5 on the probability heatmaps. A larger probability interval implies the need to sample more probability points for fine-grained rendering, which increases precision but significantly increases computational redundancy in the inference process. Conversely, a too small probability interval, while speeding up inference, may lead to ineffective fine-grained rendering of many tiny cracks and boundary details, seriously affecting the final recognition accuracy.

Specifically, two probability parameters need to be set: the critical probability value α between background pixels and boundary regions on the coarse prediction probability map, and the critical probability value β between boundary regions and crack pixels. For the critical probability value α between background and boundary regions, this study sets three different probability parameters: 0.2, 0.3, and 0.4. Similarly, for the critical probability value β between boundary regions and crack pixels, three different probability parameters are also set: 0.6, 0.7, and 0.8. By defining these nine different probability interval ranges based on the two types of critical probability values, the boundary regions are categorized. Table 2 provides statistics on the inference results of the models that use these 9 different probability intervals for sampling on the test dataset.

From Table 2, it can be seen that the models from Set No. 4, 5, and 6 (i.e., background region probability range between 0 and 0.3) achieved relatively better accuracy than other models. This is because, compared to the sampling groups with the background region probability range set between 0 and 0.4, the sampling methods under these three parameter settings encompass a wider background sampling area, which is more helpful in repairing some tiny crack details that were not detected in the background. At the same time, the sampling groups with the background region probability range set between 0.0 and 0.2 classified too many pixels originally belonging to the boundary region as background pixels, causing ambiguous boundary regions to be unable to achieve precise boundary detail repair due to insufficient sampling points, thus resulting in relatively lower mBA.

In addition, by comparing the model performance from Set No. 4 to 6, it can be observed that when the critical probability of crack internal area is set to 0.7, the model's inference accuracy is the highest, with mIoU, Dice, and mBA reaching 87.23%, 93.56%, and 89.59%, respectively.

Finally, the sampling parameter configuration set by Set No. 6 is adopted as the optimal inference stage sampling parameters to control the model's subsequent experiments.

Performance Comparison between CBPRN and the Traditional PointRend Model

Considering that the CBPRN is built based on the original PointRend model and is specifically designed for crack segmentation, with the main improvement being the introduction of a fine-grained boundary point rendering sampling method based on the fine-grained probability heatmap in the inference phase. In order to further demonstrate the effectiveness of this approach, which involves fine-grained point sampling guidance based on the probability heatmap in the inference phase, in comparison to the traditional approach of fine-grained point sampling guidance based on coarse segmentation, this section compares the segmentation results of the original PointRend using different sources of coarse segmentation on high-resolution images collected in Section 3.1 with the corresponding results obtained by the method proposed in this study.

Specifically, the authors selected five mainstream deep learning segmentation architectures with varying levels of segmentation accuracy, including FCN-18, UNet, DeepLabV3+, PSPNet, and RefineNet, as the generating networks for the coarse segmentation required when the original PointRend architecture performs predictions. In contrast, the method proposed in this study uses probability heatmaps proposed based on enhanced crack features extracted by the encoder and the boundary guidance branch to perform boundary point sampling guidance for the fine-grained prediction head based on point rendering.

It is worth noting that all coarse segmentation architectures and fine-grained segmentation networks were trained with default optimal parameters in the same deep learning framework with the same configuration. Additionally, when using the trained coarse segmentation models for prediction, all high-resolution images were proportionally scaled down to have a long edge of 900 pixels to avoid issues related to GPU memory overflow caused by excessively high original resolutions.

The experimental results are shown in Table 3. From the table, it can be observed that there are significant differences in the prediction results generated by different coarse segmentation mask generation architectures, with differences in mIoU, mBA, and Dice ranging from 2.92%, 3.06%, to 2.36%, from the lowest accuracy of FCN-18 to the highest accuracy of RefineNet. However, after applying the original PointRend model for the refinement process, the differences between the fine-grained prediction results become less pronounced, with mIoU, mBA, and Dice for all five experimental groups fluctuating within the intervals of 83.30±0.09%, 87.14%±0.16%, and 85.27±0.06%, respectively. These results indicate that the PointRend refined segmentation method is indeed independent of specific coarse segmentation masks and exhibits good robustness to coarse-grained features from different coarse segmentation architectures. However, when comparing the final experimental results with the best-performing coarse segmentation results based on RefineNet within the PointRend group, it can be observed that the method guided by probability heatmaps further improves the segmentation accuracy. It can be noted that among the three improved metrics, mBA shows the most significant improvement, which is more than twice the improvement in mIoU and Dice, reaching 6.27%. This outstanding robust performance is largely attributed to the introduction of the edge-guided branch in the feature extraction stage of this study, which preserves sufficient information about small cracks and crack boundaries in the enhanced crack features used to generate probability heatmaps. This enables the pixels in these detail areas to be detected during inference and finely represented through point-wise dense rendering. Additionally, it should be noted that CBPRN, by avoiding the guidance of coarse segmentation from external sources and based on a customized non-uniform inference point sampling method, surpasses the original PointRend in inference speed by more than twice on average, achieving 13.45 FPS. To further demonstrate the effectiveness of the above conclusions, Figure 7 provides visualizations of the test results for five randomly selected high-resolution crack images collected in the field. It is evident that the inference model guided by probability heatmaps proposed in this study outperforms any fine-grained rendering method guided by coarse segmentation masks in terms of crack edge recognition accuracy and sensitivity to tiny cracks.

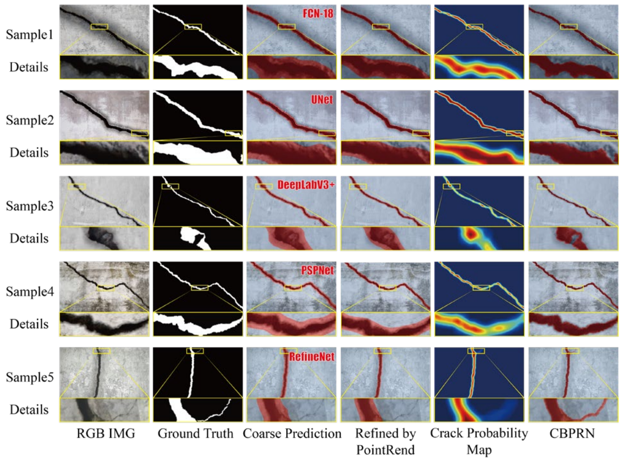


Figure 7: Visualization of fine-grained segmentation results of the PointRend architecture guided by different coarse segmentation masks and CBPRN guided by probability heatmaps

Conclusions and Future Work

In this study, a HR crack image fine-grained segmentation architecture named CBPRN is proposed. For the first time, rendering techniques from computer graphics are introduced into HR crack image segmentation tasks. Through three customized improvements, the originally designed point rendering technique for natural scene objects is adapted to effectively perform crack segmentation with fine-grained boundaries. The proficiently trained CBPRN attains a remarkable inference speed of 13.45 FPS, yielding mIoU and mBA scores of 87.23% and 93.56%, respectively, along with a Dice score of 89.59%. This performance was demonstrated on crack images exceeding 2K resolution, thereby establishing CBPRN as the current state-of-the-art benchmark in this domain.

In future, the implementation of model pruning and quantization techniques will be advanced to facilitate the lightweight deployment of CBPRN on the UAV, aiming to provide bridge maintenance departments with a more reliable and secure method for conducting bridge crack detection in practical engineering scenarios. Additionally, this method can be adopted to the hydropower projects, similarly to detect the defects of dam structures.

Table 1: Performance comparison of traditional decoding architecture and the proposed point-rendering-based fine-grained prediction head in models trained with different parameterized feature point sampling strategies

Set No.	Decoding architecture	Sampling point extraction method for the training phase	Dilating coefficient	Width of the boundary area after dilating	IoU(%)	mBA(%)	Dice(%)
1	Traditional convolution and upsampling operations	Uniform sampling	/	/	85.54	88.67	87.31
2	Fine-grained prediction head based on point rendering	Boundary guided sampling	0	1	85.98	89.60	87.76
3			1	3	86.12	89.94	88.03
4			2	5	86.49	90.52	88.54
5			3	7	85.76	89.09	87.50

Table 2: Comparison of inference performance obtained by different boundary probability ranges

Set No.	Probability range for background area	Probability range for boundary area	Probability range for crack internal area	IoU(%)	Dice(%)	mBA(%)
1	（0.0,0.2）	（0.2,0.6）	（0.6,1.0）	86.53	90.69	88.58
2	（0.0,0.2）	（0.2,0.7）	（0.7,1.0）	86.98	92.47	89.12
3	（0.0,0.2）	（0.2,0.8）	（0.8,1.0）	86.49	90.52	88.54
4	（0.0,0.3）	（0.3,0.6）	（0.6,1.0）	86.78	91.12	88.80
5	（0.0,0.3）	（0.3,0.7）	（0.7,1.0）	87.23	93.56	89.59
6	（0.0,0.3）	（0.3,0.8）	（0.8,1.0）	86.61	90.98	88.91
7	（0.0,0.4）	（0.4,0.6）	（0.6,1.0）	86.51	90.58	88.49
8	（0.0,0.4）	（0.4,0.7）	（0.7,1.0）	86.89	92.12	88.99
9	（0.0,0.4）	（0.4,0.8）	（0.8,1.0）	85.89	89.76	87.87

Table 3: Comparison of the refined segmentation results on HR images collected onsite between the proposed probability heatmap-guided method and the original pointrend architecture guided by different coarse segmentation masks

Meticulous segmentation architecture	Source of the boundary sampling guidance		Coarse segmentation accuracy (%)			Refined segmentation accuracy (%)			Total inference speed
			mIoU	mBA	Dice	mIoU	mBA	Dice	FPS
PointRend	Coarse segmentation guidance	FCN-18	78.36	79.25	81.37	83.21	86.98	85.21	7.83
		UNet	79.47	80.09	82.46	83.30	87.11	85.24	5.77
		DeepLabV3+	80.36	80.34	82.60	83.33	87.14	85.28	3.65
		PSPNet	80.65	81.79	82.88	83.37	87.26	85.31	4.21
		RefineNet	81.28	82.31	83.73	83.38	87.29	85.33	3.49
CBPRN	Probability heat map guidance	Probability interval ∈ [0.3,0.7]		/		87.23	93.56	89.59	13.45

Acknowledgements

This work was supported by the China Scholarship Council (No. 202206130068). This work is also supported by Horizon Europe projects, D-HYDROFLEX (Project No.:101122357). INHERIT (Project No.: 101123326), and UCL-ZJU Strategic Partner Fund

References

ALI, R., CHUAH, J. H., TALIP, M. S. A., MOKHTAR, N. & SHOAIB, M. A. (2021) Automatic pixel-level crack segmentation in images using fully convolutional neural network based on residual blocks and pixel local weights. Engineering Applications of Artificial Intelligence, 104, p.pp.104391.

ÇELIK, F. & KÖNIG, M. (2022) A sigmoid-optimized encoder-decoder network for crack segmentation with copy-edit-paste transfer learning. Computer-Aided Civil and Infrastructure Engineering, 37, p.pp.1875-1890.

CHENG, B., GIRSHICK, R., DOLLÁR, P., BERG, A. C. & KIRILLOV, A. (2021) Boundary IoU: Improving object-centric image segmentation evaluation. Available: https://ui.adsabs.harvard.edu/abs/2021arXiv210316562C [Accessed March 01, 2021].

CHENG, H. K., CHUNG, J., TAI, Y.-W. & TANG, C.-K. (2020) Cascadepsp: Toward class-agnostic and very high-resolution segmentation via global and local refinement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2020. Seattle, Washington, p.pp.8890-8899.

HE, K., SUN, J. & TANG, X. (2012) Guided image filtering. IEEE transactions on pattern analysis and machine intelligence, 35, p.pp.1397-1409.

KIRILLOV, A., WU, Y., HE, K. & GIRSHICK, R. (2020) Pointrend: Image segmentation as rendering. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020. Seattle, Washington, p.pp.9799-9808.

LI, S., CAO, Y. & CAI, H. (2017) Automatic pavement-crack detection and segmentation based on steerable matched filtering and an active contour model. Journal of Computing in Civil Engineering, 31, p.pp.04017045.

LIU, H., YANG, J., MIAO, X., MERTZ, C. & KONG, H. (2023) CrackFormer network for pavement crack segmentation. IEEE Transactions on Intelligent Transportation Systems, 24, p.pp.9240-9252.

MANJUNATHA, P., MASRI, S. F., NAKANO, A. & WELLFORD, L. C. (2024) CrackDenseLinkNet: A deep convolutional neural network for semantic segmentation of cracks on concrete surface images. Structural Health Monitoring, 23, p.pp.796-817.

MUNAWAR, H. S., HAMMAD, A. W., HADDAD, A., SOARES, C. A. P. & WALLER, S. T. (2021) Image-based crack detection methods: A review. Infrastructures, 6, p.pp.1-20.

PARK, S., BANG, S., KIM, H. & KIM, H. (2019) Patch-based crack detection in black box images using convolutional neural networks. Journal of Computing in Civil Engineering, 33, p.pp.04019017.

REN, Y., HUANG, J., HONG, Z., LU, W., YIN, J., ZOU, L. & SHEN, X. (2020) Image-based concrete crack detection in tunnels using deep fully convolutional networks. Construction and Building Materials, 234, p.pp.117367.

SHI, Y., CUI, L., QI, Z., MENG, F. & CHEN, Z. (2016) Automatic road crack detection using random structured forests. IEEE Transactions on Intelligent Transportation Systems, 17, p.pp.3434-3445.

YANG, F., ZHANG, L., YU, S., PROKHOROV, D., MEI, X. & LING, H. (2019) Feature Pyramid and Hierarchical Boosting Network for Pavement Crack Detection. Available: https://ui.adsabs.harvard.edu/abs/2019arXiv190106340Y [Accessed January 01, 2019].

YANG, L., BAI, S., LIU, Y. & YU, H. (2023) Multi-scale triple-attention network for pixelwise crack segmentation. Automation in Construction, 150, p.pp.104853.

ZHOU, Q., QU, Z., WANG, S.-Y. & BAO, K.-H. (2022) A method of potentially promising network for crack detection with enhanced convolution and dynamic feature fusion. IEEE Transactions on Intelligent Transportation Systems, 23, p.pp.18736-18745.

EARLY DESIGN STAGE MULTI-OBJECTIVE OPTIMIZATION FOR THERMAL REFURBISHMENT OF BUILDINGS: A CASE IN ISTANBUL

Damla Kömürcü[1], and Ecem Edis[2]
[1,2]Istanbul Technical University, Istanbul, Turkey

Abstract

To improve the energy efficiency of the large existing building stock, there is a need for decision-support to determine optimal solutions considering different objectives. In the study, it is aimed to find appropriate optimization methods through their hyperparameter optimization that will support designers at the early design stage of a building refurbishment project for environmental and economic sustainability. For this, NSGA-II, NSGA-III, C-TAEA, RVEA were investigated with a case study on residential buildings. The first three methods were observed to give limited renewal proposals, while RVEA provided diversified renovation scenarios with its models' features open to preference related to this.

Introduction

In energy efficient buildings, it is important to achieve simultaneously high performance in all aspects, even when dealing with conflicting goals. Studies for comfortable and energy efficient buildings address various objective functions such as energy consumption, thermal comfort and economic benefit indicators as single, multi or many objective/s. In addition, the differences in building functions (e.g., housing, office), decision variables (e.g., building form parameters, envelope) or constraints (e.g., budgets, thermal comfort conditions) in the optimization studies have resulted in datasets with a wide variety of features (Motlagh et al., 2021). Optimization methods suitable for problems with such different characteristics have also varied.

In comparative evaluation studies conducted to find the appropriate optimization method for the problem at hand, different methods have yielded superior performance results. To find the optimal design solutions for a passive building with a green roof for instance, six optimization algorithms [Multi-Objective Evolutionary Algorithm based on Decomposition (MOEA/D), Non-dominated Sorting Genetic Algorithm (NSGA)-II, NSGA-III, Multi-objective Particle Swarm algorithm (MOPSO), multi-objective dragonfly algorithm (MODA), Multi-objective ant lion optimization algorithm (MOALO)], were compared and MOALO algorithm was reported to lead to the best pareto front (Lin et al., 2021). In another study on minimizing the building energy consumption, CO_2 emission, and indoor thermal discomfort degree, an Adaptive Evolutionary Algorithm based on Non-Euclidean Geometry for the Many-Objective Optimization (AGE-MOEA), compared to the other four popular multi-objective optimization methods [NSGA-II, NSGA-III, MOEA/D, Constrained multi-objective optimization (C-TAEA)], was found to identify a set of pareto optimal solutions with a maximum optimization rate of 13.43% (Shen and Pan, 2023). A study to optimize the life cycle performance of the building on the other hand, compared NSGA-II, NSGA-III, and C-TAEA; and C-TAEA was the best bringing a reduction for the life cycle carbon emissions by 34.7%, for the life cycle costs by 13.9%, and for the indoor discomfort hours by 26.6% (Chen et al., 2023). In a study to be a reference for retrofit planners, the comparison of NSGA-II, MOPSO, MOEA/D, and NSGA-III showed that NSGA-III derives a comprehensive set of trade-off alternatives from possible retrofit scenarios (Son and Kim, 2018).

This study, therefore, aimed to find appropriate multi-objective optimization methods to support decision-making at the early design stage for thermal refurbishment through the building envelope to improve the energy efficiency of existing buildings considering both economic and environmental sustainability dimensions, which are sometimes conflicting. For this purpose, the optimization results of different methods were evaluated comparatively taking existing residential buildings in Istanbul as a case. In the paper, the information related to the application is given in the 'Methodology' section, and the detailed results are given in the 'Result and Discussion' section.

Methodology

The study consisted of three main steps: (1) case study and the multi-objective optimization problem definition, (2) optimization method selection, and (3) optimization study and evaluation of results (Figure 1). The studies carried out in these steps are explained in the following subsections.

Step 1: Case study and the multi-objective optimization problem definition

For the case study, the data given in Cetiner and Edis' (2011) study was decided to be used. That study was carried out to develop an environmental and economic sustainability assessment method regarding the improvements at the scale of building elements to reduce the use phase heating energy consumption of existing residential buildings in Istanbul (Cetiner and Edis, 2014). Environmental and economic sustainability scores were calculated based on the gain rate calculated by comparing the results of refurbished buildings with that of the base existing building condition obtained by EnergyPlus and SimaPro tools. Production, construction, and use periods were considered as building life cycle stages. For these phases, the environmental impact was calculated based on emissions and solid/liquid waste; the economic impact was calculated considering the cost based on the required water, material, energy (heating, cooling, transportation, and application), labor, and equipment. Design variables and objective functions, and their value ranges are given in Table 1. The distribution of the whole data by building age and by the element that would be thermally refurbished concerning their economic and environmental sustainability scores are given in Figure 2.

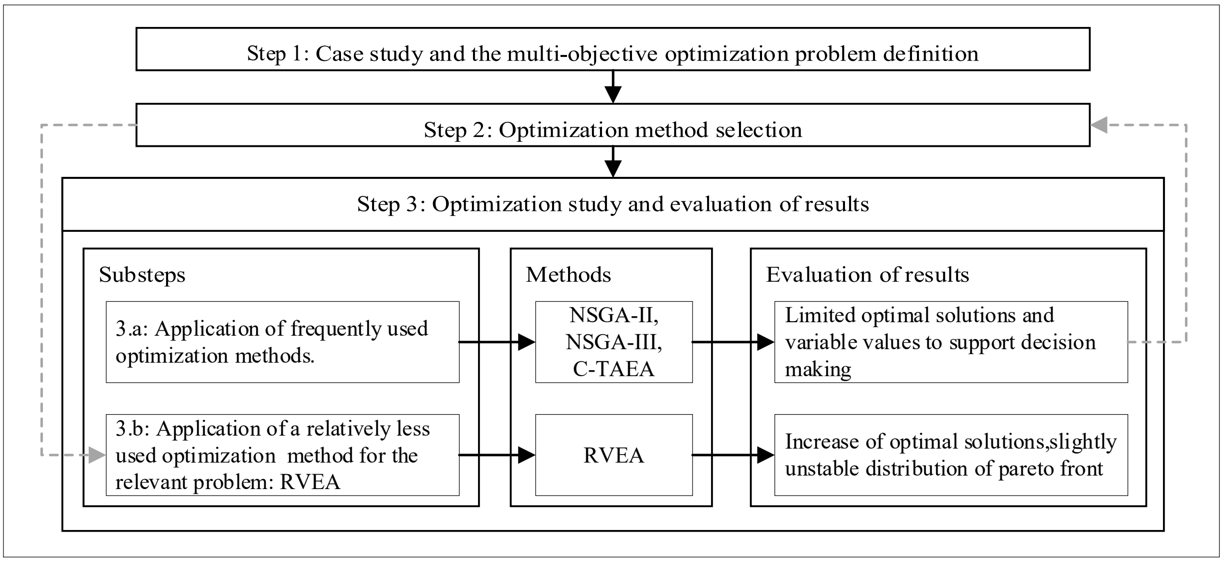
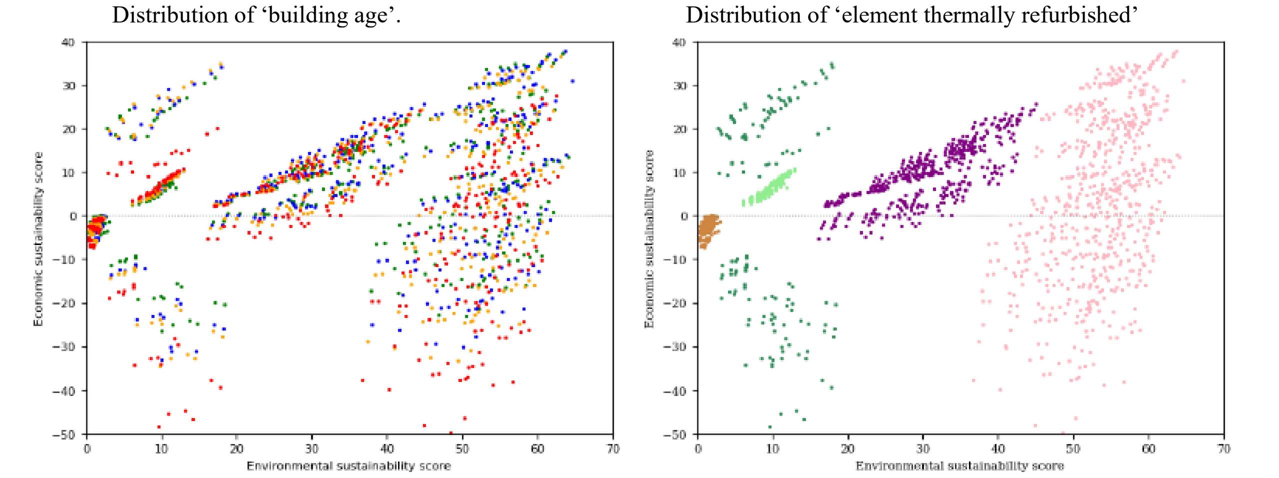


Figure 1: Research framework

Table 1: Design variables and objective functions, and their value ranges

	Variables- level 1	Variables- level 2	Ranges
Building	Building age	-	15, 20, 25, 30
	Plan type- floor area	Rectangular plan	ca. 230 sqm, and ca. 420 sqm
		Square / close to square plan	ca. 210 sqm, and ca. 400 sqm
	Orientation	Rectangular plan	Long sides facing East and West, Long sides facing North and South
		Square / close to square plan	Main orientations only
Envelope	Window-to-wall ratio	-	10%, 20%, 30%
	Window frame materials/ glass and glazing type	-	Wooden/ clear single glass, PVC/ clear double glass
	Element thermally refurbished - Insulation material and window frame (if specified) used	Floor over unheated spaces (uninsulated in base cases)	EPS, Stone wool, XPS
		Roof (uninsulated in base cases)	Glass wool
		Exterior wall and projected floor (uninsulated in base cases)	EPS, Stone wool, XPS
		Window frame (wood frame with clear single glazing in base cases)	Wood frame/double glazing, PVC frame/double glazing
		All of the abovementioned elements	
		Group 1 (wood frame with clear single glazing in the base cases)	EPS and wood frame/double glazing, EPS and PVC frame/double glazing, Stone wool and wood frame/double glazing, Stone wool and PVC frame/double glazing, XPS and wood frame/double glazing, XPS and PVC frame/double glazing
		Group 2 (PVC frame with clear double glazing in the base cases)	EPS and PVC frame/double glazing, Stone wool and PVC frame/double glazing, XPS and PVC frame/double glazing
Objective functions	Environmental sustainability score		(0, 65)
	Economic sustainability score		(-50, 38)
	Total sustainability score		(-39, 102)



Colors used for 'building age'

15 years old | 20 years old | 25 years old | 30 years old

Colors used for 'element thermally refurbished'

Floor above unheated spaces roof | Roof | Exterior wall and projected floor | Window frame and glazing | All elements mentioned on the left side

Figure 2: Distribution of the data by building age and element thermally refurbished

The multi-objective optimization study aimed to maximize environmental and economic sustainability scores for the two-objective functions; and, to maximize the total sustainability score (sum of environmental and economic sustainability scores) added to them for the three-objective functions. For the purpose of supporting the designer at the early design stage, building variables (building age, plan type, and orientation) were used as controlled variables in the study. This setup brought a total of 24 data sets with a population size of 75 for each.

Step 2: Optimization method selection

First, the methods that were observed to be frequently used in the literature were selected for the application. After the application of methods, the results were observed to be insufficient in terms of decision support performance. For this reason, another research study was conducted to select another optimization method suitable for this problem.

NSGA-II and NSGA-III are among the most frequently used methods in the studies (Pan et al., 2023; Hashempour et al., 2020; Shi et al., 2016). NSGA-II is a method that was designed based on a genetic algorithm, and a modified mating and survival selection calculation is applied, additionally (Deb et al., 2002; Blank and Deb, 2020). NSGA-III, developed from NSGA-II, specializes in many-objective optimization (Deb and Jain, 2014; Jain and Deb, 2014). C-TAEA, one of the state-of-the-art methods, is also increasingly used in studies and shows competitive performance in comparative evaluations of algorithms (Chen et al., 2023; Zhan and Huang, 2024; Picard and Schiffmann ,2021; Tian et al., 2021; Tian et al., 2022; Chen et al., 2024). It utilizes convergence-oriented archiving (CA) and diversity-oriented archiving (DA). Driving the population to the feasible area and approaching the Pareto front is primarily the responsibility of CA as the main force. To explore the area of CA development, DA is used as a supplement. The evolutionary state of CA and DA was used to select mating parents during reproduction (Li et al., 2019). For the optimization study, these three methods were selected considering their frequently usage.

Following their application and the assessment of their results, Reference Vector Guided Evolutionary Algorithm (RVEA) was then selected to be assessed as an alternative method. In RVEA method application, a scalarization approach known as angle penalized distance (APD) is proposed to measure the distance of the solutions to the ideal point and the closeness of the solutions to the reference vectors; this could be used as a diversity measure or a degree of satisfaction to the preferences (Cheng et al., 2019). An adaptation strategy is to adjust the distribution of reference vectors dynamically according to the scales of objective functions.

Step 3: Optimization study and evaluation of results

In the optimization applications with the selected methods, the pymoo framework was used (Blank and Deb, 2020). As aforementioned, due to the need for a new method search and application in relation to the limited number of refurbishment scenarios obtained with the initially selected methods, this step consisted of two sub-steps, which are detailed below. In each of these sub-steps, hyperparameter optimization and evaluation of the obtained optimal solutions in terms of decision support were carried out.

3.a: Application of frequently used optimization methods

In the study, NSGA-II method was used for two-objective functions. NSGA-III and C-TAEA methods were used both for two and three-objective functions.

The hyperparameter values that control the optimization process were selected based on the performance evaluations made in other studies. The hyperparameters evaluated in this study and their values and explanations are as follows:

- Das-Dennis and Riesz s-Energy options were selected for hyperparameter optimization for the reference directions (ref_dirs), which need to be defined for models with three-objective functions. These reference directions consist of a set of predefined reference points to guide the evolutionary search, and help to produce diverse and well-distributed solutions on the Pareto front. While Das-Dennis requires the use of more structured point sets and dimensions, Riesz s-Energy was developed to solve this problem (Ma et al., 2021).
- Termination Criterion (n_gen) is the parameter that determines when to terminate an algorithm run. For the hyperparameter optimization study, 250, 500, 1000, and 1500 values were studied (Rohit et al., 2021; Blank and Deb, 2020).
- Crossover operators were used to generate the offspring for all optimization models in the study (Katoch et al., 2021). For this purpose, Simulated Binary Crossover (SBX) was applied. SBX is a real-parameter recombination operator, and the spread of offspring solutions is determined by the operator's parameter in relation to their parent solutions (Deb et al., 2007). For all optimization studies related to this operator, the probability of SBX is set as 0.5 for two objectives, and as 1 for three objectives (Deb et al., 2007; Blank and Deb, 2020). For the hyperparameter optimization study, mutation probability (mut) values of 0.1, 0.2, and 0.3 were studied by keeping eta value constant at 30.
- Random sampling was used for unbiased representation of populations (Blank and Deb, 2020).

For all optimization studies, hyperparameters were optimized for the dataset that had 15-year-old buildings with rectangular plans of ca. 230 sqm, where long sides were facing east and west. Hyperparameter search space and settings selected accordingly are given in Table 2. For the first group of methods, mostly, values that will take less calculation time have been chosen because different parameter configurations did not produce any change in the results. NSGA-III produced mostly the same solutions as other methods, but for certain datasets, they were less than others. Therefore, hyperparameter setting determination criteria for this method were decided to be as giving the optimal solutions being the same as the optimal solutions obtained in all other methods.

Following the hyperparameter optimization, optimization studies were done for all data sets with the selected hyperparameter values. These optimization studies done with the selected methods led to limited optimal solutions and limited design scheme proposals to support decision-making (Figure 4), and these are given in detail and discussed in the 'Results and Discussion' section.

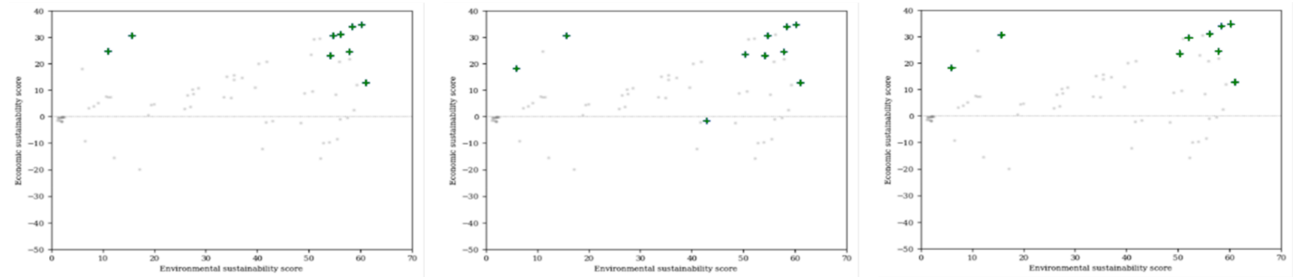
Table 2: Hyperparameter search space and parameter values selected from it for optimization models (Note: values shown in bold provided better results, others made no difference in solutions with different settings)

Objective numbers and optimization methods		Hyperparameter search space, and the parameter values chosen for the related methods.		
		Generation	Mutation	Reference direction
		250, 500, 1000, 1500	0.1, 0.2,0.3	Das-Dennis, Riesz s-Energy (energy)
2 objectives	NSGA-II	250	0.1	-
	NSGA-III	250	0.1	energy
	C-TAEA	250	0.1	energy
3 objectives	NSGA-III	250	**0.1**	**energy**
	C-TAEA	250	0.1	energy
	RVEA	**500**	**0.2**	**energy**

3.b: Application of a relatively less used optimization method: RVEA

Because of the limited optimal solution obtained in the previous step, RVEA, in which the Pareto front result can be determined depending on the user preference, was chosen to be used.

The hyperparameter search space and parameters were chosen the same as in the application of previous methods. Compared to the results of all previous optimizations in this study, changes in the hyperparameter settings were observed to have a significant impact on the Pareto front distribution for RVEA. In response to this, hypervolume (mean and standard deviation) and Pareto front were used as criteria for the hyperparameter value selection in the RVEA application. The Pareto front was evaluated by expecting to have a distribution that is concentrated as close to the ideal point as possible depending on the objective functions without any outlier in the distribution. Regarding this, hyperparameter value settings and corresponding hv values (mean and standard deviation) obtained by RVEA are given in Table 3 and three Pareto front graphs obtained by specific hyperparameter settings are given in Figure 3. Considering pareto front, n_gen is more effective during hyperparameter optimization, since it operates for the optimization of the distance of the solutions to the ideal point and the closeness of the solution for APD in RVEA. Additionally, compared to the other methods in Step 3.a, an increase in optimal solutions was also observed. Yet, the Pareto front distribution was slightly unstable.



Hyperparameter settings: n_gen: 500, mut:0.2, ref_dirs: energy	Hyperparameter settings: n_gen: 1000, mut: 0.3, ref_dirs: energy	Hyperparameter settings n_gen:1500, mut:0.1, ref_dirs: das-dennis

Figure 3: Pareto front obtained by RVEA with specified hyperparameter settings for the dataset that had 15-year-old buildings with rectangular plans of ca. 230 sqm, where long sides were facing east and west only

Table 3: Hyperparameter value settings and corresponding hv values (mean and standard deviation) obtained by RVEA

n_gen	mut	ref_dirs	hv(mean)	hv(std)
250	0.1	das-dennis	3.349152	0.0
		energy	3.349152	0.0
	0.2	das-dennis	3.349152	0.0
		energy	3.354702	8,90E-10
	0.3	das-dennis	3.349152	0.0
		energy	3.349152	0.0
500	0.1	das-dennis	3.349152	8,89E-10
		energy	3.356645	8,89E-10
	0.2	das-dennis	3.349152	8,89E-10
		energy	3.34602	0.0
	0.3	das-dennis	3.349152	8,89E-10
		energy	3.34915	8,89E-10
1000	0.1	das-dennis	3.349152	8,89E-10
		energy	3.34602	8,89E-10
	0.2	das-dennis	3.349152	8,89E-10
		energy	3.349152	8,89E-10
	0.3	das-dennis	3.349152	8,89E-10
		energy	3.356645	8,89E+05
1500	0.1	das-dennis	3.349152	0.0
		energy	3.356645	0.0
	0.2	das-dennis	3.349152	0.0
		energy	3.356645	0.0
	0.3	das-dennis	3.349152	0.0
		energy	3.356645	0.0

Results and discussion

The optimization methods applied within the scope of the study gave different results in terms of the diversity and stability of optimal solutions.

The first group of methods that is commonly used in literature, together with the dataset with a limited Pareto front solution caused by the data distribution depending on the objective functions in this study, gave very precise results that did not show any difference with different hyperparameter values in the optimization of the methods. The Pareto front obtained by this first group of methods resulted in two optimal solutions for each of the 24 datasets one of which is shown in Figure 4, and these results provided only three design options as presented in Figure 5.

These design options were all for the building envelopes with a 10% window-to-wall ratio (WWR) and wooden window frame, except for an option with a 20% WWR, and only the renewal by thermally refurbishing 'all elements' came up for all. When evaluated through the graph, it can be concluded that these methods give accurate results regarding the optimum renewal solution. However, those results provided no other alternative solutions for refurbishment projects for instance with a limited budget where an option for renewing a single building element would therefore be more preferable. This situation was considered to be an issue that would limit the designer’s decisions considerably.

To obtain more alternative design options, the use of RVEA, where user preferences are effective, has been tried. The optimum results obtained with that method not only include the Pareto front but also include feasible options further inside of the periphery (Figure 4). Results were obtained above the average of the economic sustainability score, and the environmental sustainability score ranged from a minimum of 3 to a maximum of 64.

Regarding the decision-support for the designer at the early design stage, optimal solutions obtained by RVEA provided diversified thermal refurbishment options. Design schemes corresponding to these optimal solutions given in Figures 5 and 6 show that:

- Renewal through all elements was the most proposed design by RVEA, same with the first group of methods because this option was the only one where the corresponding results for the values regarding the economic sustainability scores are the highest (Figure 2). Yet, depending on the budget, options for thermally refurbishing an individual element (i.e., ‘roof’, ‘window’, or ‘exterior wall and projected floor’) came up too.
- Options for 10% WWR and wooden frames were recommended more frequently. However, when the WWR of the building with a wooden window frame increased to 20% or 30% and refurbishment of all elements was preferred, replacing wood with a PVC frame was the only recommendation.
- More options were proposed for rectangular planned buildings than square alternatives.

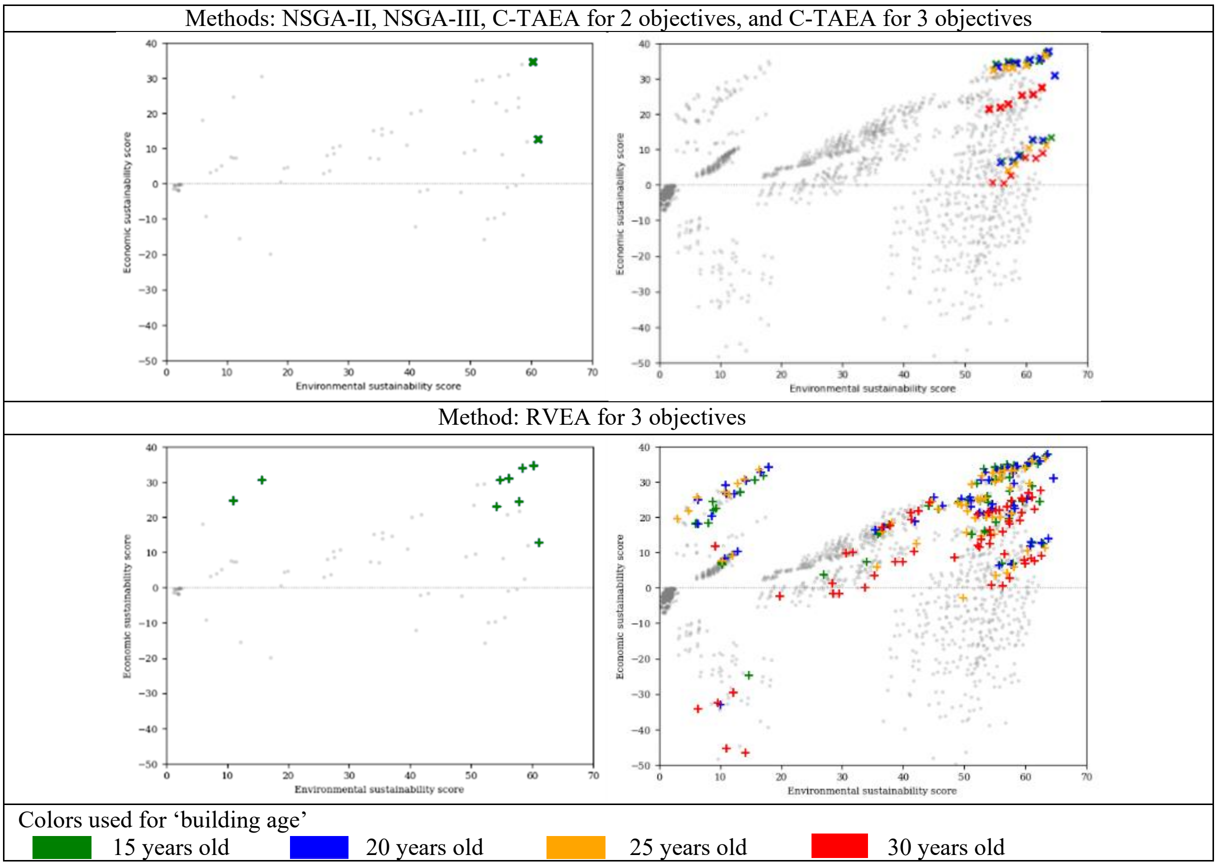


Figure 4: Optimal solutions for the dataset that had 15-year-old buildings with rectangular plans of ca. 230 sqm, where long sides were facing east and west only (left), and all optimum solutions obtained with all sets colored according to building age (right)

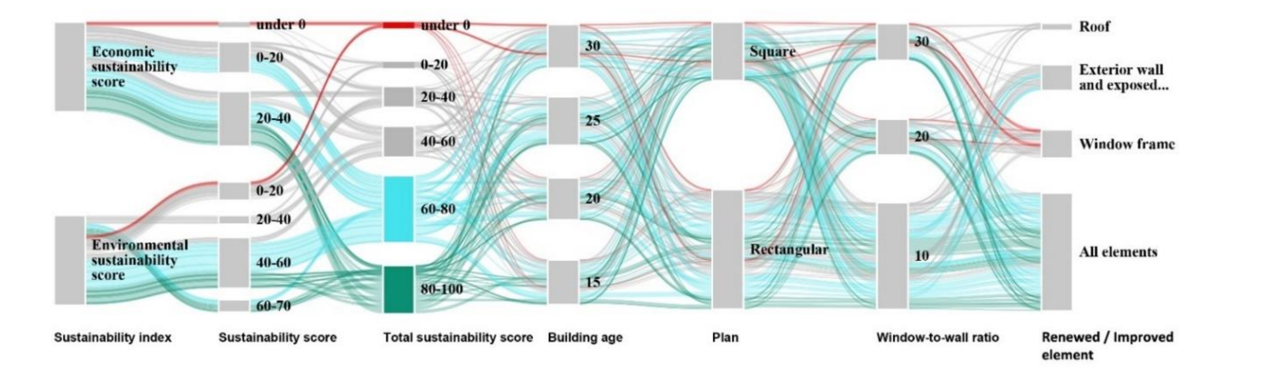
Window-to-wall ratio	Window frame materials/ glass and glazing type	Element thermally refurbished - Insulation material and window frame (if specified) used															Optimization methods
		Floor over unheated spaces			Roof	Exterior wall and projected floor			Window frame		All elements						
		EPS	Stone wool	XPS	Glass wool	EPS	Stone wool	XPS	Wood frame	PVC frame	EPS and Wood frame	EPS and PVC frame	Stone wool and wood frame	Stone wool and PVC frame	XPS and wood frame	XPS and PVC frame	
10	Wood/ clear single glass														◆ green ◆ blue ◆ orange ◆ red ▯■ (n=22**)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=24-all)	Group 1*
							◆ red ▯ (n=2)	◆ blue ◆ red ▯■ (n=4)		◆ green ◆ blue ◆ orange ▯■ (n=4)	◆ blue ◆ orange ◆ red ▯■ (n=4)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=24-all)		◆ green ◆ blue ◆ orange ◆ red ▯■ (n=8)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=20)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=24-all)	RVEA
	PVC/ clear double glass				◆ green ◆ orange ▯■ (n=3)		◆ orange ◆ red ▯■ (n=7)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=14)				◆ green ◆ blue ◆ orange ◆ red ▯■ (n=10)		◆ orange ▯ (n=1)		◆ green ◆ blue ◆ orange ◆ red ▯■ (n=14)	RVEA
20	Wood/ clear single glass															◆ blue ■ (n=1)	Group 1*
									◆ red ▯■ (n=3)	◆ green ◆ blue ◆ orange ◆ red ▯■ (n=10)		◆ green ◆ blue ◆ orange ◆ red ▯■ (n=11)		◆ green ◆ blue ◆ orange ▯■ (n=4)		◆ green ◆ blue ◆ orange ▯■ (n=12)	RVEA
	PVC/ clear double glass				◆ green ▯ (n=1)		◆ green ▯ (n=1)										RVEA
30	Wood/ clear single glass								◆ green ◆ blue ◆ red ▯■ (n=4)	◆ green ◆ blue ◆ orange ▯■ (n=14)		◆ green ◆ blue ◆ orange ▯■ (n=11)		◆ green ◆ blue ◆ red ▯■ (n=5)		◆ green ◆ blue ◆ orange ▯■ (n=5)	RVEA
	PVC/ clear double glass				◆ blue ◆ orange ▯ (n=4)												RVEA

Colors used for ‘building age’
15 years old (green) 20 years old (blue) 25 years old (orange) 30 years old (red)

Symbols used for ‘plan type’
▯ Rectangular plan ■ Square / close to square plan

(n=…): the number of data sets for which the corresponding configuration is recommended
*: Group 1 methods include those used for 2 objectives (i.e., NSGA-II, NSGA-III, C-TAEA), and 3 objectives (i.e., C-TAEA).
**: the related design schemes are optimum options for 18 datasets with NSGA-III for 3 objectives.

Figure 5: Variable values corresponding to optimal solution



Colors by ‘Total sustainability score’

	Score between ‘80-100’		Score between 60-80
	Score between ‘0-60’		Score under 0

Figure 6: Optimum design scheme with RVEA (colored depending on the value of ‘Total sustainability score’)

- For 30-year-old buildings, designs for 20% and 30% WWRs have been less commonly recommended even with RVEA.
- In addition to XPS, which was the only insulation material recommended by the first group of methods, EPS has become a frequently recommended option with the use of RVEA in the cases of replacement of all elements for all building ages and in all plan types. However, it was not recommended in the case of the replacement of any single element.
- There were some solutions which RVEA gave as the optimal solution to maximize objective functions, but they correspond to low economic values (values specified as under 0 in Figure 6), and therefore they can be considered incorrect. These were the options in which the window frames of 30-year-old buildings with 20% and 30% WWR were renewed in by wooden window frames. These options were clustered in a corner opposite to the ideal point completely.
- Regarding the refurbishment option of renewing the floor above the unheated spaces on the other hand, which are clustered far from the ideal point similarly, no suggestions came up.

Conclusion

The objective of this study was to find appropriate multi-objective optimization method(s) to support decision-making during early design for building thermal refurbishment through the building envelope. In this manner, four evolutionary algorithm methods were compared for a case in Istanbul, Turkey concerning residential buildings.

As a result of the studies conducted on the dataset used, NSGA-II, NSGA-III, C-TAEA methods brought limited solutions that were highly close to the Pareto front; and, setting the hyperparameters differently didn't change or expand these limited solutions in almost all individual optimization cases. The reason for this is due to the characteristics of the dataset studied: the distribution of the data in a way that the Pareto front does not spread widely and the fact that the dataset is in a discrete distribution. As a solution to these undesirable situations from the decision support perspective, RVEA, which allows user preferences to bring alternative design schemes, was evaluated. As a result, RVEA provided more solutions in number and variety of designs with reasonable stability regarding the study related to hyperparameter optimization.

Further research is planned on the application of optimization methods to other problems to evaluate the effect of problem characteristics on optimization performance.

References

Blank, J. and Deb, K. (2020) Pymoo: Multi-Objective optimization in Python. IEEE Access, Access, IEEE, 8, pp. 89497–89509.

Cetiner, I. and Edis, E. (2011) İstanbul’ daki mevcut konut stoğunun bina elemanlari ölçeğinde kullanim süreci çevresel ve ekonomik sürdürülebilirliğinin değerlendirilmesi ve katki sağlayacak iyileştirme önerileri geliştirilmesi. Available at: https://research.ebsco.com/linkprocessor/plink?id=8bf21cbc-e7ad-3c99-bd93-16b51f58b1ce (Accessed: 27 January 2024).

Cetiner, I. and Edis, E. (2014) An environmental and economic sustainability assessment method for the retrofitting of residential buildings. Energy & Buildings, 74, pp. 132–140.

Chen, R., Tsay, Y.-S. & Zhang, T. (2023) A multi-objective optimization strategy for building carbon emission from the whole life cycle perspective. Energy, 262.

Chen, R., Samuelson, H., Zou, Y., Zheng, X. & Cao, Y. (2024) Improving building resilience in the face of future climate uncertainty: A comprehensive framework for enhancing building life cycle performance. Energy & Buildings, 302.

Cheng, R., Jin, Y., Olhofer, M. & Sendhoff, B. (2016) A Reference Vector Guided Evolutionary Algorithm for many-objective optimization. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 20(5), 773–791.

Deb, K., Pratap, A., Agarwal, S. & Meyarivan, T. (2002) A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 6(2), 182–197.

Deb, K., Sindhya, K. & Okabe, T. (2007) Self-adaptive simulated binary crossover for real-parameter optimization. Proceedings of the 9th annual conference on Genetic and evolutionary computation, pp. 1187–1194.

Deb, K. and Jain, H. (2014) An Evolutionary Many-Objective Optimization Algorithm Using Reference-Point-Based Nondominated Sorting Approach, Part I: Solving Problems with Box Constraints. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 18(4), pp. 577–601.

Hashempour, N., Taherkhani, R. & Mahdikhani, M. (2020) Energy performance optimization of existing buildings: A literature review. Sustainable Cities and Society, 54.

Jain, H. and Deb, K. (2014) An evolutionary many-objective optimization algorithm using reference-point based nondominated sorting approach, part II: handling constraints and extending to an adaptive approach. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 18(4), pp. 602–622.

Katoch, S., Chauhan, S.S. & Kumar, V. (2021) A review on genetic algorithm: past, present, and future. Multimedia Tools and Applications: An International Journal, 80(5), pp. 8091–8126.

Li, K., Chen, R., Fu, G. & Yao, X. (2019) Two-Archive Evolutionary Algorithm for Constrained Multiobjective Optimization. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 23(2), 303–315.

Lin, Y., Zhao, L., Liu, X., Yang, W., Hao, X. & Tian, L. (2021) Design optimization of a passive building with green roof through machine learning and group intelligent algorithm. Buildings, 11(5), 192.

Ma, H., Wei, H., Tian, Y., Cheng, R. & Zhang, X. (2021) A multi-stage evolutionary algorithm for multi-objective optimization with complex constraints. Information Sciences, 560, 68–91.

Motlagh, S. F. M., Sohani, A., Saghafi, M. D., Sayyaadi, H. & Nastasi, B. (2021) The road to developing economically feasible plans for green, comfortable and energy efficient buildings. Energies, 14(3), 636.

Pan, Y., Zhu, M., Lv, Y., Yang, Y., Liang, Y., Yin, R., Yang, Y., Jia, X., Wang, X., Zeng, F., Huang, S., Hou, D., Xu, L., Yin, R. & Yuan, X. (2023) Building energy simulation and its application for building performance optimization: A review of methods, tools, and case studies. Advances in Applied Energy, 10(100135-).

Picard, C. and Schiffmann, J. (2021) Realistic constrained multiobjective optimization benchmark problems from design. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 25(2), pp. 234–246.

Dwivedula, R., Madhuri, R., Srinivasa Raju, K. & Vasan. A. (2021) Multiobjective optimisation and cluster analysis in placement of best management practices in an urban flooding scenario. Water Science and Technology, 84(4), 966–984.

Shen, Y. and Pan, Y. (2023) BIM-supported automatic energy performance analysis for green building design using explainable machine learning and multi-objective optimization. Applied Energy, 333.

Shi, X., Tian, Z., Chen, W., Si, B. & Jin, X. (2016) A review on building energy efficient design optimization rom the perspective of architects. Renewable & Sustainable Energy Reviews, 65, 872–884.

Son, H. and Kim, C. (2018) Evolutionary many-objective optimization for retrofit planning in public buildings: A comparative study. Journal of Cleaner Production, 190, 403–410.

Tian, Y., Zhang, T., Xiao, J., Zhang, X. & Jin, Y. (2021) A coevolutionary framework for constrained multiobjective optimization problems. IEEE Transactions on Evolutionary Computation, Evolutionary Computation, IEEE Transactions on, IEEE Trans. Evol. Computat, 25(1), 102–116.

Tian, Y., Zhang, Y., Su, Y., Zhang, X., Tan, K. C. & Jin, Y. (2022) Balancing objective optimization and constraint satisfaction in constrained evolutionary multiobjective optimization. IEEE Transactions on Cybernetics, Cybernetics, IEEE Transactions on, IEEE Trans. Cybern, 52(9), 9559–9572.

Zhan, J., He, W. & Huang, J. (2024) Comfort, carbon emissions, and cost of building envelope and photovoltaic arrangement optimization through a two-stage model. Applied Energy, 356.

EVALUATING THE CAPABILITIES OF SURROGATE MODELING TECHNIQUES IN PREDICTING HOURLY BUILDING ENERGY CONSUMPTION

Shivram Krishnan[1], Anagha Divakaran[1], Omprakash Ramalingam Rethnam[1], Elin Markarian[2], Seif Qiblawi[2], Albert Thomas[1], and Elie Azar[2]
[1] Indian Institute of Technology Bombay, Mumbai, India
[2] Carleton University, Ottawa, Canada
Corresponding author, e-mail address: albert@iitb.ac.in

Abstract

Building energy simulation models have strong energy prediction capabilities but suffer from high computational costs, which could be reduced through surrogate modeling approaches. Existing surrogate models predict energy consumption on an annual resolution, however, for strategizing net-zero measures, granular predictions are necessary. This paper evaluates the ability of four state-of-the-art machine learning algorithms to predict building energy consumption on an hourly basis. The results indicate that Random Forest Regression is the most suitable predictive model due to the high R^2 value of 0.94. The proposed framework can be further expanded to test net-zero energy retrofits at minimal computational costs.

Introduction

The operation phase of buildings accounts for nearly 30% of global emissions (IEA 2021) which necessitates the need for innovative strategies for reducing emissions. The International Energy Agency estimates that in the next 30 years, global building floor area will grow by 75%, with around 80% of the increase in emerging markets such as India. The building sector in India accounts for over 30% of total electricity consumed (BEE 2018) and the commercial building sector in India consumed 8.31% of end-use electricity in the year 2021 (MoSPI 2022). Previous studies on energy consumption show that the availability of an accurate building energy forecasting system can save about 30% of the total energy consumption in buildings (Olu-Ajayi et al. 2022). Hence, predicting the energy consumption pattern of buildings is of significance to multiple stakeholders as it can help improve buildings' energy efficiency and aid in energy conservation measures for buildings.

Building energy simulations for energy prediction are typically done by creating a prototype building model using an energy modeling tool such as EnergyPlus, developed by the U.S. Department of Energy based on the proposed design inputs (Chan et al. 2022). Data-driven approaches using surrogate machine learning (ML) based models help overcome the challenges of building an energy simulation approach, which requires high computational costs for performing the energy simulations. The ML models identify the relationship between input features impacting building performance such as weather parameters and output variables such as energy consumption. Such a trained ML model develops the capability to predict the energy consumption of buildings eliminating the need for multiple energy simulation runs for varying simulation scenarios (Olu-Ajayi et al. 2022).

The development of a surrogate ML model consists primarily of four steps: data collection, data preprocessing, model training, and model testing. Data collection involves collecting available past data, such as weather data and energy consumption data of buildings. Data preprocessing is done by integrating the required data into a comprehensive dataset, which will be used further for model training. The model training step involves developing and deploying a suitable ML model for energy prediction studies. Model testing involves using standard model evaluation measures to check the accuracy of the prediction model (Amasyali & El-Gohary, 2018).

Energy prediction using machine learning models is commonly done using algorithms such as regression models and artificial neural networks (ANN). Some of the commonly employed regression models in energy prediction include multiple linear regression (MLR), polynomial regression (PR), and random forest regression (RFR) (Sun et al. 2020). In a study based out of buildings in Italy, multiple linear regression is used as a decision support tool for assessing the preliminary energy demand. This was used as an alternative to assessing energy demand by building energy modeling. The correlation of the results shows that the MLR model can predict the buildings' heating, cooling, and total energy demand with a high degree of reliability (Ciulla et al. 2019). Meanwhile, the polynomial regression model is more flexible than multiple linear models as it can fit data in a wide range of curvature, making it suitable for problems where the relationship between input and output variables is non-linear, as in the case of building energy loads. In a study conducted on 17 buildings in Europe, polynomial regression was used to predict the heating demand based on climatological conditions and the architecture of the building. The results showed that the polynomial regression model prediction is best suited for predicting fast, early-stage energy consumption (Tiberiu 2012).

Another ML model that can capture non-linear relationships between input and target variables is Random Forest Regression. In a study conducted on a dataset of 5 existing buildings from a building data genome project (Miller et al. 2017), short-term energy consumption in hourly resolution is predicted to propose effective solutions to help buildings' owners and facility managers understand building energy consumption patterns for enhancing energy efficiency in buildings. The RFR model was selected for the study as it can reduce predictive error while solving regression problems, and it considers all decision trees in the prediction correlation (Pham et al. 2020). The results from the study showed that RFR is an effective technique in predicting hourly building energy consumption, which helps stakeholders take adequate steps to reduce the energy consumption of buildings.

Artificial neural network (ANN) is also a commonly used ML model for predicting the energy consumption of buildings besides RFR. In a study of 243 commercial buildings in Southeast Asia, the cooling loads of buildings obtained from the physics-based building energy model were compared to an ANN model. ANN was used due to its ability to solve non-linear problems dependent on numerous variables. The features for developing the ANN model were general building information, building envelope, and internal loads. The results showed that the cooling load prediction by the ANN model agrees well with the physics-based building energy simulation at a fraction of the time required for the latter (Ngo 2019).

Overall, existing studies have largely focused on predicting the 'annual' energy consumption of buildings, and energy prediction for 'hourly' resolution and 'one-year' time horizon has not been addressed thoroughly. The current studies are also limited to using meteorological information from a single climatic zone, and the impact of variation of varied climatic conditions on energy consumption is not explored in detail. Hence, this study aims to address this critical gap by identifying the best ML model for predicting building energy consumption in hourly resolution under varying climatic conditions for an annual time horizon. The developed framework is expected to reduce the computational effort required by a traditional building energy modeling approach for hourly energy predictions, even under varying climatic conditions, by eliminating the need for multiple energy simulation runs. This approach can further be used in optimization studies to inform design strategies, such as renewable energy generation, that support the transition to net-zero energy performance for existing buildings.

Methodology and Experiment

A four-stage methodology to identify the best predicting ML model for hourly energy consumption is proposed as follows:

Stage 1: Generating hourly energy simulation using Energy Plus

Stage 2: Compiling dataset for ML training

Stage 3: Developing ML Model and evaluating

Stage 4: Predicting energy consumption using ML model

The flow chart of the hourly energy prediction framework is given in Figure 1 below.

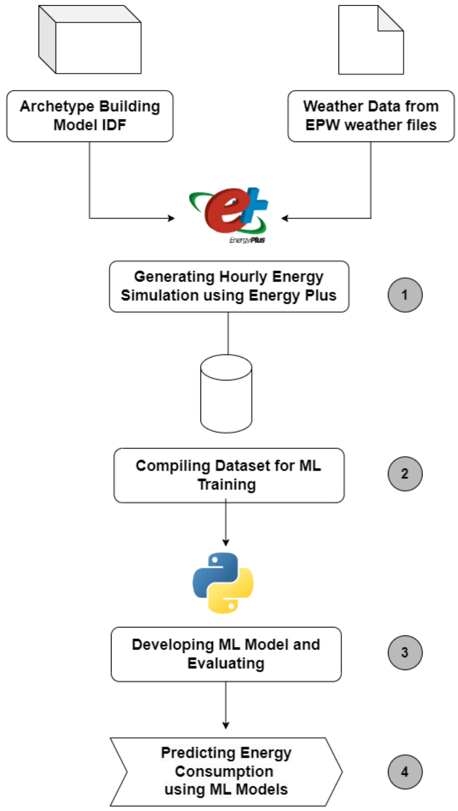


Figure 1 Framework for hourly energy prediction

Stage 1: Generating Hourly Energy Simulation Using Energy Plus

The dataset used for ML training is the collection of meteorological data from weather files and hourly building energy consumption data generated from energy simulations. The process is explained in detail below.

Hourly building energy consumption data based on annual simulation of a standard reference building model is used for the ML model development. A validated prototype reference commercial building model, based on a previous study of 230 commercial buildings in India (Bhatnagar et al. 2019), is adopted to simulate hourly energy consumption, and the archetype-building model used for this study is shown in Figure 2 below.

Archetype Building Model

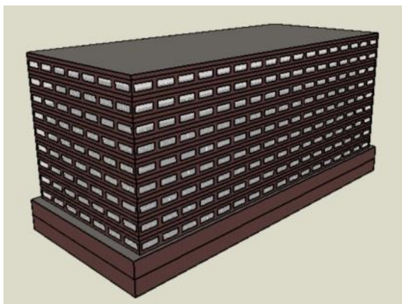


Figure 2 Archetype reference commercial building model for an 8-hour working schedule

Building features of the archetype reference building used in this experiment are mentioned in Table 1 below.

Table 1 Archetype reference building features

Feature	Value
Total floor area (m2) (excluding basement)	31,381
Percent of conditioned Area (%)	76
Basement floor area (m2)	7,689
Floor Area	3,487
Number of floors	9
Building shape	Rectangular
Aspect ratio	2.3
Window-to-wall ratio	30%
Window U-factor (W/m2-K)	2.05
Window SHGC	0.27
Exterior wall	Brick wall (200mm width) with plaster
Exterior wall U-factor	1.46
Roof	Roof tile + Concrete (200 mm) + Expanded polystyrene (50mm) + Plaster
Roof U-factor	0.64
Average lighting power density (W/m2)	Office: 8.32; Basement parking: 1.81
Daylighting controls	No
Occupancy controls	No
Exterior Light (kW)	8.12
Cooling type	Screw Chiller
Distribution and terminal	Variable air volume AHU with motorized dampers
COP	5.6
Supply Fan Fan Power (W/l/s)	1.25
Economizers	No
Energy recovery	No

Weather data

The weather data in 'epw' format, developed by the U.S. Department of Energy, is used for this study. Representative cities used for this study are shown in Table 2. The input variables (X-variables) used for ML model development, are listed in Table 3. The weather file input parameters are also taken as per studies conducted previously. In addition to the weather parameters, occupancy fraction is also considered as one of the input variables for energy consumption prediction as the occupancy schedule has been shown to significantly influence energy consumption (Li et al. 2018).

Table 2 Representative cities for ML model development

Sl No:	Climatic Zone	Representative City
1	Hot-Dry	Ahmedabad
2	Hot-Humid	Kolkata
3	Temperate	Bengaluru
4	Composite	Lucknow

Table 3 List of input variables for ML model development

Sl.No.	Input Variables (X)	Variable Name
1	Dry bulb temperature	X1
2	Dew point temperature	X2
3	Relative humidity	X3
4	Global horizontal radiation	X4
5	Direct normal radiation	X5
6	Diffuse horizontal radiation	X6
7	Wind direction	X7
8	Wind speed	X8
9	Occupancy	X9

Stage 2: Dataset for ML Training

The simulation software 'EnergyPlus,' version 8.7, developed by the U.S. Department of Energy, is used to generate hourly energy consumption for an annual simulation with 8,760 hourly energy data points. Out of the total 8,760 data, an 8-hour working schedule, from morning 09:00 to evening 17:00, is considered. The days considered were from Monday to Friday, excluding 10 public holidays. The resulting dataset has 2,000 hourly data points of input (X-variables) and output variables (Y-variable) as given in Table 4.

Table 4 Data points considered based on an 8-hr schedule

Climate Zone	City Name	Total Simulation Data	Data points based on 8-hr schedule
Hot-Dry	Ahmedabad	8,760	2,000
Hot-Humid	Kolkata	8,760	2,000
Temperate	Bengaluru	8,760	2,000
Composite	Lucknow	8,760	2,000

Stage 3: ML Model Development and Evaluation

The dataset is further split into training and test sets. The 'X-variables' and 'Y-variables' from the dataset are split in an 80:20 ratio for all the subsequent ML models used to develop a predictive model. Four different ML algorithms, such as MLR, PR, RFR, and ANN, are used to investigate the best-performing model in predicting hourly energy consumption. MLR, PR, and RFR algorithms are developed using the 'scikit-learn' library, and the ANN model is developed using the 'tensorflow' library with Python as the programming language. ML model development is given in more detail in the subsequent sections.

Multiple Linear Regression

From the 'scikit-learn' Python library, the 'LinearRegression' function is imported. The split dataset is further given as input to the linear regression function. The 'predict' function is then used with the test set of 'X-variables' to predict the hourly energy consumption.

Polynomial Regression

Similar to the multiple linear regression, 'PolynomialFeatures' and 'LinearRegression' functions are imported from the 'scikit-learn' library. The training dataset for 'X-variable' is then transformed using 'fit_transform'. The 'predict' function is further used with the test set of 'X-variables' to predict the hourly energy consumption.

Random Forest Regression

For RFR, the 'RandomForestRegression' function is imported from the 'ensemble' module from the 'scikit-learn' library. The 'RandomForestRegression' function is then called on the split dataset with 'n_estimators' defined as 10 and 'random_state' set as 0. The 'predict' function is further used with the test set of 'X-variables' to predict the hourly energy consumption.

Artificial Neural Network

ANN model differs slightly from the other three models used. For ANN, the 'Tensorflow' python library is used for analysis. A multi-layered feed-forward ANN model with one input layer, one hidden layer, and one output layer is used based on trial and error for prediction accuracy. The number of 'epochs' is set as 100. The 'predict' function is further used with the test set of 'X-variables' to predict the hourly energy consumption.

Following the ML model development, model performance evaluation techniques are further deployed on each ML model to select the best predicting model. The model performance evaluation is conducted separately for the four representative cities.

Stage 4: Energy Prediction using ML Model

Some of the standard model performance evaluation metrics used for evaluating the test set prediction results are coefficient of determination (R^2), Mean Squared Error (MSE), Mean Bias Error (MBE), Root Mean Squared Error (RMSE), and coefficient of variance (CV) (Sun 2020). All these model performance parameters are estimated based on the equations below, and the results are discussed in further sections.

The coefficient of determination (R^2), as mentioned in Equation 1, measures how close the predicted value is to the actual value, with a value of 1 indicating a perfect fit.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(PEi - AEi)^2}{\sum_{i=1}^{n}(AEi - \overline{AEi})^2} \quad (1)$$

The Mean Squared Error (MSE), mentioned in Equation 2, calculates the variance between actual and predicted values.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(AEi - PEi)^2 \quad (2)$$

The Root Mean Squared Error (RMSE), in Equation 3, represents the square root of the quadratic mean of the differences between predicted and expected values.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(AEi - PEi)^2} \quad (3)$$

The Mean Bias Error (MBE), mentioned in Equation 4, is the average error or bias of each value between the actual and predicted values in the sample space.

$$MBE = \sum_{i=1}^{n}\frac{(AEi - PEi)}{n} \quad (4)$$

The Coefficient of Variation (CV), mentioned in equation 5, indicates how well a model approximates the real data points, which measures the model's degree of predictability.

$$CV = \frac{\sqrt{\frac{\sum_{i=1}^{n}(PEi - AEi)^2}{n-1}}}{\overline{AEi}} * 100 \quad (5)$$

where PEi is the predicted energy, AEi is the actual energy at the hour 'i', $\overline{AEi}$ is the mean of actual energy, and n is the total number of samples.

Based on the results obtained from the model performance evaluation conducted for each representative city, the best predicting ML model is selected for further scenario analysis.

Scenarios considered using the best predicting ML Model

The best ML algorithm, selected based on model performance evaluation, is further deployed in two different scenarios. The two different scenarios considered in the study are explained below.

Scenario 1

In Scenario 1, the X and Y variables from all four representative cities are considered and combined to train the ML model. The combined dataset, which contains 8,000 data points, is then split into an 80:20 ratio. The hourly energy consumption of the building using the test set is then predicted.

Scenario 2

Scenario 2 demonstrates the energy forecasting capacity of the ML model for a dataset it is not trained with. In this case, the data points from three climatic zones were used, and the combined dataset contained 6,000 data points. In this study, the training set included X and Y variables of Ahmedabad, Lucknow, and Bengaluru, and the hourly energy consumption of Kolkata is predicted.

Results and Discussion

The experiment to predict the hourly energy consumption using ML models was conducted and the results are given and discussed as follows.

Model Performance Evaluation

The results of the model performance evaluation conducted for the four ML models MLR, PR, ANN, and RFR using Lucknow's dataset are shown in Figure 3 and Table 5 below. Similarly, the model performance evaluation is conducted for the Kolkata, Ahmedabad, and Bengaluru datasets, and the results are shown in Table 6 to Table 8, respectively.

Table 5 Model performance evaluation for Lucknow with MLR, PR, ANN, and RFR models

Model performance	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
MLR	0.89	1169.6	34.20	-1.27	9.6
PR	0.83	1857.6	43.10	-4.58	10.9
ANN	0.88	1296.2	36.00	-3.20	10.1
RFR	0.94	636.39	25.22	-2.88	9.7

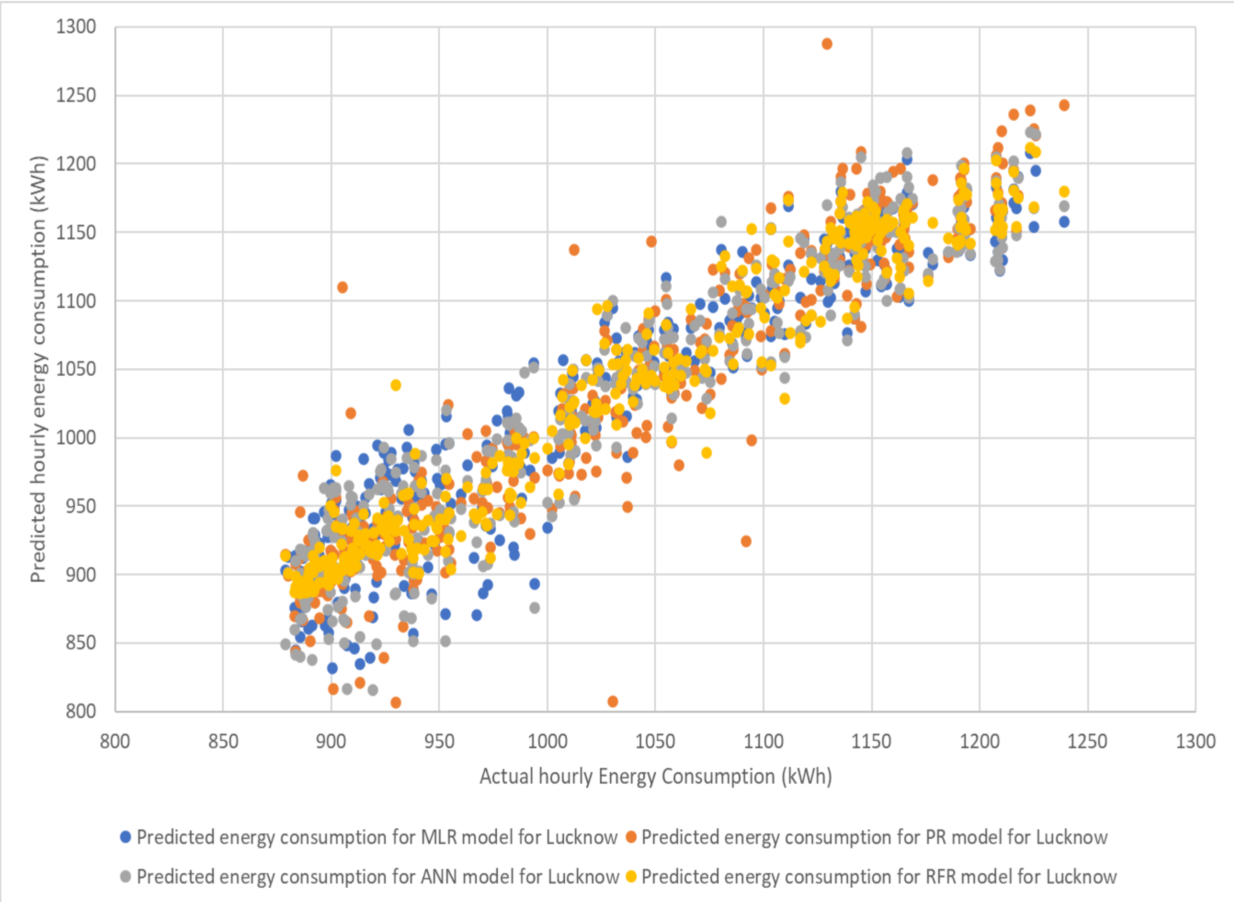


Figure 3: Graph showing the predicted vs actual hourly energy consumption for the four ML models tested for Lucknow

Table 6 Model performance evaluation for Kolkata with MLR, PR, ANN, and RFR models

Model performance	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
MLR	0.86	1208.1	34.75	1.72	8.9
PR	0.56	4055.1	63.67	4.31	10.4
ANN	0.82	1574.4	39.67	7.54	9.7
RFR	0.91	748.4	27.37	1.55	8.8

Table 7 Model performance evaluation for Ahmedabad with MLR, PR, ANN, and RFR models

Model performance	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
MLR	0.75	1637.88	40.47	5.80	7.0
PR	0.42	3846.89	62.02	-1.02	9.3
ANN	0.75	1667.55	40.83	-5.89	8.0
RFR	0.88	739.46	27.19	1.69	7.4

Table 8 Model performance evaluation for Bengaluru with MLR, PR, ANN, and RFR models

Model performance	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
MLR	0.61	559.18	23.64	1.110	3.4
PR	0.25	1071.9	32.73	-1.105	4.6
ANN	0.50	709.73	26.64	-3.126	3.9
RFR	0.70	434.28	20.83	1.904	3.4

The graphs shown in Figure 3, represent predicted and actual energy consumption for the four ML models MLR, PR, ANN, and RFR for the city of Lucknow. The X-axis represents actual hourly energy consumption and the Y-axis represents predicted hourly energy consumption in kWh. In the case of the RFR model, in Figure 3, the predicted energy consumption closely follows the actual energy consumption. However, in the case of MLR, PR, and ANN, the predicted energy consumption shows variations. In some instances of MLR, PR, and ANN, the predicted energy consumption value is higher than actual energy consumption. Thus, the RFR model shows better performance when compared to MLR, PR, and ANN.

The results observed from Table 5 to Table 8 show that the RFR model outperforms other models due to its high R^2 values and lower MSE, RMSE MBE, and CV values. Hence, the RFR model is chosen as the prediction model for further scenarios 1 and 2.

Energy Prediction Results for Scenario 1 and 2

Scenario 1

Hourly energy consumption is predicted using the RFR model, trained with the dataset of all four representative cities combined. The results are shown in Table 9 below.

Table 9 RFR model performance in the test phase of scenario 1

Scenario	Training points	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
I	6,400	0.90	837.8	28.94	2.21	8.70

The performance evaluation of the test data in the test phase of scenario 1 gives an R^2 value of 0.90 and a CV value of 8.70%, indicating that the RFR model is reliable for hourly building energy prediction (Yang et al. 2015).

Further, to understand the results figuratively, Figure 4 shows the predicted versus actual hourly energy consumption for the test data of scenario 1. The scatter plot is generated from the test set of scenario 1, i.e., with 1600 data points (20% of 8000 data points) combined from all four cities. The concentration of the scatter plot points along the 45-degree line indicates a high degree of correlation between actual and predicted values.

Further to test the hourly energy prediction capability of the RFR model with changing weather conditions scenario 2 was conducted and the results are given in the following section.

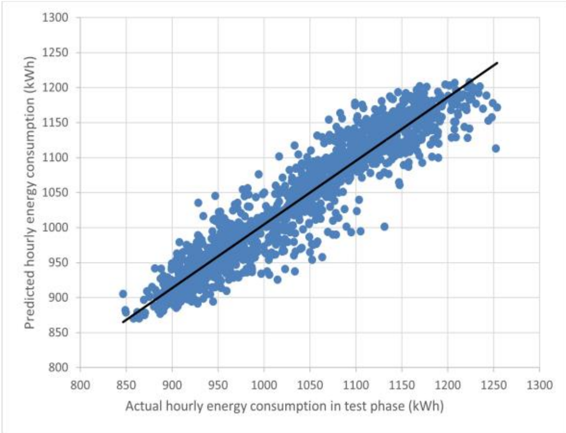


Figure 4: Scatter plot of predicted vs actual hourly energy consumption in the test phase of scenario 1

Scenario 2

Hourly energy consumption is predicted using the RFR model trained with a dataset of three cities: Ahmedabad, Bengaluru, and Lucknow, and the energy consumption of Kolkata is predicted. The performance of the RFR model is given in Table 10.

Table 10 RFR model performance in test phase of scenario 2

Scenario	Training points	R^2	MSE	RMSE (%)	MBE (%)	CV (%)
II	4,800	0.80	1825.1	42.72	8.96	8.80

The model performance evaluation shows an R^2 value of 0.80 and a CV of 8.80%. The results indicate that the prediction accuracy of the RFR model for scenario 2 is less than that for scenario 1. To understand the results figuratively, Figure 5 shows the predicted versus actual hourly energy consumption for test data of scenario 2.

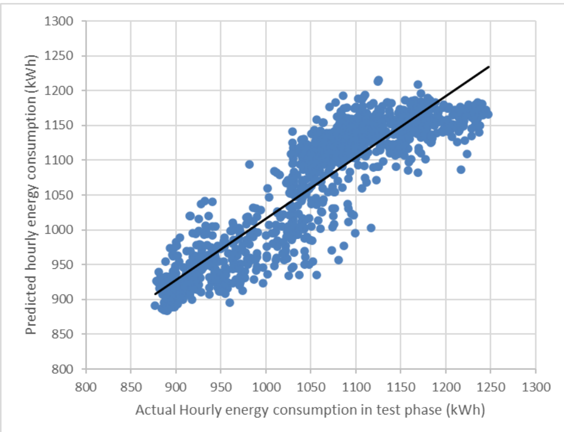


Figure 5 Scatter plot of predicted vs actual hourly energy consumption in the test phase of scenario 2

The graphs generated for the test data set of scenario 2 predicted from 1200 data points using the weather file of Kolkata to generate the scatter plot. Although the scatter plot concentrates along the 45-degree line, the distribution is not uniform, like in scenario 1. The observation is as per expectation because the ML model was tested for a dataset it was not trained with. However, the results are promising since MBE falls within the permissible limits of 10% and CV less than 30% (Yang et al. 2015). This shows that there is scope for further evaluating the use of ML models such as RFR for predicting energy consumption for hourly resolution under changing climatic conditions.

Scenario 1 and Scenario 2 results indicate that the RFR model is a suitable ML model for predicting energy consumption on an hourly basis due to an R^2 value of more than 0.8 in both cases and CV values of less than 10%. The RMSE value is 42.7% in scenario 2, while that in scenario 1 is 28.9%. This indicates a greater spread of the data points in scenario 2 from the mean value Figure 5, corresponding to scenario 2, also indicates the same. The MSE and MBE values for Scenario 2 are also greater than Scenario 1, indicating greater prediction accuracy in Scenario 1.

Conclusions

The study was conducted to demonstrate ML models' capability to predict commercial buildings' energy consumption in hourly resolution for annual time horizons under varying climatic conditions. An archetype reference commercial building model for an 8-hour working schedule representative of the Indian office buildings was used for this study. A methodology was also developed to generate a dataset and predict hourly energy for changing climatic conditions. Four ML models were used: MLR, PR, RFR, and ANN. The model performance evaluation on these four models identified RFR as the most accurate predictive ML model for hourly energy predictions. Thus, the RFR model was used for further evaluation in Scenario 1 and Scenario 2. Scenario 1 was conducted to test the holistic prediction capability of the RFR model and the results of Scenario 1 indicate that the RFR model is suitable for hourly energy prediction. Scenario 2 was conducted to check the capability of the ML model to predict hourly energy consumption for test data for which it was not trained. Model performance evaluation results for scenario 2, although less accurate than scenario 1, are promising. This indicates that there is scope for using ML models such as RFR for predicting hourly energy consumption and changing weather conditions.

The comparison of Scenarios 1 and 2 results indicates that the relationship between input weather features and building energy output needs further study. Further studies in this direction can consider ensemble-based methods, which combine the output of multiple machine-learning algorithms to enhance the prediction performance of a single data-driven model. Commonly used ensemble methods such as bagging, boosting, and stacking may be deployed to improve the hourly energy prediction for changing weather conditions such as the one considered in this experiment in scenario 2. This will further aid in improving the efficiency of building operation systems and support the transition of existing buildings to achieve net-zero energy.

Acknowledgments

The authors acknowledge the funding support from IC-IMPACTS (India-Canada Centre for Innovative Multidisciplinary Partnerships to Accelerate Community Transformation and Sustainability) and DST (Department of Science and Technology), Government of India, grant number DST/IC/IC-IMPACTS/2022/P-7, as well as the Canada Research Chairs Program.

References

Bhatnagar, M., Mathur, J., & Garg, V. (2019). Development of reference building models for India. Journal of Building Engineering, 21, 267–277. https://doi.org/10.1016/j.jobe.2018.10.027

Bhatnagar, M., Mathur, J., Garg, V., & Iqbal, J. (2018). Development Of a Method for Selection of Representative City in A Climate Zone. www.ashrae.org

Chan, J., Shen, J., Valdes, O., Drope, L., & Frisque, A. (2022). An Automated Approach to Parametric Design Analysis for Large-scale Design Projects. Building Simulation Conference Proceedings, 3372–3378. https://doi.org/10.26868/25222708.2021.31080

Ciulla, G., & D'Amico, A. (2019). Building energy performance forecasting: A multiple linear regression approach. Applied Energy, 253. https://doi.org/10.1016/j.apenergy.2019.113500

Wang, L., & Liu, X. (2017). Prediction of the impacts of climatic change on energy consumption for a medium-sized office building with two climate models. Energy and Buildings, 157

Eco-Niwas Samhita. (2018). Energy Conservation Building Code for Residential Buildings, Bureau of Energy Efficiency (BEE), Government of India

India Energy Outlook (2021) International Energy Agency (IEA). World Energy Outlook Special Report. www.iea.org/t&c/

National Building Code of India, Volume 2 (2016) Bureau of Indian Standards (BIS)

Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modeling and energy performance prediction: A review. In Renewable and Sustainable Energy Reviews (Vol. 23,pp.272–288). https://doi.org/10.1016/j.rser.2013.03.004

García-Martín, E., Rodrigues, C. F., Riley, G., & Grahn, H. (2019). Estimation of energy consumption in machine learning. Journal of Parallel and Distributed Computing, 134, 75–88. https://doi.org/10.1016/j.jpdc.2019.07.007

Kumar, S. (2016). Performance Based Rating and Energy Performance Benchmarking for Commercial Buildings In India. https://www.researchgate.net/publication/266225743

Miller, C., & Meggers, F. (2017). The Building Data Genome Project: An open, public data set from non-residential building electrical meters. Energy Procedia, 122, 439–444. https://doi.org/10.1016/j.egypro.2017.07.400

Mosavi, A., Salimi, M., Ardabili, S. F., Rabczuk, T., Shamshirband, S., & Varkonyi-Koczy, A. R. (2019). State of the art of machine learning models in energy systems, a systematic review. In Energies (Vol. 12, Issue 7). MDPI AG. https://doi.org/10.3390/en12071301

Ngo, N. T. (2019). Early predicting cooling loads for energy-efficient design in office buildings by machine learning. Energy and Buildings, 182, 264–273. https://doi.org/10.1016/j.enbuild.2018.10.004

Olu-Ajayi, R., Alaka, H., Sulaimon, I., Sunmola, F., & Ajayi, S. (2022). Building energy consumption prediction for residential buildings using deep learning and other machine learning techniques. Journal of Building Engineering, 45. https://doi.org/10.1016/j.jobe.2021.103406

Pham, A. D., Ngo, N. T., Ha Truong, T. T., Huynh, N. T., & Truong, N. S. (2020). Predicting energy consumption in multiple buildings using machine learning for improving energy efficiency and sustainability. Journal of Cleaner Production, 260. https://doi.org/10.1016/j.jclepro.2020.121082

Energy Statistics India. (2022). Ministry of Statistics and Program Implementation (MoSPI), Government of India

Sun, Y., Haghighat, F., & Fung, B. C. M. (2020). A review of the state-of-the-art in data-driven approaches for building energy prediction. In Energy and Buildings (Vol. 221). Elsevier Ltd. https://doi.org/10.1016/j.enbuild.2020.110022

Tiberiu Catalina, V. I. B. C. (2012). Multiple regression model for fast prediction of the heating energy demand.

Yang, Z., & Becerik-Gerber, B. (2015). A model calibration framework for simultaneous multi-level building energy simulation. Applied Energy, 149, 415–431. https://doi.org/10.1016/j.apenergy.2015.03.048

Li, K., Xie, X., Xue, W., Dai, X., Chen, X., & Yang, X. (2018). A hybrid teaching-learning artificial neural network for building electrical energy consumption prediction. Energy and Buildings, 174, 323–334. https://doi.org/10.1016/j.enbuild.2018.06.017

Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. In Renewable and Sustainable Energy Reviews (Vol. 81, pp. 1192–1205). Elsevier Ltd. https://doi.org/10.1016/j.rser.2017.04.095

CONSTRUCTION DATA SPACE FOR BUILDING PERMIT MANAGEMENT

Gonzalo Gil[1], Francisco Javier Diez[1], Ricardo Romero[1], Ignacio Lazaro[1] and Izaskun Fernandez[1]
[1]Tekniker, Basque Research and Technology Alliance (BRTA), Eibar, Spain

Abstract

Building permit management involves complex compliance checking processes that today are mainly performed by sharing documents between stakeholders and reviewing them manually. This causes long delays in obtaining building permits. Based on the DigiPLACE Reference Architecture Framework for digital construction platform, this paper presents the DigiChecks Framework, a novel approach to digitize and automate building permit management based on the orchestration of compliance checks and a data space, a distributed and open infrastructure for interoperable, trusted, and sovereign data sharing between platforms of different stakeholders. The DigiChecks Framework reduces data sharing and compliance checking efforts, while it ensures strict compliance with existing regulations.

Introduction

The management of building permits is a process that requires numerous compliance checks based on the information shared between different stakeholders, including architects, engineers, construction companies, subcontractors, designers, certification companies, suppliers, and public authorities. The current low level of digitization in the construction sector leads to resource-consuming compliance checks, based mainly in documents sharing and manual review. This causes considerable delays in obtaining building permits.

The challenges in digitizing and automating the management of building permits are significant (Beach et al., 2013). Among all of them, the following three stand out.

1. For different countries, but also regions or municipalities, there is a need to verify compliance with different regulations (Rezgui et al., 2011).
2. To check compliance, coordination issues should be tackled by combining building data stored in different proprietary formats (Rezgui et.al., 2013).
3. The reluctance to share data needs to be avoided by providing control over the data (Jarke et al., 2019).

To address these challenges, based on the DigiPLACE Reference Architecture Framework (RAF) (David et al. 2021), this paper presents the DigiChecks Framework, a novel approach to digitize and automate compliance checks by combining an orchestration of compliance checks and a data space. The orchestration provides a modular and scalable approach to manage the different compliance checks to be performed based on current regulations. The data space, following the path for the digital transformation of the construction sector (Buhler et al., 2023), facilitates the necessary collaboration between stakeholders through an open and distributed infrastructure for interoperable, trusted, and sovereign data sharing between their different platforms.

The rest of the paper is structured as follows. First, the related work is presented. Second, based on observed needs concerning building permit management, the main principles of the DigiChecks Framework are stated. Third, based on the related work and considering the main principles of the DigiChecks Framework, the conceptual architecture of the DigiChecks Framework is described. Fourth, the DigiChecks Data Space which, as part of the DigiChecks Framework, provides an open and distributed infrastructure to enable the collaboration between platforms based on interoperable, trusted, and sovereign data sharing, is detailed. Finally, conclusions about this work are drawn.

Related Work

The DigiPLACE RAF sets common guidelines for building digital platforms for the construction sector. Considering them, the DigiChecks Framework proposes a novel approach to digitize and automate compliance checks by combining an orchestration of compliance checks and a data space. In what follows, the DigiPLACE RAF and the work carried out on the definition and design of data spaces is described.

DigiPLACE Reference Architecture Framework

The RAF is based on two guidelines: “Interoperability and Common Processes” and “Data Control”. They are described below.

“Interoperability and Common Processes” is divided into data interoperability, data and process management and governance. Data interoperability establishes the definition of data vocabularies based on open standards, common data formats and models, and standardized data access. Data and processes management sets the collaboration between stakeholders through common data environments and data lifecycle management based on digital twins. Governance proposes open accessibility to standards and community engagement to improve them. Figure 1 represents this.

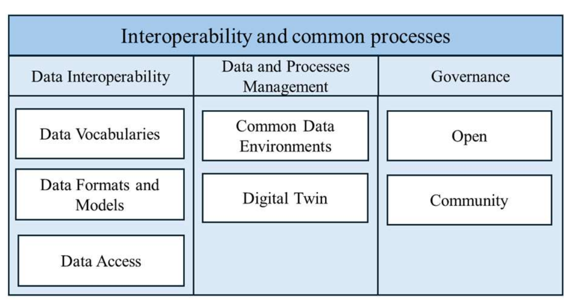


Figure 1: Interoperability and Common Processes

“Data Control” is divided into data security, data ownership, data trust and data availability and sustainability. Data security covers the security of the data, data sovereignty and data sharing within data spaces. Data ownership sets compliance with the GDPR, the ownership of the data and data transparency. Data trust encompasses data certification and value. Finally, data availability and sustainability establish compliance with FAIR principles as well as the sustainability of the data. Figure 2 represents this**.**

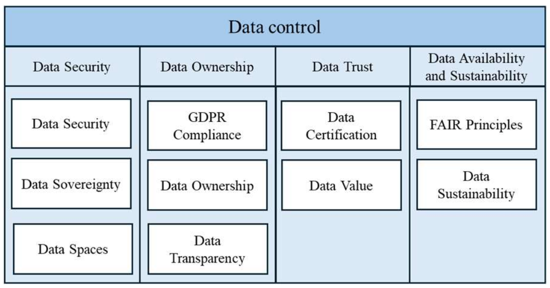


Figure 2: Data Control

Data Spaces

Nowadays, different initiatives are working worldwide on the definition and design of data spaces: The OPENDEI project, the Data Spaces Business Alliance (DSBA), GAIA-X and the International Data Spaces Association (IDSA) stand out. This large number of initiatives makes it difficult to clarify which building blocks are needed to build a data space and the technical components required to support them.

Recently, the design principles for data spaces resulting from the OPENDEI project (Nagel and Lycklama, 2021) and the common framework proposed subsequently by the DSBA (Gronlier et al., 2021) agree on the building blocks required to build a data space. These are detailed in Figure 3.

Figure 3: Data Space Building Blocks

In this context, GAIA-X and IDSA are positioned as the reference initiatives for the specification of the technical components that must be developed to support these building blocks. While GAIA-X defines the GAIA-X Technical Architecture 23.10 (GAIA-X European Association for Data and Cloud, 2023), a set of technical components for a governance framework for a federated infrastructure of services that complies with European standards and values of digital sovereignty, IDSA defines the IDSA Reference Architecture Model (RAM) 4.0 (International Data Spaces Association, 2023), a set of technical components for an interoperable, and sovereign data sharing framework between services.

On this basis and considering that the data space building blocks meet the DigiPLACE RAF core guidelines to build digital platforms in the construction sector, we consider the GAIA-X Technical Architecture 23.10 and the IDSA RAM 4.0 as the technical foundations to build the DigiChecks Data Space that acts as the core of the DigiChecks Framework.

DigiChecks Framework Principles

For the design of the DigiChecks Framework, a review of the main needs concerning building permit management has been made. These needs are summarized below:

1. Automated compliance checking cannot be achieved with one-step solution, but by combining multiple tools in the form of services provided by different organizations that gradually makes compliance checking more mature.
2. The wide variety of existing processes, tools and contexts means that there will be no one-size-fits-all solution, but a solution that enables the interaction with multiple services.
3. Individual services communicating with each other need a common language so that they can understand each other. The common language should follow a formal, explicit specification of a shared conceptualization.

These needs lead to the following three principles of the DigiChecks Framework.

1. The framework is a distributed infrastructure, rather than a centralized, monolithic application.
2. The framework is an open infrastructure with a modular and scalable approach, where services can be included and swapped easily.
3. The framework uses a shared set of conceptual information models that align with existing domain ones. Information models are formalized based upon open and widely accepted standards.

DigiChecks Framework Conceptual Architecture

Considering the three principles of the DigiChecks Framework, the DigiChecks Framework Conceptual Architecture is build based on the following four architecture principles:

1. The DigiChecks Framework Conceptual Architecture is a modular and distributed architecture in which the components have the flexibility to be deployed across multiple platforms managed by different stakeholders.

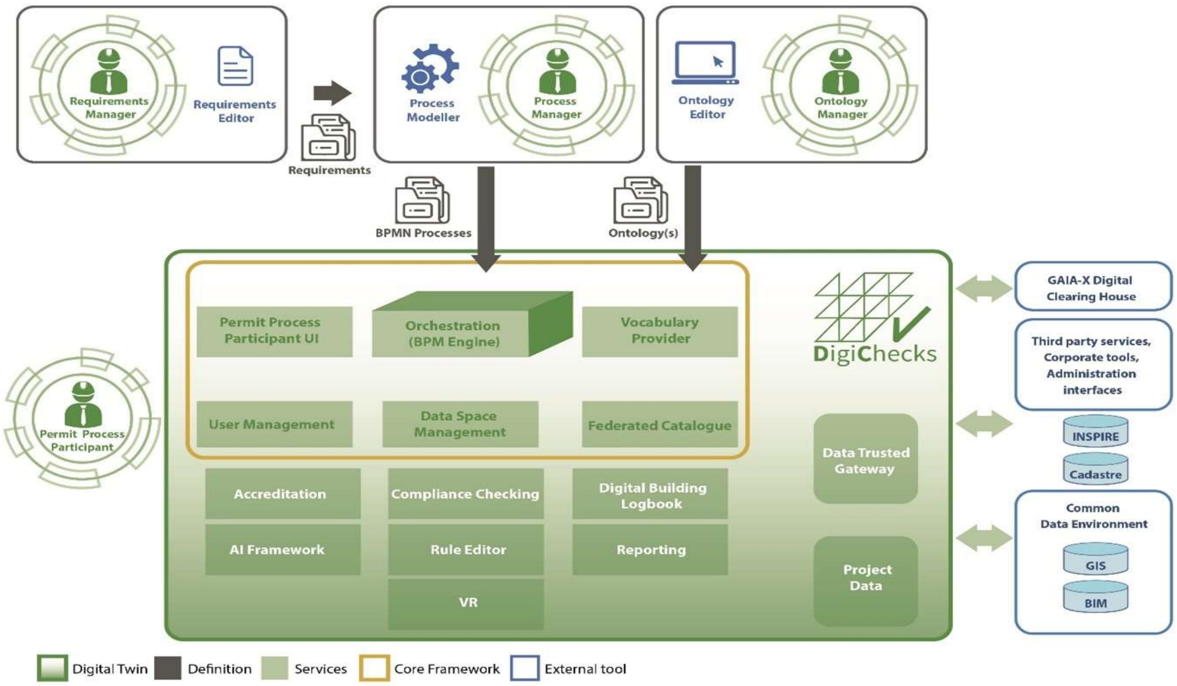


Figure 4: DigiChecks Framework Conceptual Architecture

2. The DigiChecks Framework Conceptual Architecture is based on the DigiChecks Data Space to enable the collaboration between the platforms od different stakeholders based on interoperable, trusted, and sovereign data sharing.
3. The DigiChecks Framework Conceptual Architecture provides Core Components that play an essential role in the orchestration of compliance checking processes and the management of the DigiChecks Data Space.
4. The DigiChecks Framework Conceptual Architecture can be extended with new components that provide additional functionalities for specific use cases.

Considering these four architecture principles, Figure 4 represents the DigiChecks Framework Conceptual Architecture. It can be divided into three main blocks:

- **Design tools:** They are services that support the definition of the main elements involved in the building permit management.
- **Core Framework:** It consists of unique services that provide core functionalities to manage building permits.
- **Digital Twin:** It is composed of the Core Framework and additional services that provide specific functionalities that in conjunction with the Core Framework enable the performance of modular and scalable compliance checks.

All these main blocks and their components are further described below.

Design Tools

The Design Tools are divided into the Requirements Editor, the Process Modeler, and the Ontology Editor.

The **Requirements Editor** assists the *Requirements Manager* in the definition of the requirements to be involved in compliance check processes. These requirements can be defined manually or extracted from structured documents and defined in a structured format.

The **Process Modeler** enables the *Process Manager* to define the processes in a Business Process Model and Annotation (BPMN) 2.0 standard format, defining who is responsible for the process, what requirements must be checked, and the data required to do so, what tasks and when must be performed and how they must be performed, either manually or automatically against and specific service.

The **Ontology Editor** supports the *Ontology Manager* to design the necessary semantic information model in form of ontology to represent the data to be shared in the DigiChecks Framework.

Core Framework

The Core Framework is divided into the Orchestration, the User Interface, the User Management, the Data Space Management, the Vocabulary Provider, and the Federated Catalogue.

The **Orchestration**, which is a Business Process Model (BPM) Engine, is the core component in charge of managing the full process of turning a BPMN process in an executable version. It orchestrates and manages the state of the tasks required.

The **User Interface** offers to the *Permit Process Administrator* a user interface to manage and monitor the processes and workflows of the defined permitting processes. Also, provide specific forms and screens to enable the *Participant* to provide the necessary information for any step in the permitting process, be able to see at any time the status of the process, to provide manual inputs and get the final resolution approval or rejection.

The **User Management** enables administrators to manage the access of users to the different services that compose the DigiChecks framework.

The **Data Space Management** provides the services required to manage the DigiChecks Data Space. These are the Dynamic Attribute Provisioning Service (DAPS) and the Clearing House. The DAPS ensures trusted data sharing within the DigiChecks framework. The Clearing House provides a service to log data transactions. As they are part of the DigiChecks Data Space, they are further described in the following Section.

The **Vocabulary Provider** manages the semantic information model designed through the Ontology Editor to ensure interoperable data sharing within the DigiChecks Framework. As part of the DigiChecks Data Space, it is further described in the following Section.

The **Federated Catalogue** provisions a catalog of services that have been certified according to European guidelines of digital sovereignty. As part of the DigiChecks Data Space, it is further described in the following Section.

The Digital Twin

The Digital Twin is composed of several services. Among others, the Compliance Checking, the Rule Editor, and the Interfaces with External Tools stand out. They are described below.

The **Compliance Checking** ensure the adherence of digital permitting processes to expert system rule-driven systems. It plays a crucial role in maintaining regulatory compliance and enforcing predefined rules and standards. By allowing operators to introduce and update rules, Compliance Checking empowers organizations to adapt their processes dynamically and stay aligned with evolving regulations.

The system leverages advanced algorithms and automation to evaluate the current state of the permitting process against the established rules. It constantly monitors and verifies whether the process complies with the required criteria, minimizing the risk of non-compliance and associated penalties. This automated approach significantly reduces the burden of manual verification, enabling organizations to streamline their operations efficiently.

The **Rule Editor** enables *Rule Managers* to define, in an easy-to-use graphical interface, the rules to be executed by Compliance Checking services to validate compliance with specific requirements.

The **Interfaces with External Tools** integrate external third-party services, corporate tools, and data sources with the DigiChecks Framework through the DigiChecks Data Space. Within this category are the Data Trusted Gateway and Project Data. While the Data Trusted Gateway is a generic integration component, the Project Data is specific to manage the Common Data Environment (CDE) supporting the integration of Building Information Model (BIM) and Geographic Information System (GIS) data.

It should be mentioned that several external data sources are of interest for automating the management of building permits. Cadastre and Inspire Geoportal external databases are integrated in the DigiChecks Framework to provide Industry Foundation Classes (IFC) models and other relevant geographical information. Also, the Common Data Environment (CDE) is integrated into DigiChecks Framework to provide building data involved in the compliance checking services and to supply advanced visualization of the BIM models through external services and tools such as navigate on the building model pieces or parts, or by enriched 3D visualizations.

DigiChecks Data Space

Based on the GAIA-X Technical Architecture 23.10 and the IDSA Reference Architecture Model 4.0, Figure 5 represents the DigiChecks Data Space. It can be divided into the technical components required for the role of the Core Participant and the role of the Intermediary that is provided in the DigiChecks Framework through the Data Space Management component within the Core Framework. These roles are described below.

Intermediary

Intermediaries act as trusted entities which assume a central role that creates benefit for Core Participants in the DigiChecks Data Space ensuring interoperability, providing trust, and creating new business models.

The technical components assigned to this category are the GAIA-X Digital Clearing House (GXDCH), the Federated Catalogue, the DAPS, the Vocabulary Provider, and the Clearing House.

The **GXDCH** operates and runs the services of the GAIA-X Framework. The GAIA-X Framework sets the rules that define whether organizations and the services they provide comply with European standards and values of digital sovereignty.

The services of the GXDCH are operated by GAIA-X Federators. Therefore, they are not really provided within

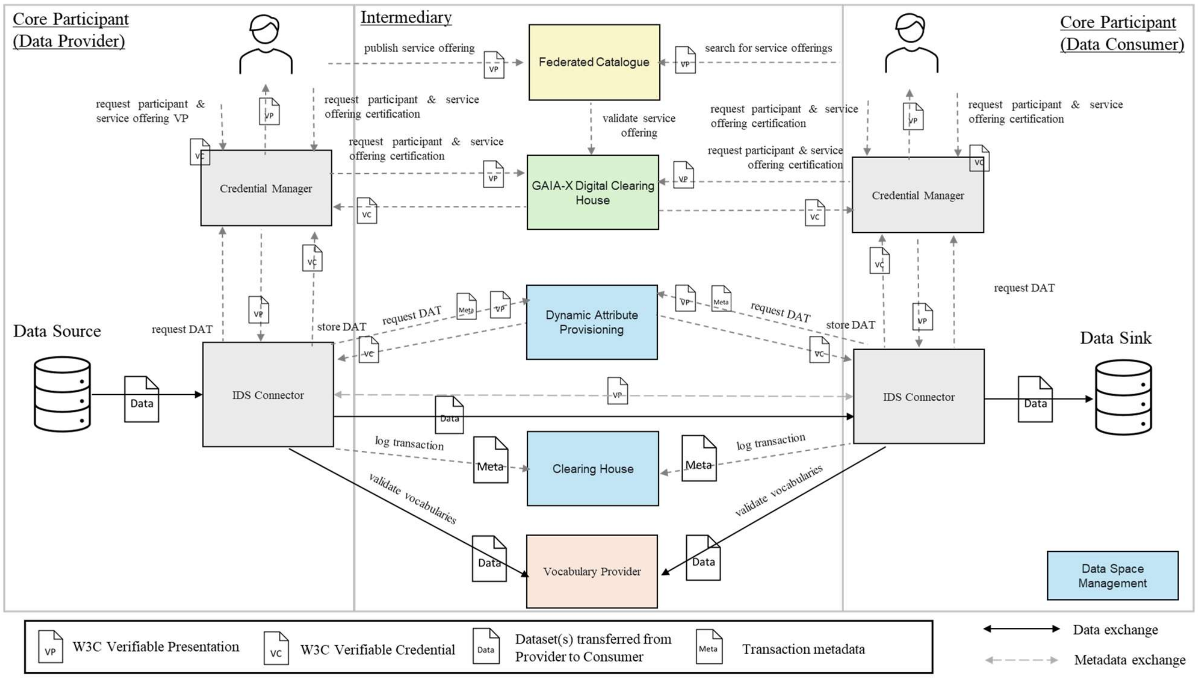


Figure 5: DigiChecks Data Space

the DigiChecks Data Space. These services are accessible to anyone to automatically verify compliance with the GAIA-X Trust Framework and obtain GAIA-X compliance attestations.

It is important to mention that the GXDCH relies on a user-centric and self-sovereign approach for identity management called Self-Sovereign Identity (Maier and Pohlmann, 2022) that is based on W3C recommendations of Decentralized Identifiers (DIDs)[1] and Verifiable Credentials (VC)[2]. Based on the information provided by Core Participants, the GXDCH verify compliance of either the organization or the services they provide with the GAIA-X Trust Framework. If satisfied, the GXDCH issues to the Core Participant a VC containing the GAIA-X compliance attestation. This VC is self-managed by the Core Participant. To enable other services to validate the integrity, the authenticity, and the validity of the issuer of the VC, the VC contains a cryptography proof.

The goal of the **Federated Catalogue** is to enable Core Participants within the DigiChecks Data Space to find best-matching service offerings, which comply with European standards and values of digital sovereignty.

To publish service offerings, the Federated Catalogue requests and validates VCs attesting compliance of the organization providing the service and the service itself with the GXDCH. To search for service offerings, it controls the access to the services available in the Federated Catalogue to organizations which comply with the GXDCH.

The **DAPS** sets the rules that define how an IDS Connector, the technical component of a Core Participant responsible for data sharing, manages data control. In other words, it states the level of trust of an IDS Connector. The IDS Connector is further described later.

As the trust level is higher, the requirements are higher, from a basic trust level where only interoperability is guaranteed, to higher levels that implement secure communication protocols, usage control etc.

As the level of trust can change in an operational environment over the time, once an IDS Connector is initially manually certified and deployed, the DAPS sets the operational trust level considering the one initially certified and operational environment data.

The DAPS, as the GXDCH, relies on DIDs and VCs. Therefore, it issues VCs that attest the trust level of an IDS Connector. This VC is self-managed by the Core Participant. It is used by the IDS Connector to access to other IDS Connectors.

The **Vocabulary Provider**, as part of the Data Space Management, hosts, maintains, and publishes the semantic information model that is designed using the Ontology Editor, describing the relevant concepts (e.g., “building”, “office space”, “toilet space” and “size”) shared by IDS Connectors in the DigiChecks Data Space to manage building permits in the construction domain. Furthermore, it implements Shapes Constraint Language

[1] Decentralized Identifiers (DIDs) v1.0 (w3.org)

[2] Verifiable Credentials Data Model v2.0 (w3.org)

(SHACL)[3] validations that are required by the IDS Connector to verify whether the expected data is received.

It is important to mention that the Vocabulary provider only includes the model. It does not include its instantiation (specific data) and therefore doesn't provide information about specific instances. The instances should be managed and provided by Core Participants through the IDS Connectors.

The **Clearing House**, as part of the Data Space Management, records all the relevant information related to data sharing between the IDS Connectors of Core Participants. This includes contract agreements and data usage information for clearing, settlement and billing services.

Core Participant

A Core Participant can be assumed by any organization that provides or consumes data within the DigiChecks Data Space. Roles assigned to this category are the Data Provider and the Data Consumer.

To share data between the Data Provider and the Data Consumer, both uses a software component that is compliant with the IDSA RAM. This is the IDS Connector. To manage the VCs issued by the GXDCH and the DAPS, the Data Provider and the Data Consumer use a software component called Credential Manager.

The **IDS Connector** is the technical component responsible for interoperable, trusted, and sovereign data sharing between Core Participants within the DigiChecks Data Space.

To enable a modular and scalable development of the services, the DigiChecks framework is based on a microservices approach (De Lauretis, 2021). For a system that provides a specific service, while the user interface layer enables participants to manage the service itself, the IDS Connector provides those layers to share data within the DigiChecks Framework in an interoperable, trusted, and sovereign way. Figure 6 represents these system layers.

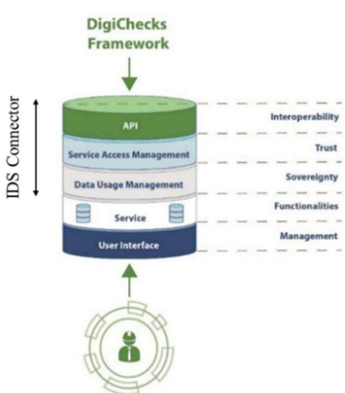


Figure 6: DigiChecks Framework Microservices System Layers

The Application Programming Interface (API), the Service Access Management and Data Usage Management layers are described below.

The API ensures interoperable data sharing. To this end, it implements the Dataspace protocol v1.0 (International Data Spaces Association, 2024) and several Data transfer protocols.

The Dataspace protocol is a set of specifications that define the interfaces, protocols, and schemas to expose a data catalog, to negotiate contracts and access negotiated data. To describe the data catalog, the Dataspace protocol proposes the Data Catalog Vocabulary (DCAT)[4]. It is a vocabulary designed to facilitate interoperability between data catalogs. It should be noted that DCAT is domain-agnostic. To provide a domain-specific description of the data, DCAT is extended with domain-specific vocabularies, which are provided by the Vocabulary Provider. Once negotiated data is accessed, data sharing is delegated to data transfer protocols.

Data transfer protocols are traditional protocols, such as HTTP or MQTT, that will be used for data transfer after data access. They are also integrated with the Vocabulary Provider. As the Vocabulary Provider implements SHACL validations, it can be ensured that the expected data is received.

The Service Access Management ensures trusted data sharing. It controls the access to the IDS Connector based on the VC that attests the level of trust of a requesting IDS Connector. For authentication purposes, it validates the integrity, the authenticity, and the validity of the issuer of the VC. For authorization purposes, it analyses the attributes of the VC.

The Data Usage Management ensures sovereign data sharing. Data sovereignty is referred as the self-determination of organizations regarding the usage of their data (Jarke et al., 2019). To grant data sovereignty, the IDS Connector implements Distributed Usage Control (Gil et al., 2023). It is a particularization of Usage Control for data sharing scenarios that extends Access Control to control what must happen to data through its life cycle (Jung and Dorr., 2022).

The **Credential Manager** abstracts Core Participants from the complexity of managing VCs.

On the one hand, it provides Core Participants with the interfaces to request the certification of the organization itself and the services it provides to the GXDCH, store the VC attesting the result and retrieve them.

On the other hand, it provides the IDS Connector with the interfaces to store retrieved VC from the DAPS and retrieve them for authentication purposes.

Conclusions

This paper addresses the main challenges regarding the automated management of building permits. These are the lack of standards between regulations in different

[3] Shapes Constraint Language (SHACL) (w3.org)

[4] Data Catalog Vocabulary (DCAT) - Version 3 (w3.org)

countries and regions and municipalities, interoperability between data in different formats and the control of the data.

Considering the DigiPLACE RAF, which sets the common guidelines for building digital platforms for the construction sector, and with the aim of addressing the challenges identified, this paper presents the DigiChecks Framework, a novel approach to digitalize and automate building permit management that combines an orchestration of compliance checks and a data space. While the orchestration provides a modular and scalable approach to verify different compliance checks depending on existing regulations, the data space facilitates the collaboration between stakeholders through interoperable, trusted, and sovereign data sharing between their different platforms. Thus, the DigiChecks Framework reduces data sharing and compliance checking efforts, while it ensures strict compliance with existing regulations.

Future Work

The conceptual architecture of the DigiChecks Data Space, which is the main development within the DigiChecks Framework and the focus of this paper, is based on existing standards defined by GAIA-X and IDSA. In this regard, the defined technical components are built from open-source tools that have already been developed following these standards or from new tools developed from scratch. The application of these tools to real-world scenarios presents two major challenges. First, as data spaces is a novel concept, these standards are continuously evolving. As the piloting stage approaches in the DigiChecks project, there is a need to make decisions on the versions of the standards to be used for two main purposes. For those tools already developed, to select the version to be used. For those new tools required, to select the version to be followed in the development. Second, different pilots lead to different requirements. Although existing tools are developed to be 100% agnostic to the use case requirements, the truth is that they should be adapted in about a 10% to the use case. This may lead to weaknesses in our approach that should be addressed by providing extensions to existing tools. For example, in the case of the IDS Connector, specific data transfer protocols may be required for the different pilots. If they are not implemented on existing tools, they must be implemented on top of an existing tool.

Having said that, at the end of the piloting phase, with the tools already selected and adapted, the DigiChecks Data Space technical architecture is expected to be published considering the version of the standards followed, the tools selected, the specific challenges addressed for their application in real-world scenarios and how they have been addressed.

Acknowledgments

This work has received support from Horizon Europe project DigiChecks (Digital environment for management of permits and compliance in building and construction) with gran agreement number 101058541.

References

Beach, T.H., Kasim, T., Li, H., Nisbet, N., Rezgui, Y. (2013). Towards Automated Compliance Checking in the Construction Industry.

Bühler, M., Nübel, K., Jelinek, T., Riechert, D., Bauer, T., Schmid, T., and Schneider, M. (2023). Data cooperatives as catalyst for collaboration, data sharing and the digital transformation of the construction sector.

David, A., Zarli, A., Mirarchi, C., Naville, N., Perisich, L. (2021). DigiPLACE: Towards a reference architecture framework for digital platforms in the EU construction, pages: 511-518

De Lauretis, L. (2019). From monolithic architecture to microservices architecture. 2019 IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW), pages 93–96.

GAIA-X European Association for Data and Cloud (2023). GAIA-X Architecture Document 23.10.

Gil, G., Arnaiz, A., Higuero, M. (2023). Advances in Sovereign Data Sharing: Identification and Assessment of the Main Features of Distributed Usage Control Solutions and Improvements in the Policy Quality.

Gronlier, P., Hierro, J., and Steinbuss, S. (2021). Data Spaces Technical Convergence v2.0.

International Data Spaces Association (2023). IDS Reference Architecture Model 4.0.

International Data Spaces Association (2024): Dataspace protocol v1.0.

Jarke, M., Otto, B., and Ram, S. (2019). Data sovereignty and data space ecosystems. Bus Inf Syst Eng, 61:549–550.

Jung, C. and Dörr, J. (2022). Data Usage Control, pages 129–146. Springer International Publishing, Cham.

Maier, B., Pohlmann, N. (2022). GAIA-X secure and trustworthy ecosystems with Self Sovereign Identity

Nagel, L. and Lycklama, D. (2021). Design Principles for Data Spaces

Rezgui, Y., Miles, J. (2011). Harvesting and Managing Knowledge in Construction.

Rezgui, Y., Beach, T., & Rana, O. (2013). A governance approach for bim management across lifecycle and supply chains using mixed modes of information delivery.

IDENTIFING REFERENCE DISTRICT BYMEANS OF MACHINE LEARNING AND OPEN-SOURCE DATA

Thi Thu Ha Dam[1], Maxim Shamovich[1], Julia Schlueter[2], Jannes Kruse[2]
Elisabeth Beusker[2], Jérôme Frisch[1] and Christoph van Treeck[1]
[1]Institute of Energy Efficiency and Sustainable Building (E3D), RWTH Aachen University, Germany
[2]Teaching and Research Field for Real Estate Development (IPE), RWTH Aachen University, Germany

Abstract

This study addresses neighborhood-level sustainability in urban development as a critical lever in Germany's energy transition goals. It explores identifying reference districts through Geographic Information System (GIS)-based machine learning (k-means algorithm) and public data, utilizing clustering methods to analyze spatial and socio-infrastructural metrics. The methodology yields significant insights into district definition and characterization, integrating technical and human understanding of urban dynamics. The findings highlight the importance of attribute selection in neighborhood classification and extend beyond mathematical validation to include social context comprehension. The developed technique is applied to a case study involving the city of Aachen.

Introduction

In the energy transition, particularly within the building sector, renewable energy systems are crucial for a climate-neutral energy supply, with heating and hot water consumption constituting a significant portion of Germany's building energy use (BMWK-Bundesministerium für Wirtschaft und Klimaschutz, 2022). The traditional, fossil-based energy infrastructure often limits the implementation of innovative, demand-driven thermal solutions (Acksel et al., 2017). District solutions are pivotal in optimizing energy potential and achieving economies of scale through sector coupling (Reicher et al., 2020). This approach integrates energy-saving strategies, efficiency, and renewable use, considering building characteristics, user behavior, and technical infrastructures (Többermann and Yu, 2021).

However, municipalities' current practices in urban-scale energy system planning are fragmented, often leading to slow and inefficient implementation (Boenigk et al., 2019). To address this, the paper advocates for establishing 'reference districts' as a means to streamline and standardize energy solutions, facilitating their application within and across various municipalities. Therefore, the paper discusses the challenges in defining district boundaries, especially in existing, evolving neighborhoods, and proposes using GIS and machine learning techniques on public datasets to identify and classify districts based on diverse attributes. This approach aims to enhance the scalability and effectiveness of energy solutions within existing infrastructures.

The concept of a neighborhood or district, lacking a universal definition, is interpreted variably across urban planning, cultural, and social perspectives (Feldmann, 2009). In urban planning, districts are defined by physical attributes such as building density, types, ages, and locations, including aspects like energy-efficient urban aggregation (Reicher, 2013; Reicher et al., 2020). Socially, they are important spaces for daily life, social interaction, and identity formation, embodying a range of physical to historical dimensions, thus influencing individual and communal life (Schnur, 2008). Neighborhood sizes vary, ranging from large housing estates to small settlements, with manageability and social identification as key criteria, typically not exceeding 20,000 residents (Bundesinstitut für Bau-, Stadt- und Raumforschung, 2012; Mehnert and Kremer-Preiß, 2014). Neighborhood attributes encompass a diverse spectrum including functional, spatial, socio-demographic, technical, economic, cultural-historical, and legal aspects, with differentiating factors such as natural boundaries, architectural typology, social structure, and community spaces (Malottki et al., 2013).

In case of newly planned districts, specific demarcation and energy zoning are simple, as they are planned in advance and meet current energy standards. In the case of existing districts, however, the question of boundaries arises because they are continuously developed, changed, and have grown over time. The question is, how can boundaries be defined in order to utilize energy potential in the context of the district?

For the practical application of district energy solutions, it is vital to recognize the diversity among district types. This paper puts forth a strategy for establishing neighborhood boundaries by employing georeferenced information combined with machine learning techniques on accessible datasets. Such techniques enable the amplification of energy solutions in pre-existing infrastructure, by creating reference districts defined by specific attributes. These can range from singular features like the type of buildings to an amalgamation of several traits. Consequently, this aids in the convenient replication of district energy solutions in comparable neighborhoods, resulting in a wide spectrum of reference districts, each marked by its distinct set of characteristics.

Related Research

This literature review explores statistical approaches, machine learning, and GIS in urban planning, emphasizing their role in energy evaluation and spatial analysis. It covers various approaches to mapping urban energy properties, classifying urban areas, and developing tools for sustainable urban development, highlighting the advancements and applications of these technologies in urban morphology analysis.

In the field of urban energy evaluation, several key projects and methodologies have emerged. The UrbanReNet project by Dettmar et al. (2020) systematically mapped the energetic properties of urban areas, analyzing prototypical elements of urban architecture and spaces for energy and structural values. This approach, combining urban morphology and land use, allows for a representation of Germany's building stock. It integrates qualitative and quantitative methods for evaluating energy generation, storage, and networking, leading to the development of mathematical models and software tools for neighborhood-level energy supply concepts.

GIS-based analyses have been pivotal in understanding energy dynamics in urban areas. März (2016) employed GIS-based multi-criteria decision analysis to locate neighborhoods at risk of energy poverty, focusing on space heating. Alpagut et al. (2021) and Quénéhervé et al. (2018) used GIS to optimize land use for solar production and its local energy impact.

In identifying and classifying settlement areas, Jochem et al. (2018) utilized geo-based vector data and machine learning to classify residential settlements in Afghanistan, while Gonzalez et al. (2020) applied deep learning to identify urban building typologies. Arribas-Bel et al. (2021) and Perez et al. (2020) explored urban cluster analyses using modified DBSCAN and Bayesian Networks, respectively, to analyze building types and urban functions.

Regarding city planning and administration, considering neighborhood dynamics, Photis (2012) developed the SPiRAL algorithm for redistricting electoral districts. López-Moreno et al. (2022) introduced a GIS-based approach for classifying residential areas in Madrid, aiding in energy-efficient urban renewal strategies. Similarly, Kelm et al. (2019) applied a semi-automatic approach using official geobase data in North Rhine-Westphalia (NRW) to create block structures for city planning and administration, demonstrating the growing importance of technology and data in urban planning and energy management.

Our methodology's choice of attributes for analysis was guided by this literature review on urban neighborhood characteristics, which highlighted key socio-infrastructural and energy-related factors to include in our study.

Methodology

Data Sources

In the process of identifying various reference districts in Aachen, NRW, a robust data foundation is essential. To conduct a comprehensive analysis, various public data sources were utilized in this study, with the main focus on the city of Aachen, located in western Germany. An overview of the freely available data sources used can be found in Table 1.

A central dataset titled ALKIS Real Estate Cadastre provides detailed information about plots and buildings in NRW. According to the Surveying and Cadastre Law of the state of NRW (VermKatG NRW), it offers extensive information about the structure and identification of plots and buildings. The cadastre, in which these data are recorded, offers a wealth of information, including geometric and geographical data, usage, size, and development of each plot.

Additional datasets were sourced from the OpenGeodata.NRW portal, which includes, among others, information on general and psychiatric hospitals, was last updated in June 2023. Complementarily, locations of childcare centers (KiTas) and schools were also incorporated into the analysis, as they represent crucial social spaces within a neighborhood.

Energy-related data, such as the location of renewable energy sources or potential areas for renewable energy installations, is sourced from the Energieatlas NRW. It is collected and provided by the State Agency for Nature, Environment, and Consumer Protection of NRW.

The Approach

In this section of our study, we describe the methodology adopted for analyzing reference districts through a bi-fold approach focusing on socio-infrastructural factors and energy sources. The objective was to leverage the clustering of various public datasets to gain insights into neighborhood structures.

Multiple variables are considered to provide a comprehensive view of the various characteristics and dynamics within a neighborhood. This is achieved by first creating a so-called feature matrix. A feature matrix is a table where each row corresponds to an object (in this case, a building), and each column represents a feature or characteristic of that object (e.g., the distance to churches). In this context, the features represent urban planning factors. Each entry in this matrix denotes the value of the corresponding building feature. Once the feature matrix is established, a clustering algorithm can be applied. The first clustering process aims to identify districts that show similarities in terms of the features captured in the matrix. After the initial clustering, in which districts are identified, the dataset is enriched with additional information regarding the building type and specific and absolute heating demands from the Wärmekaster NRW dataset. Each district is characterized by the total amount of contained buildings, the distribution of building types, and the district's aggregated heating demand. Districts with similar feature values are grouped into the same cluster, resulting in a reference district, while those with significantly different values are classified into different clusters. This approach enables a multidimensional examination of the neighborhoods, allowing us to uncover hidden relationships and patterns between neighborhoods that might be overlooked in isolated analyses. Two different feature sets are considered and are described in the following sections.

Table 1: Overview of the utilized freely available data sources

Dataset	Description	Source
ALKIS Real Estate Cadastre	Buildings, Plots	Bezirksregierung Köln (2024)
Geoportal.NRW	Locations of Hospitals and Schools	Lanuv NRW (2024b)
Wärmekataster	Building Category, Building Type, Heat Demand	Lanuv NRW (2024c)
Open Street Map	Streets, Religious Institutions, Parks, Supermarkets, Car-sharing, Restaurants	OpenStreetMap Contributors (2024)
Solarkataster	Locations of Rooftop and Open-space PV	Lanuv NRW (2024a)
Energieatlas NRW	Biomass, Wind power plants, Hydropower plants, Lignite, Natural gas, Sewage gas, Mineral oil, Hard coal, Mine gas, Landfill gas	Bezirksregierung Köln (2024)

Socio-Infrastructural Approach

Various socio-infrastructural data were analyzed to examine a neighborhood's social structure. These will be described in the following.

In this study, building density is understood as the distance between individual buildings, which can provide insights into population density and the intensity of residential developments in a neighborhood. Typically, buildings in urban areas are constructed in close proximity to each other, whereas in rural areas, the distance between buildings tends to be greater and more variable. For instance, town or terraced houses often exhibit uniform distances from one another.

Considering the distance of buildings to religious institutions might reveal social and urban structural patterns. This concept dates back to ancient and medieval times, when churches, temples, and even city halls evolved as focal points of urban planning (Kaupp (2022)).

Furthermore, green spaces enhance the quality of life for adjacent residents, offering areas for both social interaction and individual recreation. It is anticipated that the proximity to parks influences the likelihood of residents utilizing these areas for leisure activities, potentially shaping local social and cultural boundaries.

In the state of NRW, according to the statewide hospital plan, accessibility to hospitals within a 20-minute drive should be guaranteed for 90% of its citizens (Ministerium für Arbeit Gesundheit und Soziales des Landes Nordrhein-Westfalen, 2023). In this context, the question arises whether the location of buildings and, thus, their proximity to hospitals plays a role in defining neighborhoods. Consequently, distance was selected as a relevant attribute for investigating this relationship. Additionally, the locations of schools and childcare centers might play a role in the social delineation of neighborhoods. Children often attend the school or childcare center closest to their residence. Areas with such facilities are typically characterized by family-friendly infrastructure, including playgrounds and similar amenities. Hence, the distance to playgrounds is also considered a parameter for investigation.

Moreover, local businesses could influence the character of neighborhoods. The distance to restaurants, for instance, might indicate the liveliness of an area. Accordingly, distances to restaurants and supermarkets were included in the study.

The spatial proximity to office buildings can indicate places of work and potentially offer new urban planning insights. Office buildings, often part of non-residential structures, tend to concentrate in specific areas along with other non-residential structures.

Furthermore, the structure of districts is often shaped by the traffic route network. Examining this factor could reveal how street layouts influence the formation of neighborhoods. In urban areas, car-sharing services are increasingly spreading. Analyzing the locations of such car-sharing facilities could provide indications of neighborhood boundaries.

In calculating distance as a feature, the shortest distance to the destination, for example, the nearest school, was determined. For streets with a speed limit of 50 km/h, the shortest orthogonal distance from a building to the street was chosen as the feature.

Table 2: Clustering attributes for Socio-infrastructural approach

Building density	Car-sharing	Religious institutions
Office buildings	Residential buildings	Childcare centers
Playgrounds	Restaurants	Supermarkets
Schools	Parks	Hospitals
Roads (max. 50 km/h)		

The complete set of attributes relevant to the Socio-Infrastructure Model is detailed in Table 2.

Energy Approach

In addressing the energy aspect of our study, we focus on the significant shift from fossil fuels to renewable energy sources, a strategic move by the Federal Government to counteract climate change. This transition is particularly pertinent in NRW, Germany's most populous state, which boasts a wide array of energy sources. These range from traditional fuels like lignite and mined gas to innovative renewable technologies, including biomass and photovoltaic (PV) energy generation. The pivot towards renewable sources signifies the phasing out of fossil fuels, underscoring the need to ensure continuous access to energy sources for the future. This is vital for maintaining an efficient energy supply chain. The crux of our energy approach is the development of an urban model designed to map out the distribution of various neighborhoods based on energy sources within urban settings. A key aspect of this model involves analyzing the proximity of buildings to diverse energy carriers, thereby identifying potential areas for the integration of renewable energy solutions based on the density of energy sources. By mapping the city's proximity to these energy sources, our model not only highlights Aachen's current energy infrastructure but also provides a blueprint for enhancing renewable energy uptake. This approach aligns with national efforts to transition towards a more sustainable and environmentally friendly energy mix, reflecting a commitment to reducing carbon emissions and combating climate change. The following attributes were considered for the clustering:

- Roof-surfaces PV
- Open-space PV
- Wind power plants
- Hydroelectric power plants
- Sewage gas
- Biomass availability

Integrating these two approaches enables us to construct a multidimensional description of urban neighborhoods, highlighting both challenges and opportunities in socio-infrastructural development, energy efficiency, and renewable energy utilization. Subsequently, districts from both approaches are enriched with data on building types and heating demand, resulting in a comprehensive dataset encompassing all identified districts and their attributes. An extra clustering step is then conducted to define reference districts, enabling the assignment of district types from the social infrastructure and energy perspectives.

After careful consideration of various algorithms, the k-means clustering method was chosen based on its efficiency in processing large datasets and its simplicity in implementation. Furthermore, the main hyperparameter k, the number of clusters, is highly interpretable, allowing expert knowledge to inform the algorithm tuning process. While DBSCAN was considered for its density-based clustering capabilities, it is earmarked for future work; however, it was discarded for the first proof of concept due to its complex parameter tuning.

Further, to evaluate the effectiveness of the k-means algorithm in producing meaningful clusters, we employ the Silhouette method. The Silhouette method is a popular evaluation technique, offering a quantifiable measure of cluster cohesion and separation, providing a robust justification for the selected number of clusters compared to other methods like the Elbow Method. It provides a succinct graphical representation of how well each object has been classified. The Silhouette score ranges from -1 to +1, where a high value indicates that the object is well-matched to its own cluster and poorly matched to neighboring clusters. This method is particularly useful in determining the optimal number of clusters and in assessing the quality of the clusters formed by the algorithm. A hyperparameter study was conducted using the Silhouette method for k-values ranging from 6 to 70. The minimum number of clusters was determined based on the number of the seven main administrative districts in the city of Aachen. The choice of the upper limit of 70 was also informed by the results of the Social Development Plan of the city of Aachen (Stadt Aachen, 2020).

Through this methodological approach, we aim to achieve a robust clustering of the given data to find definable reference districts, providing valuable insights into the underlying patterns and relationships within the dataset.

The study used the High-Performance Computing Cluster (HPC) at RWTH Aachen University for intensive computations. Due to QGIS's limitations in processing large datasets and tuning hyperparameters, it was replaced by Python for analysis, while QGIS was used only for visualization. The research employed Python libraries like scikit-learn for machine learning, GeoPandas and Pandas for data management and feature selection, and Fiona for loading geo-referenced data.

Results

Socio-infrastructural Approach

The methodology focussing on the socio-infrastructural attributes enabled us to delineate 54 unique districts, each representing a specific typology. The resulting cluster map can be seen in Figure 1. It can be observed that the larger clusters tend to be near and around the city center, which can be explained by the higher density of social infrastructure in the historical urban core.

Our analysis revealed significant differentiation among the clusters, particularly in terms of their spatial distribution and proximity to key urban infrastructures. For instance, significant variations were observed when examining the distribution of average distances of all clusters to residential buildings. Some clusters were characterized by notably shorter distances to residential areas, suggesting a predominance of residential use within those districts. This observation hints at the potential functional specialization of neighborhoods, where some are more residentially-focused while others may serve different urban functions.

Further examination of the average distance to religious in-

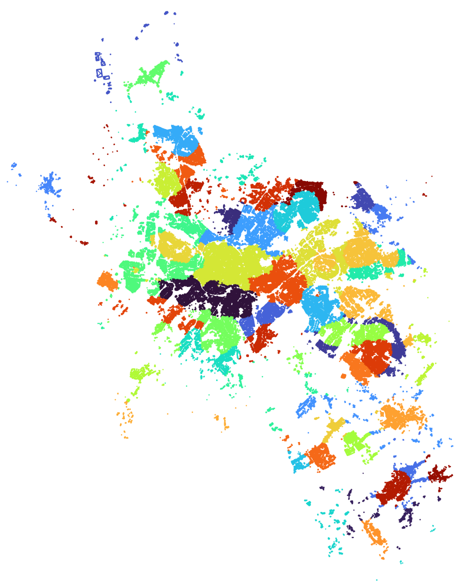


Figure 1: Clustering result of the Socio-Infrastructural-Model: 54 districts were identified

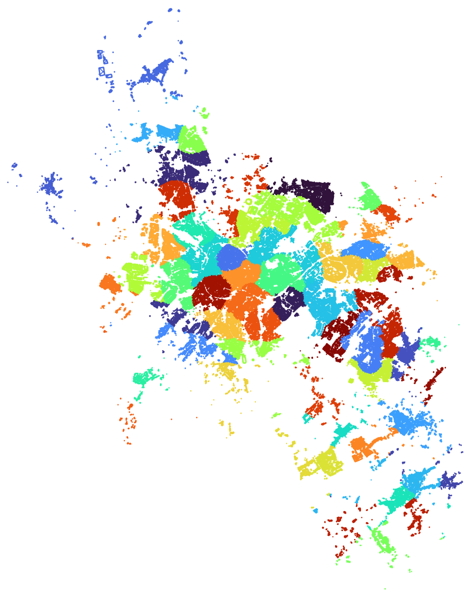


Figure 2: Clustering result of the Energy-Model: 58 districts were identified

stitutions across clusters also highlighted distinct patterns. Certain clusters demonstrated closer proximity to these institutions, reflecting their potential role in the social and cultural fabric of those neighborhoods. This differentiation in proximity to key urban elements underscores the diverse character of urban districts and their unique contributions to the city's overall structure.

Additionally, the analysis of the distribution of average distances for various attributes revealed nuanced differences between clusters. While some attributes showed a relatively linear distribution across clusters (e.g., distance to healthcare facilities), others exhibited significant variability (e.g., distances to childcare centers).

The observed variation indicates that specific urban characteristics might play a more significant role in distinguishing neighborhood types. There's a notable disparity in the number of buildings per cluster; for instance, clusters in the city center encompass over 6000 buildings, whereas those on the city's outskirts may have as few as 65 buildings. Additionally, there's a trend where the proportion of non-residential buildings in a cluster tends to rise with its distance from the city center. Meanwhile, the heating demand seems to be proportional to the cluster size.

Energy Approach

While looking at the different locations of energy resources in the city of Aachen, a total of 58 different clusters were found, with a Silhouette value of 0.38. Figure 2 shows the different clusters based on the location of various energy sources.

The clusters, each representing a unique combination of energy source proximity, were visually represented to provide insights into the spatial distribution of energy infrastructure across the urban fabric. This visualization underscores the variance in access to different energy sources across neighborhoods, such as biomass producers and PV installations. This examination revealed notable differences in how closely neighborhoods are situated to renewable energy sources, highlighting areas with potential for further development of sustainable energy solutions.

In particular, the study found varying degrees of proximity to biomass energy producers and rooftop PV systems across the clusters. Some clusters demonstrated close proximity to biomass sources, suggesting an emphasis on bioenergy utilization, while others were characterized by greater distances, indicating potential areas for increased biomass energy integration. Similarly, the examination of distances to PV installations provided a snapshot of the penetration of solar energy within urban areas, identifying clusters that could benefit from enhanced solar energy deployment.

In contrast to the first method, the range of building distribution per identified cluster is significantly narrower, with the highest count being 2334 buildings and the lowest at 70. Additionally, the total heating demand correlates less with the number of buildings and more with the types of buildings present. The distribution of building types per cluster also exhibits a less steep gradient compared to that observed in the first approach.

This clustering analysis sheds light on the intricate relationship between urban form and energy infrastructure, offering a foundation for targeted urban planning and energy policy development. By pinpointing districts with specific energy characteristics, policymakers and planners can tailor strategies to optimize energy efficiency and expand renewable energy use within urban ecosystems. This approach not only contributes to the sustainability of urban areas but also supports broader goals of energy transition and climate change mitigation.

Reference Districts

Integrating districts from socio-infrastructure and energy methods into a unified clustering process resulted in the identification of six unique reference district types. These types were defined by key characteristics: the total number of buildings in each cluster, the variety of building types, and the total aggregated heating demand. This comprehensive overlay of districts from both methods facilitated their categorization into one of the six reference types, as depicted in Figure 3. The clustering effectiveness, measured by the Silhouette score, was 0.433 for k=6.

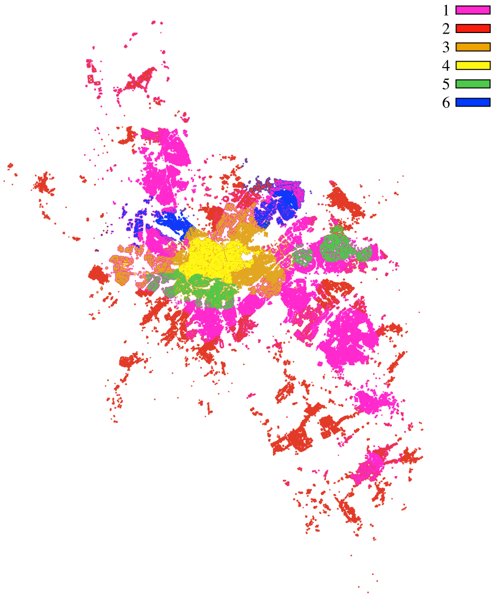


Figure 3: Results of the Reference Districts Clustering: Overlay of Socio-Infrastructural and Energy Model Districts

On the map, it is clearly visible that the reference district types form an orbicular pattern around the city center. Table 3 summarizes the mean distributions of each building type and specific heating demand for the six reference district types.

Cluster number four represents Aachen's urban center and is the city's most densely populated region, which is evident from the high number of buildings in this reference district. Characterized predominantly by large multi-family housing (BMFH), this cluster has the highest specific heating demand among all clusters, significantly exceeding others. In contrast, cluster two, encompassing the city's outskirts, has the lowest heating demand. Clusters one and six share striking similarities, not only in residential building distribution but also in heating demand and location, as depicted in Figure 3. However, there are significant differences in their absolute heating demands. The total demand of cluster six is nearly 1.5 times higher than that of cluster one.

A detailed examination of Table 3 highlights the primary emphasis of the clustering process. It shows that each identified reference district type is predominantly characterized by a single building type.

Discussion

The clustering approach uses a feature matrix aimed at a holistic approach, considering multiple attributes simultaneously. In this study, while defining two base models for identifying typological neighborhoods using public datasets, it was crucial to select relevant attributes, as irrelevant ones could distort outcomes. Selecting specific attributes for clustering still remained a subjective challenge. The approach incorporated biases related to socio-infrastructural aspects to provide insights into patterns that may be prevalent across other communities within NRW and thereby acknowledging the influence of preconceived constructs of the attributes. The results demonstrated that not all attributes cluster effectively, with some showing little variance across neighborhoods, highlighting the need for careful attribute selection in urban area analysis due to the dynamic nature of urban development and infrastructure changes. Additionally, increasing dimensions complicated clustering without compromising quality, often resulting in indistinct clusters.

The chosen k-means algorithm necessitates predefining the number of clusters (k), with determining the ideal number being a significant challenge. Computation efficiency requires setting a specific range for k, as demonstrated in the Aachen case study. However, estimating this range becomes increasingly complex for larger or more diverse geographical areas, such as the whole state of NRW, which encompasses multiple cities and rural areas, the latter being outside the scope of this study. The utilization of the

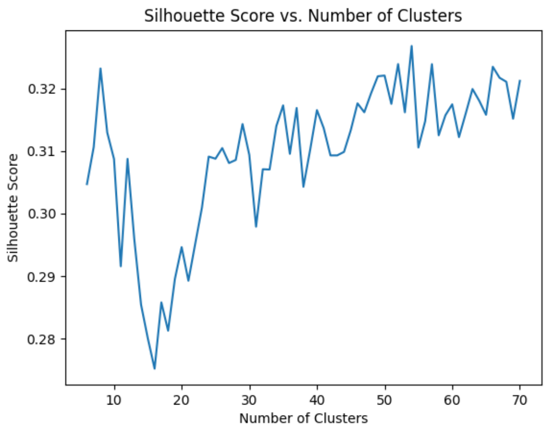


Figure 4: Silhouette score for k-means clustering of Socio-Infrastructural-Model for different k

Silhouette score to identify the optimal 'k' for k-means clustering was a critical component of our analysis. This evaluation method, grounded in mathematical principles, provides a quantifiable measure of cluster coherence and separation. However, it is essential to note the distinction between the mathematical assessment of cluster quality and the perceived social characteristics of neighborhoods. While the mathematical evaluation offers a systematic ap-

Table 3: Mean distribution of cluster attributes: Single-Family-Dwelling (SFD), Big-Multi-Family-House(BMFH), Terraced House (TH), Multi-Family-House (MFH), Non-Residental Buildings (Non.-Res), specific heating demand (demand)

Cluster	SFD	BMFH	HR	MFH	TH	Non-Res.	demand $kWh\,m^{-1}\,a^{-1}$
1 pink	490.2	106.17	0.5	80.66	314.37	169.79	156305.82
2 red	190.22	33.87	0.1	34.66	67.00	93.06	56945.96
3 orange	232.90	728.60	1.30	205.30	257.90	535.00	269755.08
4 yellow	411.0	2451.00	4.0	825.0	545.00	2370.00	956452.06
5 green	787.50	746.00	1.0	351.00	915.00	587.50	472010.96
6 blue	355.75	164.25	5.25	101.75	178.25	353.25	157519.04

proach to determining cluster quality, it may not capture all aspects relevant to interpreting these clusters effectively. Moreover, it was observed that the graph displayed comparably high Silhouette scores with significantly fewer clusters (Figure 4). This phenomenon suggests that a more compact cluster formation can provide a similarly coherent and well-separated structure as configurations with a larger number of clusters. This finding is particularly illuminating, as it implies that optimal clustering does not necessarily equate to maximizing the number of clusters but rather to identifying a configuration that balances cluster coherence with the complexity of the model. This discrepancy underscores the importance of not solely relying on quantitative metrics for interpreting cluster results. An understanding of the urban structure, historical developments, and other pertinent factors should also be incorporated. Such a holistic approach ensures that the interpretation's significance is not merely confined to numerical values but is enriched by a comprehensive context of neighborhood structures.

Conclusion and Future Work

This study yielded significant insights into the feasibility of neighborhood classification studies while highlighting opportunities for future research. It showcased an innovative approach to urban planning and energy management by integrating socio-infrastructural and energy data to identify reference districts within Aachen, NRW. The cluster analysis conducted for Aachen revealed specific structural and characteristic patterns within its neighborhoods. By employing clustering techniques on public datasets, the research illuminated the complex interplay between urban form, energy infrastructure, and social dynamics, contributing valuable insights into sustainable urban development and the energy transition. The methodology demonstrated the potential of using machine learning and GIS technologies to dissect and understand the multifaceted characteristics of urban areas. However, the study also highlighted several challenges, including the subjectivity in selecting clustering attributes, the difficulty in managing and integrating large and diverse datasets, and the computational demands of processing extensive urban data. Applying the same analysis to the entire state of NRW would likely yield different results due to its diverse composition of cities, rural areas, and villages, each with unique structures and attributes, requiring spatial understanding and various validation methods. Investigating various validation methods could also comprehensively assess cluster result quality. A systematic examination of the weighting of individual characteristics may enable a more nuanced, context-specific classification, acknowledging the varying impacts of different features on neighborhood formation. These challenges underscore the need for further refinement of the methodology and the exploration of more advanced machine learning techniques to enhance the precision and applicability of the findings. Additionally, integrating socio-economic or environmental indicators could better represent neighborhood characteristics and contribute to developing more comprehensive models. Including the performance and impacts of energy producers on the power and heating networks and building energy inputs in the study of energy neighborhoods could offer a more holistic view.

Acknowledgments

We gratefully acknowledge the financial support by BMWK, promotional reference 03EWR010B.

References

Acksel, D., Huenges, E., and Kastner, O. (2017). Wärmewende am Beispiel Quartier: Ein Beitrag zur Energiewende.

Alpagut, B., Lopez Romo, A., Hernández, P., Tabanoğlu, O., and Hermoso Martinez, N. (2021). A GIS-Based Multicriteria Assessment for Identification of Positive Energy Districts Boundary in Cities. Energies, 14(22):7517.

Arribas-Bel, D., Garcia-López, M.- ., and Viladecans-Marsal, E. (2021). Buildings and cities: Delineating urban areas with a machine learning algorithm. Journal of Urban Economics, 125:103217.

Bezirksregierung Köln (2024). ALKIS NW Grundrissdaten. Accessed: 2023-05-01.

BMWK-Bundesministerium für Wirtschaft und Klimaschutz (2022). Unsere Energiewende: sicher, sauber, bezahlbar. Accessed: 2024-01-29.

Boenigk, N., Cantos, E., Eyerich, L., Kradolfer, A., Müller, A., Müller, I., and Schwalbe, A. (2019). Komm:mag: Das Jahresmagazin zu erneuerbaren Energien in Kommunen.

Bundesinstitut für Bau-, Stadt- und Raumforschung (2012). Bestand und städtebauliche Bedeutung - neue Stadtquartiere.

Dettmar, J., Drebes, C., and Sieber, S., editors (2020). Energetische Stadtraumtypen. Fraunhofer IRB Verlag.

Feldmann, P. (2009). Die strategische Entwicklung neuer Stadtquartiere: unter besonderer Berücksichtigung innenstadtnaher oder innerstädtischer, brachgefallener Industrieareale, volume 53. Immobilien-Manager-Verl.

Gonzalez, D., Rueda-Plata, D., Acevedo, A. B., Duque, J. C., Ramos-Pollan, R., Betancourt, A., and Garcia, S. (2020). Automatic detection of building typology using deep learning methods on street level images. Building and Environment, 177:106805.

Jochem, W. C., Bird, T. J., and Tatem, A. J. (2018). Identifying residential neighbourhood types from settlement points in a machine learning approach. Computers, Environment and Urban Systems, 69:104–113.

Kaupp, A. (2022). Corporate urban responsibility–Kirche in der Stadtentwicklung. In CSR und Kirche: Die unternehmerische Verantwortung der Kirchen für die ökologisch-soziale Zukunftsgestaltung, pages 261–272. Springer.

Kelm, T., Schonlau, M., Pitz, N., and Klein, U. (2019). Semiautomatisches Verfahren zur Ableitung von Baubloecken. Selbstverlag des Vereins CORP - Competence Center of Urban and Regional Planning, Wien, Österreich. Meeting Name: REAL CORP.

Lanuv NRW (2024a). Dachflächen-Solarthermie: Potentialdaten Solarkataster NRW. Accessed: 2023-08-01.

Lanuv NRW (2024b). Krankenhäuser. Accessed: 2023-06-01.

Lanuv NRW (2024c). Raumwärmebedarfsmodell NRW. Accessed: 2023-07-01.

López-Moreno, H., Núñez-Peiró, M., Sánchez-Guevara, C., and Neila, J. (2022). On the identification of Homogeneous Urban Zones for the residential buildings' energy evaluation. Building and Environment, 207:108451.

Malottki, C., Koch, T., and Vaché, M. (2013). Anforderungen an energieeffiziente und klimaneutrale quartiere (eq). Werkstatt: Praxis. Bundesministerium für Verkehr Bau und Stadtentwicklung. Bonn, Institut für Wohnen und Umwelt.

März, S. (2016). Identifikation kleinräumiger Hotspots der Energiearmut : Gis-gestützte Analysen zur Vulnerabilität von Quartiersbewohnern am Beispiel Oberhausen. In Schmitt, H. C., editor, Raummuster : Struktur, Dynamik, Planung, pages 101 – 119. Klartext-Verl., Essen.

Mehnert, T. and Kremer-Preiß, U. (2014). Ist-Analysen im Quartier. Handreichung im Rahmen des Förderbausteins 3.1. 1 „Projekte mit Ansatz zur Quartiersentwicklung "des Deutschen Hilfswerks.

Ministerium für Arbeit Gesundheit und Soziales des Landes Nordrhein-Westfalen (2023). Krankenhausplan Nordrhein-Westfalen 2022. die Strukturen müssen für die Menschen da sein, nicht die Menschen für die Strukturen! Accessed: 2023-08-21.

OpenStreetMap Contributors (2024). Open Street Map Daten Nordrhein-Westfalen. Accessed: 2023-07-01.

Perez, J., Fusco, G., Araldi, A., and Fuse, T. (2020). Identifying building typologies and their spatial patterns in the metropolitan areas of Marseille and Osaka. Asia-Pacific Journal of Regional Science, 4(1):193–217.

Photis, Y. N. (2012). Redefinition of the greek electoral districts through the applicaton of a region-building algorithm. European Journal of Geography, Volume 3(Issue 2):72–83.

Quénéhervé, G., Tischler, J., and Hochschild, V. (2018). Energiewende im Quartier–ein Ansatz im Reallabor. Bausteine der Energiewende, pages 385–405.

Reicher, C. (2013). Das (Stadt) Quartier, vom Umgang mit dem gebauten Raum und seinen dynamischen Parametern. Stadtquartiere. Sozialwissenschaftliche, ökonomische und städtebaulich-architektonische Perspektiven. Klartext, Essen, pages 197–209.

Reicher, C., Schmidt, A., and Hangebruch, N. (2020). Energieeffizienz und Quartier: Herausforderung Energieeffizienz im Quartier. Handbuch Energieeffizienz im Quartier: Clever versorgen, umbauen, aktivieren, pages 1–16.

Schnur, O. (2008). Quartiersforschung im überblick: Konzepte, Definitionen und aktuelle Perspektiven. Quartiersforschung: Zwischen Theorie und Praxis, pages 19–51.

Stadt Aachen (2020). Sozialentwicklungsplan Aachen. Accessed: 2023-08-24.

Töbermann, I. J.-C. and Yu, I. Y. J. (2021). Energiewendepotentiale von Quartieren und Quartiersgrössen. Forschungsberichte, page 69.

METADATA EXTRACTION OF RFIs USING NATURAL LANGUAGE PROCESSING AND MACHINE LEARNING ALGORITHMS

Ceyhun Ozogul[1], Esin Ergen[1]
[1]Istanbul Technical University, Istanbul, TURKEY

Abstract

Accurately assigning metadata of RFIs plays an important role in the analysis and management of RFI documents. However, these metadata are manually entered in the RFI management system, which results in loss of time and incorrect entries. This study aims to demonstrate that metadata of RFI documents can be extracted and assigned automatically using natural language processing and machine learning algorithms. To achieve this aim, the performance of Naïve Bayes and K-Nearest Neighbor algorithms are evaluated and compared. The results show that machine learning models perform well in automatically extracting the metadata of RFIs and, the performance of machine learning models for each label varies. The findings of this study can be used to develop an artificial intelligence based RFI management system by integrating natural language processing and machine learning models into the system.

Introduction

In construction projects, documents such as drawings and specifications may not address all aspects of the structures to be built or may contain deficiencies, uncertainties, and overlaps. These issues need to be clarified through Request For Information (RFI) documents (Shim et al., 2016). Hanna et al., (2012) defines RFI as " *a formal written procedure initiated by the contractor seeking additional information or clarification for issues related to design, construction , and other documents*". RFIs are a communication tool between the design team and the construction team (Bhat, 2017). Even though it holds significance, RFI, as a communication channel, is frequently perceived negatively within the project (Aibinu et al., 2019). This is due to the effort involved in initiating an RFI, as well as the process of reviewing and generating responses (Afzal et al., 2023). The response time for RFI can become critical and have a negative impact on progress in the field (Kelly and Llozor, 2020). Moreover, a delay in responding to the RFI, or the absence of a response altogether, can lead to a sense of frustration and mistrust among project team members (Philips-Ryder et al., 2012; Afzal et al., 2023).

Utilizing common data environments, such as Aconex and Procure, assists in overcoming the challenges associated with Request for Information (RFI) management. These platforms facilitate online communication and provide a means to track information effectively (Das, Tao and Cheng 2020; Afzal et al., 2023). However, metadata, which is a classification output, is typically entered manually into common data environments and the users are reluctant to fill out the related metadata, since this process requires additional time and effort. This results in inaccurate or incomplete metadata, which is problematic because metadata holds significant importance when analysing the unstructured RFI text data to obtain valuable insights and lessons learned to manage future projects in a more effective way. The erroneous metadata entry directly impacts the data quality, consequently influencing the outcomes of analysis studies.

This paper proposes to use Natural Language Processing (NLP) and machine learning models to automatically extract discipline code, which is one of the RFI metadata. To achieve this, NLP processes were applied to the RFI documents and Naïve Bayes (NB), K-Nearest Neighbor (KNN) algorithms were used to classify the RFI documents according to their disciplines. This approach enables extraction of the metadata of RFIs in an automated way. It reduces the time required for filling out the metadata of RFIs and results in effective management and analysis of RFIs by reducing erroneous or incomplete metadata.

Literature Review

Automatic extraction of metadata from documents is associated with the automatic classification of documents. Documents can be classified according to their metadata by using NLP and machine learning models. In the early 2000s, studies were conducted on the automatic extraction of metadata from texts in different domains, such as marketing, information-document management, educational sciences (Paik et al., 2001; Han et al., 2003; Yılmazel and Finneran, 2004). In these studies, rule-based methods were combined with NLP. Later, with the increase in the processing power of computers, machine learning models have been integrated with NLP instead of rule-based approach to automatically extract metadata.

Valdez et al. (2016) developed a natural language processing-based ontology for extracting metadata from texts in the biomedical domain. In addition, a natural language processing-based ProvCaRe-NLP tool was developed to perform clinical text analysis and information extraction. Some other studies used NLP and ML on metadata stored in a computerized maintenance management system (CMMS). Zhang et al. (2020) employed the information contained in failure notifications to forecast failure codes and assess the precision of the labels. Texts are converted into word counts after NLP preprocessing. The SMOTE (Synthetic Minority Oversampling Technique) algorithm is utilized to balance the classes. In the final stage, one of the eight supervised machine learning models that are used in the classification problem, such as Support Vector Machine (SVM), NB, or Logistic Regression (LR), was applied.

Arif-Uz-Zaman et al. (2017) sought to enrich the metadata in a Computerized Maintenance Management System (CMMS) by classifying downtime notifications as either resulting from a failure or indicating no failure. In this study, the vectors obtained after applying NLP operations to the text data are used as input for SVM and NB algorithms. The SVM algorithm is noted to achieve the highest performance.

Tanguy et al. (2020) applied the SVM algorithm to categorize fault notifications in their study. The study demonstrated that incorporating all 3-character substrings along with word stems resulted in a slight enhancement compared to using only words. This improvement is likely attributed to the inclusion of abbreviations, compound terms, and alternative writing forms.

Deloose et al. (2023) have made a recent contribution to the study of CMMS metadata, in which natural language processing techniques and artificial models are utilized to forecast and rectify CMMS metadata. The performance of shallow machine learning models and deep machine learning models using multiple labels were compared. Random forest algorithm (RF), which is a shallow machine learning model, and Recurrent Neural Networks (RNN), which are deep learning methods, gave better results than other algorithms. It was also identified that the RNN algorithm performed slightly better than the RF algorithm.

In the field of construction management, studies on automatic classification of documents focuses on the accident analysis/occupational safety management, contract management, building information modeling (BIM), and facility management. In the studies related to the accident analysis and occupational safety management, accident reports were analyzed and the performances of machine learning models in the automatic classification of accident reports were compared. Tixier et al. (2016) used Stochastic Gradient Tree Boosting (SGTB) and Random Forest (RF) algorithms. The SGTB algorithm performed better than the RF algorithm. In Goh and Ubeynarayana (2017), the performances of six different machine learning models were compared, and the SVM algorithm showed the best performance.

In the area of contract management, studies involving automatic classification were conducted using NLP and ML (Yilmaz and Dikbas, 2013; Yilmaz, 2013; Candas, 2022; Eken, 2022). In Yilmaz and Dikbas (2013), dispute decision documents were classified using four machine learning algorithms. The algorithm with the highest accuracy was decision tree (DT). Eken (2022) conducted classification work to automatically review construction contracts. Five different machine learning algorithms are combined with different vectorization techniques. Models that perform well were selected and combined with ensemble learning. Candas (2022) conducted research on the multi-obejctive semantic analysis of construction contracts. NLP and machine learning algorithms were applied for the purposes of the automatic classification of contract clauses according to departmental relevance and accurately predicting the presence of ambiguity in contract clauses. The SVM and DT algorithms demonstrated the best performance.

The classification studies were also carried out for analysis of RFIs, but these classification studies were mostly performed manually (Tilley et al.,1997; Morales et al., 2022). In their studies on the analysis of RFIs, the researchers manually classified RFI documents under three main headings according to type, cause, discipline.

The manual classification of texts is acknowledged as time-consuming and error-prone. The process of manually categorizing thousands of RFI documents inevitably prolongs the time required to finalize the analysis. Additionally, misclassifications significantly impact the accuracy and reliability of the analysis outcomes concerning RFI documents. Consequently, the adoption of an NLP-based approach to analyze RFI documents offers potential benefits, such as shorter analysis times and more precise analysis results.

Methodology

Python programming language, Natural Language Toolkit and Scikit-learn library were used for natural language processing and machine learning methods applied to RFI documents. The original dataset is the RFI documents extracted from the common data environment of a project. In order to improve the quality of the original dataset, RFI documents with a lot of noise and very short descriptions were excluded. The metadata to be automatically retrieved from the RFIs was the selected as discipline (i.e., architectural, electrical, mechanical, and structural). The dataset comprises 25 RFI documents for each discipline. In total, 100 RFI text documents were analyzed. The RFI documents retrieved in .pdf format and each RFI document was turned into .txt format.

Implementation workflow for automatic extraction of RFI discipline metadata is shown in Figure 1. It has three main steps: Label assignment, NLP and machine learning. In the first step, each RFI document was labeled in terms of discipline. In the second step, NLP techniques were applied. NLP techniques were performed using the Natural Language Toolkit (NLTK) in Python. These blocks, also known as modules in NLTK, aid in tasks such as tokenization, stemming, lemmatization, and data classification. First, RFI documents were tokenized. Tokenization involves breaking a sequence of text into smaller units, referred to as tokens. These tokens can vary in size, ranging from individual characters to entire words

(Manning et. al., 2009). Second, punctuation marks and stop words were removed from each RFI document. Punctuation removal is often performed to diminish the dimensionality of the data and to eliminate elements that may not carry significant semantic meaning for certain NLP tasks. Thirdly, stop word removal was performed. It is a common preprocessing step in NLP where frequently occurring words, known as stop words, are removed from a text corpus. These words are often common (i.e., 'the', 'a', and 'in') and do not contribute significantly to the understanding of the content. Finally, RFI text data was converted into vectors via vectorization process. Vectorization techniques transform text data into a form that a computer can make sense of. In this study, Bag of Words (BOW) vectorization technique, one of the mostly used vectorization techniques (Qader et al., 2019), was used. BoW was utilized to represent text based on the number of word occurences, using a fixed-length vector created from a vocabulary. The BoW technique was implemented using the Scikit Learn Python library.

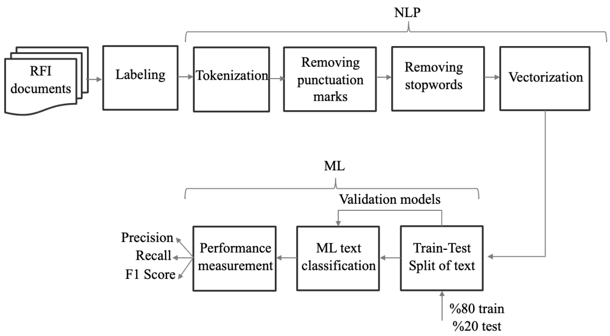


Figure 1: Implementation workflow for automatic extraction of RFI discipline metadata

In the third step, machine learning algorithms were applied. The automatic extraction of metadata from RFI documents is a classification problem. Therefore, supervised machine learning models need to be trained, and two classification algorithms, Naïve Bayes and K-Nearest Neighbor, were selected. Bayes' rule is used to calculate the probabilities of the classes. The NB algorithm is based on Bayesian Decision Theory (Alpaydin, 2014) and the NB model is the most widely used Bayesian network model in machine learning. The term "naïve" is used because the NB algorithm is based on the assumption that attributes are conditionally independent of each other (Russell and Norvig, 2010). It is extensively used in text categorization tasks, including document classification and spam e-mail detection.

The K-Nearest Neighbor classifier categorizes the input into the class with the highest frequency among the k neighbors of the input. Each neighbor holds an equivalent vote, and the class with the highest count of votes among the k neighbors is chosen. Ties are resolved arbitrarily or through a weighted vote. Typically, k is chosen as an odd number to reduce ties, especially when confusion occurs between two adjacent classes (Alpaydin, 2014).

First, the data was split into 80/20, where the first percentage represented the proportion of data allocated to the training set, and the second percentage represented the testing set. Validation was performed before testing the model within training set. The selected algorithms are trained to automatically extract RFI discipline metadata. Then the KNN and NB algorithms were trained and tested by using the output of the vectorization process.

Traditionally, two metrics are commonly used for measuring the performance of machine learning models: Precision and Recall. Precision measures the proportion of truly relevant documents in the result test. Recall measures the ratio of all relevant documents in the corpus that are included in the result set. It is possible to achieve a balance between precision and recall by adjusting the size of the returned result set. The F1-Score is the harmonic mean of sensitivity and recall (Russell and Norvig, 2010).

Precision, recall and F1Score values are calculated by substituting the parameters shown in Table 1 into the equations (1) to (3).

$$Precision = \frac{TP}{(TP+FP)} \quad (1)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (2)$$

$$F1Score = \frac{2\ x\ Precision\ x\ Recall}{Precision+Recall} \quad (3)$$

Table 1: True and false positives and negatives

	Relevant	Irrelevant
Retrieved	True Positives (TP)	False Positives (FP)
Not retrieved	False Negatives (FN)	True Negatives (TN)

Findings

In the BoW vectorization process, word counts were performed for each discipline and they are used as input when applying the BoW technique. Twenty most common words related to the architectural discipline were determined and used in the vectorization process. The first five most repeated words out of the 20 most common words are given in Figure 2.

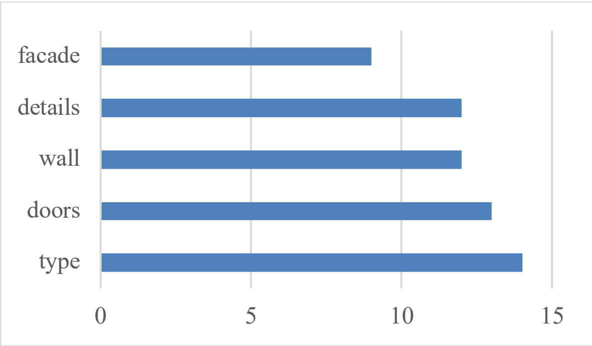


Figure 2: First five most common words in RFI documents related to the discipline of architecture

Similar to the architecture discipline, twenty most common words related to structural, mechanical and electrical disciplines were identified. The first five most repeated words out of the 20 most common are reported in Figure 3 to 5 for structural, electrical, and mechanical disciplines respectively.

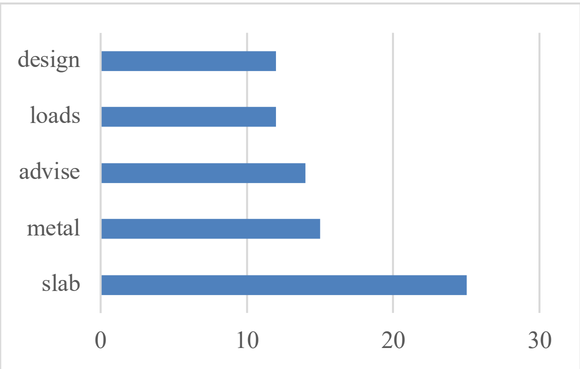


Figure 3: First five most common words in RFI documents related to structural discipline

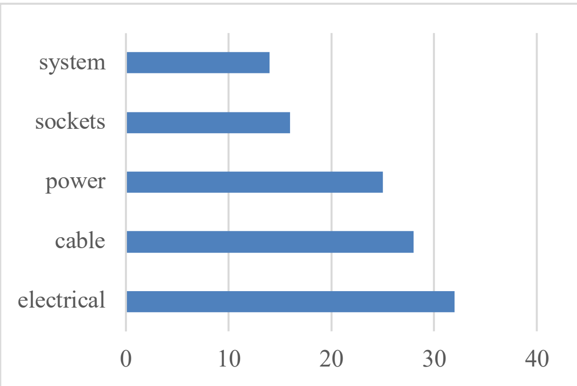


Figure 4: First five most common words in RFI documents related to electrical discipline

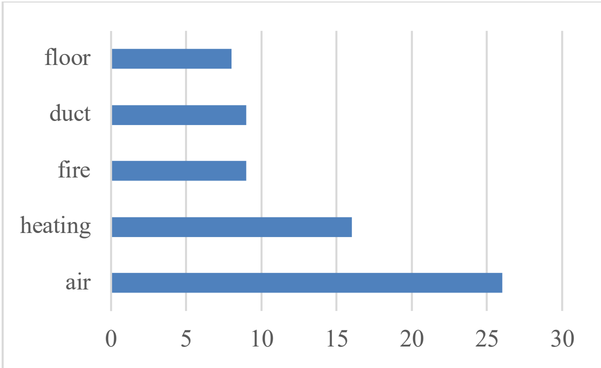


Figure 5: Some of the most common words in RFI documents related to mechanical discipline

After the vectorization process, the NB and KNN models were trained. %80 percent of the data was used for training and %20 percent for testing. The content of the eighty percent training set was changed ten times and validation work was carried before the test phase. At different iterations, the performance fluctuation of the trained algorithm was checked. The performance result of the trained NB algorithm is shown in Figure 6. After ten iterations, the Average Precision, Recall, and F1 Score values are approximately 85%, 82% and 82% respectively. The NB algorithm proves to be highly successful in automatically extracting the discipline metadata of RFIs.

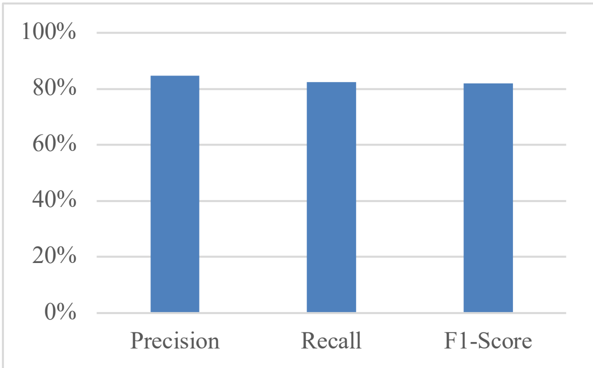


Figure 6: Performance results of the NB algorithm

The tuning of the hyperparameter k corresponding to the number of nearest neighbors is important for enhancing the performance of the KNN algorithm.

The performance results for different k numbers are determined and compared as shown in Figure 7. The results illustrated that KNN demonstrated the best performance when k was set to 5. Precision, Recall and F1 Score values are approximately %75, %60 and %62, respectively.

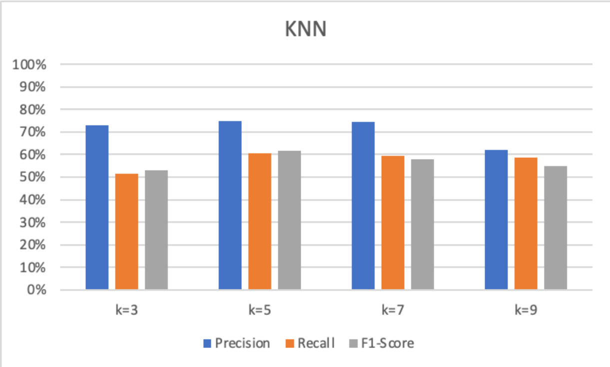


Figure 7: Performance results of KNN algorithm for different k values

Comparative performance results of NB and KNN algorithms are given in Figure 8. NB algorithm showed much better results than KNN algorithm according to precision, recall and F1-Score criteria. Overall, the findings have illustrated that one of the metadata of RFIs, which is discipline, can be successfully extracted automatically.

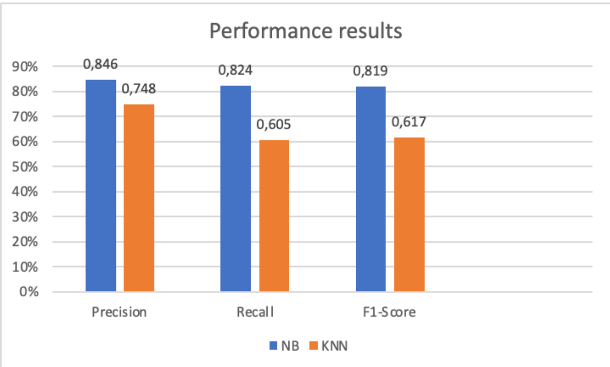


Figure 8: Comparative performance results of the algorithms

Conclusion

In this study, an NLP and machine learning-based model is developed to automatically extract the discipline, which is one of the metadata of RFIs. The performances of the two machine learning models, NB and KNN algorithms, are compared. The average precision of the NB algorithm is around 85%, while for the KNN algorithm, it is approximately 75%. In the case of KNN, it was demonstrated that the KNN algorithm gave better results for k=5.

The model proposed in this study has the potential to prevent time losses during the manual entry of metadata for RFIs in the common data environments, and it can contribute to a reduced number of incorrect entries.

An artificial intelligence based RFI management system can be developed by integrating NLP and machine learning models into the system. As a result, RFI management systems will be more effectively used. The utilization of emerging technologies and data-driven analytics can be harnessed to augment the RFI analysis process, leading to improvements in overall project efficiency (Afzal et al., 2023).

In future studies, we plan to conduct a more comprehensive analysis by expanding the dataset and diversifying the algorithms. In addition to training shallow machine learning models, it may be beneficial to explore the development of deep learning models.

References

Afzal, M., Wong, J.K.W, Fini, A.A.F. (2023) Unlocking insights: analysing construction issues in request for information (RFI) documents with text mining and visualisation. IEEE 19th International Conference on Automation Science and Engineering (CASE).

Aibinu, A.A., Carter, S., Francis, V., Vaz-Serra, P. (2019) Request for information frequency and their turnaround time in construction projects. Built Environment Project and Asset Management, Vol. 10, No.1, pp. 1-15.

Alpaydin, E. (2014) Introduction to Machine Learning, Third Edition. London, England, The MIT Press.

Bhat, A.C. (2015) Data visualization of requests for information to support construction decision-making (Master Thesis). The University of British Columbia.

Candas, A.B. (2022) Multipurpose semantic analysis of construction text documentation. A Thesis Submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University. Ankara.

Das, M., Tao, X., Cheng, J.C.P. (2020) A secure and distributed construction document management system using blockchain. Proceedings of the 18th International Conference on Computing in Civil and Building Engineering (ICCCBE).

Deloose, A., Gysels, G., Baets, B.D., Verwaeren, J. (2023) Combining natural language processing and multidimensional classifiers to predict and correct CMMS metadata. Computers in Industry, 145, 103830.

Eken, G. (2022) Using natural language processing for automated construction contract review during risk assessment at the bidding stage. A Thesis Submitted to the Graduate School of Natural and Applied Sciences of Middle East Technical University. Ankara.

Goh, Y.M., Ubeynarayana, C.U. (2017) Construction accident narrative classification: an evaluation of text mining techniques. Accident Analysis and Prevention, 108, 122-130.

Han, H., Giles, C.L., Manavoğlu, E., Zha, H. (2003) Automatic document metadata extraction using support vector machines. Proceedings of the Joint Conference on Digital Libraries.

Hanna, A. S., Tadt, E.J., Whited, G.C. (2012) Request for information: benchmarks and metrics for major highway projects. Journal of Construction Engineering and Management, 138(2), 1347-1352.

Kelly, D. and Llozor, D.B. (2020) Performance outcome assessment of the integrated project delivery (IPD) method for commercial construction projects in USA. International Journal of Construction Management, 1-9.

Manning, C.D., Raghavan, P., Schütze, H. (2009) An information to information retrieval. Cambridge University Press, Cambridge, England .

Morales, F., Herrera, R.F., Rivera, F.M.-L., Atencio, E., Nunez, M. (2022) Potential application of BIM in RFI in building projects. Buildings, 12, 145.

Paik, W., Yılmazel, S., Brown, E., Poulin, M., Dubon, S., Amice, C. (2001) Applying natural language processing (NLP) based metadata extraction to automatically acquire user preferences. K-CAP'01, October 22-23, Victoria, British Columbia, Canada.

Qader, W.A., Ameen, M.M., Ahmed, B.I. (2019) An overview of bag of words; importance, implementation, applications, and challenges. International Engineering Conference.

Philips-Ryder, M., Zuo, J., Jin, X.H. (2012) Evaluating document quality in construction projects – subcontractors' perspective. International Journal of Construction Management 13 (3) 77– 806 94.

Russell, S. and Norvig, P. (2010) Artificial Intelligence: A Modern Approach, Third Edition. Upper Saddle River, New Jersey 07458, Pearson.

Shim, E., Carter, B., Kim, S. (2016) Request for information (RFI) management: a Case Study. 52nd ASC Annual International Conference Proceedings.

Tanguy, L., Tulechki, N., Urieli, A., Hermann, E., Raynal, C. (2016) Natural language processing for aviation safety reports: from classification to interactive analysis. Comput. Ind. 78, 80–95.

Tilley, P.A.; Wyatt, A.; Mohamed, S. Indicators of design and documentation deficiency. In Proceedings of the IGLC-5, Fifth Annual Conference of the International Group for Lean Construction, Gold Coast, Australia, 16–17 July 1997; pp. 137–148.

Tixier, A.J.P., Hallowell, M.R., Rajagopalan, B., Bowman, D. (2016) Automated content analysis for construction safety: a natural language processing system to extract precursors and outcomes from unstructured injury reports. Automation in Construction, 62, 45-56.

Tixier, A.J.P., Hallowell, M.R., Rajagopalan, B., Bowman, D. (2016) Application of machine learning to construction injury prediction. Automation in Construction, 69, 102- 114.

Uz-Zaman, K.A., Cholette, M.E., Ma, L., Karim, A. (2017) Extracting failure time data from industrial maintenance records using text mining. Adv. Eng. Inform. 33, 388–396.

Valdez, J., Rueschman, M., Kim, M., Redline, S., Sahoo, S.S. (2016) An ontology- enabled natural language processing pipeline for provenance metadata extraction from biomedical text. OTM Conferences, LNCS 10033, 699-708.

Yılmaz, İ.C., Dikbaş, A. (2013) Türk kamu inşaat projelerinde yaşanan uyuşmazlıklara yönelik bir veri madenciliğI yaklaşımı. Online Academic Journal of Information Technology, Cilt: 4, Sayı:13, doi: 10. 5824/1309-1581.2013.4.005.x.

Yılmaz, İ.C. (2013) İnşaat sözleşmelerinde hak talebi yönetimi: kamu projeleri için öneri model (Doktora Tezi). İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü.

Yılmazel, O., Finneran, C.M., Liddy, E.D. (2004) MetaExtract: an NLP system to automatically assign metadata. Proceedings of the 2004 Joint ACM/IEEE Conference on Digital Libraries (JCDL'04).

Zhang, T., Bhatia, A., Pandya, D., Sahinidis, N.V., Cao, Y., Flores-Cerrillo, J. (2020) Industrial text analytics for reliability with derivative-free optimization. Comput. Chem. Eng. 135, 106763.

DEVELOPMENT OF NATURAL LANGUAGE APPLICATION FOR CONSTRUCTION SITE PROCESSES SAFETY SUPPORT

Alessandra Corneli1, Francesco Spegni1, Berardo Naticchia1 and Leonardo Messi1
1DICEA, Polytechnic University of Marche, Ancona, Italy

Abstract

Efficiency in safety management at construction sites is an extremely present issue despite the many technological advances that could represent valuable means of support. This is because safety is a complex, multifactorial field in which the human factor plays a central role. It is precisely for the latter reason that finding systems that integrate technology into processes in as natural a way as possible is a challenge to be met. In this context, this research aims to develop a Digital Twin that takes use of Natural Language Processing for supporting safety during construction site operations through communication improvement.

Introduction

Construction is a complex process that requires decisions at every stage. Decision makers are inevitably influenced by past experience, situational knowledge, and other human factors. This often results in operational inefficiencies or safety and environmental problems (Saini et al., 2022). Safety management on construction sites has always been a critical issue in the construction sector since, as is well known, among all industries this is the one with the highest number of fatal accidents even today (Mihic et al., 2019). In this context, the support that could be given by innovative technologies still fails to be incisive (Yap et al. 2022). What makes it extremely difficult to go to work on incorporating new technologies into safety management worksites is that it is a complex issue. The explanation for this lies in multiple factors (Giri, 2020):

- the construction site is a workplace that evolves (rapidly) over time.
- construction procedures are not as precisely defined as operations within factories.
- the centrality in many cases of workers in implementing safety procedures and thus the human factor (Haupts et al., 2019).
- health and safety risk assessments, especially in the drafting of safety plans, which often use standard prescriptions instead of identifying, on a site-by-site basis, the main critical issues.

All of this contributes to the fact that what happens on a construction site, including from a safety management perspective, is difficult to predict in advance and therefore needs appropriate tools. This complexity has also led over time to a misguided approach of focusing on certain types of hazards while neglecting an integrated approach (Mihic et al., 2019).

This has resulted in the paradigm recently promoted for the management of domains characterized by high complexity, namely the concept of digital twin (DT) (Saini et al., 2022). Due to the characteristics of real-time interaction and self-evolution, DT is considered the best technology that can contain all the required information (Zhao et al., 2019). One aspect that is often overlooked in DT development, however, is the circularity of information that on the one hand is collected from the real environment but on the other hand should return to it so that the agents involved can learn about the information and possibly self-correct. This approach from a safety perspective is called safety II. In order, however, to communicate the information needed for safety management on the job site, one aspect to consider is that this information be easily usable and integrated into the processes. The use of natural language represents an optimal method of communication on the construction site for the purpose of human-machine interaction that is as natural as possible. The innovation proposed in this research concerns the integration of several innovative technologies (localization using UWB and RTK sensors, natural language, BIM approach for representing the built environment) in order to develop an application to support safety in construction processes. One of the innovative aspects is the use of NLP in spoken form, as opposed to the more widely used automatic text analysis, is the communication improvement on the construction site. This together with the early detection of dangerous situations through the development of DTs represents a method to apply a safety II approach with the aim of guiding workers toward recommended attitudes rather than pursuing an exclusively prescriptive approach.

Literature review

The construction industry is recognized for its dynamic and complex nature. Safety planning in construction environments is therefore challenging. Moreover, due to the lack of standardization of production processes it is very complicate to incorporate new methods and technologies into workflows without disrupting the entire industry. Among new technologies, BIM is the most cited approach in the literature. This is usually used in conjunction with other innovative technologies or the basis for others (e.g., AR) (Khudhair et al., 2021). This is also because if BIM means IFC this represents only a part of all the knowledge underlying the processes of the construction industry, namely the objects of which a building is composed. Missing from this representation are a number of crucial parts including the representation of processes, workings, economic aspects and temporal aspects. Furthermore, in the construction industry textual information (including BIM) is the dominant data type that exists in every stage of construction management, with over 80% of it being unstructured (Shamshiri et al., 2024; Wu et al., 2022). This information format is not very convenient to use on the construction site. It is precisely this high percentage of written text in construction management that has caused much research over the years to introduce natural language as a tool for handling this amount of data. Some approaches to Natural Language Processing (NLP) involve its use for interpreting documents. In (Fang et al., 2020) they developed a novel approach for classifying text contained with safety reports specifically in the case of near misses using NLP specifically to gather information in a smart way. Similarly (Baker et al., 2020) test two state-of-the-art deep learning architectures for NLP, providing a method for identifying (after training) the textual patterns. Again, the application of NLP is focused on automation in document interpretation. (Martinez et al., 2020)'s retrieval systems have been proposed for retrieving accident cases and supporting health and safety plan preparation. (Tian et al., 2023) have developed an information retrieval system that can answer questions about the best procedures to follow regarding safety during operations. Shamshiri et al., 2024 highlight how the use of NLP applications interests' safety management much more than the other domains of interest. Despite this, most text classification and information extraction applications have been widely used for the specific task of construction site accident and injury reports.

However, one aspect lacking in safety management is communication, and so natural language voice interaction (rather than written) could be a real breakthrough in workplaces where the use of paper and digital texts is not very smooth. There are two different types of interactions that can be imagined on a construction site between human agents and a Vocal Virtual Assistant (VVA). The first one requires the user to begin an interaction (passive VVA). The second one requires the system to monitor the human and artificial agents on the construction site, and proactively sends voice commands or requests to specific human actors on the field (pro-active safety VVA).

In the present work we focus on pro-active safety VVAs because assuming that workers would actively begin an interaction with a safety VVA would be equivalent to assuming that they continuously have a high degree of awareness level for safety hazards on their working environment. Unfortunately, field experience and the vast literature on construction site risk management and hazards teach us that repetitive and tiring tasks such as those characterizing the construction industry induce the workers to lower their safety hazards awareness levels. This motivates the development of a methodology for implementing an efficient pro-active safety VVA that would either provide useful safety hints to the workers on site or ask questions that aim to raise the level of attention of the worker potentially at risk.

Methodology

In Fig. 1 a methodology is introduced for driving the integration of the required technologies (see Fig. 1). The methodology spans two of the main stages of a construction project, viz. system set up and construction phase. During system set up, the first step concerns the digital representation of agents that are relevant for safety reasons (workers, vehicles, equipment, ...). Next, a connection between the real environment and the digital environment is established and maintained along time installing and configuring a network of sensors capable of localizing the described agents in the construction site (the technology for the sensor network is chosen based on the specific construction site, and especially it depends on whether it is outdoor, indoor, or mixed). In the Natural Language Processing step, the NLP microservice of the platform is fed with a set of sentences that should be recognized or reproduced by the platform itself in order to interact with the workers when it is needed. The construction phase is divided in two main steps: construction site monitoring and interaction. The former observes the positions of the agents in the construction site and detects whether some of the rules should be activated. The latter takes care of the voice interaction with the worker following the specific rules that are activated during the monitoring step, if any.

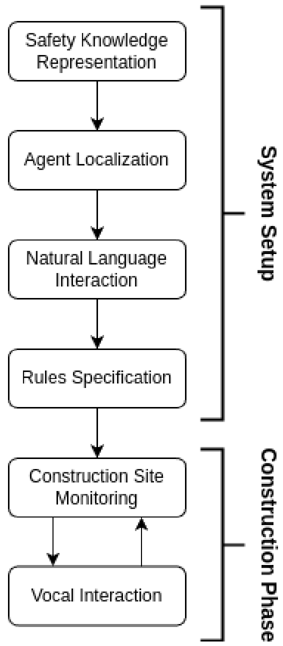


Figure 1 - Methodology

Safety knowledge digital representation

Construction site safety management is a complex issue because it requires the knowledge and integration of different information. Indeed, aspects that are part of safety concern workers, equipment, operations, hazards and spaces to name a few. Multiple efforts to represent aspects concerning safety can be found in the literature. From methods to classify hazards (Mihic et al., 2018; Fang et al., 2020; Johansen et al., 2023) to integrations with existing built classification such as IFC (Farghaly et al., 2022) contexts (Doukari et al., 2024; Gao et al., 2022) and rules (Shen et al., 2022).What is clear from these studies, however, is that to create an ontology that can collect everything that is needed for construction site safety monitoring is almost impossible.

This is because the domains to which the various parts to be taken into account refer are very different, and in the perspective of digitization of construction these differences become even more evident. In fact, if one thinks of BIM as a building model this involves a digital transposition of information based on objects, not on processes and workings that for example are central to making reasoning in the field of safety.
For this reason, the method chosen to deal with knowledge was to integrate different representations. The ontology that was taken as the basis is the one presented by (Zhang et al., 2015). Since the primary focus in this work is on monitoring resources (workers and machinery) in the chosen ontology we focused on the structure of resources not using the other parts for the time being. A handbook from the National Institute for Insurance against Accidents at Work (INAIL, 2019) was referenced to identify which components within the classification made sense to consider. This manual was compiled following the 2016 earthquake that struck central Italy. In the aftermath of that tragic event a very large number of construction sites were opened at the same time and in areas (urban centers, often including historic ones) that provided even more risks and aspects to be taken into account. In light of the great turmoil for reconstruction that is still unfinished today these guidelines identify the necessary human resources, machinery and equipment for each job. These items were used to fill within the system the categories identified by the chosen ontology.
As for the spaces that refer to the different construction site areas, we chose to model them with a BIM authoring software and then export information in IFC. The construction site layout is a mandatory document in Italy forming part of the safety documents. In this first implementation of the system, a single construction site layout was modelled with BIM and inserted into the DT platform to be used as a basis of the spatial reasoner. The site spaces planned for machine allocation, maneuvering, material storage, and different work areas were modeled in BIM as IFCSpace to identify their geometry and dimensions in digital space. In subsequent developments of the system, it could be envisaged to have more than one BIM model of the construction site with identified spaces and their changes in the duration of the construction.

Agent localization

Agents' localization on the construction site for the purpose of which the installation of a sensor network is planned. The real-time localization system has the task of attributing spatial coordinates to any entity, equipment or person, moving around the construction site.
The first requirement of the network for agent localization in order to be effectively implemented is high-precision localization capabilities in both open and closed environments for work taking place outside buildings and inside. The second one is real-time data sending to a platform for DT management.
The technologies selected for tracking are Ultra Wide Band tags for indoor and GPS-RTK for outdoor (Corneli et al., 2023). The two technologies can also be integrated into one tag which would allow both technologies to be exploited simultaneously in the case of mixed indoor/outdoor environments. As for UWB technology this involves the use of battery-powered mobile tags (to be provided to workers and attached to machinery) and requires the placement of anchors and gateways, two powered components. For this reason, UWB technology is intended for use in indoor environments where wiring the system is not an added risk to the worksite compared to outdoor environments. The accuracy for UWB localization is 10 cm.
On the other hand, for outdoor use, the technology used is GPS-RTK, which provides an accuracy of 1,5 cm. This can be implemented through an antenna inside the tags, which are provided to personnel and hooked onto the machinery. In both cases, localization information is sent by the sensor network to the Digital Twin Platform by means of intermediary service implementing a publish-subscribe information exchange pattern. This is often the case of services such as MQTT or SignalR.

Natural Language Interaction

The vocal interaction is a natural means of communication among agents in the construction site. Technological advancements in recent years make feasible to apply vocal interaction using natural language even on miniaturized and wearable devices. Cloud computing and high speed internet allows to displace the task of "reading" a given sentence or recognizing a vocal interaction on remote computers that are likely to have higher computing capabilities than those on the construction site. Neural networks allow to automatically transcribe the voice input onto its textual representation with high accuracy.
In our methodology, a pro-active safety VVA contains a collection of sentences that can be reproduced to specific group of people on the field, under suitable conditions. Each sentence is a template, mixing raw text with expressions that are computed if and when the sentence is reproduced. For instance, take the following sentence:

{{ A.name }}, you are in danger: equipment {{ B.name }} is too close!

In it {{ A.name }} and {{ B.name }} denote expressions that should be evaluated in order to deduce, respectively, the name of the worker at risk and the name of the equipment that represents a danger for the worker. A sentence with expressions in double curly braces is called an open-sentence, while sentences that do not contain expressions in curly braces are called closed-sentences. The key difference between the two kinds is that the latter can easily be converted to an audio file passing it as input of one of the several available text-to-speech AI driven cloud services, while the former need a pre-processing stage responsible for replacing the expressions with meaningful pieces of text. How this pre-processing phase happens is explained in the last stage of the methodology. The fact that sentences contain parameters allow the system to be more explicit and address specific workers calling them by name or referring to their close context.

This contributes in increasing the level of awareness in workers that perhaps do not even notice that their context is changing around them in a way that is becoming too risky to continue their task.

Parameters in template sentences are associated with one among three types of entities from the selected safety ontology: Human Resource, Equipment, or Space.

In general, the Safety Manager can devise sentences that aim to:

- raise the level of awareness in workers that are closer to dangerous or risky areas, equipment, or materials
- ask the worker for confirmation that he/she is wearing the required personal protective equipment (PPEs) or had the proper training;
- ask the worker to stay out of a delimited area.

It is evident that such goals can only be achieved by a pro-active safety VVAs.

Rules Specification

In order for the VVA to exhibit a pro-active behavior, the Digital Twin Platform must have the tools to express how to react when agent positions change, possibly denoting a risky situation on the field. To this aim, the methodology provides a stage for specifying rules that will activate the vocal interactions of the system with the workers on the field. In this stage the Safety Manager can model the configuration of agents locations that will initiate a vocal interaction of the Digital Platform, and what workers should be involved in the interaction. To this aim, each rule allows to specify two core pieces of information: a preconditions; an NLP sentence; one or more workers. The precondition expresses under what circumstances the rule should trigger a vocal interaction (Fig. 2). The NLP

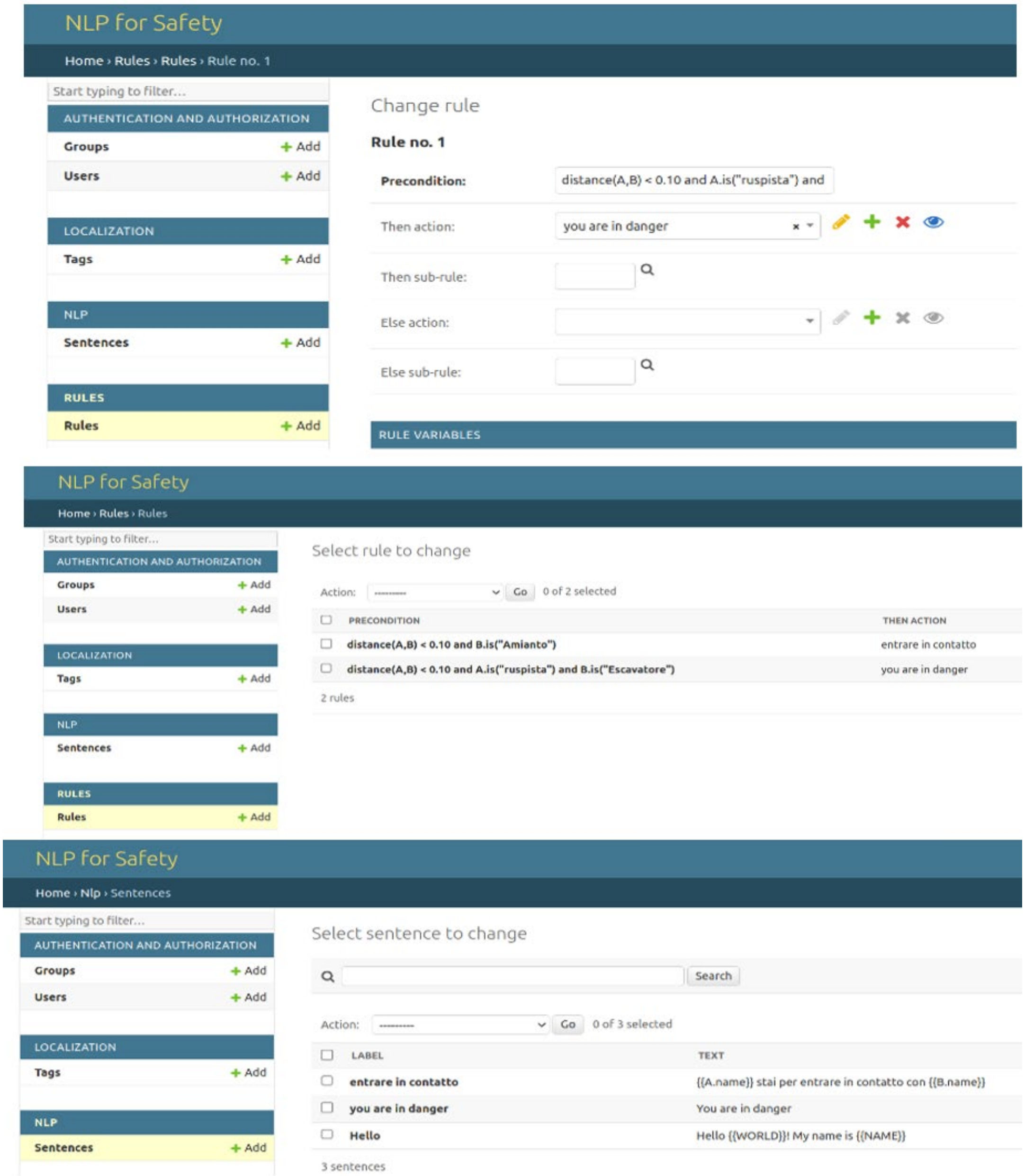


Figure 2 - NLP implementation: from above 1. rule definition; 2. rule-vocal action combination; 3. vocal action composition.

sentence specifies what audio message the Digital Twin Platform should send to the workers.
The rules preconditions are boolean formulas that can contain:

• variables;
• rational numbers;
• boolean operators (and, or, not);
• arithmetic operators (+, -, *, \);
• comparison operators among rational numbers (>, <, ==, !=); and
• function names.

From a formal point of view, the rule precondition is a first-order, existentially quantified, boolean formulas over rational numbers. An example of rule precondition is the following:
\exists A:Human Resource, \exists B: Equipment. (B.identifier == "BridgeCrane" and distance(A,B) < 0.5 and A.identifier == "BridgeCraneOperator")
The intended meaning of such a precondition is to detect when any worker qualified as a bridge crane operator is "close enough" (less than 50 cm) to the controller of the bridge crane on the construction field.
The target of the rule, i.e. the recipient of the vocal message, can be the involved worker, in case the aim of the Safety Manager is to raise the attention of the worker itself, or it can be the Safety Manager, in case she/he wants to be notified of certain relevant events happening on the construction field.

Construction Site Monitoring and Vocal Interaction

Construction Site Monitoring constitutes of a digital process continuously receiving updates concerning the agents locations on the field, and at each location update all the rule preconditions in search for those rules that should be considered enabled.
In order to detect that a rule is enabled, the Construction Site Monitoring contains a pattern matching logic searching for agents in the construction site that are compatible with the variables required by each rule, and selecting only those configurations of agents that satisfy the rule preconditions. Each configuration of agents that satisfy a rule is called an environment satisfying that rule. Note that the same rule can be satisfied by several environments, meaning that the same rule can generate multiple vocal interactions targeting different agents at the same time (for instance, to raise their attention or to ask for confirmation about the task they are executing on the field).
The latter stage, Vocal Interaction, involves generating an audio file whenever a rule is found enabled during the Construction Site Monitoring stage (Fig. 2). In case the NLP sentence associated with the rule is an open-sentence with parameterized expressions in it, the Vocal Interaction stage receives from the Construction Monitoring Site stage the environments satisfying the rule, and by binding the parameters of each such environment to the open-sentence a closed-sentence is obtained and converted to an audio file using text-to-speech cloud services. Finally, such audio file is routed to the workers associated with the enabled rule. It should be specified that rule activation has no effect on the environment except for voice interaction.

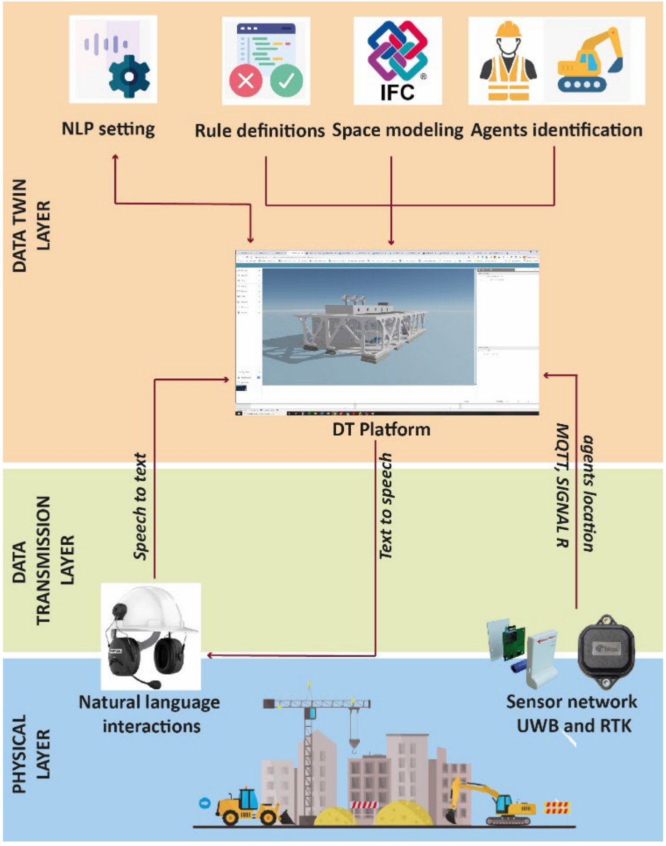


Figure 3 - System Architecture

System architecture

The architecture of the proposed system is shown in Fig. 3. The physical layer, typical of any DT, is represented in our case by the construction site, including resources (workers and machinery), definition of spaces and hazards. The connection between the real world and the digital world is represented by the Data transmission layer. In this DT it is not possible to speak only of Data acquisition layer because the system focuses on communication at the construction site through natural language. Therefore, the data transmission layer has on the one hand the task of acquiring data from the construction site through a sensor network, on the other hand, the information from the reasoning part applied to the data is poured back into the construction site through voice interactions performed by the system.
From data transmission layer we move to what we have called DT layer. In this layer of the system there are several components. Into the DT platform goes all the relevant information for site management and it is the platform that invokes the natural language micro-service. The necessary information also already covered in the methodology concerns the definition of the resources involved (workers and means), the rules to be applied and the spaces. The latter are the spaces of the construction site thus work areas, material storage areas, machinery maneuvering spaces, internal roadway areas, etc.

Application development

A simulated construction site has been recreated in order to test a prototype implementing the proposed methodology. The experiment was conducted within the DC3 laboratory of the Polytechnic University of Marche. The laboratory is equipped with a bridge crane and the agents in the simulated construction site were monitored using an UWB sensor network composed of five UWB anchors mounted along the perimeter of the laboratory and seven UWB tags attached to the agents moving in the simulated construction site. The gateway of the UWB sensor network published the coordinates of every tag once per second on a configured MQTT broker, while the web application subscribed the topics of the same broker in order to receive the most updated position for each tag and store it on the database. To enable audio and vocal interaction with the VVA each worker was equipped with a Bluetooth headset and a battery powered Raspberry PI mini PC connected to a local Wi-Fi internet connection. A web application implemented the aforementioned architecture to drive the Safety Manager along the steps of the proposed methodology. The basic entities at the core of the Digital Twin data-model are Equipment, Protective Equipment, Space, and Human Resource. On a different entity type named Construction Site, the Safety Manager creates instances representing each construction site he/she is managing, specifying for each the relevant attributes such as an identifying label, the location of the construction site itself, the date when the works began and the date when they are supposed to end.

Finally, entity type Item collects instances of each family of equipment, protective equipment, space or human resource, and put it in relation with the specific construction site. For example, two entities named “Jane Doe” and “Tom Cobley” respectively can be linked at a single entity labeled “Scraper operator” of type Human Resource, and at the entity of type Construction Site with label “Gotahm City Central Hospital”. This example clearly identifies two (working) agents in a specific construction stie. They can also be linked to two different entities of type Human Resource, such as “Team A” for “Jane Doe” and “Team B” for “Tom Cobley”. As a different example, the Safety Manager may add an entity of type Item named “Parking Area” that is linked to an entity of type Space labeled “Team A Working Area”, while a second Item may be labeled “Warehouse” and be linked to an entity of type Space named “Team B Working Area”. In the web application, an entity of type Tag is stored on the database for each of the UWB tags, together with their relevant data, viz. the tag identifier, a label for it, the last known position, and the time when the position was detected by the UWB sensor network. Each Tag entity is linked to the entities of type Item mentioned earlier. The next task of the Safety Manager is to model the sentences the VVA should pronounce. Here is the list of template sentences and questions used for our simulated construction site:

S1: {{ A.name }}, the bridge crane requires the following personal protective equipment: {{ PPEs }}. Can you confirm you are wearing all of them?

S2: {{ A.name }}, it appears you are operating on the Bridge Crane {{ B.name }} as though you are not assigned to this task. Are you sure you are allowed to do it?

S3: Watch out {{ A.name }}! You are in danger because the crane is above you

S4: {{ A.name }}, it appears you are working outside the area you are assigned to.

Finally, the Safety Manager can specify the rules that defines the set of events triggering a sentence or a question. Rules are formalized as a logic precondition, associated with a sentence that is to be spoken by the VVA.

Rules are associated with typed variables and the following sentences have been associated to actual implementations:

- distance : it takes two 3-dimension coordinates and computes the linear distance between them
- above : it takes two 3-dimension coordinates and computes the linear distance on the XY plane, thus checking if one of the objects is close or above the other

Below are examples of rules tested in the simulated environment:

- R1: \exists A:Human Resource, \exists B: Equipment. (B.identifier == “BridgeCrane” and distance(A,B) < 0.5 and A.identifier == “BridgeCraneOperator”)
- R2: \exists A:Human Resource, \exists B: Equipment. (B.identifier == “BridgeCrane” and distance(A,B) < 0.5 and A.identifier != “BridgeCraneOperator”)
- R3: \exists A:Human Resource, \exists B: Equipment. (B.identifier == “BridgeCrane” and above(A,B))
- R4: \exists A:Human Resource, \exists B: Space. (B.identifier == “Team A Working Area” and A.identifier == “Team B”)

Rules are associated to sentences as follows: (R1, S1), (R2, S2), (R3, S3), and finally (R4, S4).

The most expensive task of the implemented system, in computational terms, is the construction monitoring phase, re-evaluating rules preconditions in search for environments satisfying the rules themselves and required in order to transform open-sentences onto closed-sentences ready to be reproduced on the workers’ headsets.

Results

In order to carry out an initial verification of the proposed framework, three tests were planned:

- TEST 1: identification of interference risk situation between overhead crane and worker and subsequent voice warning;

Figure 4 - DC3 lab tests, from left to right: test 1, test 2 and test 3.

- TEST 2: verification of possession of the qualifications for handling a construction site machinery (represented by the overhead crane) and vocal communication of any non-possession of the above requirements;
- TEST 3: verification of possession of permits to access areas of the construction site.

Figure 4 and 5 show the setting of the tests performed. In the first scenario, the hook of the overhead crane from which the load was hung was instrumented. The worker, on the other hand, was instrumented with a badge attached to the vest. The minimum distance below which the system sent a warning was set equal to 50 cm, and the warning sent to the worker via Bluetooth earpiece said "Alessandro (worker name) you are about to make contact with the overhead crane (equipment). In the second case, the worker (who always wears the sensor) holds the remote control of the bridge crane, which is also instrumented with a sensor for tracking. In this case, the minimum distance at which the warning starts is 20 cm, and the warning received by the worker is "Alessandro you are/are not in possession of a license to drive the overhead crane." Finally in the last case tested, the cones are instrumented with UWB sensors to detect location and the worker also has the sensor for location. In this case the area identified by the cones is connected with an area defined in BIM for which access rules have been introduced in the system. In the latter test in case the worker sensor overlaps with the identified area, the system reports "Alessandro you do not have permission to access this area."

Figure 5 - Localization tag inside the DC3 lab during tests.

Conclusions

Safety management at construction sites is a complex problem in which recent technologies still struggle to establish themselves. In addition to reasons such as the economic investment required and the low perception of the usefulness of the proposed innovations, there is also the fact that changing processes or standardizing them in an industry that has never worked on rationalizing procedures is seen as a difficulty in addition to the others already mentioned. For this reason, Safety II proposes an approach to safety that, rather than imposing standard procedures, strives to find ways to make them follow naturally and integrated into operations. With this in mind, this research work aims to integrate natural language within processes through collaborative monitoring techniques. Initial tests have demonstrated the feasibility of the system and the effective development of this initial integration.

The next steps involve studying the best tool for voice communication on the construction site. A helmet with integrated headphones could be a solution whose feasibility should be investigated. Further testing should also be done at construction sites since an important limitation of the proposed approach could be the noise present at the construction site, which could adversely affect message comprehension. Finally, a systematization of knowledge about safety could be implemented for the purpose of testing the scalability of the system to all site operations. Another main limitation of the current research lays in the non-optimized implementation of the rule engine, that at the moment is based on Python and capable of handling only small construction sites. Nevertheless, state-of-the-art and freely available SAT-solvers such as Chaff or Z3 are highly optimized to handle existentially quantified first-order boolean formulas over rational numbers and find environments that satisfy them in systems made of thousands of actors, thus making the methodology feasible even for real-world scenarios.

References

Baker, H., Hallowell, M. R., & Tixier, A. J. P. (2020). Automatically learning construction injury precursors from text. Automation in Construction, 118, 103145.

Corneli, A., Naticchia, B., Vaccarini, M., Carbonari, A., & Spegni, F. (2023). APPLICATION OF DIMINISHED REALITY FOR CONSTRUCTION SITE SAFETY MANAGEMENT. In Proceedings of the 23rd International Conference on Construction Applications of Virtual Reality.

Doukari, O., Wakefield, J., Martinez, P., & Kassem, M. (2024). An ontology-based tool for safety management in building renovation projects. Journal of Building Engineering, 108609.

Fang, W., Luo, H., Xu, S., Love, P. E., Lu, Z., & Ye, C. (2020). Automated text classification of near-misses from safety reports: An improved deep learning approach. Advanced Engineering Informatics, 44, 101060.

Fang, W., Ma, L., Love, P. E., Luo, H., Ding, L., & Zhou, A. O. (2020). Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. Automation in Construction, 119, 103310.

Farghaly, K., Soman, R. K., Collinge, W., Mosleh, M. H., Manu, P., & Cheung, C. M. (2022). Construction safety ontology development and alignment with industry foundation classes (IFC). Journal of Information Technology in Construction, 27, 94-108.

Gao, S., Ren, G., & Li, H. (2022). Knowledge management in construction health and safety based on ontology modeling. Applied Sciences, 12(17), 8574.

Giri, O. P. (2020). Factors causing health and safety hazards at construction sites. Technical Journal, 2(1), 68-74.

Haupt, T. C., Akinlolu, M., & Raliile, M. T. (2019, November). Emerging technologies in construction safety and health management. In International Conference on Innovation, Technology, Enterprise, and Entrepreneurship (ICITEE) (pp. 413-420).

INAIL, Raccomandazioni di sicurezza, CANTIERI POST-SISMA, (2019), https://www.inail.it/cs/internet/docs/alg-pubblicazione-cantieri-post-sisma.pdf LAST ACCESSED Feb. 2024

Johansen, K. W., Schultz, C., & Teizer, J. (2023). Hazard ontology and 4D benchmark model for facilitation of automated construction safety requirement analysis. Computer-Aided Civil and Infrastructure Engineering.

Khudhair, A., Li, H., Ren, G., & Liu, S. (2021). Towards future BIM technology innovations: A bibliometric analysis of the literature. Applied Sciences, 11(3), 1232.

Martinez-Rojas, M., Antolín, R. M., Salguero-Caparrós, F., & Rubio-Romero, J. C. (2020). Management of construction Safety and Health Plans based on automated content analysis. Automation in Construction, 120, 103362.

Mihić, M., Cerić, A., & Završki, I. (2018). Developing construction hazard database for automated hazard identification process. Tehnički vjesnik, 25(6), 1761-1769.

Mihic, M., Vukomanovic, M., & Završki, I. (2019). Review of previous applications of innovative information technologies in construction health and safety. Organization, technology & management in construction: an international journal, 11(1), 1952-1967.

Saini, G. S., Fallah, A., Ashok, P., & van Oort, E. (2022). Digital Twins for Real-Time Scenario Analysis during Well Construction Operations. Energies, 15(18), 6584.

Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. Automation in Construction, 158, 105200.

Shen, Q., Wu, S., Deng, Y., Deng, H., & Cheng, J. C. (2022). BIM-Based Dynamic Construction Safety Rule Checking Using Ontology and Natural Language Processing. Buildings, 12(5), 564.

Tian, D., Li, M., Ren, Q., Zhang, X., Han, S., & Shen, Y. (2023). Intelligent question answering method for construction safety hazard knowledge based on deep semantic mining. Automation in Construction, 145, 104670.

Wu, C., Li, X., Guo, Y., Wang, J., Ren, Z., Wang, M., & Yang, Z. (2022). Natural language processing for smart construction: Current status and future directions. Automation in Construction, 134, 104059.

Yap, J. B. H., Lam, C. G. Y., Skitmore, M., & Talebian, N. (2022). Barriers to the adoption of new safety technologies in construction: A developing country context. Journal of Civil Engineering and Management, 28(2), 120-133.

Zhang, S., Boukamp, F., & Teizer, J. (2015). Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA). Automation in Construction, 52, 29-41.

Zhao, Y., Wang, N., & Liu, Z. (2022). An Established Theory of Digital Twin Model for Tunnel Construction Safety Assessment. Applied Sciences, 12(23), 12256.

RE:STOCK INDUSTRY: DIGITAL FRAMEWORK FOR THE CIRCULAR REUSE OF EXISTING STRUCTURES FOR VERTICAL PRODUCTION

Julia Reisinger[1], Maria Antonia Zahlbruckner[1], Peter Kán[2], Iana Podkosova[2], Hannes Kaufmann[2] and Iva Kovacic[1]
[1] Institute of Building and Industrial Construction, TU Wien, Austria
[2] Institute of Visual Computing and Human-Centered Technology, TU Wien, Austria

Abstract

Despite numerous existing vacant industrial buildings and vast unused brownfield land, new production settlements are rarely integrated within existing structures. This paper presents the RE:STOCK INDUSTRY framework, aiming to develop a method for assessing the reuse potential of industrial structures, focusing on vertical expansion and circular economy principles. Utilizing scanning and photogrammetry data, a novel AI method will generate as-built FEM models for structural analysis and prediction of retrofitting measures under vertical expansion. An augmented reality app will aid visualization and decision-making on-site. Innovatively combining AI, computational structural design and AR, the project advocates sustainable reuse, reducing reliance on recycling or disposal.

Introduction

Industry accounts for 30% of Austrian economic growth (Bolen, 2022), with over 1,000 new industrial enterprises annually (Statistik Austria, 2022a). There are more than 82,200 industrial and warehouse buildings, with over 1,000 new construction permits granted each year (Statistik Austria, 2022b,c). The significant land sealing resulting from horizontal production and storage processes is a problem: Austria has approximately 20,000 hectares of unused, derelict industrial and commercial areas (Enzinger, 2017), and up to 6,000 commercially or industrially pre-used buildings with potential for reuse (Janitsch, 2022). Companies and municipalities that want to relocate businesses are faced with the challenge of securing space for the expansion or relocation of production, warehousing and logistics (Lasi et.al, 2014, Vajna, 2014), but too rarely integrate these into existing buildings. Cost- and time-efficient concepts for integrating new production and logistic uses into existing buildings and solutions for vertical production processes must also be developed to relocate production back to Europe while minimizing new land sealing.

The potential of a building for reuse and vertical upgrading depends heavily on the actual condition of the load-bearing structure, as it must be able to cope with the new loads and space requirements. Information on the material and structural composition, the condition, the pollutant and contaminant content, and, consequently, the effective load-bearing capacity and serviceability of the existing structure for conversion and new use is often lacking. In addition, most building owners do not see their existing building stock as a valuable material resource bank: existing structures are usually demolished and rebuilt, which increases the volume of waste and landfill, leaves valuable substances unused and results in increased resource and energy consumption.

The current challenge in evaluating the existing structural stock regarding its reuse capability is to accurately digitize, model and analyze the actual structural building stock and make it accessible to the life cycle-oriented economy. The creation of digital analytical as-built models using the finite element method (FEM) for structural analysis of existing structures is currently a predominantly manual, time-consuming, and error-prone process.

This paper deals with the question: "Which methods and technologies are suitable for digitally recording and documenting the structural properties and material composition of the existing building stock and evaluating its reuse, modernization and upgrading potential for vertical industrial uses from a circular economy perspective?"

The research is conducted within the research project RE:STOCK INDUSTRY, funded by the Austrian Research Promotion Agency FFG. The overall goal of the project is to reuse the existing load-bearing structures of industrial buildings in a targeted manner for new settlements, expansions, or conversions instead of resorting to recycling or disposal processes. Through innovative vertical production concepts and their integration into structural analysis, new soil sealing is to be avoided and the service life extended by reusing the building structure. In the research, missing automated methods for scan to FEM using artificial intelligence (AI) algorithms will be developed and innovative approaches for vertical retrofitting of industrial buildings and customized concepts for vertical production and storage processes in existing buildings researched. An interactive augmented reality (AR) application will enable the visualization of reuse concepts with real-time feedback directly at the construction site and it will aim at motivating planners and building owners to upgrade instead of demolishing and new construction.

This paper presents the RE:STOCK INDUSTRY methodology and framework. The paper is structured as follows: A state-of the art review presents existing work on Scan to FEM, Vertical production, Re-use assessment models and AR technologies for retrofitting purposes. The next section describes the research design and methodology followed by the presentation and description of the RE:STOCK INDUSTRY framework. Within the conclusion potentials and challenges of the research are discussed and future research steps highlighted.

State-of-the-art

Scan to FEM

Structural analysis models are essential for the reuse of industrial buildings. The level of knowledge about the load-bearing capacity and serviceability of existing buildings is currently achieved through extensive plan studies, former available structural calculations, and on-site inspections, followed by the creation of simplified digital twins or 2D plans. Digital as-built building modeling often focuses on scan-to-BIM approaches. Researchers are investigating the reconstruction of 3D building models for material quality analysis (Paral et al., 2021), structural analysis (Alfio et al., 2022) or reconstruction from point cloud data (Poullis, 2013). However, methods for the parametric representation of 3D models for FEM analysis out of scans are lacking. Scan to FEM is often investigated in architectural cases for historical buildings (Alfio et al., 2022, Barazzetti et al., 2015). Others include manual identification of properties for a procedural model to reconstruct geometric structures from point clouds (Funari et al., 2021), estimate elastic parameters of beams from point clouds (Riveiro et al., 2018), or detect cracks in reinforced concrete to ensure stability (Yu et al., 2021). Automatic methods for the recognition of cylinders in point clouds for the reconstruction of plants (Liu et al., 2013), FEM mesh structures for the numerical analysis of tunnels from point clouds (Cui et al., 2023), or the assignment of a B-spline representation to a point cloud for the purpose of FEM analysis (Xu and Neumann, 2020) are being investigated. Semi-automated methods for converting point clouds to BIM and using them for structural analysis (Rolin et al., 2019) or machine learning for monitoring the structural condition of buildings (Mishra, 2020) are used. The combination of point clouds with image data allows parts of point clouds to be segmented and classified as known classes of structural elements (Barrile et al., 2019). Most of the existing methods require manual intervention or work semi-automatically.

Vertical production

A third of Vienna's added value comes from production. STEP2025 addresses the future of the productive sector and advocates sufficient space for trade and industry (Rosenberger et al., 2017). Vertical production, involving the multi-storey use of buildings for the production of goods, can reduce the space required for trade and industry and requires efficient material flows across several levels (Hompel et al., 2018). VERTICALurbanFACTORY explores stacked functions and vertical production for efficient space utilization (Haselsteiner et al., 2019). Haselsteiner et al. (2020) show utilization scenarios for different building types to implement vertical production. Vertical production and logistics solutions are already established in densely populated regions such as Singapore and Hong Kong (Kuznetsova et al., 2018, Low et al., 2015). Vertical production concepts are being investigated in algorithmic optimization of material flows in a layout (Ahmadi et al., 2017, Karateke et al., 2022). Others are researching adaptive reuse of industrial buildings regarding critical factors (Vardopoulos, 2019) or for residential purposes (Glumac and Islam (2020). The integration of 3D utilization concepts into structural analysis can be enabled by spatial zoning methods, reaction grammars and parametric planning methods (Claessens et al., 2020, Boonstra et al., 2020, Reisinger et al., 2022). To the best of our knowledge, a methodology that considers the effects of multi-story production in structural analysis has not been widely researched.

Re-use assessment models and resource passports

The uncertainty in predicting the future use of materials and buildings is a challenge in circular economy approaches (De Wolf et al., 2020). BIM4eco is developing a web tool that automatically imports BIM information of a building model in early phases into a life cycle assessment program, records it component by component and calculates environmental impacts (SOLID, 2022). Reuse-Life Cycle Assessment (Escamilla, 2023) analyzes the potential for reducing the environmental impact of reusing materials in buildings. In Build-Re-Use (Lead AEE, 2022), basic principles for the construction and dismantling of buildings with short utilization cycles - supermarkets, office buildings, interim buildings, sanitary facilities - are developed. The buildings are constructed with reusable components and can be dismantled and returned after use. Gebäudepass (Umweltbundesamt, 2013) is developing the basis for the standardization of building passports as building material information systems. M-DAB (2019) uses BIM-based digital technologies to digitize, analyze and sustainably manage the city's material resources. BAMB, (2020) aims to involve the construction industry in the circular economy and increase the value of used materials. PlattformCB'23 (2023) has created a guideline for the circular economy, circular tenders, and future reuse.

Existing digital material register platforms (Madaster, 2024, Concular, 2024, Rosen, 2024, EcoDesign Circle, 2023) document the products and materials used in buildings in terms of quality and quantity and enable the creation of building resource passports. The focus is on circularity, the environment, life cycle assessment and financial evaluation. Raw material utilization is mapped in Madaster and status descriptions are possible in the form of certificates and archive documents. The RhinoCircular (Heisel, 2021) and Phoenix (EPFL, 2024) tools deal with structural components and parameters relevant to the circular economy. There are few approaches for documenting and assessing the load-bearing capacity and serviceability of existing buildings for new use. Re-use assessment and documentation methods often focus on materials, elements, compounds, and chemical constituents. Structural aspects are neglected, and the focus is primarily on recycling and deconstruct ability. RE:STOCK INDUSTRY is investigating ways of evaluating and documenting structural information relating to the reusability of the building.

Augmented Reality for retrofitting measures

One of the main advantages of AR is situated visualization, where the simulated content is closely linked to the environment and decision-making processes are supported by facilitating the calculation and understanding of domain knowledge (Martins et al., 2022). In the construction industry, AR is used in education for on-site training (Tan et al., 2022) and quality testing of HVAC systems (Schranz et al., 2021). GAMMA AR (Gramma, 2024) is used for construction site monitoring and documentation with BIM. The combined use of BIM and AR is gaining increasing interest. AR enables the display of BIM elements directly in the building to make annotations or show hidden installations (Chai et al., 2019, Urban et al., 2019, Hugo Silva et al., 2021). When displayed on AR glasses, the addition of AR data enables hands-free exploration of the built structure. The use of AR for interactive finite element analysis has also been investigated (Huang et al., 2017, Huang et al., 2023). The research methods in which conversion and retrofitting options of buildings can be visualized as scenarios on site within AR are sparse.

Research Design and Methodology

The research objective is to develop a framework for the end-to-end digital recording, modeling, and analysis of the load-bearing stock of industrial buildings for circular re-use and vertical expansion. The potential for reuse, modernization and upgrading in vertical extensions, considering circular economy aspects, should be determined, and recorded by documenting structure-specific information in resource passports and material registers. Through efficient visualization stakeholders should be able to foresee vertical retrofitting possibilities including modernization measures directly on the construction site and to have an assessment model which evaluates the planned retrofitting measures in comparison to demolition and new construction through ecological and economic assessment.

Figure 1 shows the research design of the project for the method and framework development. The data collection and method development are based on industrial buildings in steel and reinforced concrete skeleton construction, as 80% of the industrial building stock consists of these types of structures. Ten buildings in Austria with different geometries and structural systems are defined as use cases. The existing buildings will be captured using scanning and photogrammetry, on-site inspections by experts with knowledge of deconstruction and structural retrofitting and supplemented by the analysis of planning documentation. During the inspection, contaminated elements and damages will also be digitally recorded and localized.

In the first step, two novel methods for automated generation of the analytical as-built FEM model by coupling point clouds and image data with AI algorithms will be researched and developed: a.) Reconstruction of the parametric FEM model from 3D point clouds by heuristic optimization methods. The reconstructed model will contain information on the geometry, material and condition of the existing structure. A digital process enables the integration of additional information on damage assessment, detailed material properties such as material grades and contaminants. b.) Development of a deep neural network (DNN) for the automated recognition of structural joint connections from geometric and image data. Real as well as synthetic data from existing use cases and BIM models will be used to train the DNN.

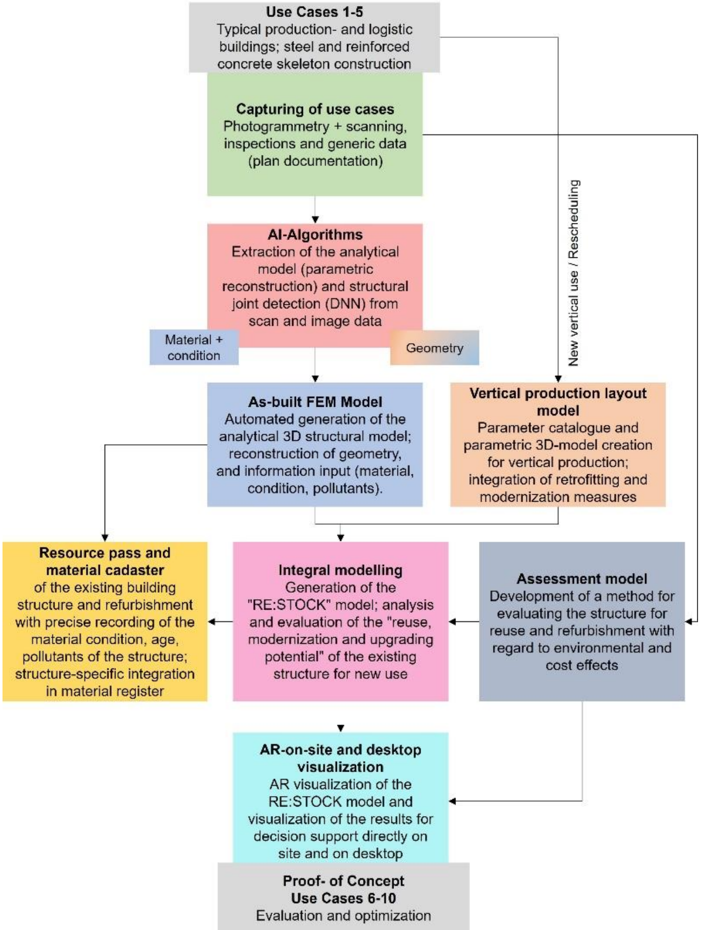


Figure 1: The research design and methodology of RE:STOCK INDUSTRY

Furthermore, vertical production and modernization concepts for integration into structural analysis processes are researched to quickly determine the load-bearing and strengthening capacity of the existing building for new use: By analyzing use cases, best practice, literature and interviews, requirements and processes for vertical production, the dependencies with the supporting structure and modernization concepts for vertical strengthening are defined, documented in a catalog of requirements and transferred into parametric planning rules. Developments from the authors previous research (Reisinger et al., 2022), in which a concept for the integration of 2D horizontal production processes into structural design was developed, serve as the basis for the development of the parametric RE:STOCK model, which couples the as-built FEM model with 3D vertical production concepts.

To integrate circular economy aspects into the analysis of existing structures two methods are being researched: a) Development of a method for evaluating the reuse and retrofitting of load-bearing structures. Clear evaluation criteria for reuse and retrofitting, such as environmental and cost effects, quality, functionality, and technical feasibility, are defined and integrated into existing methods of ecological and economic life cycle analysis in order to take specific reuse requirements into account. The aim is to use the RE:STOCK model to compare the existing structure including retrofitting with demolition and new construction scenarios and to compare their environmental impacts and costs. b) Development of a method to efficiently document the digitized inventory and retrofitting planning in building resource passports and material registers. This is done by analyzing available databases and cadasters and standardizing data and interfaces with the RE:STOCK model. As a result, detailed structural information such as material composition, quantity, load-bearing capacity, condition, year of construction and pollutant content can be efficiently recorded.

Finally, a method for integrating AR technology as a visual decision support tool for planners and clients into the early retrofitting process will be researched. The aim is to visualize the effects of reusing the structure, the vertical strengthening and the ecological and economic evaluation feedback directly on the construction site. The proof-of concept will be conducted on real use cases and validated with experts.

RE:STOCK INDUSTRY framework

The RE:STOCK INDUSTRY framework will integrate all of the aforementioned research developments into a digital platform to guarantee a holistic planning process and decision support tool for early re-use, expansion and retrofitting processes. The framework aids to automatically generate accurate FEM models of the building structure inventory from point clouds and photogrammetry data and document the structure-specific information in resource passports. A computational tool, the RE:STOCK model, will generate re-use plans of the existing structure for vertical upgrades and evaluate them from a circular economy perspective. By linking an AR method to the integral RE:STOCK model, the visualization of possible expansion and reinforcement measures will be displayed directly on site in the existing building.

Figure 2 shows the framework of RE:STOCK Industry and the most relevant developments with the parametric *RE:STOCK Model (6)* as the central part:

1 and 2 - Scan Data to as-built FEM Model: An as-built FEM model will be generated out of received point clouds and photogrammetry data and enriched with the exact geometry and material properties. Optimization-based techniques will be used to reconstruct the main shape of vertical and horizontal structural elements and their parameters (e.g. size, thickness). The elements will be initialized by heuristic search for vertical and horizontal point beams and the parameters of the model will be optimized to fit the scanned point cloud. Deep learning will be used with geometry and image data to classify the joint connections of the structural elements and describe their properties in relation to the connected elements. The neural network will be trained with real and synthetic data from existing use cases and BIMs. The recognized joint connections and status descriptions will be added to the parametric model to enable the structural analysis in the as-built FEM model. The proposed algorithms will significantly accelerate the as-built analysis of skeleton constructions. This model can be used not only for traditional structural analysis but also for cycle-oriented documentation and evaluation of the re-use capability of the structure.

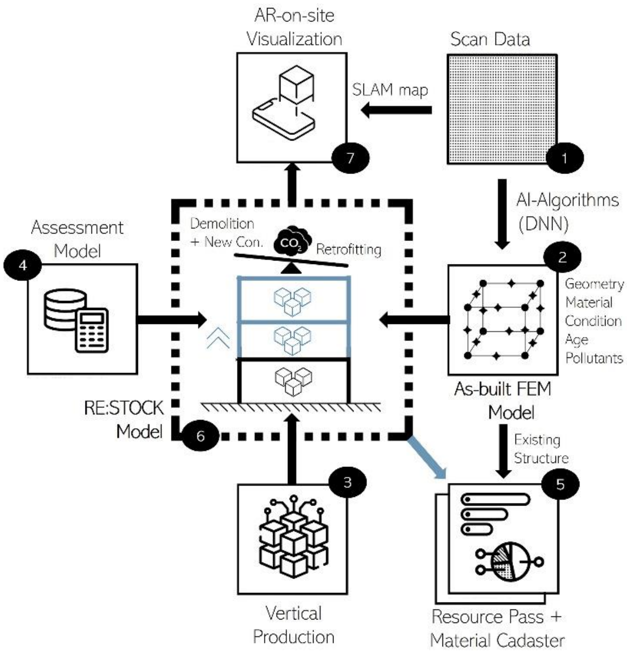


Figure 2: The RE:STOCK INDUSTRY framework

3 - Vertical Production: The goal is to create a 3D-layout generator for vertical production planning. Requirements and processes for vertical production as well as modernization concepts for structural vertical upgrading are defined, documented in a catalog of requirements and transferred into parametric planning rules. Parametric methods from the previous developed 2D production layout generator (Reisinger et al., 2022) are used to develop the vertical 3D layout model. For the integral RE:STOCK model, the as-built FEM model is extended with the vertical production layout model through parametric and generative design methods.

4 and 5 - Assessment Model and Resource Pass + Material Cadaster: The interfaces to existing material registers and resource passports are examined so that relevant structural information from the RE:STOCK model can be transferred and recorded in material passports and building cadasters. An ecological-economic method will be developed to evaluate the reusability of the structure in terms of retrofitting vs. demolition/new construction. For this purpose, qualitative and quantitative sustainability principles (life cycle assessment, circular construction, adaptability, preservation of existing buildings, CO2 storage capacity, reusable structure, legal and technical feasibility) are defined. The method integrates ecological and economic life cycle analyses to determine the gray energy of the existing building, CO_2 savings through structural preservation, and the evaluation of retrofitting measures to promote reuse instead of demolition/new construction. Argumentation aids and proposed solutions. such as CO_2 pricing and climate impact costs, will be included, so that not primarily the option with the lowest costs is chosen.

7 - AR-on-site Visualization: For this purpose, the RE:STOCK model will be integrated into a specially developed situated AR application. On site, the model will be aligned with the real building by locally comparing the initial point cloud scan - from which the RE:STOCK model was generated - with the sparse point cloud provided by the AR device. The registration between original scanned point cloud and the point cloud provided by the AR device will be achieved using iterative closest point algorithm (Besl and McKay, 1992). Once the RE:STOCK model has been aligned with the real building, the AR view is correctly updated in perspective as the user walks around using Simultaneous Localization and Mapping (SLAM).

Conclusion

The presented digital framework for the circular reuse of existing industrial building structures for vertical production addresses industrial companies, builders and architects, production and structural engineers or life cycle management stakeholders. By digitally capturing the structural stock and assessing and visualizing the retrofit potential using the AR visualization approach, stakeholders should be motivated to reuse the building stock as it is, increase its value through vertical retrofit measures and extend the life of their buildings. The research promotes vertical production and logistic processes, optimizes operations and resource utilization, and maximizes business efficiency and productivity. The automation and standardization of the data basis for the structural stock of industrial buildings is essential for the long-term optimization and permanent monitoring of building waste management, whereby the central linking and provision of structural data with building geometry information plays a key role for numerous other fields of research and applications. The methods to be developed aim to avoid manual, error-prone processes in inventory recording and modeling as well as lengthy iterations in interdisciplinary planning processes. Integral analysis and evaluation can be expected to lower the inhibition threshold for stakeholders to make changes and investments in their buildings and production facilities. The methods will enable sustainable structural modernization under vertical use and lead to an extension of the building's service life and avoidance of soil sealing. Land recycling enables the resettlement and return of production companies from surrounding areas and abroad. By upgrading populated areas, economic added value is achieved, and social structures and workplace situations are expected to be improved. The methods should reduce the ecological impact by up to 50% and waste volumes by up to 70%, this will be examined in a case-study of real industrial buildings.

In the context of the presented study several limitations need attention in future scientific research. The framework's applicability is currently focused on the specific building type of industrial skeleton construction and the geographic region of Austria, necessitating further research to evaluate its generalizability across diverse contexts and construction practices. Furthermore, the framework relies heavily on emerging technologies such as AI, AR, and novel structural analysis methods, whose interoperability and integration into existing systems need to be fully explored. The quality of input data—scanning, photogrammetry, material properties, and structural details—directly influences the framework's output accuracy, highlighting the need for rigorous validation and sensitivity analyses in future. Economic viability, encompassing lifecycle cost analysis need to consider dynamic market conditions, which poses to be challenging. Regulatory, legal, and policy barriers present additional impediments to practical adoption, which highlights the importance of in-depth analysis and close work with municipalities to identify ways for regulatory reform and policy advocacy. A holistic environmental impact assessment extending beyond the CO_2 emissions and waste reduction of the structural system but to include environmental impacts of the whole building systems as facades and built in components will be crucial for understanding the full environmental footprint. The project solely focuses on the resource efficiency of the structure to avoid CO_2-equivalent emissions; however, energy efficiency is not addressed. Lastly, the technological and methodological advancements proposed necessitate a change in traditional practices among stakeholders in the planning and construction industry. The adoption of such new technologies and methods is influenced by behavioral factors, including perceived utility, ease of use, and cultural aspects and will be respected in user studies with practitioners and students on real use-cases.

With a focus on industrial and logistic buildings, the results of this project will be also beneficial for numerous planning areas that require (automated) recording and evaluation of existing buildings for new use in long term.

Acknowledgments

The authors would like to acknowledge the support by the Austrian funding institution FFG (Österreichische Forschungsförderungsgesellschaft) for the research project "RE:STOCK INDUSTRY" within the "Kreislaufwirtschaft" (3rd call for tenders) programme. We would particularly like to thank our project partners.

References

Ahmadi, A., Pishvaee, M.S, Jokar, M.R.A. (2017) A survey on multi-floor facility layout problems. Computers & Industrial Engineering, Volume 107, 2017, Pages 158-170, ISSN 0360-8352. https://doi.org/10.1016/j.cie.2017.03.015

Alfio, V.S., Costantino, D., Pepe, M., Garofalo, A.R. (2022) A Geomatics Approach in Scan to FEM Process Applied to Cultural Heritage Structure: The Case Study of the Colossus of Barletta. Remote Sensing, 14, 3. https://doi.org/10.3390/rs14030664

BAMB (2020) Buildings as material banks. European Union's Horizon 2020: Grant agreement No. 642384. https://www.bamb2020.eu/

Barazzetti, L., Banfi, F., Brumana, R., Gusmeroli, G., Previtali, M., Schiantarelli, G. (2015) Cloud-to-BIM-to-FEM: Structural simulation with accurate historic BIM from laser scans, Simulation Modelling Practice and Theory, Volume 57, 2015, Pages 71-87, https://doi.org/10.1016/j.simpat.2015.06.004

Barrile, V., Candela, G., Fotia, A. (2019) Point cloud segmentation using image processing techniques for structural analysis. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences. doi:10.5194/isprs-archives-XLII-2-W11-187-2019

Besl, P. J. and N. D. McKay (1992) A method for registration of 3-D shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence 14(2): 239-256.

Bolen, L. (2022) Alles über die Industrie in Österreich. Industriemagazin on 12.08.2022, Online-Access: 27.06.2023.

Boonstra, S., van der Blom, K., Hofmeyer, H., Emmerich, M.T.M. (2020) Conceptual structural system layouts via design response grammars and evolutionary algorithms. Automation in Construction, Volume 116, 2020, 103009, ISSN 0926-5805. https://doi.org/10.1016/j.autcon.2019.103009

Chai, C., Mustafa, K., Kuppusamy, S., Yusof, A., Lim, C.S., Wai, S.H., (2019) BIM Integration in Augmented Reality Model. International Journal of Technology, Volume 10(7), pp. 1266-1275.

Claessens, D.P.H., Boonstra, S., Hofmeyer, H. (2020) Spatial zoning for better structural topology design and performance. Advanced Engineering Informatics, Volume 46, 2020, 101162, ISSN 1474-0346. https://doi.org/10.1016/j.aei.2020.101162

Concular GmbH (2024) Access on 19.02.2024. https://concular.de/

Cui, L., Zhou, L., Xie, Q., Liu, J., Han, B., Zhang., T., Luo, H. (2023) Direct generation of finite element mesh using 3D laser point cloud. Structures, 47. https://doi.org/10.1016/j.istruc.2022.12.010

De Wolf, C., Hoxha, E., Fivet, C. (2020) Comparison of environmental assessment methods when reusing building components: A case study. Sustainable Cities and Society, Volume 61, 2020, 102322, ISSN 2210-6707. https://doi.org/10.1016/j.scs.2020.102322

EcoDesign Circle (2023) Circular Design Tools. Interreg - Baltic Sea Region & European Regional Development Fund (EU). https://circulardesign.tools/

Enzinger, S. (2017) Bodenverbrauch gefährdet Lebensgrundlage der nächsten Generationen. Environmental Agency Austria, Access on 27.06.2023. https://www.umweltbundesamt.at/aktuelles/presse/news2017/news-170612#:~:text=

Escamilla, E. Z. (2023) Reuse-Ökobilanz: Identifizierung des Potenzials zur Verringerung der Umweltauswirkungen der Wiederverwendung von Materialien bei Schweizer Gebäuden. Bundesamt für Energie (BFE), (2021-2023). https://sc.ibi.ethz.ch/forschung/research-projects/reuse-_lca--identification-du-potentiel-de-reduction-des-impacts.html

Funari, M.F., Hajjat, A.E., Masciotta, M.G., Oliveira, D.V., Lourenco, P.B. (2021) A parametric scan-to-FEM framework for the digital twin generation of historic masonry structures. Sustainability. 13, 19. https://doi.org/10.3390/su131911088

Glumac, B., Islam, N. (2020) Housing preferences for adaptive re-use of office and industrial buildings: Demand side. Sustainable Cities and Society, Volume 62, 2020, 102379, ISSN 2210-6707. https://doi.org/10.1016/j.scs.2020.102379

Gramma Technologies S.á r.l. (Luxemburg) (2024) GRAMMA-AR. Access on 19.02.2024. https://gamma-ar.com/

Haselsteiner, E., Madner, C. (2019) Kriterien, Potenziale und innovative Konzepte der vertikalen Verdichtung von Produktion und Stadt (2017-2019), Stadt der Zukunft, FFG. https://nachhaltigwirtschaften.at/en/sdz/projects/vertical-urban-factory.php

Haselsteiner, E., Madner, V., Frey, H., Grob, L.M., Laa, B., Winder, M. Schwaigerlehner, K. (2020)

VERTICAL urban FACTORY – Innovative Konzepte der vertikalen Verdichtung von Produktion und Stadt. Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie (BMK), Vienna, https://nachhaltigwirtschaften.at/resources/sdz_pdf/schriftenreihe-2020-9a-verticalurbanfactory.pdf

Heisel, F. (2021) RhinoCircular. Cornell AAP Architecture Art Planning - Circular Construction Lab. Access on 19.02.2024 https://labs.aap.cornell.edu/ccl/rhinocircular

Hompel, M.; Schmidt, T.; Dregger, J. (2018) Materialflusssysteme. Förder- und Lagertechnik. 4. Auflage. Berlin, Heidelberg: Springer Vieweg (VDI-Buch).

Huang, J.M., Ong, S.K., Nee, A.Y.C. (2017) Visualization and interaction of finite element analysis in augmented reality, Computer-Aided Design, Volume 84, 2017, Pages 1-14, ISSN 0010-4485. https://doi.org/10.1016/j.cad.2016.10.004

Huang, J.M., Ong, S.K., Nee, A.Y.C. (2023) An Augmented Reality Platform for Interactive Finite Element Analysis. In: Nee, A.Y.C., Ong,S.K. (eds) Springer Handbook of Augmented Reality.

Hugo Silva, A.C, Gaber, M. & Dolenc, M. (2021) Using Augmented Reality in Different BIM Workflows. 10.5772/intechopen.99336

Janitsch, C. (2022) Brachflächen-Dialog des BMK – neue Impulse für Flächenrecycling [Webinar on 18.05.2022]. WKÖ Fachverband Immobilien-und Vermögenstreuhänder, https://immowebinar.at/brachflaechen_dialog

Karateke, H., Şahin, R., Niroomand, S. (2022) A hybrid Dantzig-Wolfe decomposition algorithm for the multi-floor facility layout problem. Expert Systems with Applications, Volume 206, 2022, 117845, ISSN 0957-4174. https://doi.org/10.1016/j.eswa.2022.117845

Kuznetsova, E., Ng, T.S., Cardin, M.A., HE, Z. (2017) A Stochastic Programming Approach for the Design of Multi-Story Recycling Facility. IISE Annual Conference.

Lasi, H., Fettke, P., Kemper, HG. Feld, T., Hoffmann, M. (2014) Industrie 4.0. Wirtschaftsinf 56, 261–264. https://doi.org/10.1007/s11576-014-0424-4

Lead AEE intec (2022) Build-Re-Use. FFG (2022-2024). https://www.ibo.at/forschung/referenzprojekte/data/build-re-use

Liu, Y.J., Zhang, J.B., Hou, J.C., Ren, J.C., Tang, W.Q. (2013) Cylinder detection in large-scale point cloud of pipeline plant. IEEE Transactions on Visualization and Computer Graphics, 19, 10. https://doi.org/10.1109/TVCG.2013.74

Low, S.P., Gao, S. and Tiong, K.L. (2015) Applying lean production principles to facilities design of ramp-up factories. Facilities, Vol. 33 No. 5/6, pp. 280-301. https://doi.org/10.1108/F-11-2013-0086

Madaster Austria GmbH (2024) Access on 19.02.2024. https://madaster.at/

Martins, N.C., Marques, B., Alves, J., Ariújo, T., Dias, P., Santos, B.S. (2022) Augmented reality situated visualization in decision-making. Multimed Tools Appl, 81, 14749–14772. https://doi.org/10.1007/s11042-021-10971-4

M-DAB (2021) Energie der Zukunft, Technical University of Vienna, FFG call SdZ 6. https://projekte.ffg.at/projekt/3307471

Mishra, M. (2020) Machine learning techniques for structural health monitoring of heritage buildings: A state-of-the-art review and case studies. Journal of Cultural Heritage. 47. 10.1016/j.culher.2020.09.005.

Paral, A., et al. (2021) A deep learning-based approach for condition assessment of semi-rigid joint of steel frame. Journal of Building Engineering, 34: 101946. https://doi.org/10.1016/j.jobe.2020.101946

Platform CB'23 (Netherlands) (2023) https://platformcb23.nl/index.php?option=com_content&view=article&id=39&Itemid=265

Poullis, C. (2013) A framework for automatic modeling from point cloud data. IEEE Transactions on Pattern Analysis and Machine Intelligence. 35. https://doi.org/10.1109/TPAMI.2013.64

Reisinger, J.; Kugler, S.; Kovacic, I.; Knoll, M (2022) Parametric Optimization and Decision Support Model Framework for Life Cycle Cost Analysis and Life Cycle Assessment of Flexible Industrial Building Structures Integrating Production Planning. Buildings, 12, 162. https://doi.org/10.3390/buildings12020162

Riveiro, B., Cubreiro, G., Conde, B., Cabaleiro, M., Lindenbergh, R., Soilán, M., and Caamaño, J. C. (2018) AUTOMATED CALIBRATION OF FEM MODELS USING LIDAR POINT CLOUDS, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-2, 969–974. https://doi.org/10.5194/isprs-archives-XLII-2-969-2018

Rolin, R., Antaluca, E., Batoz, J.L., Lamarque, F., Lejeune, M. (2019) From point cloud data to structural analysis through a geometrical hBIM-oriented model. Journal on Computing and Cultural Heritage. 12, 2. https://doi.org/10.1145/3242901

Rosen, A (2024) Urban Mining Index (Germany). Access on 19.02.2024; https://urban-mining-index.de/

Rosenberger, M., Braumann, A., Emrich, H. (2017) Fachkonzept produktive Stadt. Wien: Stadtentwicklung Wien, Magistratsabteilung 18 -Stadtentwicklung und

Stadtplanung, https://www.digital.wienbibliothek.at/wbrup/content/thumbview/3954866

Schranz, C., Urban, H., Kaufmann, H., Schönauer, C., Rattenberger, J., O´Brien, P., Ozeraitis, L., & Jaritz, P. (2021) AR-AQ-Bau − Einsatzvon Augmented Reality zur Abnahme und Qualitätssicherung auf Baustellen. http://hdl.handle.net/20.500.12708/40511

SOLID - Wirtschaft und Technik am Bau (2022) Das BIM4eco-Modell von Delta. Access on 27.06.2023. https://solidbau.at/fokus/das-bim4eco-modell-von-delta/

Statistik Austria. (2022a) Amount of industrial companies in Austria 2012 to 2020 [Graph]. In Statista, Access on 27.06.2023, https://de.statista.com/statistik/daten/studie/1180095/umfrage/industrieunternehmen-in-oesterreich/

Statistik Austria (2022b) Datenpakete Gebäude- und Wohnungsregister. Access on 27.06.2023. https://www.statistik.at/datenbanken/adress-gebaeude-und-wohnungsregister/adress-gebaeude-und-wohnungsregister/daten-des-gebaeude-und-wohnungsregisters/bestandsdaten

Statistik Austria. (2023) Anzahl der bewilligten neuen Nichtwohngebäude in Österreich nach Gebäudearten im Jahr 2022 [Graph]. In Statista. Access on 27.06.2023. https://de.statista.com/statistik/daten/studie/421471/umfrage/anzahl-der-bewilligten-nichtwohngebaeude-in-oesterreich-nach-bauweise/

Statistik Austria (2022c) Treibhausgas-Emissionen in Österreich nach Sektor von 1990 bis 2020 [Graph]. In Statista, Access on 27.06.2022. https://de.statista.com/statistik/daten/studie/961605/umfrage/treibhausgas-emissionen-in-oesterreich-nach-sektor/#

Tan, Y.; Xu, W.; Li, S.; Chen, K. (2022) Augmented and Virtual Reality (AR/VR) for Education and Training in the AEC Industry: A Systematic Review of Research and Applications. Buildings 2022, 12, 1529. https://doi.org/10.3390/buildings12101529

Umweltbundesamt Austria (2013) GEBÄUDEPASS - als Gebäudematerialinformationssystem (2013 - 2014). AG. Bundesministerium für Land- und Forstwirtschaft, Umwelt und Wasser (BMLFUW). http://www.rma.at/node/1580

Urban, H., Schranz, C., & Gerger, A. (2019) BIM auf Baustellen mit Augmented Reality. Bauaktuell, 10(5), 192–196. http://hdl.handle.net/20.500.12708/143144

Vajna, S. (2014) Integrated Design Engineering. Ein interdisziplinäres Modell für die ganzheitliche Produktentwicklung. Springer Vieweg Berlin, Heidelberg, https://link.springer.com/book/10.1007/978-3-642-41104-5

Vardopoulos, I (2019) Critical sustainable development factors in the adaptive reuse of urban industrial buildings. A fuzzy DEMATEL approach. Sustain, Cities Soc. 2019, 50, 101684.

Xu, W. and Neumann, I. (2020) Finite element analysis based on a parametric model by approximating point clouds. Remote Sensing, 12, 3. https://doi.org/10.3390/rs12030518

Yu, K., Zhang, C., Shooshtarian, M., Zhao, W., Shu, J. (2021) Automated finite element modeling and analysis of cracked reinforced concrete beams from three dimensional point cloud. Structural Concrete, 22, 6. https://doi.org/10.1002/suco.202100194

AUTOMATED ON-SITE HAZARD IDENTIFICATION USING DIGITAL AND THERMAL IMAGERY

Mohamad F. Abbas[1], Bahaa Eddine Mneymneh[1], and Hiam Khoury[1]
[1]American University of Beirut, Beirut, Lebanon

Abstract

A major component of construction safety management processes entails identifying potentially on-site hazardous areas. However, traditional methods are manual, labor-intensive, and time-consuming. In an attempt to automate hazard identification, researchers have used digital imagery and computer vision at large but failed to explore thermal imaging analysis. As such, this paper introduces an automated vision-based hazard identification approach using both digital and thermal imagery. Several experiments were conducted and results highlighted the effectiveness of analyzing both digital and thermal images to reduce false positive detections and rapidly identify on-site hazardous areas.

Introduction

The construction sector has always been responsible for a high rate of fatal injuries. As such, regular safety inspections are of paramount importance and help management ensure that safe work practices are being maintained on jobsites. However, current inspection methods still rely on safety officers patrolling the site and identifying unsafe conditions, or on the workers' capabilities in identifying hazards. The former method is manual, labor-intensive, and time-consuming, while the latter approach is considered inadequate and not very dependable. Therefore, there is a need to automate construction safety inspections. Initial research efforts have targeted automating the safety inspection processes and monitoring the site conditions and construction personnel using digital imagery and computer vision (Yang et al. 2010, Brilakis et al. 2011, Park and Brilakis 2012). Other studies have focused on actively detecting PPE-wearing from digital images and videos using computer vision (Shrestha *et al.* 2015, Park et al. 2015, Abbas et al. 2016, Mneymneh et al. 2017, Mneymneh et al. 2018, Mneymneh et al. 2019, Nath et al. 2020, Yang et al. 2020). In a nutshell, extensive work has been done in the past years targeted at automating safety inspection processes while building on recent developments in the field of computer vision (Guo et al. 2021, Maali et al. 2024).

In an attempt to further automate safety inspection processes, several other research efforts evaluated the applicability of Unmanned Aerial Vehicles (UAV) or camera-equipped drones in construction safety (Gheisari et al., 2014, Abbas et al., 2016, Mneymneh et al. 2016, Melo et al. 2017, Kim et al. 2019, Martinez el el. 2020, Maali et al. 2024). Results revealed the importance of adopting drones in safety applications as they can provide safety personnel with real-time visual access to jobsites. However, no prior work has analyzed thermal imagery or coupled it with digital imagery to identify construction hazardous areas. Therefore, the objective of this paper is to enhance construction safety inspections by developing an automated vision-based hazardous identification system using both digital and thermal imagery. As such, real-time images and videos of indoor construction sites are captured using digital and thermal camera-equipped drones or UAVs.

Methodology

The hazardous situations considered in this study were extracted from a previously published survey conducted with construction personnel (Abbas et al. 2018). It was found that the top ranked hazardous situations are unprotected openings, fire, steel grinding, and steel welding. Only fire, steel grinding, and steel welding were considered in this study as the exact location of openings can be obtained from the as-planned drawings. It is worth noting that the routine activities of steel grinding and welding can be safely conducted. However, they may pose risks if adequate precautions are not observed and appropriate protective gear is not utilized. Hence, detecting these activities serves as the initial step towards implementing necessary health and safety measures.

Given the nature of the aforementioned three hazardous scenarios, it was decided to opt for digital imagery analysis using object detection and color-based segmentation as well as thermal imagery analysis. It is worth noting that the latest deep learning vision algorithms commonly used for object detection and segmentation (e.g. Faster RCNN, Mask RCNN, YOLO, etc.) were not adopted in this study, as they require voluminous training data and huge computational power. In this case, simplicity, interpretability, and computational efficiency were prioritized. As such, the chosen vision tools applied on digital imagery were the cascade object detector and RGB-CIE color-based segmentation. Besides, these can be complemented by a thermal imaging analysis to enhance the overall results.

A cascade object detector based on HOG features was used to detect objects having an appearance that does not change significantly (Alionte and Lazar 2015). This technique involves training the detector by creating a database containing images of relevant objects from several angles and views. The detector then uses a cascade classifier to decide whether the window sliding over a certain image contains the object of interest. In this research work, the cascade object detector is used to detect hazardous areas displaying repeating patterns.

On the other hand, the potential hazardous scenarios of fire, grinding, and welding have particular colors that differentiate them from the background. For this reason, color-based image segmentation was used to detect the presence of these hazards. The segmentation was performed using both RGB and CIE LAB color spaces. The acceptable interval of color values that represent each hazard was obtained following several trials and experiments. In other words, hazardous pixels were extracted from real positive images, and their color ranges were identified. If the number of detected hazardous pixels in any image is greater than a significant threshold specified according to the image resolution, then the developed software reveals a hazardous situation detection. This procedure was applied to fire (orange), steel grinding (bright yellow), and to steel welding (bright white).

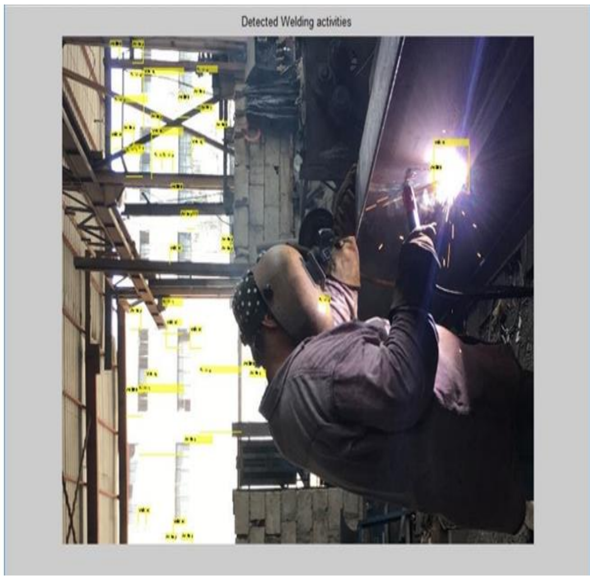
In addition, all considered hazardous scenarios are known to generate high heat when occurring. As such, analysis of thermal images captured by a heat camera was deemed necessary in order to confirm the presence and the detection of high heat hazardous scenarios. In general, heat cameras or thermal imaging cameras are devices that usually detect infrared radiation above 9,000 nanometers and create images of that radiation.

Experiments and Results

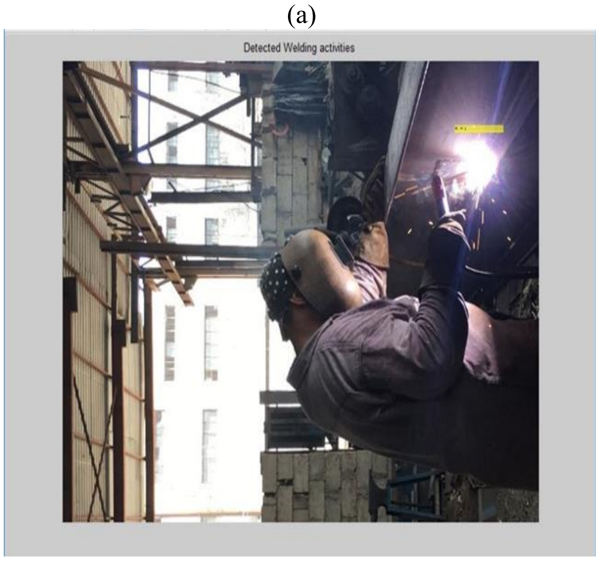
In order to test and demonstrate the feasibility and performance of the three proposed methods, numerous experiments were conducted in several indoor environments under different scenarios. In this case, standard resolution images and videos (640x480p) were captured from different construction sites and construction-like environments using digital and thermal camera-equipped drones.

The first experiment involved training a cascade object detector for the welding activity. The cascade detector is designed to detect objects according mainly to the shape, features, and the contrast difference between the object and its background. Generally, welding is characterized by an approximately circular shape and a white color but may be captured by cameras in other random shapes due to reflectance. When training the detector, it is important to include challenging negative images having shapes similar to welding, including bright windows and light reflections, to reduce the chance of false positive detections. To assess the quality of training images, two detectors were trained, whereby the first was trained with a random set containing non-challenging negative images, and the second was trained with a set of negative images captured from real construction sites and locations where welding activities may occur. Both detectors were trained with a number of stages equal to 7 and a false alarm rate equal to 0.05. Results highlighted the importance of carefully selecting the negative instances to avoid false positive detections, as tests conducted using the first detector revealed a large number of false detections captured in the background (Figure 1a). False negative detections were greatly reduced when using the second detector, as shown in Figure 1b.

One major limitation of the proposed algorithm is that, in case of the presence of objects that reflect sunlight, it may detect this reflection as a welding activity due to the resemblance of shape and shining light between the two. Furthermore, fire and steel grinding have various irregular shapes while the cascade object detector works best when used to detect objects with fixed shapes. Therefore, in order to address the aforementioned limitations, another vision tool was tested which is color-based segmentation.



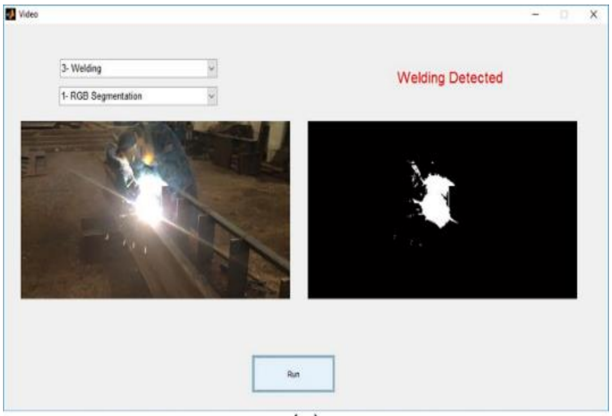
(a)



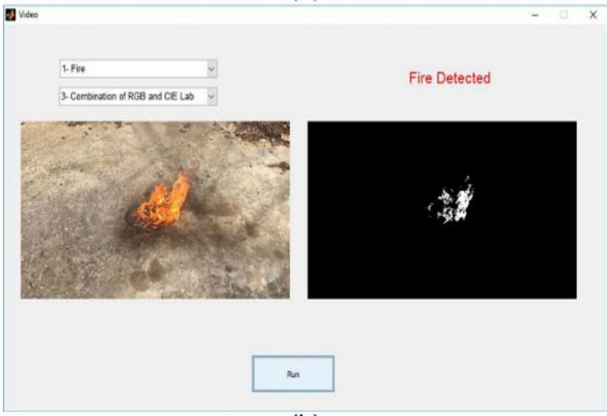
(b)

Figure 1: Cascade object detection examples: (a) welding detection using random training images and (b) welding detection using challenging training images

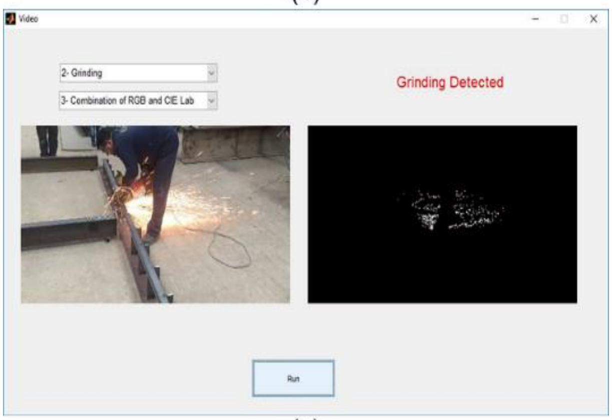
Color-based image segmentation was implemented using the RGB color space, CIE Lab color space, and a combination of both. In this case, the acceptable color ranges need to be pre-defined and calibrated according to the used digital camera for all three color segmentation tools and while considering the three hazardous situations. To test the proposed algorithms, a video-based graphical user interface was implemented whereby the user is first prompted to choose among the three hazardous scenarios then select a particular color segmentation algorithm to test. Figure 2 displays examples of true positive detections for all hazardous cases.



(a)



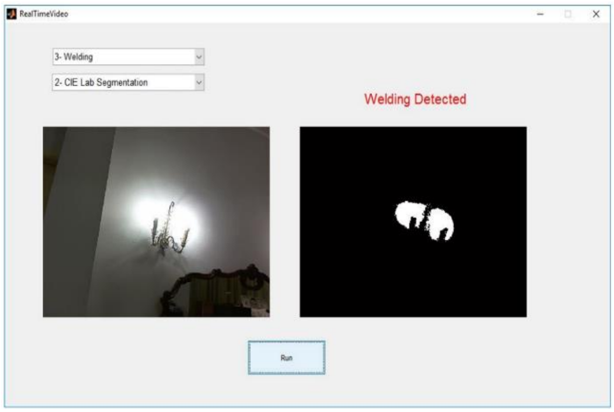
(b)



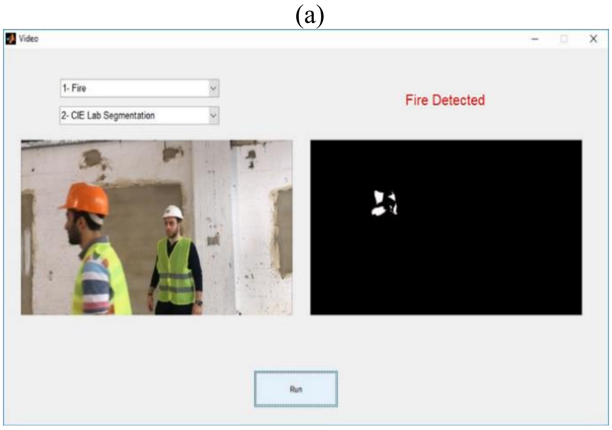
(c)

Figure 2: True positive detections using color-based segmentation: (a) welding, (b) fire, and (c) grinding

Although color-based segmentation was tested on negative images and displayed convenient true negative detections, false positive detections may occur when an object satisfying the algorithm conditions exists. For example, objects displaying a bright white color may be mistakenly classified as welding (Figure 3a) and objects displaying a bright orange color can be wrongly classified as fire (Figure 3b). The same is true for grinding whereby objects displaying a bright yellow color are falsely detected as grinding.



(a)



(b)

Figure 3: Examples of false positive detections using color-based segmentation: (a) welding and (b) fire

One way to enhance the performance of this tool is to benefit from the variation of the shape of the hazardous activities from one frame to another. The intersection of positive pixels in two consecutive frames can be obtained by multiplying their binary matrices. If the ratio of intersected positive pixels over the minimum number of detected positive pixels in both frames is less than a reasonable value, then the change of shape is considered significant and a positive detection occurs. This value was obtained following several trials performed on fire, welding, and grinding videos, and was set equal to 0.4. Experiments conducted highlighted the ability of this frames intersection method in accurately detecting hazardous situations while reducing the occurrence of false positive scenarios (Figures 4, 5, and 6).

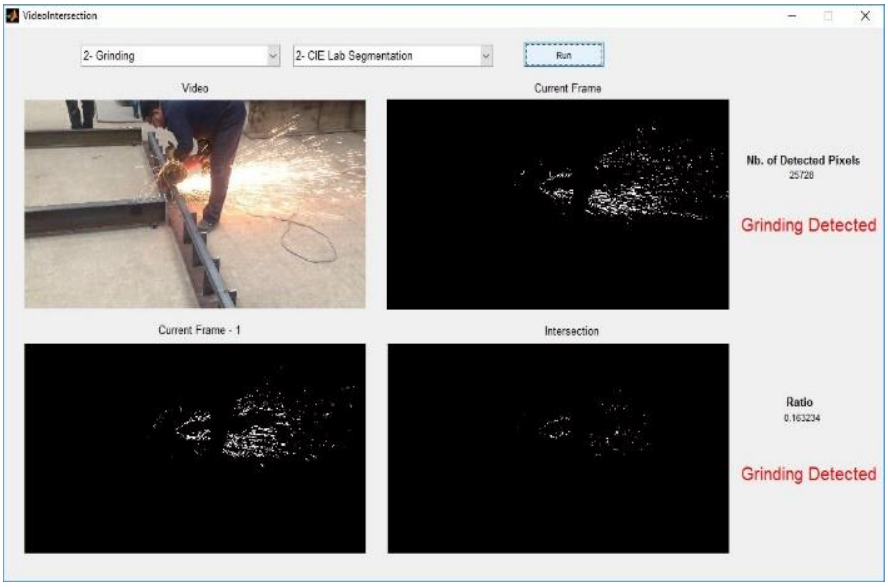


Figure 4: True positive detection of grinding using the frames intersection algorithm

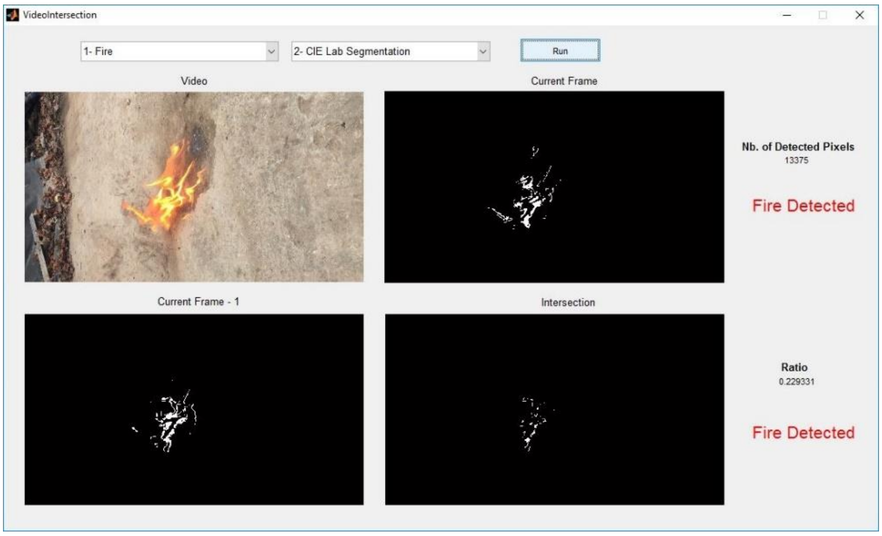


Figure 5: True positive detection of fire using the frames intersection algorithm

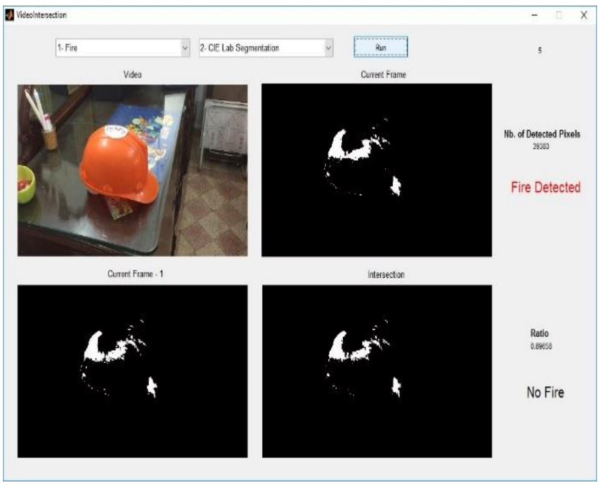


Figure 6: True negative detection of fire using the frames intersection algorithm

Accordingly, this frames intersection tool can potentially reduce false detections resulting from static objects that may have similar color properties of fire (i.e. orange hardhat placed on ground, signs and cones), and are no longer detected as fire since they do not vary remarkably from one frame to another (Figure 6). However, this tool can exhibit some limitations whenever an object having color properties that match one of the pre-defined hazardous situations is actually moving on a construction site (e.g. mobile construction worker wearing an orange hardhat). In this case, the ratio of intersection pixels between two consecutive frames might be lower than the threshold and a false detection may then occur. A false detection can also occur if the camera itself was moving. It was thereby decided to opt for thermal imaging analysis to reduce false positive detections whenever a high heat scenario is not witnessed.

As a matter of fact, the hazardous scenarios of fire, grinding, and welding are known to generate high heat when occurring. Therefore, the analysis of images captured by a thermal imaging camera may then be applied to confirm the presence and detection of high heat hazardous scenarios. Accordingly, several non-processed 14-bit TIFF still images of fire, grinding, and welding scenarios as well as random negative images were captured from several construction sites using the Flir Vue Pro heat camera. These images were then assessed using the software 'Flir ResearchIR' and processed and analyzed later on in order to detect whether the hazardous or high heat scenarios are indeed occurring or not. It was found that the maximum heat values for negative instances did not exceed a value of 9,200 nm (Figure 7), while images containing the three hazardous scenarios (i.e. fire, grinding and welding) displayed high heat values ranging between 12,000 and 16,000 nm (Figure 8).

Figure 7: Heat values for negative images

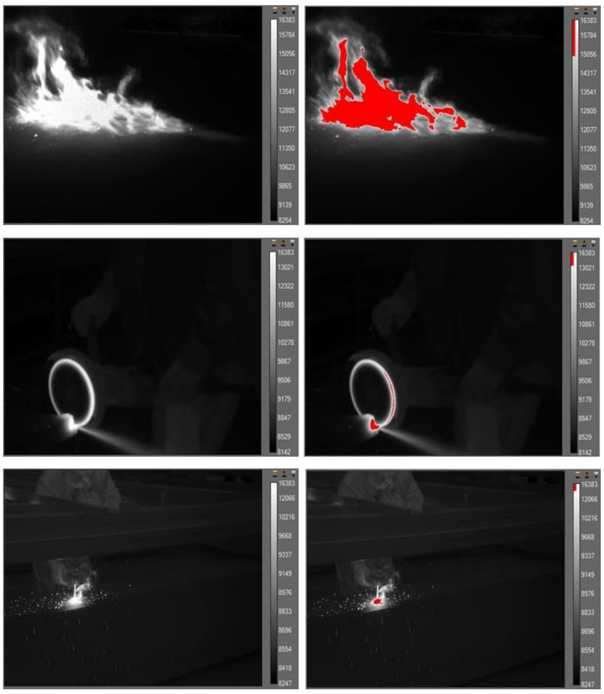


Figure 8: Heat values for positive images

Figure 9 displays an example of a fire thermal image where a high heat was detected.

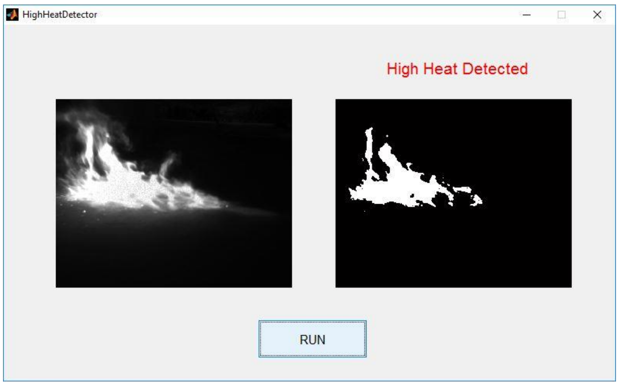


Figure 9: Detection of high heat using thermal imaging

The results in Figures 8 and 9 further highlight the potential of thermal imaging in recognizing high heat-related scenarios on jobsites and can, as such, complement the object detection or color-based segmentation algorithms and reduce false positive detections. For instance, in the case of the indoor light in Figure 3a, only a thermal analysis can confirm a non-welding case since the light with its circular shape and color properties can be mistakenly detected as welding when using the cascade object detector and color segmentation respectively. Similarly, orange objects found on construction sites (e.g. hardhat, cone, sign, etc.) may satisfy the color segmentation algorithm's conditions and can be falsely detected as fire if exposed to certain lighting situations. In this case, capturing and analyzing thermal images of these objects can greatly deny the presence of a high heat scenario and, in turn, eliminate the false positive detection resulting from the color segmentation tool. Figure 10 shows an example of a thermal image captured from a construction site that contains an orange hardhat and displays, as a result, no high heat.

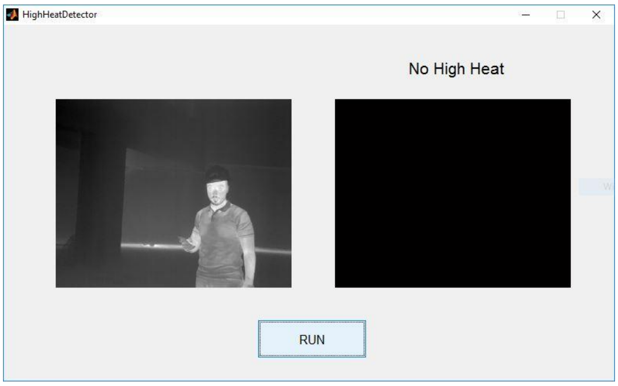


Figure 10: No detection of high heat using thermal imaging

It is worth noting that not only thermal imaging analysis is capable of complementing the two other vision tools but these latter tools can also complement the thermal imaging analysis, as it might be hard to differentiate among the three hazardous scenarios when similar heat values are registered under certain circumstances. Therefore, all three presented tools can complement each other to accurately identify and visualize each of the three hazardous scenarios. Table 1 summarizes the best use for each of the investigated methods for the hazardous scenarios.

Table 1: Summary of Hazardous situations and investigated detection methods

Hazard	Cascade Object Detector	Color Segmentation	Heat Analysis
Welding	Detection	Detection	Reduce False Positive Detections
Grinding	Not Analyzed	Detection	Reduce False Positive Detections
Fire	Not Analyzed	Detection	Reduce False Positive Detections

Conclusion and Future Work

Construction is one of the most dangerous industries whereby many hazardous tasks and conditions occur, which may pose injuries, risks and fatalities to the construction personnel. Thus, safety inspections must be carried out to maintain a safe construction environment. Many research efforts have targeted automating the safety inspection processes using digital imagery but failed to do so using thermal imagery. Hence, this study presented an automated vision-based hazardous identification system using both digital and thermal imagery. The results, obtained from several experiments conducted on construction sites, highlighted the effectiveness of analyzing both digital and thermal images to reduce false positive detections and rapidly identify on-site hazardous areas, in particular fire, grinding and welding scenarios. More specifically, the cascade object detector worked best with certain shapes. Color-based segmentation using RGB or CIE LAB color spaces, coupled with the proposed frames intersection method proved effective in detecting hazardous situations and reducing the number of negative detections. On the other hand, high heat scenarios were accurately identified using thermal imaging, thereby eliminating false positive detections when using the other vision tools.

Future work aims at enhancing and fully automating hazard identification on construction sites. This will be achieved by: (1) detecting a larger set of potential hazardous scenarios besides the three presented ones, (2) acquiring thereby a larger training data set and resorting to deep learning vision tools for analysis, (3) experimenting further with thermal imaging, (4) adjusting the implemented algorithms to read data from the construction tasks' schedule and accordingly detect the associated risks or possible hazardous scenarios, and (5) rapidly locating workers and alerting them of a nearby hazardous area, activity or scenario.

Acknowledgments

The presented work is supported by the AUB's University Research Board (URB). Any opinions, findings, conclusions, and recommendations expressed by the authors in this paper do not necessarily reflect the views of URB.

References

Abbas, M., Mneymneh, B. E., & Khoury, H. (2018) Assessing on-site construction personnel hazard perception in a Middle Eastern developing country: An interactive graphical approach. Safety Science, 103, pp.183-196.

Abbas, M., Mneymneh, B., E., and Khoury, H. (2016) Use of unmanned aerial vehicles and computer vision in construction safety inspections. Proceedings of the Third Australasia and South-East Asia Structural Engineering and Construction Conference (ASEA-SEC-3), Kuching, Sarawak, Malaysia, DOI: 10.14455/ISEC.res.2016.183.

Brilakis, I., Park, M. W., & Jog, G. (2011). Automated vision tracking of project related entities. Advanced Engineering Informatics, 25(4), pp.713-724.

Gheisari, M., Irizarry, J., & Walker, B. N. (2014) UAS4SAFETY: The potential of unmanned aerial systems for construction safety applications. In Construction Research Congress 2014: Construction in a Global Network, pp. 1801-1810.

Guo, B., Zou, Y., Fang,Y., Miang Goh, Y., Zou,P. (2021) Computer vision technologies for safety science and management in construction: A critical review and future research directions, Safety Science, 135,105130, https://doi.org/10.1016/j.ssci.2020.105130.

Ham, Y., Han, K. K., Lin, J. J., & Golparvar-Fard, M. (2016) Visual monitoring of civil infrastructure systems via camera-equipped Unmanned Aerial Vehicles (UAVs): a review of related works. Visualization in Engineering, 4(1), 1.

Han, K., Lin, J., & Golparvar-Fard, M. (2015) A formalism for utilization of autonomous vision-based systems and integrated project models for construction progress monitoring. In Proc., 2015 Conference on Autonomous and Robotic Construction of Infrastructure.

Lin, J. J., Han, K. K., & Golparvar-Fard, M. (2015) A framework for model-driven acquisition and analytics of visual data using UAVs for automated construction progress monitoring. In Computing in Civil Engineering, pp. 156-164.

Maali,O., Ko,C.H., & Nguyen,P.H.D.(2024) Applications of existing and emerging construction safety technologies, Automation in Construction, 158, 105231,https://doi.org/10.1016/j.autcon.2023.105231.

Martinez, J. G., Gheisari M., Alarcón L. F. (2020) UAV Integration in Current Construction Safety Planning and Monitoring Processes: Case Study of a High-Rise Building Construction Project in Chile. Journal of Management in Engineering, 36(3), ASCE, https://doi.org/10.1061/(ASCE)ME.1943-5479.0000761.

Melo, R.R.S, Costa, D.B., Álvares, J.S., Irizarry, J. (2017) Applicability of unmanned aerial system (UAS) for safety inspection on construction sites. Safety Science, 98, Elsevier, pp. 174-185.

Mneymneh, B. E., Abbas, M., & Khoury, H. (2019) Vision-based framework for intelligent monitoring of hardhat wearing on construction sites. Journal of Computing in Civil Engineering, 33(2), ASCE, 04018066.

Mneymneh, B. E., Abbas, M., & Khoury, H. (2018) Evaluation of computer vision techniques for automated hardhat detection in indoor construction safety applications. Frontiers of Engineering Management, doi: 10.15302/J-FEM-2018071.

Mneymneh, B. E., Abbas, M., & Khoury, H. (2017). Automated hardhat detection for construction safety applications. Procedia Engineering, 196, Elsevier, 895-902.

Mneymneh, B. E., Abbas, M., and Khoury, H. (2016). A UAV-Based Image Processing System for Identifying and Visualizing Construction Hazardous Areas. Proceedings of the 16th International Conference on Construction Applications of Virtual Reality (CONVR), Hong Kong.

Nath, N. D., Behzadan, A. H., & Paal, S. G. (2020) Deep learning for site safety: Real-time detection of personal protective equipment. Automation in Construction, 112, 103085.

Park, M. W., & Brilakis, I. (2012) Construction worker detection in video frames for initializing vision trackers. Automation in Construction, 28, pp.15-25.

Park, M. W., Elsafty, N., & Zhu, Z. (2015) Hardhat-wearing detection for enhancing on-site safety of construction workers. Journal of Construction Engineering and Management, 141(9), 04015024.

Shrestha, K., Shrestha, P. P., Bajracharya, D., & Yfantis, E. A. (2015) Hard-hat detection for construction safety visualization. Journal of Construction Engineering, 2015,https://doi.org/10.1155/2015/721380.

Yang, J., Arif, O., Vela, P. A., Teizer, J., & Shi, Z. (2010) Tracking multiple workers on construction sites using video cameras. Advanced Engineering Informatics, 24(4), pp.428-434.

Yang, X., Yu, Y., Shirowzhan, S., & Li, H. (2020) Automated PPE-Tool pair check system for construction safety using smart IoT. Journal of Building Engineering, 32, 101721.

Data Integration Methods

ADDRESSING THE CONCEPTUAL CHALLENGES IN THE PROCESS OF UPDATING THE GEOMETRY OF DIGITAL TWINS OF BUILT ASSETS

Iryna Osadcha[1], Andrius Jurelionis[1], Vishal Singh[2]
[1]Kaunas University of Technology, Kaunas, Lithuania
[2]Indian Institute of Science, Bangalore, India

Abstract

Maintaining up-to-date Digital Twin information for building or infrastructure assets is critical to its effective use. The virtual model, which reflects the real geometry of an asset, serves as an essential component of the Digital Twin and requires continuous reliability assurance and high fidelity. However, the link between physical structures and their digital counterparts remains under-researched due to the complex nature of construction and the associated technical and conceptual challenges of updating geometric data. This paper aims to identify and systematically analyse these conceptual issues in order to formulate a comprehensive framework that can be applied in practice. Design Science Research (DSR) methodology is used as a basic approach, resulting in the formulation of a theoretical framework. This framework will be further applied to specific case studies to improve the understanding of when, why, and how geometry updates affect the short- and long-term performance of the Digital Twin.

Introduction

The adoption of Digital Twins for buildings is rapidly increasing in the construction industry. Although Digital Twins are primarily used for large and complex structures, advances in data capture and modelling technologies, particularly real-time data collection, are extending the applicability of Digital Twins to a broader range of assets. Regardless of the specific application, the architecture of a Digital Twin is typically complex, involving a physical object, its corresponding virtual model, data capture and processing tools, and a continuous exchange of information between the physical entity and its virtual counterpart (Boje et al., 2020). Ideally, Digital Twin allows the state of a physical object to be tracked and managed based on its virtual model, including real-time monitoring of its condition and prediction of its future states and behaviour based on simulation at different scales and time intervals. To ensure optimal Digital Twin performance, accurate data analysis and simulation results, it is essential that the high fidelity of the Digital Twin is assured and that it contains all the necessary spatial information (Lu et al., 2020). This information does not need to be exhaustive but should be sufficient for specific needs.

Considering the evolution of the Digital Twin concept since 2003, when the vision of applying a digital model of a physical process or object was first proposed, the Digital Twin paradigm was generalised to a cyber-physical system and began to contain a virtual component (Grieves, 2003). As technologies evolved, the complexity and functionality of Digital Twin started to grow, leading to the evolution from a digital model reflecting only a physical object to a digital shadow with unidirectional automatic data flow and further evolution to Digital Twin with bidirectional automatic data flows (Fuller et al., 2020).

To date, several taxonomies describe the maturity level and the analytical or intelligent capacity of Digital Twin (Agrawal et al., 2022). By applying these taxonomies to the Digital Twin of a built asset, it is possible to define general characteristics and requirements for the components of the virtual model. For example, based on a hierarchy proposed by Gartner (Gartner, 2013) and describing different levels of digital analytics capabilities (descriptive, diagnostic, predictive, and prescriptive), at the lowest level (descriptive), building geometry data can be organised as a database or common data environment containing all relevant information such as 2D drawings, images, visualisations, component specifications, etc. starting from the design stage and further accumulated according to the life cycle stages, but without generation a comprehensive 3D model. This method of data organisation is acceptable but inefficient, as it requires extensive human involvement in data analysis and management, and significantly limits the capabilities of Digital Twin in terms of simulation and task automation.

As the level of Digital Twin automation improves, the need for a 3D model increases. However, for some tasks, such as simple visualisation, it may be sufficient to have only 2D data on the building geometry. Simultaneously, considering Digital Twin as a complex of interconnected systems that collect and analyse data on the current state of the building based on information from different sources (building technical systems, structural health monitoring systems, indoor climate control, etc.), a 3D model of the building geometry becomes not only a visualisation, but also a component involved in data analysis, simulation, and prediction. The role of the geometry of a 3D model becomes even more important in the event of a breakdown or emergency, speeding up the process of identifying the location of the problem, its repair, and the analysis of the consequences.

The geometric representation of a building is the basic component of the Digital Twin's virtual counterpart. With a high level of maturity, accuracy, and comprehensive semantic data, it becomes a valuable resource for various applications, including analysis, simulation, and forecasting. At the same time, the challenge of maintaining the geometry of a virtual building model, particularly on a global scale (not just its local elements), remains under-studied. This problem is compounded by the frequent and incorrect interchangeability of BIM and

Digital Twin terminology in the construction industry, often oversimplified to 'as-built' geometry.

Certainly, several factors contribute to the gaps in Digital Twin geometry updating research. First, the geometry of a building is typically complex, including structural and non-structural components, openings, and highly ornate elements. Additionally, in most cases, the geometry of a building remains unchanged for a long period after its construction stage, unless there are major renovations, reconstructions, emergencies affecting the geometry of the building, or the detection of deformations due to external factors. This stability means that changes to building geometry information during the operational phase are rare, reducing the attractiveness of investing in this technology. However, when updates to the geometry become necessary, the costs and technical requirements associated with this process can be significant. In addition, updating geometry is a very case-specific process, depending on the type of building, the stage of the life cycle, the required frequency of updates (regular, periodic, one-off), the source data available (relevance to the BIM model) and the type of data collected (static, dynamic, multi-source, real-time). Consequently, the approach to updating geometry will be customised to meet these specific requirements.

When the geometry update process is considered from a system thinking perspective, it becomes clear that there are conceptual challenges that go beyond the consideration of tools, equipment, and technical methods. These challenges include a deeper understanding of the geometry update process itself, its purpose, timing, and the critical elements of the building involved. It also involves recognising the impact that maintaining the current state of the geometric model has on the overall performance of the Digital Twin. This is particularly important in the context of the rapid development of advanced technologies that enable optimised and enhanced data collection and processing techniques. In this scenario, a technology-driven approach carries inherent risks, where new technology is seen as the driving force, and the challenge is to identify the appropriate problem to address (Agrawal et al., 2022). These risks include the potential disruption of existing processes in the pursuit of optimisation. There is also the possibility that optimisation will remain elusive due to fundamental constraints, resulting in the technology's potential not being fully realised. There are also risks associated with implementation costs, such as time, financial resources, and potential productivity reductions. In this scenario, the process of updating geometry should adopt a problem-solving approach that involves identifying current needs or problems and then finding appropriate technologies to address them. In certain situations, it may also involve aligning both approaches to achieve the desired outcome.

The objective of this study is to develop a comprehensive framework for identifying critical areas, referred to as 'hot spots', within the geometry update process. The framework operates at two levels: the entities that undergo update and the updating process itself. It also includes the main components involved, such as physical and virtual parts, and data. This framework aims to facilitate the analysis of hot spots during the planning phase of the Digital Twin geometry update process, with the ultimate goal of devising a more efficient set of actions. Additionally, the framework is intentionally designed to be sufficiently abstract, ensuring its flexibility and applicability across various real-world scenarios.

Background and context

In addressing the main conceptual challenges of updating the geometry of the Digital Twin, it is essential to consider the complexity of the Digital Twin, not only from a technological perspective but also in terms of conceptual complexity (Singh, 2016). This must include consideration of the diversity, fragmentation, and interdisciplinary nature of the knowledge associated with the Digital Twin. To date, there is no universally accepted definition of a Digital Twin that fits all construction projects. The concept of a Digital Twin covers a wide spectrum, ranging from a basic digital representation of a building to a sophisticated model that incorporates various engineering systems, data inputs, and predictive capabilities (Agrawal et al., 2022). Moreover, due to the complexity of real-world scenarios, often characterised by numerous unknown variables, it is reasonable to expect that a virtual model will never perfectly mirror real-world asset behaviour. To address this, it is important to assess the components of a Digital Twin, including its subcomponents and related requirements, the connection between the physical object and its virtual counterpart, as well as the supporting system that facilitates their interaction. The aim is to identify both functional and non-functional requirements as well as the desired performance metrics. As the subcomponents of the physical and virtual models require different levels of support within the system (such as software, hardware, data traffic protocols, etc.), it is worth identifying in advance generic requirements for their evaluation with further adaptation to the specific needs. The general directions outlined in Figure 1 should be considered independently of specific case applications.

The fuzziness around the concept of a Digital Twin includes various classifications related to its level of maturity, autonomy, adaptability, data and process standardization, cost-effectiveness, and the balance between computational load, execution time, and efficiency. Agrawal et al. (2023) proposed that a Digital Twin can play multiple roles, including an observer (performing tasks such as representation, visualization, description, sensing, etc.), an analyst (performing analysis, monitoring, pattern recognition, interpretation, prediction, etc.), a decision maker (engaged in optimization, simulation, forecasting, planning, etc.), and an action executor (responsible for actuation, communication, control, etc.). However, when it comes to

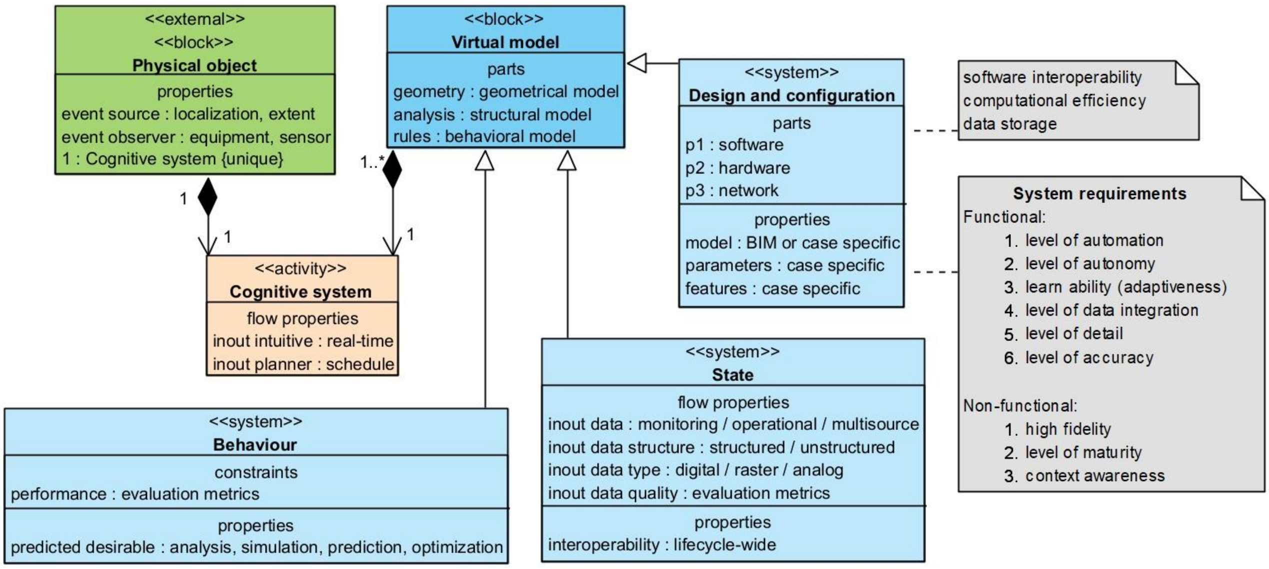


Figure 1: Digital Twin Component Model

deciding whether to update the geometry of the virtual model, it may not be wise to rely entirely on the Digital Twin itself, even though the system may typically operate at the agent level (as a decision-maker or action executor) for other tasks. Nevertheless, the Digital Twin can generate new data through analysis and simulation and serve as a valuable source of information for assessing the need to updating process (Gabor et al., 2016).

In many cases, there is no need for ongoing and constant updates to the geometry of the building model. Changes to the physical structure typically occur in specific, localised areas rather than affecting the whole building. As a result, it is not necessary to remodel the geometry of the entire building regularly, but only to address specific areas (Kaiser et al., 2022). Considering the Digital Twin as the data source in this scenario, changes in the building geometry can be analyzed in two different ways. First, point anomalies can be detected, where individual data instances are identified as deviating from their expected normal state. Second, contextual anomalies can be detected when a data instance is considered anomalous within a specific contextual scenario. In this case, the data instances under consideration could be information obtained from wireless sensor networks (WSNs) responsible for real-time monitoring of the building's condition. These sensors, such as inclinometers, accelerometers, or strain gauges, can detect and signal changes in the structures of the building in real-time. Simultaneously, for this approach to be effective, it is essential to establish the criteria for normal operating conditions that can be derived from historical or simulation data relating to the behaviour of the building over time.

Research Methodology

The Design Science Research (DSR) methodology was chosen as the basic approach for developing the framework (Singh et al., 2020). DSR is designed to methodically generate knowledge about a specific problem and its potential solutions within a research context. Figure 2 illustrates the steps of the methodology and provides a detailed description of the process. These steps include theoretical analysis techniques to identify the problem and define research objectives, followed by iterative design and development of the framework. In addition, the methodology outlines steps for future research efforts aimed at validating the developed approach and assessing its practical applicability.

During the initial phase of the theoretical research, a systematic literature review was conducted with the following objectives (Osadcha et al., 2023):

- analyse the relationship between Digital Twin, BIM, and the building lifecycle, focussing on updating building geometry data.
- examine the prevalent equipment and techniques used to collect, process, and integrate building geometry data.
- explore key areas where the geometry of building models is of significant importance, such as structural health monitoring applications, prefabricated construction, and simulations of building responses to extreme events.

Furthermore, latent semantic analysis was employed to investigate the potential integration of sensing technologies, specifically Radio-Frequency Identification (RFID), as an additional tool for capturing building geometry data in conjunction with conventional techniques such as laser scanning, photogrammetry, and regular sensors. Additionally, practices for maintaining the accuracy of model geometry in other industries and manufacturing sectors were explored. Based on the results of the theoretical research phase, the main conceptual challenge has been identified, and the current state of

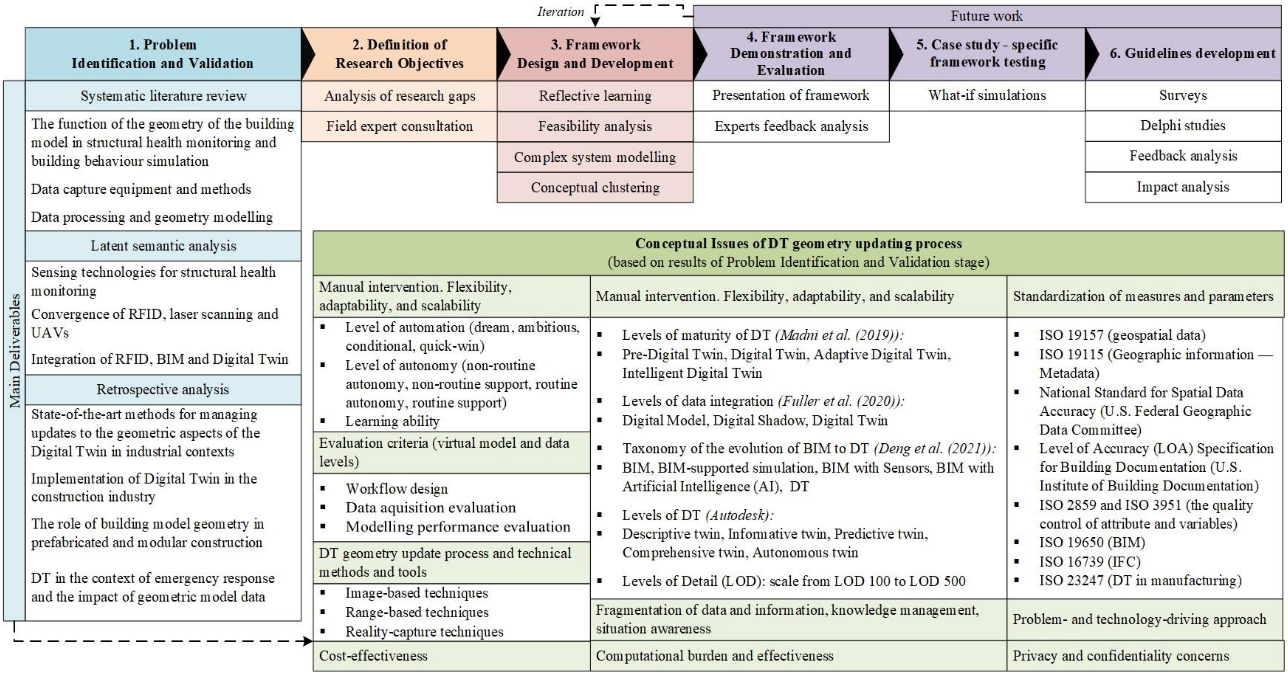


Figure 2: Research methodology stages

knowledge on these issues has been assessed to provide a basis for the development of a framework.

Results and Findings

From the standpoint of modelling cyber-physical systems, the Digital Twin of a building is a scientific model, i.e. it describes and reflects the behaviour of an existing physical object, as opposed to an engineering model, which defines what the behaviour of a physical object should be once it has been created (Lee, 2018). It is crucial to highlight that models are unable to comprehensively represent all of the properties and characteristics of complex physical objects. Therefore, the list of properties of interest must be established beforehand and assumptions about the behaviour of the model should be made. A scientific model is considered complete when it encompasses all of the properties of interest of the physical object; i.e., what is true of the physical object is also true of the model. Using this logical framework for the Digital Twin of a building, we solely focus on the virtual component of the model, particularly its geometric aspect. This means that all the geometric properties of the building should be reflected in its virtual model. To achieve this, the geometrical imperfections of the model should be reduced to a tolerable point so that the model can reflect the physical object as fully as possible while maintaining an acceptable level of computational burden. Considering that we are not taking into account the particular Digital Twin at this stage, the levels used to describe the geometric model should be sufficiently abstract to cover as many variables as possible. In this stage, the aspects related to the updating of the geometry of a virtual model are examined from two perspectives: the entity level, which deals directly with the geometric model, its sub-components and functions, and the process level, which addresses the procedures related to its updating. It is worth noting that this approach focusses only on the geometric part of the Digital Twin and its subcomponents, as well as the process of updating and maintaining the current state of geometrical data, without reference to the dependencies between the geometry of the virtual model and other Digital Twin components, data, outcomes, or services. This allows narrowing the scope of the problems to be considered and generalising these problems at the level of geometric information only, since the other components of the Digital Twin and their interaction with the geometric model can be diverse and specific to each individual building asset. Furthermore, it is advisable to extend this approach to the levels of interaction between the geometric model of the building, its data and other components of the Digital Twin, and to verify the performance of the Digital Twin as a function of the faithful geometric model.

Digital Twin geometry update at the entity level

At the entity level, the main components of the geometric model are defined, including its geometric and semantic information, as well as the relationships between these components (e.g., Figure 3). Consequently, the components of the physical building can be represented as basic geometric shapes, complex geometric shapes, or higher-order primitives, depending on the scale of the updating and the geometric parameters of the building.

Constructive Solid Geometry (CSG) and Boundary Representation (B-rep) are the primary methods used for geometric modelling. CSG is an appropriate choice when localised geometry updates are required and when the shape of the physical object is relatively simple, i.e., it can

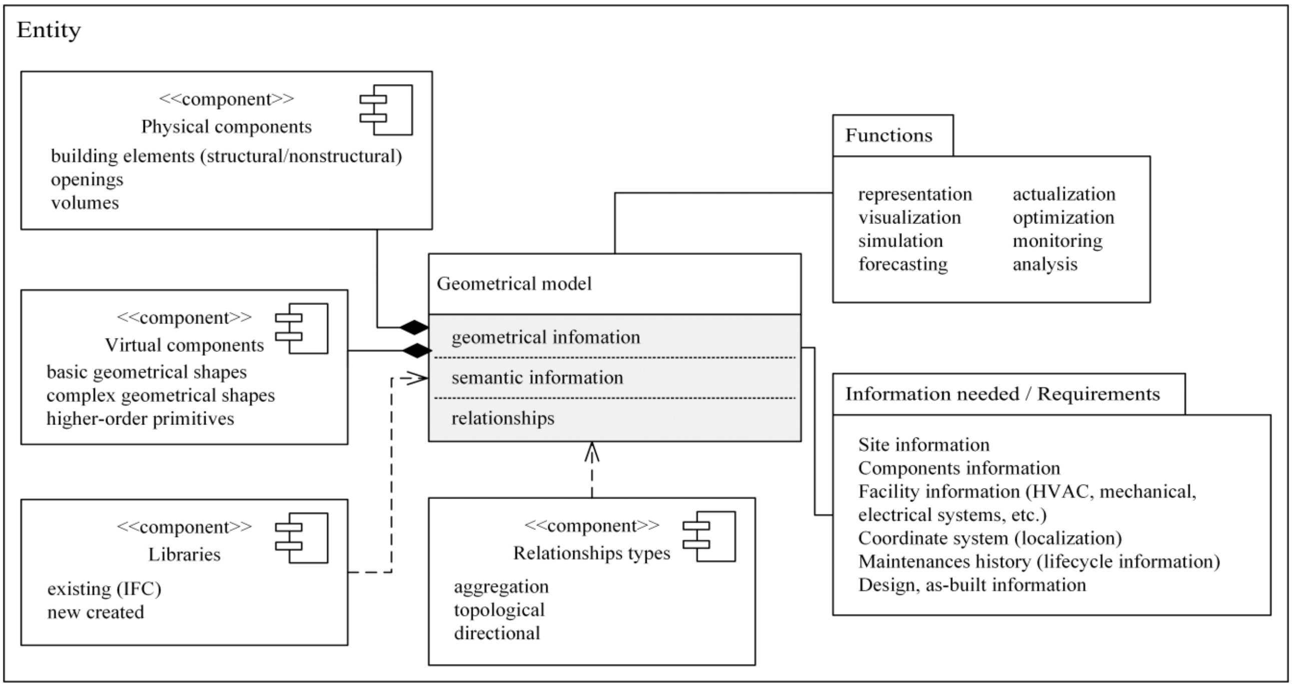


Figure 3: Geometry update process at entity level

be described using basic geometric primitives. Conversely, B-rep is a better solution for describing free-form or complex geometry. Unlike CSG, B-rep is a surface-based method that uses planar and curved surfaces to represent physical objects, making it more suitable for modelling complex architectural geometry. Although this approach is more accurate, it requires more computing power.

Semantic information in a virtual model should accurately capture the attributes associated with geometric shapes. In many cases, these attributes are determined using the Industry Foundation Classes (IFC) scheme (Masood et al., 2020). For geometric entities not covered by the IFC, it is possible to develop custom libraries or use alternative classifiers such as Unified construction classification systems (UniClass) or Construction Classification systems (OmniClass). In most cases, the BIM model serves as the basis for Digital Twin which provides support for applying IFC schema. At the same time, BIM is by nature object-oriented, and consequently, the relationships between objects are predominantly formed through parametric modeling. Non-parametric modeling techniques, such as triangular meshes, can be used as complementary tools at certain stages of the update process, particularly during point cloud processing.

Digital Twin geometry update at the process level

Figure 4 shows the main components and stages involved in the process of updating the geometry of a Digital Twin. Currently, there is no ready-to-use method for automatically updating the geometry of a Digital Twin. Although some aspects of this process can be partially automated, it remains overall complex and heavily relies on manual intervention. Building geometry data collection techniques can be categorised as either contact or non-contact. Contact techniques involve direct measurements using sensors, measuring tapes, or similar equipment. Non-contact methods include image-based techniques such as photogrammetry or videogrammetry using UAVs, range-based techniques such as laser scanning or laser measuring, and geodetic measurements using total stations, theodolites, levels, and the Global Positioning System (GPS). To achieve comprehensive coverage of building geometry data, a combination of these techniques is often necessary. To date, laser scanning and photogrammetry are the most widely used technologies, with other techniques typically serving as complementary approaches.

The data collected for updating geometry is typically in the form of a point cloud, except for WSNs and data from traditional measurement equipment. As a result, there are many techniques for the segmentation and classification of point clouds, including the most used methods such as Random Sample Consensus (RANSAC) and region-growing techniques. Point cloud segmentation and classification have traditionally relied on manual or semi-automated methods. However, there is a growing trend toward the adoption of deep learning-based approaches to automate the detection and classification of objects within point clouds, as well as the use of RFID tags to facilitate the identification of building elements within raw data.

Although deep learning methods have not yet achieved high levels of performance on large point clouds, especially compared to convolutional neural networks (CNNs) designed for image processing and applicable to photogrammetry data, there is potential for improvement. The challenges arise from limited training data and labelled datasets. However, by using synthetic data and adopting semisupervised learning architectures, these methods are likely to advance soon. In the context of

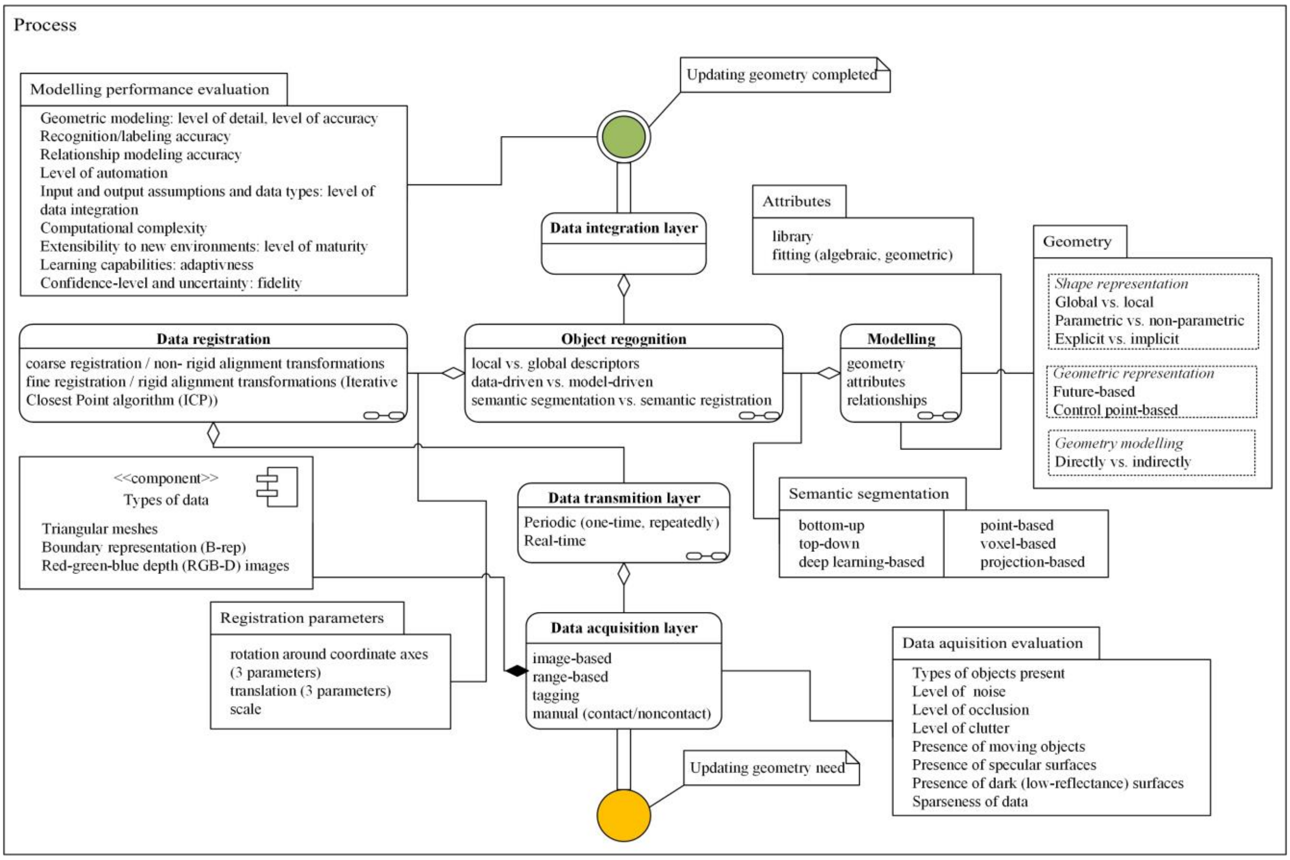


Figure 4: Process-level building geometry update

updating geometry, it is important to acknowledge that existing data processing methods are evolving rapidly, particularly with the increasing prevalence of deep learning techniques. This evolution will lead to greater automation in the short term, even for real-time data scenarios. Therefore, during the framework design the focus was shifted away from specific current technical methods and moved to the broader aspects of the process and the underlying patterns.

To develop a framework for updating the geometry of a Digital Twin several aspects were considered. This includes technical and conceptual aspects of the process, extending to the level of component entities and data (Hribernik et al., 2021). Based on theoretical research outcomes, a comprehensive roadmap for updating the geometry of a Digital Twin has been developed (Figure 5). This roadmap addresses critical research questions, including the reasons for updating, the timing of updates, and the methods involved in the updating process. It also involves identifying the current state-of-the-art, distinguishing between problems and technology-driven approaches, recognising potential knowledge gaps and associated challenges, and establishing a future research objective within this context.

Discussion and Future Work

When dealing with a complex cyber-physical system such as a building and its Digital Twin, predicting the system's behaviour when one of its components is modified can be challenging. Therefore, it is essential to assess the potential impact of changes on overall performance. These changes can occur at several levels:

1. Physical level:
   - Changes in the physical geometry of the building that are not reflected in the Digital Twin.
   - Discrepancies in the geometry of the Digital Twin, including noise, redundancy, and temporal elements, are not presented in the actual building.
2. Digital Twin data level:
   - Changes (systematic errors) in WSNs or other types of data.
   - Changes in computational capabilities from the hardware perspective.
3. Digital Twin virtual model level:
   - Immediate response: real-time updates of semantic information based on real-time data.
   - Planned response: scheduled update geometric information using 3D data capture techniques.
   - Validation: assessing how changes to the model's geometry affect other Digital Twin components, such as potential collisions with other systems (HVAC, electrical, plumbing, etc.).

Furthermore, when considering the process of updating geometry from a more systematic perspective, it is important to recognise that certain solutions aimed at maintaining data integrity may require significant

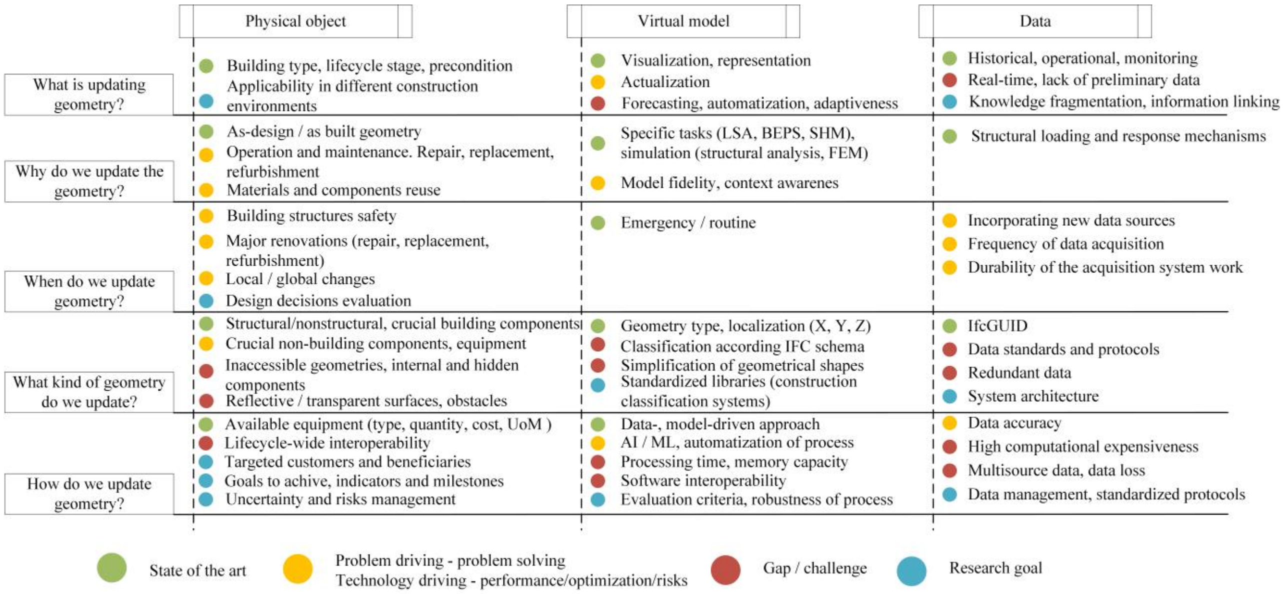


Figure 5: The roadmap for the process of updating the geometry of the Digital Twin and future research directions

investment in terms of cost, time, and effort. In other words, why do we update the geometry and what is the cost of this process in terms of benefits and risks? This requires a methodical approach to the implementation and execution of these solutions (Singh, 2016). For example, some cases may require periodic updates of geometry according to a predetermined schedule (buildings under environmental or man-made influences, unique structures (supertall, asymmetric, free-form buildings, heritage objects, etc.)). On the other hand, there are situations where a one-off effort is sufficient, such as updating the geometry in response to unexpected events that affect the physical structure. Moreover, it is crucial to acknowledge that certain decisions or actions may be irreversible and have both positive and negative effects on system performance. In such cases, alternative scenarios, including the consequences of not updating the geometry at that time, need to be considered. Developing a robust risk management system, such as a data integrity backup solution, can also be a valuable strategy in these scenarios (Reinbold et al., 2022).

Another aspect that deserves attention is the conceptualisation of the Digital Twin not as a static entity, but rather as a dynamic collection of interconnected systems that are continuously evolving. Typically, a combination of approaches is used, including both system- or model-based methods and data-driven approaches (Singh et al., 2020). This integration of approaches involves using knowledge from the original model, its constituent parts, and their technical specifications (e.g., building engineering systems specifications) and operational data generated throughout the lifecycle of the building. In this context, it is crucial to analyse the impact of planned interventions, such as the integration of new geometrical data, to prevent the disruption of existing processes and non-value-added activities that may result from a lack of awareness of the current situation or poor planning. To address these challenges, based on the analysis of the hotspots outlined in the roadmap, critical areas that require focused attention for further research can be summarised as follows:

- Identification of key scenarios for implementing the geometry update process.
- Investigation of the evolution of Digital Twin throughout the building life span and the role of influencing the geometry of the virtual model during this process.
- Addressing data interoperability issues both at the software level and in the broader context of lifecycle-wide interoperability of building information.

Conclusions

The concept of a Digital Twin is constantly evolving, and its role in the construction industry will grow significantly in the coming years. Advances in technology and the need to improve sustainability in construction will drive further digitisation. Maintaining the accuracy of a Digital Twin's geometry is a critical aspect of this evolution. It serves not only as a visual representation of a building but also as a valuable source of information for various tasks such as Building Energy Performance Simulation (BEPS) and Life Cycle Assessment (LSA). These types of analysis rely heavily on building geometry data. Without established frameworks or systematic solutions, the process can become intuitive and rely on subjective judgment. An incomplete evaluation of this process can result in missed opportunities to enhance the capabilities of the Digital Twin. Therefore, it is essential to develop a methodological approach to ensure that the geometry of the Digital Twin remains up-to-date.

Most existing publications address the geometry update process primarily from a technical perspective, discussing various data collection and processing options. However, the geometry update process is highly case-specific, which complicates the search for a universal solution, and this is one of the limitations of the current work. It is not possible at this stage to consider all the potential factors that may arise during the process. However, an analysis of Digital Twin update geometry from two angles, entity/process levels, and main components (physical, virtual, data), has resulted in the development of a roadmap that provides insights into the current landscape, identifies potential challenges, and outlines future research directions. Even at this stage, the roadmap can serve as a source of information for planning and implementing the process of updating the geometry of the Digital Twin. It provides a basis for understanding the underlying complexities and issues that may arise during the process and need to be addressed.

Acknowledgments

This study is part of the dissemination activities of the research project 'Boosting Research for a Smart and Carbon Neutral Built Environment with Digital Twins (SmartWins)' (Grant ID Number 101078997), funded under the Horizon Europe call HORIZON-WIDERA-2021-ACCESS-03-01.

References

Agrawal, A., Fischer, M., Singh, V., 2022. Digital Twin: From Concept to Practice. Journal of Management in Engineering 38, 06022001. https://doi.org/10.1061/(ASCE)ME.1943-5479.0001034

Agrawal, A., Thiel, R., Jain, P., Singh, V., Fischer, M., 2023. Digital Twin: Where do humans fit in? Automation in Construction 148, 104749. https://doi.org/10.1016/j.autcon.2023.104749

Agrawal, A., Singh, V., Fischer, M., 2022. A New Perspective on Digital Twins: Imparting Intelligence and Agency to Entities. IEEE Journal of Radio Frequency Identification 6, 871–875. https://doi.org/10.1109/JRFID.2022.3225741

Boje, C., Guerriero, A., Kubicki, S., Rezgui, Y., 2020. Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction 114, 103179. https://doi.org/10.1016/j.autcon.2020.103179

Fuller, A., Fan, Z., Day, C., Barlow, C., 2020. Digital Twin: Enabling Technologies, Challenges and Open Research. IEEE Access 8, 108952–108971. https://doi.org/10.1109/ACCESS.2020.2998358

Gabor, T., Belzner, L., Kiermeier, M., Beck, M.T., Neitz, A., 2016. A Simulation-Based Architecture for Smart Cyber-Physical Systems, in: 2016 IEEE International Conference on Autonomic Computing (ICAC), pp. 374–379. https://doi.org/10.1109/ICAC.2016.29

Gartner. 2013. "Extend your portfolio of analytics capabilities." Accessed September 16, 2021. https://www.gartner.com/en/documents/2594822 /extend-your-portfolio-of-analytics-capabilities.

Grieves M., 2003. PLM-beyond lean manufacturing. Manufacturing Engineering 130, 23

Hribernik, K., Cabri, G., Mandreoli, F., Mentzas, G., 2021. Autonomous, context-aware, adaptive Digital Twins—State of the art and roadmap. Computers in Industry 133, 103508. https://doi.org/10.1016/j.compind.2021.103508

Kaiser, T., Clemen, C., Maas, H.-G., 2022. Automatic co-registration of photogrammetric point clouds with digital building models. Automation in Construction 134, 104098. https://doi.org/10.1016/j.autcon.2021.104098

Lee E.A., 2018. Modeling in engineering and science. Communications of the ACM 62, 35-35. https://dl.acm.org/doi/pdf/10.1145/3231590

Lu, Q., Xie, X., Parlikad, A.K., Schooling, J.M., 2020. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. Automation in Construction 118, 103277. https://doi.org/10.1016/j.autcon.2020.103277

Masood, M.K., Aikala, A., Seppänen, O., Singh, V., 2020. Multi-Building Extraction and Alignment for As-Built Point Clouds: A Case Study With Crane Cameras. Frontiers in Built Environment 6.

I. Osadcha, A. Jurelionis, and P. Fokaides, 'Geometric parameter updating in digital twin of built assets: A systematic literature review', Journal of Building Engineering, vol. 73, p. 106704, Aug. 2023, doi: 10.1016/j.jobe.2023.106704.

Reinbold, A., Lappalainen, E., Seppänen, O., Peltokorpi, A., Singh, V., 2022. Current Challenges in the Adoption of Digital Visual Management at Construction Sites: Exploratory Case Studies. Sustainability 14, 14395. https://doi.org/10.3390/su142114395

Singh, V., 2016. BIM Ecosystem Research: What, Why and How? Framing the Directions for a Holistic View of BIM, in: Harik, R., Rivest, L., Bernard, A., Eynard, B., Bouras, A. (Eds.), Product Lifecycle Management for Digital Transformation of Industries, IFIP Advances in Information and Communication Technology. Springer International Publishing, Cham, pp. 433–442. https://doi.org/10.1007/978-3-319-54660-5_39

Singh, V., Mirzaeifar, S., 2020. Assessing transactions of distributed knowledge resources in modern construction projects – A transactive memory approach. Automation in Construction 120, 103386. https://doi.org/10.1016/j.autcon.2020.103386

LEVERAGING CLASSIFICATION KNOWLEDGE FOR IMPROVED DATA ACCESSIBILITY IN DIGITAL CONSTRUCTION

Jakob Martin[1,2], Klaus Linhard[2], Cornelius Preidel[1,2]
[1]Munich University of Applied Sciences, Germany
[2]Institute of Applied Building Informatics (iabi), Munich, Germany

Abstract

While the prevalent use of complex structures such as Industry Foundation Classes (IFC) in digital building models has good reasons, it also limits the easy accessibility of data and hinders a common understanding of building structures across different projects or companies. This paper presents an approach to how existing classification systems can effectively improve common understanding by representing and collecting classification rules in the machine and human-readable IDS format of buildingSMART. The formally mapped classification rationale serves as a common starting point for stakeholders, capturing and consolidating classification knowledge from project or company-specific models and breaking down technical barriers to using digital models. This initiative contributes to a more collaborative and standardized digital construction industry and promotes better data understanding and knowledge sharing between stakeholders.

Introduction

The digitization wave sweeping across the construction industry necessitates the adoption of highly structured and complex data formats like the Industry Foundation Classes (IFC) to encapsulate the rich, multi-dimensional information inherent in construction projects. (Zhang et al., 2014) However, a significant barrier arises as stakeholders, particularly those with varied disciplinary backgrounds, grapple with these complex data structures. The pressing need is to ensure the most straightforward entry point for all stakeholders into the digital realm, promoting wider adoption and effective utilization of digital methods. (Opoku et al., 2023)

Classification systems stand as a pivotal element in bridging this complexity gap. Essentially, classifications are systematic arrangements in defined categories based on shared characteristics, as ISO 14177 (International Organization for Standardization, 1994) outlines. These systems, utilized across diverse application areas and disciplines within the construction industry, provide a familiar framework for engineers working with them for decades. The essence of classification lies in its ability to streamline information retrieval, enhance communication, and promote standardization, which is fundamental in handling complex data structures inherent in digital construction models.

State of the Art

The evolution of digital construction methodologies, particularly Building Information Modeling (BIM), has catapulted the significance of classifications to a new level. (Wu and Zhang, 2019) BIM, with its ability to create rich, multi-dimensional digital representations of physical and functional characteristics of assets, necessitates a structured approach to information management. Here, classification systems serve as the backbone for organizing this voluminous data in a structured, accessible, and interoperable manner (Beetz, 2018). In the digital realm, classifications have evolved to serve as a categorization tool and a medium to enhance information flow, accessibility, and collaborative engagement. They are now employed to structure the data within digital models, facilitating more straightforward navigation, interpretation, and utilization of the information. For instance, in terms of model structuring, they aid in structuring digital models. This enables stakeholders to navigate and interact with the model contents through a familiar classification view and simplifies the engagement with complex data structures like the Industry Foundation Classes (IFC). In the context of information retrieval, classifications speed up the retrieval of pertinent information. Organizing data systematically and predictably becomes crucial for decision-making and coordination among various project stakeholders. When considering interoperability, classifications enhance this by offering a standardized framework. This facilitates efficient information exchange amongst numerous software applications and platforms typically employed in a project. Lastly, there's a shift in modern digital construction methodologies towards knowledge capture. More than ever, there's a focus on capturing and formalizing the knowledge of how components are classified. This pivotal change facilitates a shared understanding and consensus on classification standards throughout the industry.

In practice, there are two different ways of classifying components of a digital construction model: (1) Rule-based Classification and (2) Attribute-based Classification.

Attribute-based Classification describes a process in which the model authors manually add attributes to the components to assign the component according to a classification. This pragmatic approach is confirmed by its implementation in industry practices and standards, such as the guidelines of BIM Deutschland (BIM Deutschland, 2024). This initiative makes key recommendations for the use of BIM in Germany. It emphasizes the need for defined attributes to be added and maintained across all service phases for Level of Information (LoI), exemplified by the requirement to add attributes for classification according to DIN 276 manually. This customization provides flexibility in classification as attributes can be tailored to suit specific project needs. It also facilitates intuitive classification based on observable or measurable characteristics. Despite its practical application and advantages, this method involves significant manual effort, which can be cumbersome, time-consuming, and prone to errors. Additionally, a coherent understanding of the classification scheme among all stakeholders is required to ensure consistency. The manual addition of attributes leaves the underlying rationale for classification with the human operator devoid of formalization. This tacit knowledge remains unshared and unstandardized across the industry, highlighting the necessity of initiatives like BIM Deutschland in guiding the industry toward standardized practices.

Rule-based Classification is a methodology that leverages static filter-like rules within native software tools. Users implement rule logic to filter components based on inherent information. This mechanism automates the classification process, ensuring consistency and efficiency while reducing the manual effort required. However, it presents limited flexibility in adapting to varying classification needs, and the logic may become complex and challenging to manage with evolving project requirements. (Bloch and Sacks, 2018) A notable shortfall of this approach is the lack of formalization of the underlying rationale for classification, making it difficult to capture and share classification knowledge across the industry. This mechanism can also be used to harmonize different types of content from other authoring tools. For example, exterior walls that have different characteristics (e.g., name or specific property) in various models can be identified as such. Prominent examples of rule-based classification in practice are seen in tools like *Solibri Office* and *BIMcollab Zoom* (Solibri, 2023; BIMcollab, 2023) (see Figure 2). In *Solibri Office*, classifications are facilitated through predefined rules, allowing users to filter and organize model components efficiently. Similarly, Smart Views in *BIMcollab Zoom* enable users to create custom views based on rule logic, aiding in the classification and visualization of model components. Figure 1 shows how users can define these rule-based classifications within *Solibri Office*. These examples showcase the practical utility of rule-based classification but also underline the inherent limitations in flexibility and formalization of classification rationale, underscoring the necessity for exploring alternative or supplementary methods to encapsulate classification knowledge in a more standardized, shareable format.

When examining the two prevalent classification methods, each brings distinct advantages. Rule-based classification stands out for its automation and consistency, offering a quicker and more standardized way of categorizing items. Conversely, attribute-based classification shines in its flexibility and intuitiveness, enabling more customized classifications.

Delving into the disadvantages, rule-based classification reveals certain limitations. Its adaptability to varying classification needs is initially restricted, posing a challenge in dynamic projects. As these projects evolve, the rule logic might become complex, cumbersome to manage, and potentially decelerating the classification process. Most notably, the absence of clarity in the rationale behind classifications obstructs the sharing and standardization of classification logic across different projects or teams, impeding collaborative endeavors and industry-wide learning.

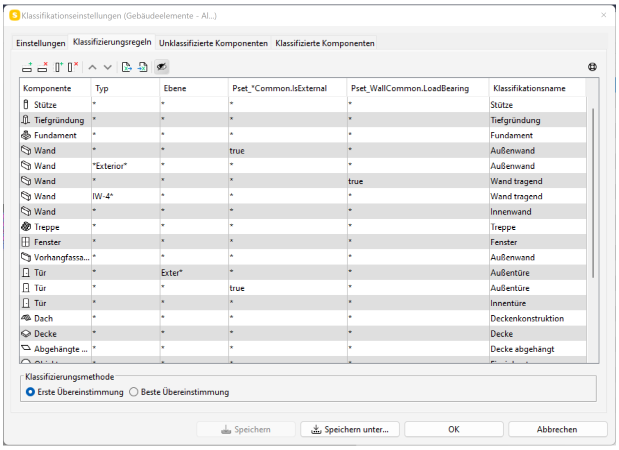


Figure 1: User interface in Solibri Office, in which a user can build up classification rules line by line

On the flip side, attribute-based classification carries its own set of hurdles. It demands manual effort in classifying items, which can be time-intensive and error-prone, particularly in large-scale projects. Additionally, a coherent understanding of the classification scheme among all stakeholders is requisite to maintain consistency, a demand that could be challenging in diverse teams or multi-stakeholder projects. Like its rule-based counterpart, attribute-based classification falls short in elucidating the reasoning behind classifications, leaving the understanding of classification logic unshared and unstandardized across the industry. This shortfall again hinders knowledge sharing and collaborative efforts.

Both methods share a common pitfall: the ambiguity in the reasoning behind classifications. This ambiguity stifles the sharing and understanding of classification logic across the industry, pinpointing an area ripe for further exploration.

In summation, while each classification method boasts its unique strengths, they both encounter significant challenges, chiefly in elucidating the reasoning behind classifications to foster knowledge sharing and industry-wide standardization. This exposition underscores the need for a refined classification approach that combines the strengths of rule-based and attribute-based classifications while alleviating their weaknesses, paving the way for enhanced sharing and understanding across the digital construction domain.

Approach

The potential benefits of formalizing this classification knowledge in a shareable, neutral format are manifold.

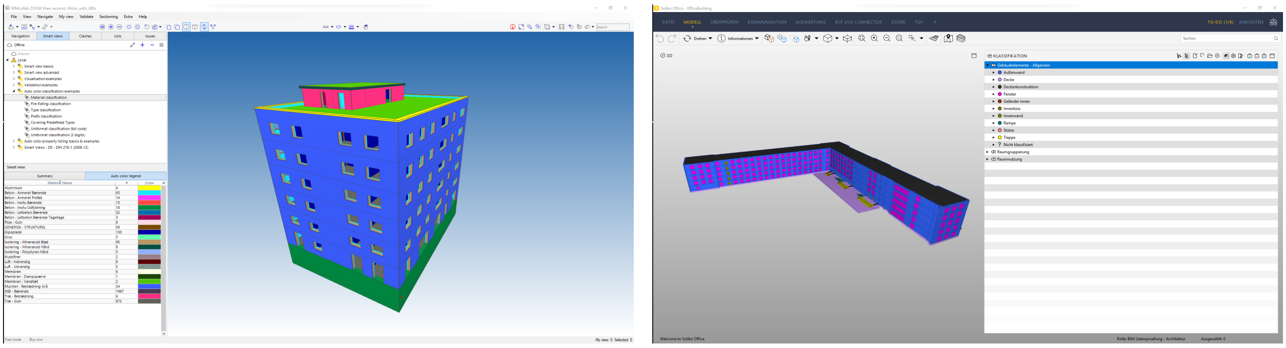


Figure 2: Showing Smart Views in BIMcollab Zoom (left) and Classifications in Solibri (right)

Firstly, it provides stakeholders with valuable insights, enabling a more comprehensive understanding and agreement on classifying components. This mutual understanding is crucial for harmonizing different stakeholders' perspectives. Additionally, a significant benefit is the reduction of manual work. By utilizing existing information about components for automated classification, the process becomes more efficient and consistent, minimizing the chances of errors that come with manual classification. A key advantage is the encapsulation of rule logic into a standardized format. This encapsulation creates a common language that explains the reasoning behind each classification, making it much easier to manage and share across various projects and teams. Overcoming the typical constraints associated with traditional classification methods, this approach fosters better collaborative engagement. A standardized format serves as a common ground, encouraging dialogue and cooperation among stakeholders, which is vital for the success of digital construction projects. Moreover, streamlined data management is a significant outcome of this approach. It simplifies the classification process and creates a more organized digital environment, making navigating and interacting with digital models easier. Often, the rationale behind a component's classification is inherently depicted by the existing information about the component. The prevalent constraints can be overcome by identifying and encapsulating this rule logic into a standardized format, paving the way for enhanced collaborative engagement and streamlined data management in digital construction. By extracting and formalizing this rule logic into a standardized form, we can transcend the current limitations, unlocking a new horizon of collaborative engagement and streamlined data management in digital construction. Through this endeavor, we aspire to intertwine the existing classification systems with complex data structures, simplifying the digital landscape for all stakeholders and nurturing a more collaborative, efficient, and user-centric digital construction ecosystem.

Standardized Formats for Formalizing Classification Knowledge

It is worth investigating existing standards that can efficiently encapsulate and share classification rule logic to formalize classification knowledge within the digital construction domain. Among the many available formats, mvdXML and Information Delivery Specification (IDS), both by buildingSMART International, emerge as suitable candidates owing to their structured frameworks and widespread recognition in the industry (Tomczak et al., 2022). A summary of the different pros and cons for our use case is shown in Table 1.

mvdXML (buildingSMART, 2016) is a technical schema accompanying Model View Definition (MVD) to formalize and validate data exchanges in construction projects. MVD is used to specify subsets of the Industry Foundation Classes (IFC) schema to facilitate interoperability in specific use cases. It can also be leveraged to formalize classification knowledge, aligning it with established industry standards. mvdXML provides a structured framework with an added validation layer, ensuring the encapsulated classification rule logic adheres to the predefined standards. Employing mvdXML to formalize classification knowledge could align the classification logic with established industry standards, promoting consistency and interoperability. Furthermore, the validation feature of mvdXML could be instrumental in maintaining the integrity of classification logic over time.

Information Delivery Specification (IDS) (buildingSMART International, 2023a) is a relatively newly introduced schema specifying a construction project's information delivery requirements. It serves as a blueprint, outlining the nature, format, and extent of information to be exchanged amongst stakeholders at various stages of a project lifecycle. By doing so, IDS augments the clarity, consistency, and efficiency of information exchange, minimizing misunderstandings and errors that could potentially derail a project.

The primary intent of IDS is to facilitate the precise articulation and fulfillment of information requirements. It acts as a bridge, ensuring that the information generated and consumed across different phases of a construction project is aligned with the defined standards, thereby promoting interoperability and collaborative engagement.

Suggested Methodology

mvdXML offers a broad framework for formalizing and validating data exchanges in construction projects across

Table 1: Comparison of mvdXML and IDS for Formalizing Classification Knowledge

	mvdXML	**IDS**
Pros	+ Powerful, structured framework for formalizing data exchanges. + Supports detailed specification of data subsets in IFC for interoperability. + Built-in validation ensuring adherence to industry standards. + Potential for aligning classification logic with established standards.	+ Simplified and straightforward encapsulation of classification rule logic. + Quickly adopted standard by buildingSMART, ensuring industry recognition. + Compatible with existing IDS editors for authoring and editing. + Facilitates interactive exploration and application of classification logic. + Supports regular expressions (RegEx) for the specification of values
Cons	– Complexity and detailed structure may challenge users. – Many features may be redundant for classification, adding to the complexity. – Low market penetration given that it was introduced years ago.	– Lacks built-in validation mechanism. – May require additional effort to capture more complex classification logic. – No existing tools for interpreting IDS in the proposed manner (before demonstrator development).

various use cases of the IFC schema. In contrast, the Information Delivery Specification (IDS) is limited to specific aspects of mvdXML, focusing on clearly articulating information delivery requirements throughout a construction project's lifecycle. While mvdXML is versatile enough to encompass the functionalities of IDS, the latter's targeted approach promises to be more user-friendly and streamlined for its intended purpose.

Despite its robustness and comprehensive technical schema, the mvdXML standard has specific challenges that could hinder its prompt adoption of our suggested approach. A primary concern is its current market penetration, which is not as widespread as desired, potentially limiting its immediate usability and acceptance for formalizing classification logic. Furthermore, the high complexity of mvdXML poses a steep learning curve for stakeholders, especially those new to digital construction standards, which could deter them from adopting this standard for classification logic formalization. On the other hand, the IDS standard, with its simplified approach, presents a lower barrier to entry, making it a more attractive option. Despite these concerns, the potential of mvdXML in this domain is acknowledged, and a comparative analysis with IDS in future studies could provide a nuanced understanding of their respective merits and limitations in formalizing classification knowledge.

Our proposed method utilizes the Information Delivery Specification (IDS) format without altering or extending its existing structure. The key idea is to employ a slightly different interpretation of certain aspects of the IDS for a distinct, constructive purpose.

The IDS file format, grounded in a standardized XML schema, is a robust mechanism for detailing information requirements for elements within Industry Foundation Classes (IFC) files. An IDS file comprises two main sections: a Header containing general metadata about the file and a list of Specifications detailing the information requirements for IFC elements.

Each Specification within an IDS file is divided into three components: Metadata, Applicability, and Requirements. Metadata provides contextual information about the Specification, Applicability defines the scope of elements the Specification applies to, and Requirements outline the actual information requirements for the objects in question.

Both the Applicability and Requirements components use a mechanism known as Facets to specify their content. In the context of IDS, Facets describes potential information an element in the IFC model might have. Six precisely defined Facet Parameters are used to make these requirements machine-interpretable, namely:

- Entity Facet
- Attribute Facet
- Classification Facet
- Property Facet
- Material Facet
- PartOf Facet

Traditionally, the IDS specification is divided into two main parts: the 'Applicability' part, which outlines the specific components the specification applies to, and the 'Requirement' part, which outlines the information requirements for those components. Our methodology retains the conventional use of the 'Applicability' part but adopts a different interpretation of the 'Requirement' part. Instead of seeing the facets defined in the 'Requirement' part as information requirements, we suggest interpreting them as classification logic. This subtle shift in interpretation allows an IDS specification to be seen as one or many applicable classification rules. These rules, encapsulated in the 'Requirement' part, apply to the components specified in the 'Applicability' part.

Figure 3 shows an example of this. In this example, the characteristics for identifying load-bearing external walls are defined in the 'Applicability' part, and then two classification requirements are listed in the 'Requirement' part. According to the introduced logic, all identified building components in a model should now be classified as *"buitenwanden; niet constructief, massieve wanden"* according to the Dutch NL-SfB 2005 and as *"332 - Nichttragende Außenwände"* according to the German DIN 276 (BIM Loket, 2023; DIN, 2018).Since an entry is stored in the buildingSMART Data Dictionary (bsDD) (buildingSMART International, 2023b) for the Dutch classification, a corresponding URI can also be stored here, which links the resulting classification to the entry.

This novel methodology facilitates the systematic classification of components based on established logic, nurturing a more structured and insightful representation of model data. The transparency of the classification logic further improves understanding and consensus among different stakeholders, fostering a more collaborative and efficient digital construction ecosystem.

Utilizing the existing IDS format in such an innovative way aligns with the industry's ethos of leveraging established standards to foster interoperability and knowledge sharing while paving the way for enhancing the practicality and usability of digital models in construction projects.

Results & Test Cases

This section examines the practical application and feasibility of the proposed methodology using selected test scenarios. The aim is to gain insights into how the approach performs when applied to real cases. For this preliminary investigation, we have focused on the example of the national classification system DIN 276 to assess its applicability and robustness. This decision marks our first effort to validate the proposed method and illustrate its feasibility straightforwardly.

IDS Authoring & Demonstrator

Creating IDS files is a crucial part of this process, and having accessible tools significantly aids in this endeavor. One significant advantage of authoring the IDS according to the suggested approach is the compatibility with existing IDS editors, whether commercial or open-source. As the IDS specification is not changed or expended, any compatible IDS editor can create, author, or edit the suggested classification logic.

For this research, the open-source *xBIM IDS Editor* (Benghi, 2023) was used to create the IDS files.

Given the absence of existing tools capable of interpreting the IDS as per our suggested methodology, a demonstrator was developed utilizing *ifcopenshell* (Krijnen, 2023) and the *IFC.js* (Viegas, 2023) viewer components. This demonstrator allows stakeholders to interact with IDS files alongside IFC files. In this setup, users can conveniently load IFC files with one or multiple IDS files, and the encapsulated classification logic is automatically applied. Consequently, users are presented with the classification systems and the corresponding classification items, alongside the count of components identified as classified components according to the rules contained. Users can interactively navigate the model by clicking and highlighting the objects based on their classification.

This demonstrator serves as a practical illustration of the utility and effectiveness of our proposed approach, facilitating an interactive exploration of classification systems and the underlying classification logic within the digital construction models. In its current form, the viewer implemented acts more like a substitute, given the lack of existing implementations capable of interpreting the IDS as per our methodology. However, a far more impactful use case for the resulting IDS would be its integration into any platform users utilize to visualize, create, or edit digital building models. Such integration would significantly enhance the user experience by allowing the contents to be immediately structured according to the classification logic represented in the IDS files. This showcases the potential of integrating the existing IDS format for classification purposes and demonstrates a path toward fostering a more collaborative and insightful interaction with digital construction data among stakeholders. Through this hands-on evaluation, the proposed methodology has showcased promise in bridging the complexity associated with digital construction models, paving the way for broader adoption and effective utilization of digital methods in the construction industry. The resulting demonstrator is shown in Figure 4.

At this point, it should be added that the resulting classification results can also be fed back into an IFC model, as the IFC schema provides explicitly for this. However, it should be noted that this content only reflects the classification results but not the classification logic itself. If there are changes in the IFC model, the logic stored in the IDS would have to be applied again, and the classification contents stored within the IFC model would have to be updated. The resulting classification data can also be added to the IFC file on demand as part of the prototype development. For example, these contents can be displayed with Solibri Office as shown in figure 5.

The entire development of the Prototype Framework and the generated IDS files are open source and can be reused (iabi, 2023).

Use Case: DIN 276

The German DIN 276 standard is a vital guideline in the construction industry for structuring and managing costs in construction projects (DIN, 2018). It categorizes costs systematically, aiding in accurate budget planning and risk reduction by standardizing cost estimation and allocation. This standard ensures financial discipline, transparency, and effective communication among stakeholders in German construction projects. We have translated significant

```xml
<ids>
  <specification name=”LoadBearing External Walls” ifcVersion=”IFC2X3 IFC4”>
    <applicability>
      <entity>
        <name pattern=”IFCWALL|IFCWALLSTANDARDCASE”/>
      </entity>
      <property name=”IsExternal” datatype=”IfcBoolean” value=”TRUE” propertySet=”Pset_WallCommon”/>
      <property name=”LoadBearing” datatype=”IfcBoolean” value=”FALSE” propertySet=”Pset_WallCommon”/>
    </applicability>
    <requirements>
      <classification uri=”https://identifier.buildingsmart.org/uri/nlsfb/nlsfb2005/2.2/class/21.11”
                value=”buitenwanden; niet constructief, massieve wanden”
                system=”NL-SfB 2005”/>
      <classification value=”332 - Nichttragende Außenwände”
                system=”DIN 276”/>
    </requirements>
  </specification>
</ids>
```

Figure 3: Example IDS File containing classification logic - this example is for illustration purposes and is not directly usable as IDS because it was simplified

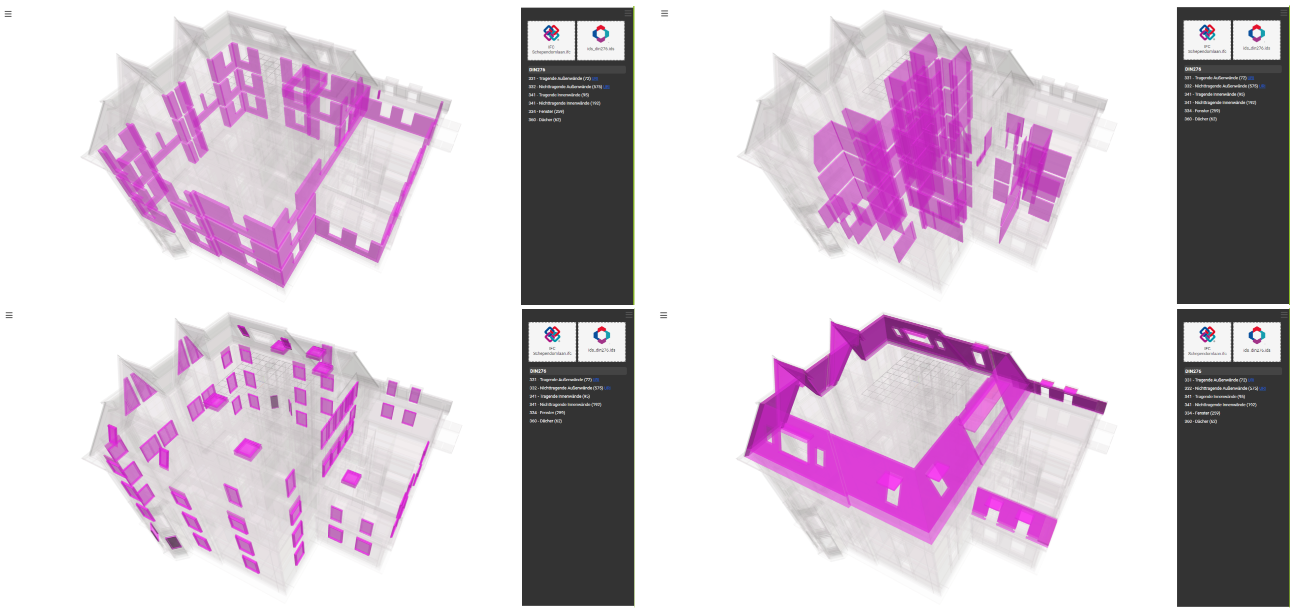


Figure 4: Resulting framework Demonstrator showing different resulting classifications processed using the IDS-representation of the German DIN 276

parts of the DIN 276 structure into IDS for demonstration purposes using the approach presented in this paper.

The resulting IDS can be used in the prototype framework to classify and structure any IFC model, showing the resulting structure on the fly. Figure 4 shows different classification results that have been successfully applied using the rules contained in the IDS. On-demand, the resulting classification can also be written back into the IFC file as references *IfcClassificationReference* objects.

Conclusions & Outlook

The exploration into formalizing classification knowledge via the IDS format reveals a promising avenue toward bridging the intricacies of digital construction data with the practical necessities of industry stakeholders. This methodology not only leverages an established standard but also proposes a nuanced interpretation that aligns classification logic with the inherent structure of construction components. The resultant framework demonstrates a tangible step towards a more collaborative, efficient, and user-centric digital construction ecosystem. The success of the applied test cases underscores the methodology's potential to facilitate more intuitive interaction with complex data structures, thereby fostering a broader understanding and engagement across disciplinary bounds.

However, our approach is only a first step and has a whole

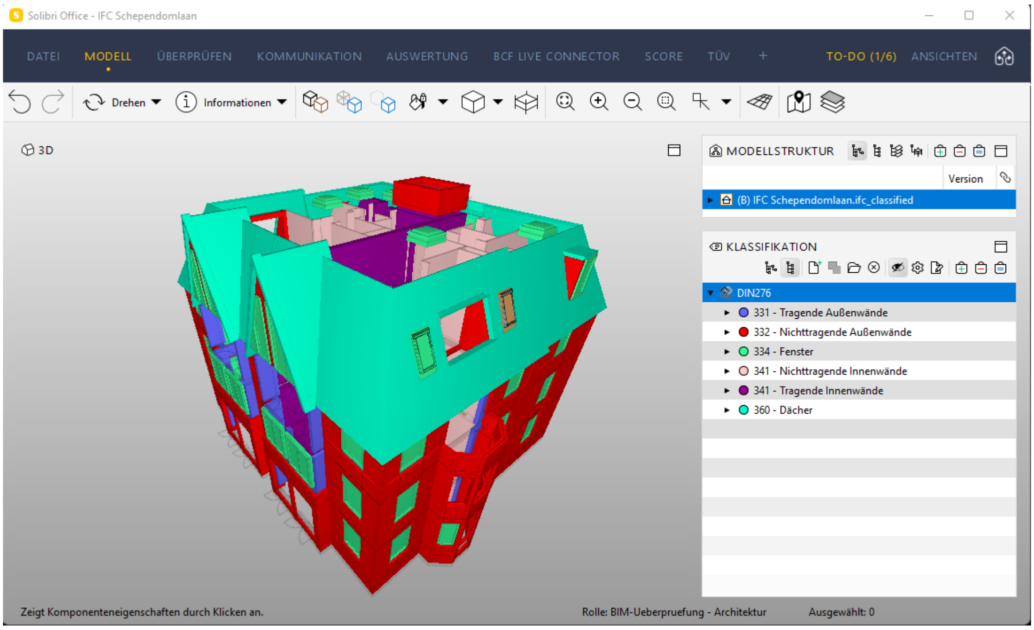


Figure 5: Resulting classification shown in Solibri Office

series of limitations that require consistent further development and refinement:

Exploration of Complex Classification Systems: In our contribution, we have initially converted elementary parts of DIN 276 as a first step to test the basic concept. Of course, there are other national or specialized classification systems worth translating into IDS to make them easily applicable to the masses. Our future research will examine more complex classification systems such as Uniclass and Omniclass. These systems are designed for more complicated and global classification requirements and will be essential for expanding the scope of our study.

Dependence on correct model content: The accuracy and precision of model content remain critical to the successful application of any classification system, underscoring the need to improve modeling procedures continuously. The effectiveness of classification systems depends mainly on users' consistent and accurate maintenance of attribute data. Our approach does not change this, but this is not the intention because, at this point, it is the task of model quality assessment and assurance mechanisms to ensure the appropriate quality.

Conversion of Native Classifications: Regardless of official classification systems, a lot of valuable classification knowledge is deeply embedded in native software solutions and remains inaccessible without targeted and sometimes laborious extraction processes. This represents a significant obstacle to using this knowledge across different systems. Classification knowledge currently only contained in systems such as Solibri Classifications and BIMcollab Smart Views needs to be extracted and converted into an open format such as IDS. This will enable wider sharing and application of this knowledge, overcoming the limitations of proprietary systems and improving interoperability and collaboration across digital construction platforms.

Our work represents an initial step towards potential improvements as we explore automated classification and model validation. It introduces possibilities for further research in model semantics, knowledge extraction, and sharing classification knowledge. Future efforts can focus on refining these methods to enhance their practicality and efficiency:

Automatic Derivation of Classification Logic: A proactive examination of existing models with classification attributes could be undertaken to derive classification logic from these models automatically. Utilizing Artificial Intelligence (AI) and Machine Learning (ML) techniques, the underlying classification logic within these models could be extracted and formalized into IDS files. This automated extraction of classification knowledge could potentially unveil a rich repository of classification logic, making it available for users across different platforms.

Model Validation: The established classification logic can be employed in its original intent, as per IDS, for validating models. This validation ensures that models adhere to defined classification schemes, enhancing the quality and consistency of digital construction data.

Semantic Enrichment of Models: The classification logic could also serve as a mechanism for semantically enriching models. When a component is identified under a specific classification, this logic could provide instructions on the data placeholders or specific values the component should possess. This semantic enrichment facilitates a more detailed and nuanced representation of construction components, enriching the digital construction model.

Community-Driven Evolution: Encouraging a community-driven evolution of this methodology could foster a collaborative environment for continuously refining and expanding the classification logic. Engaging with industry experts, academia, and software developers could collectively advance this approach, aligning it more closely with real-world needs and emerging industry standards.

The envisioned enhancements and the community-driven evolution of this methodology highlight the potential of this approach in addressing the present challenges and adapting to the evolving needs of the digital construction landscape. Building upon the established standards and engaging with the broader community could significantly contribute to the ongoing digitization efforts within the construction industry, making digital construction models more accessible, understandable, and usable for all stakeholders involved.

Acknowledgments

The authors gratefully acknowledge the support and resources that made this research possible. This work has been carried out mainly within the framework of the openDBL project, which has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101092161. We are deeply grateful for this fundamental support. Furthermore, our thanks also go to the NEMETSCHEK Innovation Foundation, whose financial contributions have been an essential basis for the progress of our research and, thus, of crucial importance.

References

Beetz, J. (2018). Structured Vocabularies in Construction: Classifications, Taxonomies and Ontologies. In Beetz, J., Borrmann, A., Koch, C., and König, M., editors, Building Information Modeling, pages 155–165. Springer International Publishing and Imprint: Springer, Cham.

Benghi, C. (2023). xBIM IDS. https://www.xbim.it/xids. Accessed: 2023-10-08.

BIM Deutschland (2024). Muster-Auftraggeber-Informationsanforderungen. https://www.bimdeutschland.de/leistungen/muster-auftraggeber-informationsanforderungen. Accessed: 2024-04-02.

BIM Loket (2023). NL/SfB. https://www.bimloket.nl/p/542/NLSfB. Accessed: 2023-10-08.

BIMcollab (2023). BIMcollab ZOOM. https://www.bimcollab.com/en/Products/Zoom. Accessed: 2023-10-08.

Bloch, T. and Sacks, R. (2018). Comparing machine learning and rule-based inferencing for semantic enrichment of bim models. Automation in Construction, 91:256–272.

buildingSMART (2016). mvdxml 1.1 final specification. https://standards.buildingsmart.org/MVD/RELEASE/mvdXML/v1-1/mvdXML_V1-1-Final.pdf. Accessed: 2023-10-08.

buildingSMART International (2023a). buildingSMART - Information Delivery Specification (IDS). https://technical.buildingsmart.org/projects/information-delivery-specification-ids/. Accessed: 2023-10-08.

buildingSMART International (2023b). buildingSMART Data Dictionary. https://www.buildingsmart.org/users/services/buildingsmart-data-dictionary/. Accessed: 2023-10-29.

DIN (2018). DIN 276:2018-12 - Building costs. https://www.beuth.de/de/norm/din-276/293154016.

iabi (2023). autoClassifier. https://gitlab.lrz.de/000000003B9C426C/autoclassifier. Accessed: 2023-10-08.

International Organization for Standardization (1994). ISO/TR 14177:1994 - classification of information in the construction industry. https://www.iso.org/standard/22703.html.

Krijnen, T. (2023). IfcOpenShell. http://ifcopenshell.org/. Accessed: 2023-10-08.

Opoku, D.-G. J., Perera, S., Osei-Kyei, R., Rashidi, M., Bamdad, K., and Famakinwa, T. (2023). Barriers to the adoption of digital twin in the construction industry: A literature review. Informatics, 10(1).

Solibri (2023). Solibri Office. https://www.solibri.com/solibri-office. Accessed: 2023-10-08.

Tomczak, A., Berlo, L. v., Krijnen, T., Borrmann, A., and Bolpagni, M. (2022). A review of methods to specify information requirements in digital construction projects. IOP Conference Series: Earth and Environmental Science, 1101(9):092024.

Viegas, A. G. (2023). IFC.js. https://bimwhale.gitbook.io/ifc-dot-js/. Accessed: 2023-10-08.

Wu, J. and Zhang, J. (2019). New automated bim object classification method to support bim interoperability. Journal of Computing in Civil Engineering, 33(5):04019033.

Zhang, X.-Y., Hu, Z.-Z., Wang, H.-W., and assem, M. (2014). An Industry Foundation Classes (IFC) Web-Based Approach and Platform for Bi-Directional Conversion of Structural Analysis Models. Computing in Civil and Building Engineering, pages 390–397.

BRIDGING INFORMATION GAPS IN AECO INDUSTRY: A PROTOTYPE FRAMEWORK FOR STANDARDIZED PRODUCT DATA PROVISION

Jakob Martin[1,2], Klaus Linhard[2], Kai Oberste-Ufer[3], Cornelius Preidel[1,2]
[1]Munich University of Applied Sciences, Germany
[2]Institute of Applied Building Informatics (iabi), Munich, Germany
[3]dormakaba Deutschland GmbH, Ennepetal, Germany

Abstract

In response to the EU's Green Deal and its emphasis on sustainable construction, this study introduces an innovative framework for the AECO industry aimed at overcoming the prevalent challenge of fragmented and non-standardized product data exchange processes. By integrating Information Delivery Specification (IDS) and buildingSMART Data Dictionary (bSDD), the proposed framework enhances the accuracy and efficiency of product data exchange. This approach addresses current inefficiencies in the sector and aligns with sustainable objectives by ensuring consistent, up-to-date information, thereby facilitating better adherence to safety, budget, and regulatory standards.

Introduction

The Architecture, Engineering, Construction, and Operations (AECO) sector critically relies on accurate and timely product data, from design specifications to material properties (Tulke and Schumann, 2018). This data is essential to ensure projects are executed smoothly, remain within budget, and comply with safety and regulatory standards. However, the sector faces significant challenges:

Fragmentation of Product Data: Varied and non-standardized data from manufacturers often lead to delays and inaccuracies in project execution.

Lack of Standardization: The absence of a unified approach in data provision hampers efficient data retrieval and application.

These challenges become more pronounced with the implementation of the European Union's Green Deal, which underscores the need for sustainable construction and stringent environmental standards (European Commission, 2023; Schunz, 2022). The Deal necessitates immediate access to detailed, accurate product data to ensure compliance. Moreover, the push towards a circular economy emphasizes the importance of understanding building materials' origin, composition, and recyclability, where data discrepancies can result in non-compliance and potential legal or reputational damage. This context sets the stage for our research, highlighting the need for a standardized and efficient method to query and retrieve product data, a crucial element for operational efficiency and reliability in the AECO sector. At the same time, a persistent challenge within the sector has been the non-uniform manner in which product data is provided by manufacturers, owing to a lack of industry-wide standardization (Kebede et al., 2022; Lucky et al., 2019). This fragmentation obstructs efficient data retrieval and utilization, inevitably leading to delays and inaccuracies in project execution. The research gap thus emerges from the necessity of a streamlined, standardized method for querying and retrieving product data from manufacturers.

State of the Art

In the dynamic domain of digital construction, the seamless exchange of product information between manufacturers and core stakeholders such as contractors, designers, and engineers is fundamental. This exchange profoundly influences various stages of the construction process, facilitating informed decision-making, accurate cost estimation, and adherence to compliance norms, among other critical facets (Aranda-Mena and Wakefield, 2006). The information regarding products is especially crucial during distinct phases such as the planning and design, tendering, procurement, construction, and operational maintenance phases. During the planning and design phase, engineers and architects heavily rely on product information to make critical design decisions, ensuring that the selected products align with the project's objectives and compliance requirements (Yogana and Latief, 2021). The data aids in evaluating the compatibility of products with the design intent and making necessary adjustments to the design. As the project transitions to the tendering stage, acquiring precise product data is indispensable, empowering contractors and planners to craft accurate bids. They necessitate detailed specifications, pricing, and the availability of products to devise competitive yet realistic tenders. A significant aspect of the tendering process is the comparative analysis of products from different manufacturers. This step helps ascertain the most suitable options in terms of cost and functionality. Procurement is the next critical phase, where accurate product data is vital for making informed purchasing decisions. At this juncture, a bulk of comparison between different manufacturers' products is carried out to secure the best value and ensure project continuity. Once the construction phase commences, access to product information is essential to ensure that suitable materials are utilized per the specifications. It also aids in addressing any on-site challenges that may arise related to product installations. Post-construction, during the operational maintenance phase, having a repository of product information proves invaluable for facility management and maintenance activities. It aids in ensuring that the products used are serviced or replaced following the manufacturer's guidelines. Traditionally, obtaining product information has been a manual process, relying on product data descriptions typically provided in PDF format by manufacturers due to industry norms. Although functional, this method is time-consuming and prone to data inconsistencies, mainly when product specifications are updated.

Platforms for the distribution of product data

In the evolving landscape of the AECO industry, the emergence of platform providers for product data has become notable. Acting as intermediaries, these platforms house various manufacturers' product data, providing a centralized hub for stakeholders to efficiently access and compare product information. Such platforms include BIMobject (BIMobject AB, 2023) and Cadenas (CADENAS GmbH, 2023).

These providers endeavor to bridge the gap between product manufacturers and various industry stakeholders, including contractors and designers, positioning themselves as intermediaries that provide access to product data that is more about availability than immediate, customizable retrieval. Such platforms are particularly valuable during phases like tendering and cost calculation, where precise product specifications are paramount. However, the adoption of these platforms presents significant challenges. The utility of such a platform is contingent upon its widespread acceptance by multiple product manufacturers, as navigating various platforms not only complicates the workflow but also introduces inefficiencies. Therefore, for contractors and other stakeholders to reap the full benefits, a unified system is essential to streamline processes and enhance efficiency across the board.

A more pressing concern is the redundancy inherent in these systems. Product manufacturers, in their typical operations, maintain detailed specifications within their proprietary databases or systems. The introduction of intermediary platforms necessitates the replication of this data onto the new platform. This duplication poses several risks:

Inconsistencies: There's always the potential for discrepancies between the data in the manufacturer's primary system and those copied onto the intermediary platform.

Data Asynchronicity: The delay between updates in the manufacturer's system and their replication on the platform can result in stakeholders accessing outdated information.

Maintenance Overhead: For manufacturers, managing and updating product data in two places can be resource-intensive and error-prone.

While these platform providers aim to streamline processes and enhance accessibility to product data, they inadvertently introduce complexities and potential data integrity issues. It underscores the need for more integrated solutions to address these challenges, ensuring stakeholders have access to consistent, up-to-date product information without the pitfalls of redundancy.

Digital labeling and classification systems

In contrast to dedicated platforms that aim to host all the information, other approaches aim to link the connections between the required data and the manufacturers' data and manage these links. A prominent example is the *UniversalTypes*, developed and owned by *ProMaterial* under the aegis of buildingSMART International (buildingSMART International, 2023c). This initiative was developed to standardize and harmonize the characteristics and properties of products in a catalog and thus significantly improve the real-time sale of building products and materials.

Initially published as a publicly available specification in the buildingSMART Data Dictionary (bSDD), the language of UniversalTypes is envisioned to be standardized and maintained for general use. Unlike previous systems that were confined to proprietary platforms, UniversalTypes is integrated with the bSDD, enabling it to leverage a wider industry acceptance and providing a robust foundation for data harmonization across the construction sector. This integration is illustrated in our study's Figure 2, where UniversalTypes are referenced as a classification system within the bSDD framework (buildingSMART International, 2024).

ProMaterial offers a dedicated platform that utilizes UniversalTypes to streamline online sales processes. Within this platform, products can be explored using specific properties defined as UniversalTypes, such as the manufacturer's product number, Data-provider product ID, and various product dimensions and material properties. These properties are essential for designers during the design phase, allowing for an in-depth analysis and comparison of products from different manufacturers.

Moreover, ProMaterial extends its functionality by providing developer documentation for API development, facilitating advanced queries and interactions with the UniversalType Center for more customized workflows. This integration enhances the efficiency and accuracy in data retrieval and analysis, making it a valuable tool for planners' workflows and embodying modern digital construction industry processes.

In summary, the classification part of UniversalTypes represents a robust approach that not only uniquely identifies individual features but also complements other classification systems. It is showcased by its integration into the bSDD, enhancing its utility and reach. However, the reliance on a dedicated platform for providing this information highlights a business case that inherits the limitations of proprietary platforms, posing challenges for achieving a truly open, vendor-neutral solution.

Product Data Templates (PDT)

A Product Data Template (PDT) is a standardized format that outlines essential and optional product attributes, such as fire rating and color, according to established product standards. PDTs offer a unified method for handling construction product data, simplifying the process for manufacturers to keep product details up-to-date. The ISO standard 23387:2020 establishes guidelines for data structures that describe construction-related entities, promoting efficient digital information exchange and operational enhancement (International Organization for Standardization, 2020b). ISO 23387:2020 is a consistent further step after ISO 23386:2020 that includes a definition of characteristic structures (International Organization for Standardization, 2020a). PDTs are supposed to maintain data consistency across a product's lifecycle and are designed to be machine-readable for easy data exchange.

A Product Data Sheet (PDS) is created when manufacturers input data into a PDT using specialized software, detailing a construction product's technical and performance characteristics according to various regulations and requirements. As a product's identifiable document, a PDS ensures the supply chain can access up-to-date and accurate information from a trusted source. Unlike static formats like PDFs, PDSs support automation by integrating with various manufacturer systems, such as Product Information Management (PIM), Digital Asset Management (DAM), or Enterprise Resource Planning (ERP), to keep data fresh. They can act as a single source of truth for automatic updates across these systems.

PDTs can potentially improve data management in construction, but they come with challenges. The complexity of PDTs, particularly their integration with the IFC schema, creates a steep learning curve that can discourage smaller firms or those without technical expertise. The specificity of PDTs to IfcTypes can also be restrictive, lacking the flexibility to handle unique data needs and sometimes leading to inaccuracies requiring manual correction. Moreover, PDTs' reliance on the IFC schema can be limiting if the IFC schema has no prominent role in the project or undergoes changes, necessitating updates to the PDTs. Integrating PDTs into existing systems not designed for them can be complex and costly. PDSs, while designed to be a standardized source of information, face issues with data consistency, mainly when multiple parties contribute to project-specific PDSs.

Filling out PDSs can be time-consuming for manufacturers, particularly for those with a wide range of products or frequent updates. Despite these obstacles, the industry is working towards improving these systems to leverage their benefits in the digitally evolving construction sector fully. The transition involves overcoming the initial difficulties associated with adopting new technologies.

Formulating requirements for successful and efficient product data exchange

The AECO sector is at a pivotal juncture where the need for accurate product data intersects with the drive for sustainable construction mandated by the EU's Green Deal. Despite the potential of existing systems like native Product Data Platforms or Product Data Templates (PDTs), their adoption faces significant hurdles. The complexity of these systems, their integration challenges, and the need for consistent, up-to-date data underscore the necessity for a more streamlined approach to product data exchange. The industry's reliance on fragmented and non-standardized data hinders project efficiency and poses risks to compliance with emerging sustainability regulations. As the sector grapples with these challenges, it becomes clear that a new paradigm is needed to address the limitations of current methodologies and set clear, actionable requirements for a robust data exchange framework. To transition from the current state of disparate data practices to a future of streamlined and standardized data exchange, the following requirements must be met: Firstly, it is imperative to have a standardized data format. A universally accepted data format ensures seamless product data integration into the myriad software tools and platforms employed across the sector. Moreover, the systems utilized by various stakeholders, be they architects, engineers, or contractors, must be interoperable. This interoperability ensures smooth data exchange devoid of any compatibility hiccups. The integration of open APIs and dedicated tools can further smoothen this process. For widespread adoption, platforms or tools designed for data exchange must feature an intuitive user experience, easing the transition for all involved parties. Additionally, these platforms should allow stakeholders to retrieve specific data subsets efficiently, negating the need to sift through entire datasets. Another cornerstone of this endeavor is real-time data access. Stakeholders should always be working with the most recent information available.

Approach & Methodology

To enable the information orderer and provider to communicate information and requirements directly, this study proposes using the newly introduced Information Delivery Specification (IDS) standard from buildingSMART in conjunction with data dictionary services, e.g., buildingSMART Data Dictionary. The main objective of this study is to optimize the information exchange between the two main stakeholders in this process, especially in phases such as the tender: the information receiver and the sender. Therefore, the central idea is to rely on existing standards and describe a process as simply as possible, allowing as few obstacles as possible between these stakeholders.

buildingSMART Data Dictionary (bSDD)

The buildingSMART Data Dictionary (bSDD) serves as a fundamental source for ensuring data consistency and interoperability between different software platforms and actors in the AECO industry. Although there are several organizations that see establishing such sources as their task (see, e.g., the BIM Portal of the Federal Government (BIM Deutschland, 2023)), in our approach, bSDD provides the standardized terminology and data structures. These structures can be referenced in the Information Delivery Specification (IDS) to effectively communicate the data requirements and ensure that the exchanged product data is correct and commonly and uniformly understood within the industry, particularly between the purchaser and supplier.

As demonstrated in Figure 2, we explicitly utilize references from bSDD to establish this connection and ensure uniformity across the data exchanged. This us-

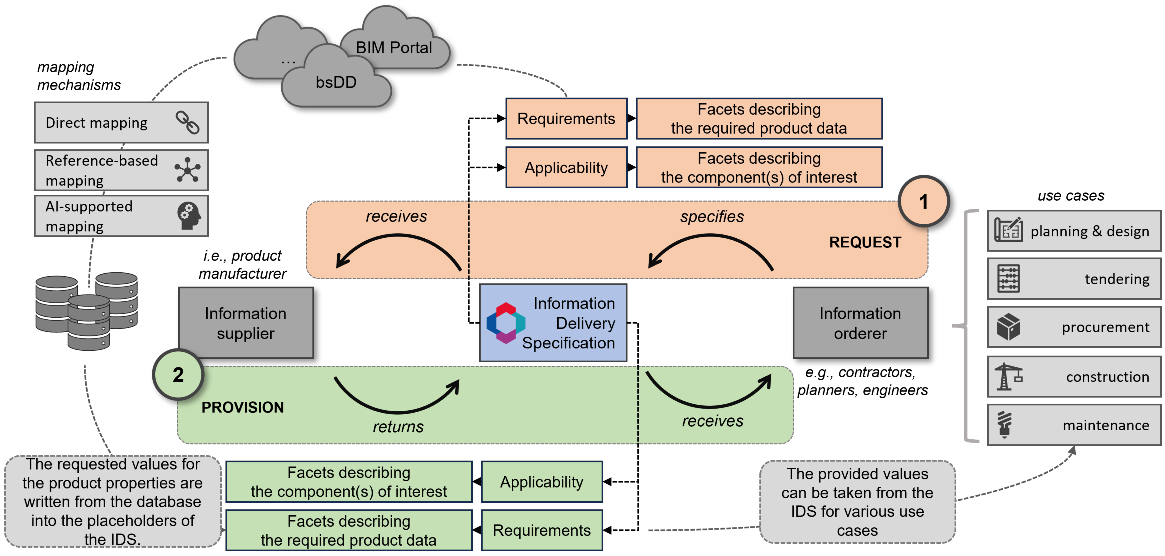


Figure 1: Schematic representation of the suggested methodology and process between the actor that orders and that provides the information

age highlights how bSDD's standardized terminology and data structures are critical in supporting the IDS's role in enhancing interoperability and clarity in communication among stakeholders.

Information Delivery Specification (IDS)

Information Delivery Specification (IDS) (buildingSMART International, 2023b; Tomczak et al., 2022) is a rather newly introduced XML-based tool to specify a construction project's information delivery requirements. It serves as a blueprint, outlining the nature, format, and extent of semantic information to be exchanged amongst stakeholders at various stages of a project lifecycle. By doing so, IDS augments the clarity, consistency, and efficiency of information exchange, minimizing misunderstandings and errors that could potentially derail a project. The primary intent of IDS is to facilitate the precise articulation and fulfillment of information requirements. It acts as a bridge, ensuring that the information generated and consumed across different phases of a construction project is aligned with the defined standards, thereby promoting interoperability and collaborative engagement. However, the structure of the format itself offers a variety of possibilities for further application areas that the introduction can open up. In addition, the format is getting a lot of attention, and many software vendors are implementing the standard. This momentum can be used to exploit the widest possible application of the format.

The following methodology has been proposed to achieve this objective and is shown schematically in figure 1:

1. Information Request

To begin with, the required product information must be collected using the IDS format. For this purpose, the features of IDS are used as follows:

- The *Applicability* elucidates which components need certain information and why. It specifically identifies which components are affected and should be considered. In this section, additional facets can add more details to describe the components, and some settings can be established as defaults. Additional design specifications could need to be set in advance during the planning phase. A more detailed description of the affected component could help identify it during the operating phase.
- The *Requirement* section further refines this by detailing the exact nature of the information needed. This can be an attribute according to the IFC schema or a user-defined property, i.e., a property in a property set with its required data type.

For the definition of the different pieces of information, a special characteristic of IDS can be used to give more robustness and reliability to the whole process: Each piece of information can be additionally provided with a URI, which refers to an external resource, e.g., to the bSDD. Similarly, other sources can be used as reference systems for attributes and properties, such as the BIM Portal in Germany (BIM Deutschland, 2023). The BIM Portal in Germany, managed by BIM Deutschland as of 2023, functions as a central resource for promoting BIM standards and practices. It offers detailed guidelines, best practices, and tools to streamline BIM adoption, supporting governmental and industry-wide efforts to enhance interoperability and efficiency in construction projects through digitization. In this way, a reference to a source can be established, which can be used for a common understanding of the same information if required. This could involve mapping to national classifications, UniversalTypes, and other standards recorded in the bSDD or any other source of information. An example of an IDS capturing requested information for external doors is shown in figure 2.

```
<ids>
  <specification name=”Example for doors” description=”Requesting Door Properties” ifcVersion=”IFC2X3 IFC4”>
    <applicability>
      <entity name=”IfcDoor”/>
      <property name=”IsExternal” datatype=”IfcBoolean” value=”TRUE” propertySet=”Pset_WallCommon”/>
    </applicability>
    <requirements>
        <property propertySet=”Pset_DoorCommon” name=”AcousticRating”
          ↪ uri=”https://identifier.buildingsmart.org/uri/buildingsmart/ifc/4.3/prop/AcousticRating”/>
        <property propertySet=”Pset_DoorCommon” name=”FireRating”
          ↪ uri=”https://identifier.buildingsmart.org/uri/.../Pset_DoorCommon/FireRating”/>
        <property propertySet=”Pset_DoorCommon” name=”DurabilityRating”
          ↪ uri=”https://identifier.buildingsmart.org/uri/.../Pset_DoorCommon/DurabilityRating”/>
        <property name=”Türschließerfunktionen”
          ↪ uri=”https://identifier.buildingsmart.org/uri/promaterial/universaltypes/1.0/prop/ST08-HODV”>
            <propertySet>
                <xs:restriction base=”xs:string”>
                    <xs:pattern value=”.” />
                </xs:restriction>
            </propertySet>
        </property>
        <property name=”Türflügelbreite”
          ↪ uri=”https://identifier.buildingsmart.org/uri/promaterial/universaltypes/1.0/prop/ST05-WWDO”>
            <propertySet>
                <xs:restriction base=”xs:string”>
                    <xs:pattern value=”.” />
                </xs:restriction>
            </propertySet>
        </property>
    </requirements>
  </specification>
</ids>
```

Figure 2: Example IDS file containing an information request for product manufacturer data for doors - for illustration purposes, the IDS has been greatly simplified and can therefore not be directly used as such

This formulated IDS can now be used as an information request and, as such, be sent to one or more information providers - i.e., product manufacturers eligible to deliver the requested information. This can happen precisely when the information is imperative, for instance, during the tendering phase.

2. Information Provision

Upon receiving the IDS-based query, a manufacturer deciphers the request, mapping it to their specific product specifications. The bSDD and, where needed, other standards like UniversalTypes aid this mapping. By doing so, the manufacturer can access its databases, where the most recent and reliable data about products is maintained. This prevents inconsistent or outdated data from being returned, and no middleman is required, which is a great help regarding liability issues. Subsequently, the manufacturer responds with the solicited information formatted in IDS. This returned IDS, while preserving its original format, now contains the values for each piece of requested data.

The planner, upon receiving the populated IDS along with the requirement specifications, has two potential courses of action:

- Manually incorporate the provided data. Given that the IDS already dictates applicability, the planner is guided on where to position the product data.
- Opt for a fully automated data integration in the model.

Using the BCF API Standard to foster the workflow

Applying the BCF API specification, as outlined by (buildingSMART International, 2023a), is recommended to facilitate prompt data exchange. Particularly, the BIMsnippet object within the BCF API framework is optimally designed for exchanging subsets of information, utilizing a harmonized schema. Particularly, the BIMsnippet object within the BCF API framework is optimally designed for exchanging subsets of information, utilizing a harmonized schema. A BIM snippet is a partial model referencing a schema, ensuring structured standards for its content. The Information Delivery Specification (IDS) XML schema is integrated with our proposed methodology. The IDS XML schema is integrated with our proposed methodology. Moreover, the pertinent attributes defined in the BCF standard are selectively employed for initiating and

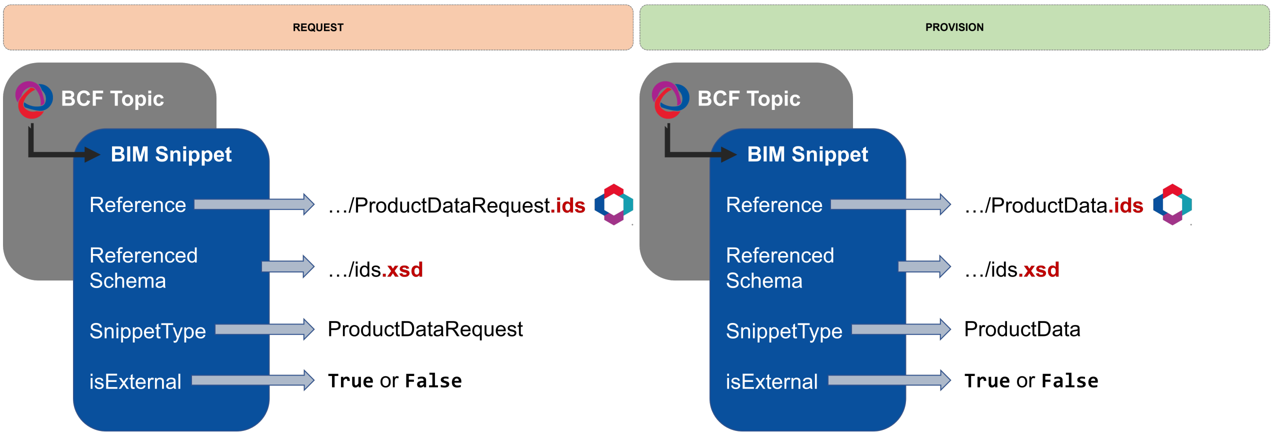


Figure 3: Schematic representation for the use of the BIMSnippet object for the BCF API for the mapping of the data request (left) and the data provision (right)

fulfilling data requests. Figure 3 illustrates the foundational principle for requesting and providing entities. The widespread adoption of the BCF API across software vendors significantly attests to its efficacy in streamlining this workflow. By employing this proposed workflow, a product manufacturer can establish a BCF API server as an access point for stakeholders seeking information. In this process, an information requester initiates a new BCF topic, embedding the request details as IDS within the BIMSnippet objects. Conversely, the supplier responds to the request by filling the same BCF topic with a BIMSnippet object that holds the information provision as IDS. This way, the request workflow can be largely automated and scaled if needed.

As an example, if multiple manufacturers are queried simultaneously via the BCF API, a mechanism for multiple suggestions could be conceived. This would allow planners to receive and compare various data samples simultaneously. Based on this comparison, the planner could decide what product to choose. To confirm that all data have been received and entered in the model as required, it can be checked very easily with the original IDS that was used for the request at the beginning - according to the original purpose of IDS, it also serves for model checking and can also be used for this purpose.

Adopting this methodology is anticipated to engender a more seamless and efficient data exchange within the AECO sector, thereby enhancing collaborative efforts and augmenting the quality of project outcomes.

Test Case: Requesting manufacturer properties of a door

To test the viability of our proposed approach in a simple test use case, we set up a common use case together with a manufacturer of door products as it is often encountered in practice: The correct specification of doors is a crucial element to ensure safe, user-friendly, and efficient buildings. At this, the key factor is the correct definition of requirements at each door aligned with local standards and regulations. Architects and planners often lack the knowledge to do this properly and are often exposed to the risk of missing important security-related features. The process proposed in this paper helps to enhance this task. In collaboration with the door manufacturer, we used our framework so that a planner generates a specification request (IDS) containing the requirement definitions, formulating a typical information request for a set of doors, mirroring common industry practices. The receiver is a domain specialist who adds the specific values to the request and sends it back to the architect. This roundtrip is repeated along the planning stages multiple times with different requirements, data, and stakeholders. The requested and transmitted response information from the product manufacturer is shown here in Table 1, together with the URI for the mapped source, e.g. bSDD:

Table 1: IDS Exchange for Door Specification with iterative planner and manufacturer responses, including URIs

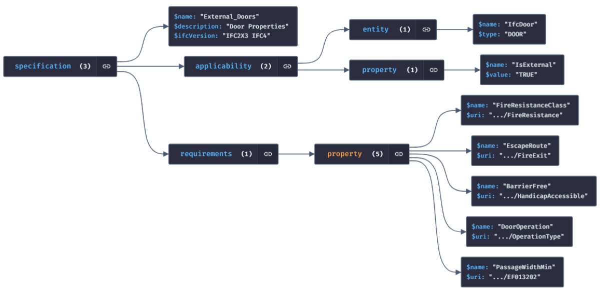
Information	URI	Responses
FireResistanceClass	🔗	T30
EscapeRoute	🔗	RW DIN 179
BarrierFree	🔗	True
DoorOperation	🔗	A (Automatic)
PassageWidthMin	🔗	900mm

The resulting IDS files for information request and provision are shown in Figure 4 - to enhance understanding further, it is shown as a graph-based illustration.

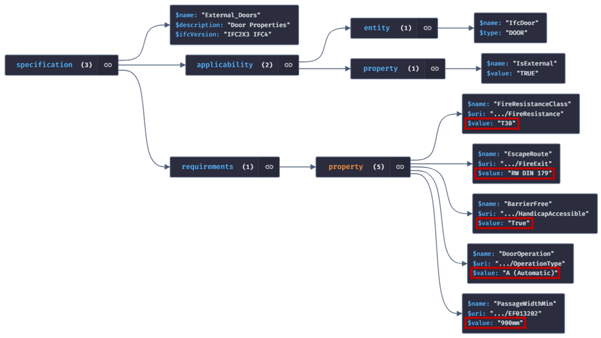
Conclusions & Outlook

The proposed method of utilizing IDS alongside bSDD and other data dictionaries presents a novel pathway toward overcoming the long-standing challenge of standardized product data provision within the AECO industry. This approach promises enhanced data accuracy and availability and fosters a collaborative ecosystem among the various stakeholders, driving the industry toward a more integrated and efficient future.

Building upon the presented research, several avenues emerge for further investigation:



Query of features in an IDS specification



Providing features by assigning values (marked red)

Figure 4: graph-based representation of a (left) query and a (right) response based on IDS-XML

Industry Adoption Analysis: Investigating the barriers to, and facilitators of, widespread adoption of this standardized data provision method within the AECO industry.

Integration with Existing Systems: Exploring this method's compatibility and integration possibilities with existing industry data management and project management tools.

Exploratory Prototype Development: In our study, we have outlined a conceptual framework designed to standardize and streamline product data exchange within the AECO industry. Moving forward, the next phase of our research will focus on evolving this conceptual framework into a more tangible model that can be rigorously tested and validated. This will involve developing an initial prototype for a broader spectrum of product data queries and processes. The prototype will be designed to demonstrate the practical applicability of our framework in real-world scenarios, allowing us to assess its effectiveness and make iterative improvements based on feedback from industry stakeholders. This progression from a conceptual framework to a prototype is crucial for bridging the gap between theoretical research and practical implementation. It will provide valuable insights into the challenges and opportunities associated with deploying such a system on a larger scale, thus paving the way for widespread adoption in the industry.

Limitations and Future Research Directions

While our current framework focuses primarily on model-based workflows typical of the AECO industry, a significant portion of the sector still operates without such models, particularly in managing existing, non-digitized assets. Given this gap, an important question for future research arises: how can the framework we propose be adapted to support non-model-based workflows effectively? This question is crucial as it aims to extend the benefits of standardized product data exchange to broader contexts within the industry, ensuring inclusiveness and broad applicability. Answering this question involves exploring methods to generalize the principles of our framework to support workflows that do not rely on structured digital models, thus expanding the scope of its utility and impact.

Additionally, it is important to note that the current implementation of IDS is limited to alphanumeric information exchanges and does not support the transfer of geometric data. While this is acceptable as an initial step, further investigation is needed since some product information exchange might also rely on geometric data, not just characteristics.

This endeavor lays the groundwork for a series of potential advancements in addressing the product data provision challenge, pushing the frontier of information management within the AECO domain.

Acknowledgments

The authors gratefully acknowledge the support and resources that made this research possible. This work has been carried out mainly within the framework of the openDBL project, which has received funding from the European Union's Horizon Europe research and innovation program under grant agreement No 101092161. We are deeply grateful for this fundamental support. Furthermore, our thanks also go to the NEMETSCHEK Innovation Foundation, whose financial contributions have been an essential basis for the progress of our research and, thus, of crucial importance. Additionally, the authors extend their sincere thanks to dormakaba Deutschland GmbH for their practical contributions to this study.

References

Aranda-Mena, G. and Wakefield, R. (2006). Interoperability of building information — myth of reality? In eWork and eBusiness in Architecture, Engineering and Construction. ECPPM 2006, page 7. CRC Press, 1st edition edition.

BIM Deutschland (2023). BIM Portal des Bundes. https://via.bund.de/bim/infrastruktur/landing. Accessed: 2023-11-02.

BIMobject AB (2023). BIMobject. https://www.bimobject.com/. Accessed: 2023-10-29.

buildingSMART International (2023a). BCF REST API. https://github.com/buildingSMART/BCF-API. Accessed: 2023-10-29.

buildingSMART International (2023b). buildingSMART - information delivery specification (ids). https://technical.buildingsmart.org/projects/information-delivery-specification-ids/. Accessed: 2023-10-08.

buildingSMART International (2023c). Universaltypes. https://www.buildingsmart.org/users/services/universaltypes/. Accessed: 2023-10-29.

buildingSMART International (2024). Türflügelbreite. https://identifier.buildingsmart.org/uri/promaterial/universaltypes/1.0/prop/ST05-WWDO. Accessed: 2024-04-21.

CADENAS GmbH (2023). Cadenas. https://www.cadenas.de/. Accessed: 2023-10-29.

European Commission (2023). The european green deal. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/european-green-deal_en. Accessed: 2023-10-29.

International Organization for Standardization (2020a). ISO 23386:2020 - Building information modelling and other digital processes used in construction Methodology to describe, author and maintain properties in interconnected data dictionaries. Technical report, International Organization for Standardization.

International Organization for Standardization (2020b). ISO 23387:2020 - Building information modelling (BIM) - Data templates for construction objects used in the life cycle of built assets - Concepts and principles. Technical report, International Organization for Standardization.

Kebede, R., Moscati, A., Tan, H., and Johansson, P. (2022). Integration of manufacturers' product data in BIM platforms using semantic web technologies. Automation in Construction, 144:104630.

Lucky, M. N., Pasini, D., and Spagnolo, S. L. (2019). Product Data Management for Sustainability: An Interoperable Approach for Sharing Product Data in a BIM Environment. IOP Conference Series: Earth and Environmental Science, 296(1):012053.

Schunz, S. (2022). The 'European Green Deal' – a paradigm shift? Transformations in the European Union's sustainability meta-discourse. Political Research Exchange, 4(1).

Tomczak, A., van Berlo, L., Bolpagni, M., Krijnen, T., and Borrmann, A. (2022). A review of methods to specify information requirements in digital construction projects. IOP Conference Series Earth and Environmental Science, 1101.

Tulke, J. and Schumann, R. (2018). BIM-Based Production Systems, chapter 27, pages 447–461. Springer International Publishing, Cham.

Yogana, E. and Latief, Y. (2021). Development of system information of building code checking in planning and permitting phase to improve building code compliance based on work breakdown structure (WBS) using building information modeling (BIM). Journal of Physics: Conference Series, 1858(1):012091.

A KNOWLEDGE GRAPH MODELING APPROACH FOR AUGMENTING LANGUAGE MODEL-BASED CONTRACT RISK IDENTIFICATION

Chunmo Zheng[1], Yinqiu Tang[2], and Xing Su[1]
[1] Zhejiang University, Hangzhou, China
[2]PowerChina Huadong Engineering Corporation Limited, Hangzhou, China

Abstract

Contract risk identification is essential for preventing disputes and losses in construction industry. Large language models (LLMs) have impacted various natural language processing tasks, offering a promising avenue for automating contract review without extensive data processing and feature engineering. However, LLMs still has difficulty in recalling facts while generating knowledge-grounded analysis, especially when related to complex domain knowledge. This paper introduces a Knowledge Graph (KG) modeling approach to enhance the LLM-based automated contract risk identification. A case study demonstrates that our approach exhibits enhanced performance on risk identification tasks compared to non-augmentation scenario.

Introduction

Legal issues, such as disputes, claims, and litigation, frequently occur in construction industry, due to the growing size and complexity of construction projects. The use of natural language in contracts has been identified as the leading cause of such issues (Zait & Zarour, 2018). Specifically, semantic vagueness and ambiguities in contracts can lead to disagreements and misunderstandings between parties, resulting in conflicts or disputes (Mahfouz et al., 2018).

Natural Language Processing (NLP) techniques offers a promising tool to automate contract text processing, thus reducing human errors and increase efficiency. NLP has demonstrated its potential in various automated text processing tasks within the construction industry, including automated contract reviewing (Lee et al., 2020), automated compliance checking (Salama & El-Gohary, 2016) and similar case retrieval (Zou et al., 2017). These techniques can be broadly categorized into rule-based, machine learning-based and deep learning-based methods. However, two limitations have been frequently highlighted in the current literature. First, existing methods rely on sufficient high-quality annotated data to perform the domain-specific tasks, which is highly data-dependent and lack in scalability. Second, current approaches are facing the criticism of only predicting fragmented risk labels without an interpretable risk reasoning.

To mitigate the aforementioned problems, researchers have proposed using large language models trained on a massive amount of unstructured data to facilitate the legal and contractual reasoning tasks. For example, LLMs have begun to reshape the practice of legal services among professionals (Fernandez, 2023). Many argues that the model's ability to learn new tasks from limited data would significantly reduce the human labor in data annotating and feature engineering. However, it still faces many challenges when adapting LLMs to downstream contract risk reasoning. They may struggle with domain-specific knowledge and may not be able to provide explanations from expert perspective, along with factual errors.

Existing works have shown that retrieving structured triple-formed knowledge from KGs can significantly improve LLM's performance on knowledge-intensive reasoning. KG is a data structure designed to accumulate and represent real-world knowledge, where nodes represent entities of interest and whose edges represent various relations between these entities (Hogan et al., 2021). The explicit knowledge representation of KGs can be provided as external knowledge to guide LLMs towards more robust and interpretable knowledge reasoning process (Yang et al., 2024).

Motivated by the identified challenges and existing KG-augmented LLMs solutions, this paper proposes a nested knowledge graph (KG)-enhanced LLM method for automated contract risk identification. A comprehensive framework of developing the ontological layer and instance layer of contract KG is introduced. The nested KG can capture the intricate and interconnected relation within contract clauses, which provide structured and accurate external knowledge for LLMs. Our approach shows improved interpretability of the generated content, and also sheds a light on the integration of LLMs and KGs in automated contract review systems and other knowledge-intensive domain reasoning tasks.

Literature Review

NLP-assisted contract management and the role of prior knowledge

NLP techniques have been widely used in addressing legal issues in construction projects (Hassan et al., 2021).

In contract review, many researchers perform the detection of requirement clauses or poisonous clauses. Lee et al. (2019, 2020) use NLP and semantic syntactic rules in both risk-prone clauses extraction and contractor-friendly clauses detection, achieving the F-score of 81.8% and 80%. Machine learning-based NLP is employed to detect risk-prone paragraphs in contract, achieving an accuracy of 94% (Chakrabarti et al., 2018). Additionally, many studies have also used machine learning algorithms to classify contract clauses into different categories to facilitate contract review process. The categories include requirement or non-requirement clauses, different categories of construction projects or different categories of risks (Candaş & Tokdemir, 2022; Hassan & Le, 2021).

Turning to automated regulation text processing, automated compliance checking aims at detecting the violation with construction laws or regulations. Semantic text classification in unstructured provisions is the first step to perform the automated compliance checking task (Beach et al., 2020). Salama et al. (2016) used a hybrid approach combining semantic analysis and machine learning to classify clauses, achieving a recall of 100%. Zhou and El-Gohary (2016) introduced a rule-based NLP method to classify environmental requirements, alongside a machine learning-based approach for regulatory codes classification.

These studies reveal that NLP techniques have achieved a promising performance in the contract document analysis. However, owing to the unique and complex nature of contract terminologies and semantic patterns, it is difficult for generic NLP models to produce equally reliable performance in construction contract domain without a proper adaptation (Hassan et al., 2021). A few recent studies have attempted to settle these issues by integrating domain-knowledge into the NLP models with the help of semantic representations such as taxonomy, ontology and KG (Lee et al., 2019; Xu & Cai, 2019; J. Zhang & El-Gohary, 2016). These studies have shown an impressive performance, highlighting the importance of integrating domain-knowledge in the future NLP-assisted contract analysis.

Knowledge representation of domain knowledge and the issue of complex knowledge modeling

Semantic web technologies enable the representation of machine-readable data on the web (Kumar, 2019). It is widely adopted in the Knowledge Representation (KR) of domain knowledge by defining domain concepts, enhancing domain information integration and performing logical reasoning (Li et al., 2022). Ontology is one of the fundamental concepts in semantic web technology. It provides a conceptual model to describe a set of concepts in a domain (S. Zhang et al., 2015). Many researchers have developed construction domain ontology for efficient knowledge management, compliance checking, risks or conflicts detection (Zhong et al., 2019; Z. Zhou et al., 2016).

Ontology can function as formal representation to create knowledge graphs (KGs). Recent years, KG has been appeared as a major trend in KR technique to serve many industrial applications (Fang et al., 2020). It presents knowledge in the form of labeled directed graph. Each entity is considered as a node and they are linked via edges which represent relations between entities. The basic unit of KG can model binary semantics with RDF triple (Duan et al., 2017).

In real world knowledge of many domains, especially contract domain, there are always additional information conveying conditional, temporal or provenance information that are beyond the modeling ability of a triple. Modeling and extracting such metadata are beneficial for efficient domain knowledge management. Semantic legal metadata extraction is also crucial for interpreting legal provisions (Sleimi et al., 2019, 2021).

Data modeling solutions for RDF metadata include standard reification, singleton property and RDF-star (Hartig & Thompson, 2021). They are able to represent additional contextual information attached to individual triples (Orlandi et al., 2021). The metadata representation in the graph structure have also been explored. Temporal KG and event KG associate triples with time or site hyper-edges, where an event triple is considered as a semantic unit (Lv et al., 2022; Park et al., 2022).

Contract KG modeling approach

Motivated by the challenges of representing complex knowledge in contract, we introduce the nested KG framework. This framework defines a triple in knowledge graphs as head and tail nodes connected by a relationship edge, while allows nodes to appear as either a single node or a fact node. A fact node is a nested entity which contains a triple inside. This allows a triple to be treated as a node, so that further constraints or statement can be link to the triple. The symbolic definition of this framework is given in Table 1. h, r, t denotes the head entity, relation and tail entity respectively. E, F, R refers to the collection of single entity, fact (nested entity) and relation. Single entities and facts can be interlinked with each other to form nested triples, namely entity-to-entity, entity-to-fact and fact-to-fact triple.

Table 1: Definition of nested triples

Triple	Definition
T_{E2E}	$T_{E2E} = \{(h, r, t) \mid h, t \in E, r \in R\}$
T_{E2F} /T_{F2E}	$T_{E2F} = \{(h, r, t) \mid h \oplus t \in \{E, F\}, r \in R\}$
T_{F2F}	$T_{F2F} = \{(h, r, t) \mid h, t \in F, r \in R\}$

By expanding the concept of entities and triples in knowledge graphs to multi-layer nested structure, the knowledge representation framework can capture deeper semantics and interconnected facts. When dealing with the complex relations in contract context such as conditional, temporal and compound relations, our framework is able to incorporate not only the scattered concepts, but relations among the concepts.

Principally, the nested structure can be infinitely nested, in order to represent more complex semantics. However, it is important to note that there exists the trade-off between expressivity and scalability of knowledge graph, where the more complex nested layer will limit the graph scalability.

Ontology layer development

In this section, a contract ontology based on the nested knowledge graph framework is developed to represent contractual knowledge. As shown in Figure 1, the ontology provides a hierarchical conceptual schema for modeling contract semantics.

contractual relations can either be conditional relations or temporal relations in the contract. Based on this, we can extract the contract rules as, for example, <Statement, if-then, Event> or <Event1, withinPeriodOf, Event2>.

Instance layer extraction

For representing the contract knowledge graph instances, RDF-star language is chosen for its efficiency in representing nested data structures efficiently and directly, without introducing redundant nodes.

Figure 2 illustrates the procedure of constructing nested knowledge graph from sample construction contract

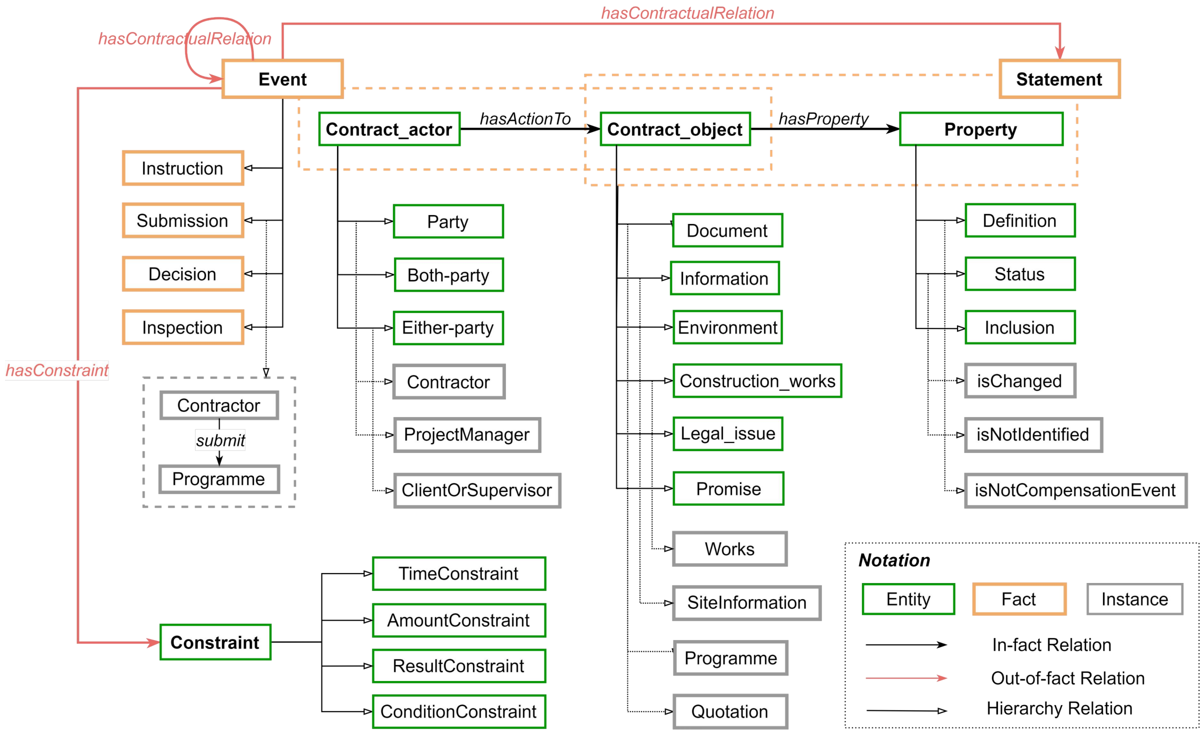


Figure 1: Ontological layer of nested contract KG

The core entities in the ontological layer of nested contract knowledge graph contains “Contract_actor”, “Contract_object”, “Property” and “Constraint”. “Contract_actor” and “Contract_object” class can be joined by the “hasActionTo” relation to form the basic “Event” class fact: <Contract_actor, hasActionTo, Contract_object>. For example, <Contractor, submit, Programme> is an instance of “Event” class, and more specifically, it is under the “Submission” subclass. Similarly, the description of the property of certain product is called a “Statement”. The fact <Programme, hasProperty, isNotIdentified> is an instance of “Statement” class.

In the contract clauses, a “Event” always has many constraints. The “Constraint” class can be attached to the “Event” class as <Event, hasConstraint, Constraint>. In real instances, the relation and tail entity should be replaced by specific constraint types such as <Event, hasTimeConstraint, beforeCompletionDate>. Finally, we have the “hasContractualRelation” edge to connect a Event and a Statement, or connect two Events. The

clause. First, entity extraction is performed by linking the identified entities to the predefined ontology classes. For example, “Contractor” and “Project Manager” belongs to Party class, while “Proposal” belongs to Document class. Accordingly, two Fact is recognized as (Contractor, make, Proposal) and (Project Manager, accept, Proposal), denoting two events in the contract. Finally, temporal relation “within two weeks of” is used to link the two facts and form the knowledge representation. The final RDF-star representation is shown in the rectangular box.

Following the ontology and construction workflow defined above, we can also develop automated approach to facilitate the knowledge graph construction. Given the in-context learning ability of LLMs, recent studies proposed using ontology prompting to generate structured knowledge representation. Motivated by such efforts, our instance development method adopts a semi-automated construction following knowledge representation from unstructured text[1]. This can reduce the intensive human efforts from constructing a knowledge graph to an acceptable level of manual intervention.

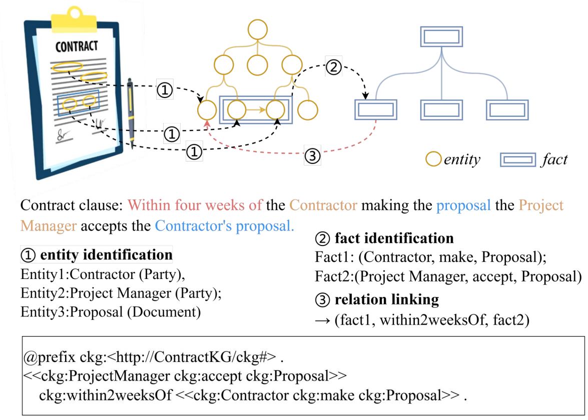


Figure 2: Schematic diagram of constructing nested KG from contract

Nested Contract KG query

The nested contract KG query allows the entity-level knowledge retrieval from the constructed domain knowledge database. This section will briefly introduce the construction of domain knowledge base and the contract knowledge query. The ontological layer of contract KG was created in Protégé, and the instance layer was stored in RDF-star language as a turtle file. The classes and instances were then imported into graph database such as GraphDB or Neo4j. Therefore, the query can be performed through SPARQL or Cypher language.

In practice, standard form of construction contracts (SFCCs) serves as a standard for regulating the responsibility and roles for stakeholders. They are the important knowledge for understanding and reviewing the contracts. We present sample queries of the SFCC KG as in Figure 3. In the first query in (a), contractor may want to know all of the time constraints for submitting documents. Through graph database, the documents and corresponding time period for submitting will be returned both in RDF and graph view. We can also explore interlinked concepts in the graph database. As shown in (b), we can perform a more complex query related to conditional relation of two "Event" class. A triple will be retrieved to indicate what a party should do in this case.

Q (TimeConstraint): What are the time constraints for documents that Contractor is responsible for submitting?

```
PREFIX ckg: <http://ContractKR/ckg#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>

SELECT ?o ?b WHERE {
BIND (<<ckg:Contractor ckg:submit ?o>> AS ?a)
?a ckg:hasTimeConstraint ?b .
}
```

(a)

Q (Behavior): What should parties do if quotation for an acceleration is accepted?

```
PREFIX ckg: <http://ContractKR/ckg#>
PREFIX rdfs:<http://www.w3.org/2000/01/rdf-schema#>

SELECT ?c WHERE {
BIND (<<ckg:QuotationForAnAcceleration
ckg:hasProperty ckg:isAccepted>> AS ?a)
?a ckg:if-then ?c .
}
```

(b)

Figure 3: Demonstration of query in contract KG

Augmenting LLMs with contract KG on automated contract review

In this section, we propose the overall framework to integrate nested knowledge graph in automated contract review using natural language prompting.

Figure 4 illustrate our contract KG-enhanced LLM contract review framework. The aim is to retrieve external knowledge in the nested contract KG for augmenting contract analysis task. Given a contract clause c, a set of natural language query $Q = \{q_1, q_2, \ldots, q_n\}$ will be generated, and subsequently transformed into SPARQL query $S = \{ s_1, s_2, \ldots, s_n \}$. External explicit knowledge $K = \{ k_1, k_2, \ldots, k_m \}$will be retrieved in RDF-star format and feed back to LLMs. The final prompt $P = \{r|c,\ k\}$ is designed to generate corresponding risk analysis result with the initial contract clause input and the retrieved knowledge.

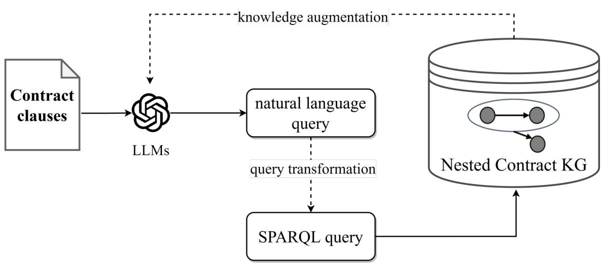


Figure 4: Contract KG-augmented LLM contract review pipeline

Specifically, in the first query generation stage, the language model is prompted to analyze the review task and output relevant queries based on the given contract clauses. The prompt is designed for asking the questions such as "What terms in this clause need further clarification?" and "What knowledge about the clause is required to determine the risk?". The output should be querying property of certain terms, constraints of contract

[1] https://bratanic-tomaz.medium.com/constructing-knowledge-graphs-from-text-using-openai-functions-096a6d010c17

events as external knowledge of the contract context. In this stage, we aim to stimulate the in-context learning ability of LLMs for identifying what external knowledge is needed in the context, to bridge the knowledge gap with experts. In the query transformation stage, the natural language to SPARQL query translation requires LLMs to map the concept in natural language query to classes in the defined knowledge graph. Multi-hop query can also be included in order to respond to the natural language query. After the knowledge retrieval from nested contract KG, the triple-formed knowledge and the initial contract clause will be input for prompting LLMs to generate the final result. This whole process performs in a retrieval-augmented fashion, providing external knowledge for the automated contract risk identification.

Case study

In this section, we validated our approach on contract clauses collected from international construction projects. The contract was thoroughly reviewed by contract experts, and the sample clauses-review pair was given as the gold standard, as shown in Table 2. The sample clause was input for risk identification and analysis in the standard LLM review (baseline) and KG-enhanced LLM contract review pipeline (our method), respectively. In the baseline setting, only contract clause was input for LLMs to identify the potential risk, while our method also provided the final output according to the process described in Figure 5. The evaluation process was manually conducted by domain experts, measuring both the risk label and the accuracy of the risk analysis.

The baseline output and the process of our approach is shown in Table 3. All the content was generated using gpt-3.5-turbo, which is a natural language understanding and generating model provided by OpenAI[2].

Table 2: Contract clause review case overview

Project clause& review demonstration
Contract clause: (a) Contract clause: The amount of the Advance Payment shall be twenty five (25%) of the Supply Price and shall be amortised proportionately against amounts recovered through interim applications for payment throughout the progress of the Supply, starting with the first Interim Payment after the Order Date. The currency shall be the same as the Supply Price. (b) The Employer shall pay the Advance Payment after: (i) receiving the Performance Bond in accordance with Clause 4.3(b), and (ii) receiving the Advance Payment Guarantee in accordance with Clause 4.3(b), provided, however, the Advance Payment shall be made within fifteen (15) days of the later of (i) and (ii). (c) The Supplier shall ensure that the Advance Payment Guarantee is valid and enforceable until the entire Advance Payment has been adjusted by the Supplier, but its amount shall be progressively reduced by the amount adjusted through proportional deductions from Interim Payment amounts paid for the Equipment under Clause 11.5 (Schedule of Payments) …
Risk review: No risk. The amount of Advance Payment is quite generous to the Contractor, which is much higher than the common practice of 10%. Requiring Performance Bond or Advance Payment Guarantee prior to paying the Advance Payment is reasonable practice to protect the interest of the Employer. The requirements for proportional deductions from Interim Payment and validity period are also common practice, which reminds the Contractor of updating the guarantee if necessary. It would be better if the contract could stipulate the Contractor's right when the release of the Advance Payment is delayed by the Employer.

[2] https://platform.openai.com/docs/models/overview

Table 3: Knowledge-enhanced LLM contract review

Automated risk analysis demonstration
Baseline output Risk. The clause introduces financial risk for the Contractor because the Advance Payment is tied to the provision of a Performance Bond and an Advance Payment Guarantee, which can be costly and administratively burdensome. Additionally, the Advance Payment must be repaid through deductions from interim payments, which could affect the Contractor's cash flow. The obligation to maintain the Advance Payment Guarantee until the Advance Payment is fully adjusted adds further financial strain and administrative responsibility on the Contractor.
Natural language query Explain Advance Payment Guarantee. [Q1] What is the condition of making advance payment? [Q2]

SPARQL query
Q1:
SELECT ?p ?actor1
WHERE{
ckg: advancePaymentGuarantee ckg:hasProperty ?p .
?actor1 ckg:submit ckg: advancePaymentGuarantee}
Q2:
SELECT ?actor2 ?c1 ?c2 WHERE{
BIND (<<?actor2 ckg:make ckg:advancePayment >> AS ?e)
?e ckg:hasTemporalConstraint ?c1 .
?e ckg:hasConditionalConstraint ?c2}

Retrieved knowledge from nested KG A1: ckg: Employ ckg:submit ckg: advancePaymentGuarantee . ckg: advancePaymentGuarantee ckg:hasProperty ckg: amountequalToAdvancePayment . ckg: advancePaymentGuarantee ckg:hasProperty ckg: currencyequalToAdvancePayment . A2: <<ckg: Employer ckg:receive ckg:PerformanceBond_and_ advancePaymentGuarantee >> ckg:within14days << ckg: Employer ckg:make ckg:advancePayment>>
KG-augmented output No risk. The clause matches the standard provision regarding the advance payment being conditional on receiving the Performance Bond and Advance Payment Guarantee. The payment timeline of 15 days, as opposed to 14 in the standard form, is slightly longer but does not significantly alter the contractor's obligations or risks. The requirement for the

Advance Payment Guarantee to be valid until the advance is fully adjusted is also standard, ensuring the employer's risk is mitigated without imposing unbalanced obligations on the contractor.

As shown in Table 3, the baseline output using standard prompting stated that the contractual risk lies in the precondition of providing performance bond and advance payment guarantee before advance payment, posing an unbalnaced obligation to the contractor. However, it is an unfounded claim as employer make payment based on receiving advance payment guarantee. Such hallucination arises due to the absence of knowledge related to relevant contract terms and provisions. In the retrieval process, the regulation of advance payment guarantee and constraints of making advance payment was retrieved as the standard provision. After knowledge-augmentation, the model answers “No risk”, with an explanation that objectively analyze the contract provision in comparison to the standard provision. The answer aligns with the expert review and showcases a more coherent and interpretable reasoning in the perspective of a contract expert.

Discussion

In this paper, we introduce the automated contract review method enhanced with contract knowledge graph. This method suggests integrating LLMs and KG to achieve a more interpretable and trustworthy automated contract review process. Although it is a feasible solution, there are some unsolved challenges and future promising directions that we would like to discuss here.

Firstly, there are also other approaches to augment LLMs, such as retrieving knowledge from websites and tools. They can enhance LLMs with real-time information or mathematics ability, which can be combined with domain knowledge graph for comprehensive knowledge augmentation.

Secondly, the current process is conducted semi-automatically, while further study is required for higher level and more accurate automation of ontology construction and query mapping. This might involve ontology learning and searching for the relevant reasoning path in KGs. It is noted that although achievements have been made in ontology extraction, current methods still struggle with complex multi-layer extraction, making human intervention for nested knowledge graph construction inevitable at the moment. The trade-off between expressivity and scalability should also be considered. More accurate and comprehensive knowledge representation comes with higher human intervention.

Thirdly, the current evaluation metrics does not consistently align with expert evaluations concerning the accuracy and coherence of risk reasoning output. This is because the existing benchmarks focus on the tasks where model learn by fine-tuning on task-specific data. Under the LLM era, constructing a benchmarks for construction contract risk analysis will be of great importance.

Conclusion

This paper presents a novel framework for augmenting large language models with structured domain knowledge to improve automated contract risk identification. A knowledge graph modeling approach is developed to capture the complex and interconnected relationships within contracts, along with the contract ontology and RDF-star representation. A framework that integrates knowledge from knowledge graph into the LLM reasoning pipeline using natural language prompting is also proposed. The case study demonstrates improved performance with our knowledge augmentation method on automated contract risk identification task.

Acknowledgement:

This research was funded by the National Natural Science Foundation of China (grant no. 71971196) and China Postdoctoral Science Foundation (No. 2023M743307).

References

Beach, T. H., Hippolyte, J.-L., & Rezgui, Y. (2020). Towards the adoption of automated regulatory compliance checking in the built environment. *Automation in Construction*, *118*, 103285. https://doi.org/10.1016/j.autcon.2020.103285

Candaş, A. B., & Tokdemir, O. B. (2022). Automating Coordination Efforts for Reviewing Construction Contracts with Multilabel Text Classification. *Journal of Construction Engineering and Management*, *148*(6), 04022027. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002275

Chakrabarti, D., Patodia, N., Bhattacharya, U., Mitra, I., Roy, S., Mandi, J., Roy, N., & Nandy, P. (2018). Use of Artificial Intelligence to Analyse Risk in Legal Documents for a Better Decision Support. *TENCON 2018 - 2018 IEEE Region 10 Conference*, 0683–0688. https://doi.org/10.1109/TENCON.2018.8650382

Duan, Y., Shao, L., Hu, G., Zhou, Z., Zou, Q., & Lin, Z. (2017). Specifying architecture of knowledge graph with data graph, information graph, knowledge graph and wisdom graph. *2017 IEEE 15th International Conference on Software Engineering Research, Management and Applications (SERA)*, 327–332. https://doi.org/10.1109/SERA.2017.7965747

Fang, W., Ma, L., Love, P. E. D., Luo, H., Ding, L., & Zhou, A. (2020). Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. *Automation in Construction*, *119*, 103310. https://doi.org/10.1016/j.autcon.2020.103310

Fernandez, J. V. (2023). Artificial Intelligence in Government: Risks and Challenges of Algorithmic Governance in the Administrative State. *Indiana Journal of Global Legal Studies*, *30*, 65.

Hartig, O., & Thompson, B. (2021). *Foundations of an Alternative Approach to Reification in RDF*. http://arxiv.org/abs/1406.3399

Hassan, F. ul, & Le, T. (2021). Computer-assisted separation of design-build contract requirements to support subcontract drafting. *Automation in Construction*, *122*, 103479. https://doi.org/10.1016/j.autcon.2020.103479

Hassan, F. ul, Le, T., & Lv, X. (2021). Addressing Legal and Contractual Matters in Construction Using Natural Language Processing: A Critical Review. *Journal of Construction Engineering and Management*, *147*(9), 03121004. https://doi.org/10.1061/(ASCE)CO.1943-7862.0002122

Hogan, A., Blomqvist, E., Cochez, M., d'Amato, C., de Melo, G., Gutierrez, C., Gayo, J. E. L., Kirrane, S., Neumaier, S., Polleres, A., Navigli, R., Ngomo, A.-C. N., Rashid, S. M., Rula, A., Schmelzeisen, L., Sequeda, J., Staab, S., & Zimmermann, A. (2021). Knowledge Graphs. *ACM Computing Surveys*, *54*(4), 1–37. https://doi.org/10.1145/3447772

Kumar, V. K. (2019). Knowledge Representation Technologies Using Semantic Web. In *Web Services: Concepts, Methodologies, Tools, and Applications* (pp. 1068–1076). IGI Global.

Lee, J., Ham, Y., Yi, J.-S., & Son, J. (2020). Effective Risk Positioning through Automated Identification of Missing Contract Conditions from the Contractor's Perspective Based on FIDIC Contract Cases. *Journal of Management in Engineering*, *36*(3), 05020003. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000757

Lee, J., Yi, J.-S., & Son, J. (2019). Development of Automatic-Extraction Model of Poisonous Clauses in International Construction Contracts Using Rule-Based NLP. *Journal of Computing in Civil Engineering*, *33*(3), 04019003. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000807

Li, X., Yang, D., Yuan, J., Donkers, A., & Liu, X. (2022). BIM-enabled semantic web for automated safety checks in subway construction. *Automation in Construction*, *141*, 104454. https://doi.org/10.1016/j.autcon.2022.104454

Lv, X., Shi, J., Cao, S., Hou, L., & Li, J. (2022). Triple-as-Node Knowledge Graph and Its Embeddings. In A. Bhattacharya, J. Lee Mong Li, D. Agrawal, P. K. Reddy, M. Mohania, A. Mondal, V. Goyal, & R. Uday Kiran (Eds.), *Database Systems for Advanced Applications* (Vol. 13245, pp. 107–121). Springer International Publishing. https://doi.org/10.1007/978-3-031-00123-9_8

Mahfouz, T., Kandil, A., & Davlyatov, S. (2018). Identification of latent legal knowledge in differing site condition (DSC) litigations. *Automation in Construction*, *94*, 104–111.

Orlandi, F., Graux, D., & O'Sullivan, D. (2021). Benchmarking RDF Metadata Representations: Reification, Singleton Property and RDF. *2021 IEEE 15th International Conference on Semantic Computing (ICSC)*, 233–240. https://doi.org/10.1109/ICSC50631.2021.00049

Park, N., Liu, F., Mehta, P., Cristofor, D., Faloutsos, C., & Dong, Y. (2022). EvoKG: Jointly Modeling Event Time and Network Structure for Reasoning over Temporal Knowledge Graphs. *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, 794–803. https://doi.org/10.1145/3488560.3498451

Salama, D. M., & El-Gohary, N. M. (2016). Semantic Text Classification for Supporting Automated Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, *30*(1), 04014106. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000301

Sleimi, A., Ceci, M., Sannier, N., Sabetzadeh, M., Briand, L., & Dann, J. (2019). A Query System for Extracting Requirements-Related Information from Legal Texts. *2019 IEEE 27th International Requirements Engineering Conference (RE)*, 319–329. https://doi.org/10.1109/RE.2019.00041

Sleimi, A., Sannier, N., Sabetzadeh, M., Briand, L., Ceci, M., & Dann, J. (2021). An automated framework for the extraction of semantic legal metadata from legal texts. *Empirical Software Engineering*, *26*(3), 43. https://doi.org/10.1007/s10664-020-09933-5

Xu, X., & Cai, H. (2019). Semantic Frame-Based Information Extraction from Utility Regulatory Documents to Support Compliance Checking. In I. Mutis & T. Hartmann (Eds.), *Advances in Informatics and Computing in Civil and Construction Engineering* (pp. 223–230). Springer International Publishing. https://doi.org/10.1007/978-3-030-00220-6_27

Yang, L., Chen, H., Li, Z., Ding, X., & Wu, X. (2024). Give Us the Facts: Enhancing Large Language Models with Knowledge Graphs for Fact-aware Language Modeling. *IEEE Transactions on Knowledge and Data Engineering*, 1–20. https://doi.org/10.1109/TKDE.2024.3360454

Zait, F., & Zarour, N. (2018). Addressing Lexical and Semantic Ambiguity in Natural Language Requirements. *2018 Fifth International Symposium on Innovation in Information and Communication Technology (ISIICT)*, 1–7. https://doi.org/10.1109/ISIICT.2018.8613726

Zhang, J., & El-Gohary, N. M. (2016). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. *Journal of Computing in Civil Engineering*, *30*(2), 04015014. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000346

Zhang, S., Boukamp, F., & Teizer, J. (2015). Ontology-based semantic modeling of construction safety knowledge: Towards automated safety planning for job hazard analysis (JHA). *Automation in Construction*, *52*, 29–41.

Zhong, B., Wu, H., Li, H., Sepasgozar, S., Luo, H., & He, L. (2019). A scientometric analysis and critical review of construction related ontology research. *Automation in Construction*, *101*, 17–31. https://doi.org/10.1016/j.autcon.2018.12.013

Zhou, P., & El-Gohary, N. (2016). *Automated Extraction of Environmental Requirements from Contract Specifications*.

Zhou, Z., Goh, Y. M., & Shen, L. (2016). Overview and Analysis of Ontology Studies Supporting Development of the Construction Industry. *Journal of Computing in Civil Engineering*, *30*(6), 04016026. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000594

Zou, Y., Kiviniemi, A., & Jones, S. W. (2017). Retrieving similar cases for construction project risk management using Natural Language Processing techniques. *Automation in Construction*, *80*, 66–76. https://doi.org/10.1016/j.autcon.2017.04.003

THE ANATOMY OF AN ARCHITECT'S ARGUMENT: FORMALLY CAPTURING SOCIALLY-ORIENTED DESIGN INTENTIONS IN THE BUILT ENVIRONMENT

Yazan N. H. Zayed[1], Anna Elisabeth Kristoffersen[1], Gustaf Lohm[1], Aliakbar Kamari[1], and Carl Schultz[1]
[1]Aarhus University, Aarhus, Denmark

Abstract

Sustainability is a large driving factor in today's society. This work focuses on assessing aspects contributing to the social sustainability of buildings. Social aspects encompass, for example, well-being and the building's impact on the occupants regarding qualities such as air quality, visual and thermal comfort, and privacy among others.

We introduce *"ProFormalize"*, a formalization framework, for capturing the structure of design arguments regarding social intentions. Moreover, presenting the formalization of design intentions and activities obtained through case-study focused interviews. In addition, the organization and formalization of design arguments are presented as preparation for software implementation.

Introduction

Building designs are often full of intentions and visions for how the building should be experienced and perceived by its users. These intentions are typically tied together with the physical aspects of the building in the form of different design choices. Examples of this can be that building users feel at home and relaxed in specific areas of the building because the area contains lounge furniture of good quality, or the building users feel a sense of ownership towards the building because they feel that the used decorations and artworks represent them. These intentions are introduced by different stakeholders through the design phases, based on the different perspectives added by stakeholders to the project (Entwistle, 2022). Throughout the changing stages of the building design process, the responsibility shifts between stakeholders (Säwén, 2023), which often leads to the risk of losing design intentions due to different stakeholders' focuses, and that these intentions are often implicit and not explicitly stated in the project documentation.

To effectively address this problem, we present the initial work of developing a formalization framework for design intentions regarding social aspects. The two main research questions in this work are:

RQ1 How can socially-oriented design arguments be captured in the digital representation of buildings?

RQ2 How can socially-oriented design arguments be represented in a way that prepares such kind of data to be implemented in software?

To address these questions, we present the first proof of concept of our framework, and the main contributions are:

C1 Setting up the overall workflow of our formalization framework *"ProFormalize"* in the form of four activity stages (elicitation, organization, formalization and implementation), and the corresponding output of each stage. This contribution is presented in "The formalization framework - ProFormalize" Section, and it addresses RQ1.

C2 Formalization of a selection of socially-oriented design arguments and presenting them in the form of First-Order-Logic (FOL) formulas and Abstract Syntax Trees (ASTs) in order to be used in further software implementation. This contribution is presented in "Analysis of case studies" Section, and it addresses RQ2.

Background

Sustainability is a large driving force in today's society and is influencing decisions and practices across numerous industries. One of these industries is building and construction, where sustainability is commonly addressed based on a large range of building certification systems (Jensen et al., 2018). Existing building certification systems consist of varying criteria, which are related to each of the three pillars of sustainability: environmental, economic and social sustainability (Jensen et al., 2018).

Out of these three pillars, the social pillar is the least developed and the most vaguely defined (Bebbington and Dillard (2009); Boström (2012); Kristoffersen et al. (2024)). The existing definitions and approaches towards social sustainability do not offer neither a widely accepted definition nor a universally accepted list of included aspects (Bebbington and Dillard (2009); Boström (2012); Kristoffersen et al. (2024)).

However, through our recent extensive survey on social sustainability, we have determined that there exists an agreement that social sustainability is focused on humans, and that definitions often include aspects such as, but not limited to, the forming of social networks, health, safety and justice (Kristoffersen et al. (2024)). This lack of a commonly agreed-upon definition can lead to a hollow shell concept of social sustainability due to its large adaptability to different contexts.

Instead of using the term *social sustainability*, it has been proposed to use the term *social values* followed by the specification of which exact value and to whom (Kristoffersen et al. (2024)). The research presented in this paper,

therefore, embraces this word choice to describe the effects on socially-oriented design intentions.
By doing that, we encourage a higher degree of transparency towards what effects are experienced in a building, and thus avoid obscuring these effects within the concept of improving social sustainability.

The formalization framework - ProFormalize

In this section we introduce our overall formalization framework "ProFormalize", and we explain the core purpose of, and activities within, each stage, as well as how the stages connect to each other. We give particular focus to how we formalize socially-oriented design intentions following our so-called "three-level approach".

Deriving the "anatomy" of architect arguments

A socially-oriented design intention argument represents a design activity that is intended to evoke a certain feeling of the occupants regarding their presence in the building or to improve their experience and perception of a specific social intention in that building.
This shows a **causal logical relationship** between "building elements" and "social intentions", that is, the design activity that architects decide to implement, influences socially-oriented design intentions and contributes to the occupants experience in the building. This relationship is illustrated in Figure 1.
For example, "installing artwork in the workplace will increase the sense of belonging to the workplace". This is a socially-oriented design intention argument. The underlying structure of this argument can be derived as follows:
The argument addresses a specific social intention which is the ultimate goal that the architect is trying to accomplish, that is, evoking a sense of belonging to the workplace among building occupants. The other part of the argument defines a design activity that is implemented in order to achieve the mentioned socially-oriented design intention, that is, installing artwork.
This anatomy of design arguments produces our "three-level argument structure", which is a core foundation of how we formalize socially-oriented architectural arguments by making them unambiguous, precise, and explicitly integrated into the digital representation of buildings.
Through the three-level approach, we define three levels of formalization; *the goal level*, *the domain level* and *the product level*. This approach is inspired by the work of Lauesen (2002), where the author introduced the *goal-design scale* that consists of four levels of requirements: *the goal level*, *the domain level*, *the product level* and *the design level*.
In our work, we focus on social intentions in the context of the built environment, that is among others, the set of emotions, behaviors, functions and experiences that people have in buildings.
The goal of this approach is to formalize the socially-oriented design intention arguments of architects and other early design stakeholders about the impact of different arrangements of objects and building elements on the social aspects experienced by building occupants.
Such aspects include sensory aspects (visibility, hearing, etc.), behavioral (movement, functions, etc.), emotional (excitement, privacy, inspiration, etc.).
We define the arguments of interest in our work as the arguments that contain information about both social intentions and the building elements used to affect them.
Consider the following arguments:

1. Installing LED lights in the working area would reduce energy consumption.
2. Organizing regular meetings with the business owners would increase the sense of belonging to the workplace.
3. Installing pieces of art in the workplace would increase the sense of belonging to the workplace.
4. Organizing regular meetings with the owners would improve business performance.

Based on the definition of arguments of interest in our work we can classify the above arguments as follows:
The first argument contains information about a type of building element (the LED lights) but the intended goal is not related to a social aspect but rather environmental and economical, hence this argument does not fit our formalization approach.
The same is true for the second argument as well, where the intended goal is related to social aspects, however the way to achieve this goal is not related to the arrangement of building elements.
The third argument is the one that fits our formalization approach. In this argument, the intended goal is to increase the sense of belonging to the place, which is a social intention, and the proposed way to achieve that, is to install pieces of art, which are building elements.
The fourth argument does not fit our formalization approach, as it does not contain any information about how a certain arrangement of building elements would achieve a social intention.
From the above definition, there are at least two basic levels of formalization that should exist for each design intention argument. These two levels are *the social intentions level* and *the building elements level*.
We use these two basic levels to form a reasonable way to initiate the structure of our formalization approach. The social intentions level is called *the goal level*, and the building elements level is called *the product level*.

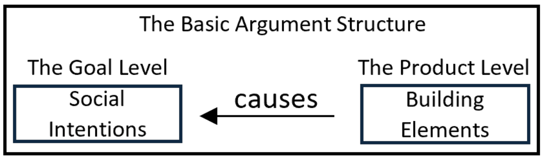


Figure 1: The causality between the two building blocks of the basic structure of a design argument: social intentions and building elements

This provides the basic structure of the formalization approach, that is to classify parts of the arguments into these two levels. The first one includes information about building elements such as walls, doors, slabs, lighting, art, furniture, etc. The second one includes information about the ultimate goal to be achieved which must be a social intention representing the feelings, emotions and behaviors of building occupants.

We argue that there is an implicit reasoning between these two basic levels of formalization, and we should introduce an intermediate level. This would be the smallest number of categories that we can work with in order to introduce a universal formalization approach that can handle a wide range of design intention arguments.

This intermediate level represents the stage between how building elements are arranged and the emotional and the behavioral experience is perceived by building occupants. It explains how changes in the design and in the arrangement of building elements will go through their senses and impact their interaction with their surroundings.

For example, considering the third argument again, "installing pieces of art in the workplace would increase the sense of belonging to the workplace". It is not enough to install pieces of art for the building occupants to have a sense of belonging to the workplace. Although the argument fits the basic approach criteria, it is important to explain how installing this art would create this experience. Where should the building occupant stand in the room to see the art and be triggered by it? What if the building occupant is visually impaired, does not that limit the area of the room from which the art is visible?

These are valid questions, and they help explain the flow of the design argument. Such questions and more are answered with the addition of the third level, *the domain level*, which has information about the sensory experience of the occupant and the functional spaces in the building.

To summarize, *the goal level* contains information about the social intention that is under consideration, and **why** architects are implementing a certain design activity. *The domain level* explains **how** the effect is going to be achieved and *the product level* explains **what** building elements are being used and controlled to achieve the goal of the socially-oriented design intention.

The workflow of ProFormalize

Figure 2 shows the four stages of the workflow and the corresponding output of each stage.

The work in this study is data-driven and case-study-based. The cases are selected related to EU funded project (Probono) of buildings located at Aarhus University Campus, and they are:

1. The Library basement.
2. The basement floor of the Molecular Biology Department.
3. Human Resources Department.
4. Center of Education Development.
5. Institute of Advanced Studies.
6. University Startup Hub (The Kitchen 2.0).

In the following, the four stages of the framework from Figure 2 are elaborated.

Stage (1) Elicitation of socially-oriented design intentions from existing projects through semi-structured interviews. The elicitation of design intentions was conducted through semi-structured interviews with early-design stakeholders, in this case, an architect who has been involved in the six projects mentioned above. The interview was planned based on the guidelines and insights by Robson and McCartan (2016) as well as Kvale and Brinkmann (2015). The interview was planned to have a narrative and anecdotal approach. This is done to encourage the interviewee to focus on what has actually happened and been done and not on what the interviewee wished for or would like to have done (Blommaert and Jie, 2020).

The interviewer developed an interview guide and updated herself on projects connected to the interviewee before the interview was conducted. A generic excerpt from the interview guide is presented in Table 1.

The approach utilised in practice in the interview has similarities with an unstructured interview since the interviewee was allowed to follow his train of thought and move freely between talking about different building projects and social intentions. In cases where the interviewee strayed away from the focus point of the interview, the interviewer utilised one of the prepared questions from the interview guide to reestablish the focus of the interview.

Table 1: Interview Guide

Question type	Question examples
Main Questions	• Which building and design projects have you been involved in? • Can you tell me a bit about the thought behind the project?
Elaborating Questions	• What was the intention? • How was the intention integrated into the project? • Has the intention been successful? • Is the intention adding the expected value to the building? • Have you gotten any response from the users?

The interview language was selected based on which languages the participants in the interview understand, and which working language the early-design stakeholders use. Whenever possible, the interviews were conducted in the same language as the main working language of the interviewee, to encourage the interviewee to use the phrases which were included in the project during the early design stages. The interview was recorded using a voice recorder on the interviewer's phone. This was done to let the interviewer and interviewee focus on the dialogue without slowing down or pausing the interview to take notes. The recording was transcribed in the interview language, either Danish or English, by first using Microsoft Word's

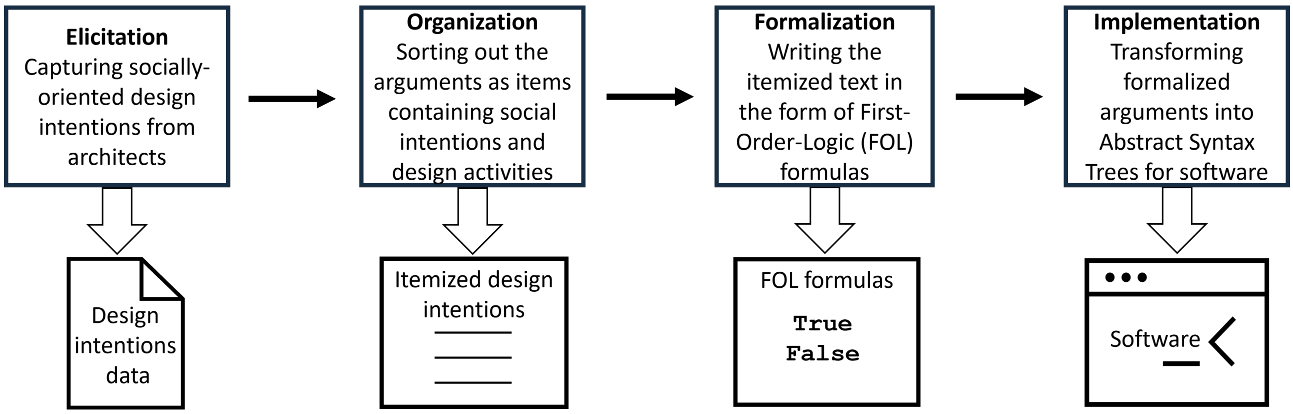


Figure 2: The workflow of ProFormalize

transcribe function, after which the interviewer edited the transcript to remove the errors originating from the transcription's automation.

Stage (2) Organization of the design intention arguments and sorting them out by identifying socially-oriented design intentions and their corresponding design activities to achieve those intentions.

In this stage we categorized the transcription from the elicitation stage into items containing two parts: **the socially-oriented design intentions** and **the design activities**.

This was done by following the four sub-stages: Identification of intents, translation, systematization/ input of data and validation.

The identification of intents was performed by reading the transcribed interview, and marking when the interviewee states a socially-oriented design intention or of an architectural or design-oriented decision. The essence or key points of the intent were then translated into English by the interviewer if the interview was conducted in Danish. After this, the intention was added to the database of the collected intentions, which is managed in an Excel sheet.

The database contains information about in which building the intention was implemented, the activity used to implement the intent, in which interview the intent was identified, and a timestamp referring to when in the interview the intent was mentioned. The references to which interview and the timestamps were recorded to make a direct connection between the database entries, the transcription and the interview recording. The identified and translated intents were sent for approval and review from the interviewee to ensure that the English translation of the intent represents and encompasses what the interviewee meant.

Stage (3) Formalization of the itemized design intention arguments.

In this stage, architects design arguments were represented explicitly, including the categories of targeted building occupants, their senses that the design activity is trying to evoke and the surrounding environment or situation.

The core of this level is to systematically identify implicit and ambiguous aspects of each design intention argument and systematically decide how to disambiguate them.

In order to do that, each statement was manually translated into first-order-logic formulas based on the three-level approach presented in "The formalization framework - ProFormalize" Section. This makes particular aspects of the argument explicit such as occupant sense-experience, the surrounding environmental conditions, etc.), moreover, natural language concepts are expressed as predicates and terms, and combined into formal logical statements.

In addition, qualitative spatial prepositions are grounded with definite interpretations that at the same time allow architects to adjust them based on their reasoning.

Hence, this approach provides architects with a structured form of building blocks to present their arguments in a valid way and shows them what different terms they can use in each level.

The three levels of formalization are defined as follows:

The goal level contains predicates about **social intentions**. Those predicates refer to the aspects that affect the experiences and the feelings of building occupants. The goal level is the only level that can take such predicates, as it is the level that represents the objective of the design activity and describes the accurate intention behind that activity.

The domain level contains predicates about **building occupants and their senses**, this level takes predicates from two major categories which are:

1. **Spatial artefacts**, those are regions of empty space that contain important semantic value because they represent the space in which an object can be seen, heard, used, etc. For example, the movement space of a room, which is the space where people can move within a room, or the visibility space and the hearing space of humans, they refer to the region of empty space where humans can see or hear the visual or acoustic effect of a building element, respectively. (Bhatt et al. (2010)), (Bhatt et al. (2012)).
2. **Spatial prepositions**, those are the terms that represent the direction, orientation, or position information of spatial artefacts and building elements, such as in front of, between, near, inside, directed at, etc.

The product level contains predicates about **building elements** that are going to be used or arranged in a specific way to affect a specific social intention.
Figure 3 demonstrates this structure and presents some examples of each predicate category.
Stage (4) Implementation of the formal design arguments as Abstract Syntax Trees for software development.
In this final stage we performed a direct one-to-one translation from the formulas into first-order abstract syntax trees to be used in software implementation.

The Goal Level	The Domain Level			The Product Level
Social value predicates	Spatial artefact predicates	Spatial preposition predicates	Building occupant predicates	Building element predicates
privacy()	visible_space()	inside()	occupant()	art()
curiosity()	hearing_space()	between()	student()	wall()
belonging()	function_space()	near()	elderly()	use_art()
.	.	.	.	.
.	.	.	.	.

Figure 3: Predicate rules of the three-level approach

Analysis of the case studies

In this section, we discuss in more detail the case study that is the focus of this research, which originates from a project conducted at the building housing the Department of Molecular Biology at university campus (institution name removed for blind review). The building houses students, researchers and other university employees. Most of the collected intents focus on how the students use and feel in the building, since students are the primary users of the areas which the interviewed architect has focused on.

Elicitation

The elicitation of socially-oriented design intentions has been conducted through semi-structured interviews in accordance with the description of the Elicitation stage in the Section "The Workflow of ProFormalize". The intentions were collected through three interviews over 14 months.
The interviewee in all the interviews was the same architect, who has been participating in the planning and implementation of the intentions into the building. The first two interviews were conducted in Danish as sit-down interviews with only the architect and the main interviewer participating. The third interview was conducted walking through the Department of Molecular Biology at the university, where the architect showed the building and the activities tied to the socially-oriented design intentions.
The interview was conducted by the same main interviewer as the previous two, with a secondary interviewer joining the interview with the goal of documenting design choices through pictures and indications on a floor plan drawing. The interview was conducted in English to accommodate the secondary interviewer. The result of this stage forms the basis for the identification, organisation and processing through the following steps.

Organization

The organization of the data consists of identifying socially-oriented design intentions in the transcribed interview, and presenting the intentions in an organized manner that can serve as a stage towards concertising the intents.
The dark red points on the floor plan shown in Figure 4 represent the locations of socially-oriented design intention items which were collected through the on-site interview and have been transcribed and organised. These points are explained in Table 2 and presented in Figure 5.
The dark blue points on the floor plan represent the locations of socially-oriented design intention items that have not been transcribed or organised yet, this includes wall art (a and c), changing wall colours (b), and specific furniture (d and e).

Figure 4: Floor plan of the basement of the molecular biology building, showing locations of organised design intention items (dark red) and the not yet organised items (dark blue)

Formalization

The three levels are represented as first-order-logic formulas, this is explained through the following example of a design intention argument about the case study of the molecular biology building: *"A delimited area of the basement in the building is being equipped with real, one-of-a-kind artworks, the intention with this space is to make an area where the students feel inspired, and actually want to be sitting"*.
Based on the three-level approach we can classify the following parts of the argument as follows:

1. The social intention under consideration is *"sense of inspiration"*, this represents *the goal level*. And this is the intention behind the architect's design decision.
2. The way to achieve the goal is to place the artworks in the area to be visible by the students, this represents *the domain level*.

Table 2: Items of design intentions and design activities

Id	Socially-oriented design intention	Design Activity
*DI*1	The art is growing and developing when you look at it and creating a space where the students feel inspired and actively want to sit.	Installing art consisting of four lamps emitting different light of different temperatures and intensities.
*DI*2	Creating an area where the students actively want to sit.	Installing a sofa that is adapted to the room size, taking accessibility into consideration.
*DI*3	Creating an area where the students actively want to sit.	Installing sockets on the sofa which are directly integrated through a conscious design choice in the sofa.
*DI*4	Creating a subspace within a larger area and providing a feeling of coziness.	Using high contrast colors of the sofa and the wall between the general area and the subspace.
*DI*5	Draw attention and encourage mobility and inspiration to explore the building.	Painting a line of color wrapped around the corner.
*DI*6	Draw attention and encourage mobility and inspiration to explore the building.	Installing a contrasting color expanding into the hallway.

3. The building elements that are under consideration are the *"artworks"*, they represent *the product level*.

The goal level formula: increasing their sense of inspiration of the occupant P in room R:

$$sense\_of\_inspiration(P,R)$$

The domain level formula: the occupant P must be inside the visible space of the artwork A to see it:

$$occupant(P) \wedge inside(P, visible\_space(A))$$

The product level formula: using artwork A in room R:

$$art(A) \wedge room(R) \wedge use\_art(A,R)$$

As our approach is based on case studies of actual buildings, this approach is under constant development and there might be a need to add more levels to achieve a different level of detail, but for the current moment, as it is unnecessary to add more levels, we use this three-level approach, and we could represent different design arguments in a reasonable way.
In the previous example, 6 predicates and 1 function have been used, this is explained as follows:

1. A predicate named "art" is used to define an artwork as a building element. It takes 1 argument, which is the variable representing the artwork.
2. A predicate named "room" is used to define a room in the building. It takes 1 argument, which is the variable representing the room.
3. A predicate named "use_art" is used to specify which art piece is used where. It takes 2 arguments, the first is the piece of artwork, and the second is the place where this art is installed.
4. A predicate named "occupant" is used to define a general building occupant. It takes 1 argument, which is the occupant.
5. A predicate named "inside" is used as a spatial preposition to describe that an element is inside a specific space. It takes 2 arguments, the first is the element itself and the second is the space that the element is inside.
6. A function named "visible_space" is used to represent the space from which a specific element is visible. It takes 1 argument, which is the element for which the visible space is needed.
7. A predicate named "sense_of_inspiration" is used to describe that an occupant has a sense of inspiration regarding a specific building element. It takes 2 arguments, the first is the occupant and the second is the building element.

There is also a term named "Var" that takes one argument only and it is used to represent a building element or an occupant as variables.

Implementation

In the implementation stage, we transform the formalized first-order-logic formulas into abstract syntax trees as a preparation for the software stage. In the abstract syntax tree representation, there are the following elements:

1. **Logical expressions**, that evaluate to either *true* or *false*. This includes logical operators (*and, or, not*) and predicates (*Pred*).
2. **Terms**, that includes functions (*Func*) that is applied to other terms and they evaluate to a new term, and variables (*Var*) that are used to represent different terms.

Figure 5: Images demonstrating the design intention arguments illustrated in Table 2

The following is an example of the abstract syntax tree of the domain level:

```
And(Pred("occupant", 1, [Var("P")]),
 And(Pred("building_element", 1, [Var("A")]),
     Pred("inside", 2,
       [Var("P"),
       Func("visible_space", 1, [Var("A")])])
))
```

Discussion and conclusions

In this paper we presented the initial stage of "ProFormalize", a formalization framework to capture socially-oriented design intentions in buildings.
The underlying structure behind architects socially-oriented design intention arguments was derived and the three-level structure was presented.
In the subsequent phase of this work, the implementation stage of the proposed framework will be further developed. One potential service that this framework provides is to query about socially-oriented design intentions that are related to specific building elements, and vice versa.
Figure 6 shows a floor plan of the basement floor of the molecular biology department highlighting a selected path that an occupant might take to reach their working area. Along the shown path, the occupant passes through several locations where architects have implemented certain design activities using various building elements in order to improve or add a specific feeling or experience.
For example, in design intention (DI5), the architect argue that painting a line of color around the corner will draw attention and encourage occupants to move around the building and explore it. This location is highlighted with the blue point with coordinates (13,3) in the floor plan.

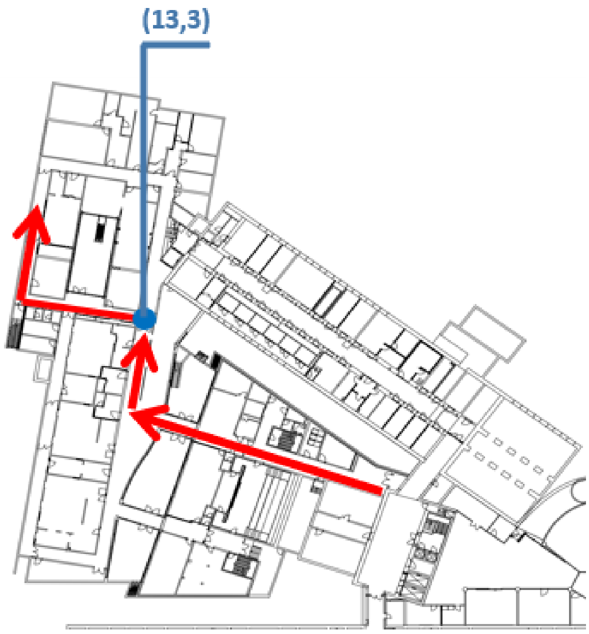


Figure 6: The floor plan of the basement of the molecular biology department, showing a selected walking path of a building occupant and the location of DI5

As an implementation of the framework, the following query example is presented:

```
>>>relation(DI5,building_element).

building_element = wall3265
building_element = paint4873
```

The predicate "relation" is referring to the relationship between building elements and design intentions. The goal of this query is to list the building elements that are used to achieve design intention 5 (DI5), those elements are listed using their unique ids as (wall3265) and (paint4873).
Another query example is to list the design intention items that are affected by a specific building element.
The following query lists the design intention items that are affected by the sofa with the unique id (sofa0012).

```
>>>relation(DI,sofa0012).

DI = DI2
DI = DI3
```

In addition, the user can query about the design intentions that will be experienced according to the argument.

For example, the path shown in Figure 6 is defined as the line segments connecting the set of 2D points:
Path = [(0,0),(10,1),(10,3),(13,3),(13,5)], then the following query can be implemented:

```
>>>social_intention(Path,DIs).
[(at((13,3)),social_intention(inspiration),id(DI5))]
```

The "social_intention" predicate takes two arguments, one is the selected path represented as the set of points and the second argument represents the design intentions that are located along the path.
The output of the predicate will be a list of elements, where each element contains a point, a social intention represented using the "social_intention" predicate, and the design intention represented using the "id" predicate.
This provides architects a way to find out which building elements affect a specific socially-oriented design intention, and reduces the risk of losing some intentions when changing the use or arrangement of the building elements through different phases of the design process.
Additional querying services might also include the visualization of the functional space of certain equipment or service areas such as printing facilities, storage rooms, working stations, among others. This space represents the area from which this equipment or service zone can be used safely, efficiently, privately, etc.
This framework was developed based on case studies of actual buildings that are mainly used by students or employees of an educational institution. Further applicability of this framework include governmental buildings (e.g., offices) and the scale of applications could be expanded to cover multiple buildings located in a close area up to a neighborhood.
The main limitation of this work was the limited amount of collected socially-oriented design arguments. As part of the future work, more locations are planned to be visited and more interviews are planned to be conducted with architects who are working on relevant projects concerning the addition or improvement of social criteria in buildings.

Acknowledgments

We gratefully acknowledge the funding support of the European Union Horizon 2020 research project PROBONO under grant agreement no. 101037075.

References

Bebbington, J. and Dillard, J. (2009). Social sustainability: An organization level analysis. In Understanding the social dimension of sustainability., pages 157–173. Taylor and Francis.

Bhatt, M., Hois, J., Kutz, O., and Dylla, F. (2010). Modelling functional requirements in spatial design. In Conceptual Modeling–ER 2010: 29th International Conference on Conceptual Modeling, Vancouver, BC, Canada, November 1-4, 2010. Proceedings 29, pages 464–470. Springer.

Bhatt, M., Schultz, C., and Huang, M. (2012). The shape of empty space: Human-centred cognitive foundations in computing for spatial design. In 2012 IEEE symposium on visual languages and human-centric computing (VL/HCC), pages 33–40. IEEE.

Blommaert, J. and Jie, D. (2020). Ethnographic fieldwork: A beginner's guide. Multilingual Matters.

Boström, M. (2012). A missing pillar? challenges in theorizing and practicing social sustainability: introduction to the special issue. Sustainability: Science, practice and policy, 8(1):3–14.

Entwistle, J. M. K., . R. M. K. (2022). Fra indsigt til forandring: Et indblik i en arkitekturantropologisk praksis. jordens folk. - text is in danish.

Jensen, K. G., Birgisdottir, H., Poulsgaard, K. S., Lind, L., Christensen, C. Ø., Skjelmose, O., Carruth, S. J., Jensen, K. K., Canera, I. O., Manbodh, J., et al. (2018). Guide to sustainable building certifications.

Kristoffersen, A. E., Schultz, C., and Kamari, A. (2024). A critical comparison of concepts and approaches to social sustainability in the construction industry. Journal of Building Engineering [Under review].

Kvale, S. and Brinkmann, S. (2015). Interviews. Sage.

Lauesen, S. (2002). Software requirements: styles and techniques. Pearson Education.

Robson, C. and McCartan, K. (2016). Real world research: A resource for users of social research methods in applied settings.

Säwén, T. (2023). Early stage architectural design practice perspectives on life cycle building performance assessment. Licentiate thesis. Chalmers University of Technology.

SINGLE POINT OF CONTACT: DENSITY AS A MAPPING PARAMETER FOR BUILDING MATERIAL DATABASES

Toivo Säwén*[1], Pierre Estève-Bourrel[23], Isac Mjörnell[1],
Alina Galimshina[12], Guillaume Habert[2], and Alexander Hollberg[1]
*Corresponding author: sawen@chalmers.se
[1]Dept. of Architecture and Civil Engineering, Chalmers University of Technology, Gothenburg, Sweden
[2]Institute of Construction and Infrastructure Management, ETH Zürich, Switzerland
[3]School of Engineering, The University of Edinburgh, UK

Abstract

Integrating building performance analysis and life cycle assessment in design requires accurate material databases interfacing different analysis engines. However, mappings of databases developed for different purposes are missing. We evaluate whether beyond material name or ID, density can be used as a mapping parameter between environmental, energy, daylight, and structural databases in two cases in Switzerland and Sweden. In total, 16 out of 19 materials could be confidently mapped based on the information available. The variability of the mapping parameter, density, was ±15 % for structural materials, and up to 64 % for insulation materials. The lasting contribution of this work is a consistent methodology for mapping between material databases.

Introduction

To deal with the major environmental impact of the construction industry, the life cycle impact of buildings needs to be considered and reduced during their design (Röck et al., 2020). This needs to be done without overlooking the functional performance of the buildings (Khoshnava et al., 2020). This calls for the integration of environmental Life Cycle Assessment (LCA) and functional Building Performance Analysis (BPA), including structural analysis, approaches into holistic assessment workflows (Säwén, 2023). Such workflows should be introduced in early design stages to achieve the maximum benefit to the final performance of the design (Hollberg et al., 2018; Meex et al., 2018).

For this purpose, several tools both for BPA (Säwén et al., 2022a) and LCA (Säwén et al., 2022b) have been developed. These tools are usually adapted for use with specific material databases. For instance, the BPA toolset Ladybug Tools includes the EnergyPlus database for thermal material properties (Sadeghipour Roudsari and Pak, 2013), whereas the parametric LCA tool Bombyx provides access to the Swiss KBOB database (Basic et al., 2019).

However, including a multitude of analysis methods during design processes quickly becomes time consuming for the designer (Purup and Petersen, 2020). First, they need to develop the geometric model for architectural purposes. Then, they need to adapt this model for each specific analysis mode considered. This includes identifying the associated material properties relevant to each analysis mode. In early design stages the extensive semantic information provided by Building Information Modelling (BIM) workflows, including material data, is usually not available (Cavalliere et al., 2019). This means the analysis process often becomes time- and resource-consuming enough that the analysis is entirely left out, or at best, carried out on a qualitative basis (Jusselme et al., 2020).

One avenue to streamline this process of modelling and data collection, is to make available a *single point of contact* in terms of material properties. By this we mean, as shown in Figure 1, that instead of selecting material properties individually for each analysis module, the material of each building component is selected only once in the user interface. Then, the material properties associated to each analysis module provided by the given analysis workflow are directly linked. In the analysis step shown in the figure, the various models necessary can then be generated based on a consistent material selection.

There are three main benefits for the designer with this approach. Firstly, the modelling and analysis process is greatly accelerated by only selecting the material once. Secondly, the risk of operator error in terms of selecting incorrect materials, or making incorrect mappings between similar materials with different properties, are sharply reduced as the number of operation required are reduced. Thirdly, the widespread use of such a consistent database would make the comparison between different studies achievable, a task which is currently difficult even within a specific analysis domain (Emami et al., 2019).

However, for this to be possible, the mapping of databases developed for different purposes is needed. The designer needs to be able to retrieve trustworthy data for a range of widely used building materials. Such a mapping does not currently exist in the market. This gap has been pointed out in number of research efforts. To bridge it, Fenz et al. (2021) developed an ontology for materials used in renovation and adapted it for use with Industry Foundation Classes (IFC), i.e., a BIM workflow (Theißen et al., 2020). Li et al. (2020) similarly define an ontology aimed at the materials design industry. Hong et al. (2019) instead propose that automated web crawling techniques could be applied to collect consistent material data and avoiding duplicated efforts by collecting data at different points in the design process.

While these studies provide useful starting point for a consistent ontology for material properties, and consistent

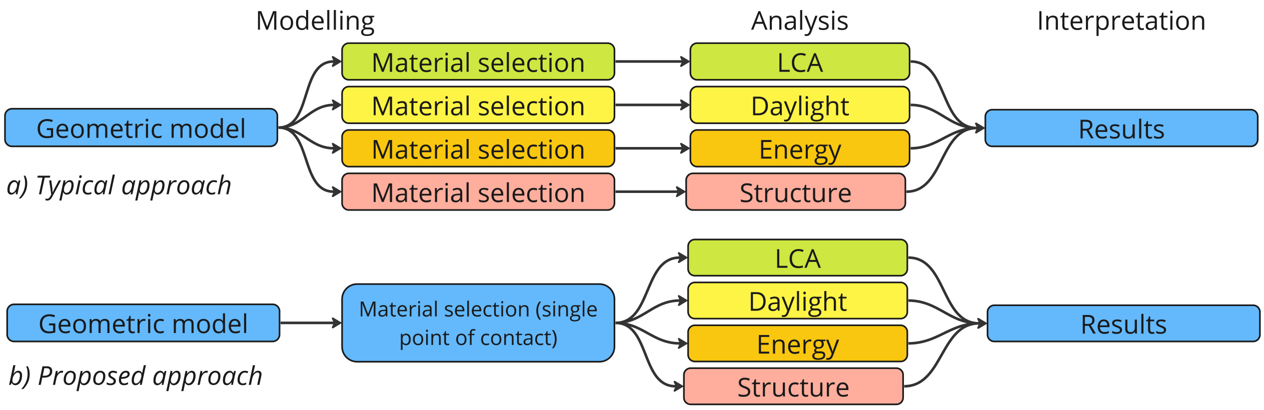


Figure 1: Typical and proposed methods of adding material data to a geometric model.

methodologies for collecting them, no studies known to the authors provide a hands-on approach for actually combining specific databases for LCA and BPA data relevant in the early design stages. We approach this by posing the following research question: *how can material databases developed for different analysis purposes be mapped using material names or IDs and density as mapping parameters?* This question is investigated through two cases, the addition of BPA data to the generic Swedish Klimatdatabas, and the holistic evaluation of an earthen house in Switzerland using KBOB, EnergyPlus, and Eurocode databases for LCA, energy, and structural analysis.

Methods and materials

The overarching method was collecting material data from a variety of sources and intended for different analysis methods, and mapping the data for each material using the density, ρ [$\mathrm{kg\,m^{-3}}$] as a mapping parameter. Both of the case studies are included in an overarching project to study the possibility of developing globally relevant databases for life cycle building performance data as compared to using locally defined databases. The specifics of each case study are described below.

Case study 1: Climate database in Sweden

The aim of the case study carried out in the Swedish context is to create a database which supports a parametric life cycle building performance assessment workflow. This workflow should contain LCA modules as well as BPA modules implemented in the parametric framework Grasshopper® (GH) (Robert McNeel & Associates, 2022).

The workflow is visualised in Figure 2. After geometry is modelled, material data is collected in a single point of contact within the GH interface. Data is collected for the three analysis modules, the analysis is carried out, and the results displayed and exported in the GH interface.

The database used for the climate assessment is the climate database (Klimatdatabas) of the Swedish National Board of Housing, Building and Planning (Boverket) (Boverket, 2023), which is available in Swedish and English. It includes the environmental impact I_{KDB} for the LCA modules A1-A5 [$\mathrm{kg\,CO_2\,eq.\,kg^{-1}}$] (Hollberg, 2016), which is mapped to geometry using the conversion factor density ρ [$\mathrm{kg\,m^{-3}}$] to transform volume of material first to mass of material and then finally to global warming potential. The data included refers to generic materials, not specific products.

There are no standardised databases for energy and daylight material information in Sweden. Instead, data was collected from a variety of sources. The following parameters were collected:

- Thermal conductivity λ [$\mathrm{W\,m^{-1}\,K^{-1}}$]
- Density ρ [$\mathrm{kg\,m^{-3}}$]
- Specific heat c [$\mathrm{J\,kg^{-1}\,K^{-1}}$]
- Roughness [-]
- Thermal absorptance α_t [-]
- Solar absorptance α_s [-]
- Visible absorptance α_v [-]

For the purposes of this article, the thermal conductivity λ and solar absorptance α_s are used for demonstration.

The climate database under consideration currently contains 171 opaque materials. Based on being representative, 10 materials from a variety of categories were selected for evaluation in this study. The mapping process, exemplified in Figure 3 for the case of "Structural steel, all sorts, primary material" in Klimatdatabas, was performed in two steps:

- Mapping of material names or IDs
- Mapping using parameter value (ρ [$\mathrm{kg\,m^{-3}}$])

The mapping of material name or IDs was done manually, by finding identical or near-identical labels for the materials included in the compared databases. When several candidate materials were present, for instance steel materials

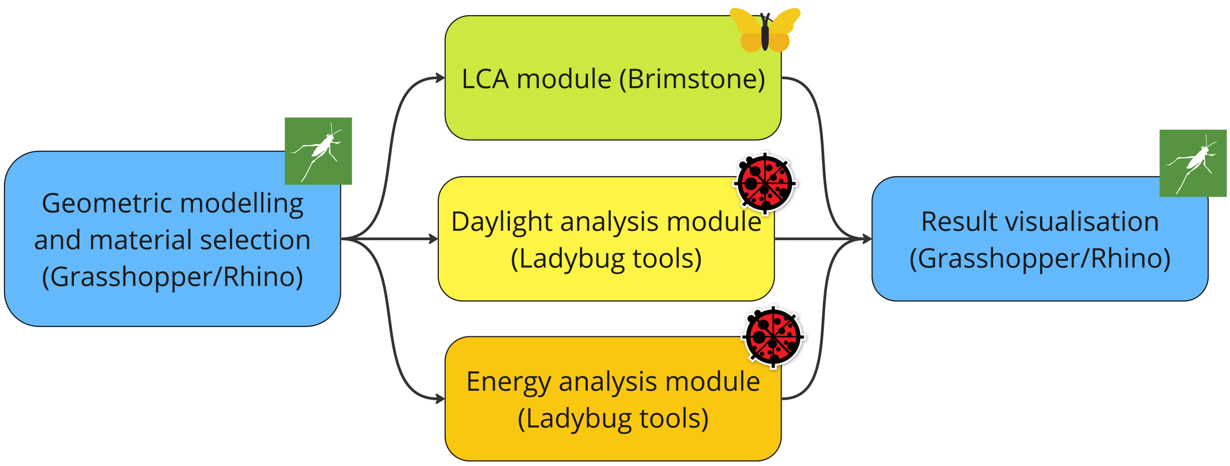


Figure 2: Workflow for parametric life cycle building performance assessment.

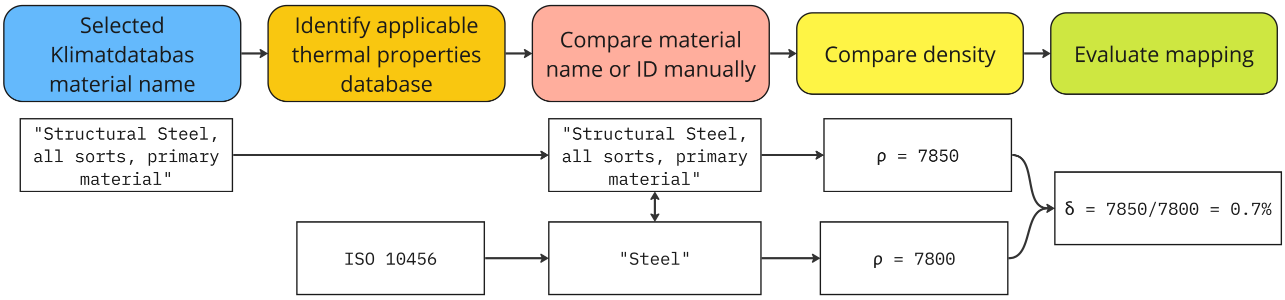


Figure 3: Example of mapping process for material "Structural steel, all sorts, primary material" in Klimatdatabas (Boverket, 2023) to thermal properties database (Pinterić, 2021).

with different structural performance, the material with the closest density was selected. Once the mapping was made, parameter values were collected and the precision of the mapping evaluated using the name or ID and density as comparative parameters to make sure that consistent data was collected.

Case study 2: Earth building in Switzerland

The aim of the Swiss case study is to combine database for an unconventional construction technique: using earthen materials. Earthen materials can be used in a variety of techniques, such as earth bricks or rammed earth, as shown in Figure 4, where rammed earth was used to build a school in the centre of Zurich. In addition to the LCA and BPA modules, the workflow of this case study integrates a Finite Element Modelling (FEM) module in GH, based on the Karamba plugin (Preisinger, 2013).

The FEM database for structural analysis is based on Eurocode materials, such as EN 1992-1-1 (European Committee for Standardisation, 2014) for concrete or EN 338 (European Committee for Standardisation, 2016) for timber. The mechanical properties of the materials are matched to the LCA and BPS database using the conversion factor density ρ [kg m^{-3}]. Although earthen materials are missing from the FEM database, the workflow in GH provides the ability to create bespoke isotropic or orthotropic materials. Assuming isotropic linear behaviour,

Figure 4: Umbau Schulpavillon Allenmoos II: school with rammed earth walls in Zurich © Boltshauser Architekten

the mechanical properties of the earthen materials are derived from the average values given in the RILEM State of the Art report (RILEM TC 274-TCE, 2022). The following parameters were collected:

- Young's modulus E [MPa]
- In-plane shear modulus G_a [MPa]
- Transverse shear modulus G_t [MPa]
- Density ρ [kg m^{-3}]
- Tensile strength σ_t [MPa]

- Compressive strength σ_c [MPa]
- Coefficient of thermal expansion α [°C^{-1}]
- Strength Hypothesis: Mises, Rankine or Tresca

The LCA database used for the Swiss context is provided by the Swiss Federal Conference of Public Project Owners for the Coordination of Construction and Building Services (KBOB, 2023). This database, available in German and French, provides information on materials, building services and systems, and transport. The Life Cycle Impact Assessment (LCIA) methodology calculates the Global Warming Potential (GWP) of materials, including manufacturing (A1-3) and disposal (C3-4). Energy consumption is assessed according to the Swiss standards SIA 2032 (Société suisse des Ingénieurs et des Architectes (SIA), 2020) for energy efficiency and embodied energy of buildings. This LCA database includes earth materials such as rammed earth, earth bricks or earth plaster. Rammed earth is taken as a reference for earth materials in the KBOB database.

The BPS database follows the EnergyPlus data developed by the US Department of Energy's Building Technologies Office (US DOE BTO, 2023). Although there is a mud material in the database suitable for modelling soils, a bespoke earth material is created against a monitored earth building (Estève-Bourrel et al., 2023).

Once the custom materials have been created and integrated into their respective database, and material names or IDs translated from French and German to English, the mapping procedure is carried out in the same way as in case study 1, using a single contact point in GH. In addition to earth materials, conventional materials such as concrete for the slab or timber for the roof are also considered.

Results

Case study 1: Climate database in Sweden

As seen in Table 1, all the materials presented were possible to map to materials in relevant databases for thermal properties, whereas consistent databases for data relevant for lighting simulations are harder to find. As seen in Figure 5, the mapping using the density ρ as a mapping parameter was generally successful, with a divergence of $\pm$15 % for most materials, notwithstanding EPS with a divergence of 36 %.

It should be noted that when collecting the data, no single database could be identified which collected all the relevant information. Instead, data was collected from varyingly from ISO standards (Pinterić, 2021), online resources (The Engineering ToolBox, 2009; Svenskt trä, 2003), and material manufacturers (Vitro Glazings, 2020; Eco Merchant, nd).

Case study 2: Earth building in Switzerland

In the case study 2, the earth building, the densities of the materials considered have a variability of less than 10%. The KBOB LCA database suggests a range in the density

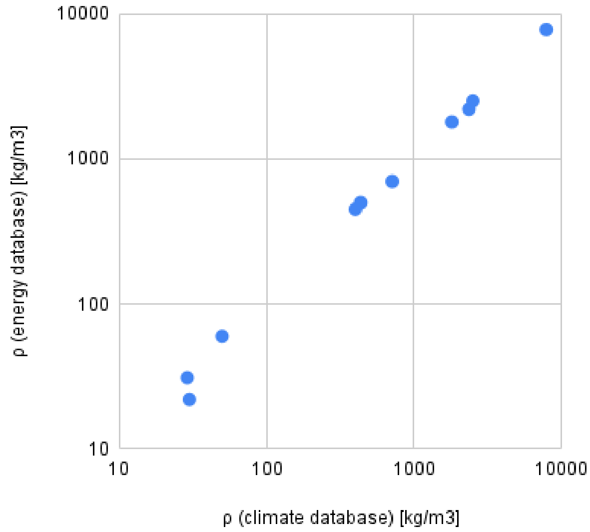


Figure 5: Overlap of parameters in baseline database and comparison database in case study 1 (Swedish climate database). X axis shows the density in the climate database for all materials, and Y axis shows the density in the respective energy databases.

of mineral wool and cellulose fibres, with values ranging from 32 to 160 $\mathrm{kg\,m^{-3}}$ and 35 to 60 $\mathrm{kg\,m^{-3}}$ respectively, which can create a greater disparity in the correct matching of densities, see Figure 6. The fact that the materials are also matched on the basis of their names or IDs confirms that they match correctly.

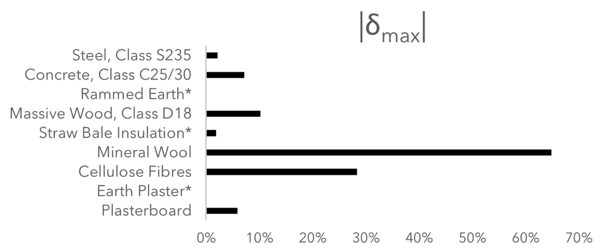


*Figure 6: Maximum absolute difference $|\delta_{max}|$ between the densities of the materials considered for the three studied databases (KBOB, EnergyPlus and Eurocodes). *bespoke material*

It is worth noting that the non-conventional construction materials chosen for this case study, shown in Table 2 such as rammed earth, straw bale or earth plaster, are included in the KBOB LCA database, but are absent from the EnergyPlus and Eurocode databases. The creation of these bespoke data is therefore based on scientific literature. The RILEM State of the Art book (RILEM TC 274-TCE, 2022) provides mechanical and thermal values for earth materials, including density matching the value in the KBOB database. The thermal characteristics were then checked against a monitored earth building. The thermal properties of straw bale were derived from experimental tests (Czajkowski et al., 2022).

Table 1: Excerpt of material mapping in case study 1 (Swedish climate database). I_{KDB} is the environmental impact (A1-A5 modules), λ is the thermal conductivity, and α_s is the solar absorptance. ρ_c is the density defined in the climate database, ρ_e is the density defined in the energy database, and ρ_d is the density defined in the reference data used for daylight material data. δ refers to the difference (in percent) between the ρ_c and ρ_e.

Material ID	I_{KDB}	ρ_c	λ	ρ_e	α_s	ρ_d	δ
Structural steel, all sorts, primary material	3.4	7850	50	7800[a]	0.65	7820	0.7
Floatglass (FG)	1.8	2500	0.94	2510[b]	0.14	2510	-0.4
Ready-mix made concrete, buildings C20/25	0.13	2350	1.65	2200[a]	0.6	2200	6.8
Bricks	0.38	1800	0.8	1800[a]	0.68	1900	0
Gypsum, standard plasterboard	0.34	710	0.21	700[a]	0.35	800	1.4
Glulam, u 12 %, spruce	0.18	434	0.13	500[c]	-	-	-13
Autoclaved Aerated Concrete, (AAC)	0.65	400	0.1	450[d]	-	-	-11
Wood fibre insulation, bats	0.43	50	0.036	60[e]	-	-	-17
EPS, expanded polystyrene	4.3	30	0.035	22[a]	-	-	36
Stone wool, bats and rolls	1.7	29	0.035	31[a]	-	-	-6.5

[a]Pinterić (2021)
[b]Vitro Glazings (2020)
[c]Svenskt trä (2003)
[d]The Engineering ToolBox (2009)
[e]Eco Merchant (nd)

Discussion

Compared to creating different models for each analysis mode, the creation of a single model which interfaces with several analysis engine would speed up the analysis process by greatly reducing the time needed for modelling and data collection. However, this hinges on the use of one database which can be used to define material properties for a variety of analysis methods. The currently available databases with life cycle impact data generally only include this data as well as conversion factors such as density to be able to map it to geometric data. This is true also in the case for the two climate databases considered in the present study: the Swedish Klimatdatabas, and the Swiss KBOB.

This means data needs to be collected from other resources to add the information needed for holistic life cycle building performance analysis including energy, daylight, and structural analyses, etc. We investigated the extent to which these resources could be mapped to the climate databases using the density as a quantitative and the resource name or ID as a qualitative parameter.

We found that the mapping to publicly available databases for energy analysis could be done based on the name or ID for most typical construction materials like concrete and steel. However, there was a variation in density ranging between $\pm$15% for most structural materials, and up to 64 % for insulation materials. This means there is an uncertainty in terms of how well the material properties are actually matched. Future work could delve into explaining these discrepancies in a deeper examination of the databases, or alternatively, evaluate what kind of uncertainties are acceptable when making design decisions based on the analysis results. It would also be relevant to study the impact of the need for translation between databases defined using different languages.

Further, we found that some important materials in the sustainable transition are missing even in widespread material databases like the ones offered by ISO (ISO 10456 regarding the hygrothermal properties of building materials and products (Pinterić, 2021)) and EnergyPlus (US DOE BTO, 2023). This includes wood materials like glulam and earthen materials like rammed earth and earth plaster, data for which instead needs to be collected from specific material manufacturers, which risks a bias. We also found that while energy data is accessible for most materials, data necessary for daylight analyses like reflectances/absorptances are more obscure and also require the use of data from specific manufacturers.

The method of using the density as a mapping parameter in addition to the name or ID of the specific resource appears useful based on our study, but it adds a layer of uncertainty to an already highly uncertain situation in the early design stages. This situation could be improved through a collective effort to improve climate databases by adding information needed also for other analysis modes. Alternatively, databases for structural, energy, and daylight analyses could be updated to also include climate information relevant to the local context. However, the constant evolution of life cycle data as more ecological approaches are adopted, and life cycle assessments are carried out in greater detail, mean that any such database needs to be in constant flux and requires regular update as new life cycle information is retrieved.

One limitation of our study is of course that only a few materials were selected for analysis. In future work, a wider overview of materials, in different national contexts, could be provided. In such an overview, the deeper analysis of the risk for false positives would be beneficial, especially if transitioning from manual to automatic mapping meth-

Table 2: Mapping of materials in case study 2 (earth building in Switzerland). I_{KBOB} is the environmental impact (A1-3; C3-4 modules), C^{bio}_{KBOB} is the biogenic carbon in the material, ρ_{KBOB} is the density defined in the climate (KBOB) database, ρ_{E+} is the density defined in the energy (EnergyPlus) database, ρ_{EU} is the density defined in the structure (Eurocode) database. $|\delta_{max}|$ refers to the absolute difference (in percent) between ρ_{KBOB}, ρ_{E+} and ρ_{EU}.

Material ID	I_{KBOB}	C^{bio}_{KBOB}	ρ_{KBOB}	ρ_{E+}	ρ_{EU}	$\|\delta_{max}\|$
Steel, Class S235	0.736	0	7850	7680	7800	2.2
Concrete, Class C25/30	0.177	0	2500	2322	2500	7.1
Rammed Earth	0.019	0	2000	2000[a]	2000[a]	0
Massive Wood, Class D18	0.129	0.413	485	540	500	10
Straw Bale Insulation	0.096	0.368	215	211[b]	-	0
Mineral Wool	1.19	0	32/160	91	-	64
Cellulose Fibres	0.280	0.404	35/60	43	-	28
Earth Plaster	0.032	0	1800	1800[a]	-	0
Plasterboard	0.301	0.018	850	800	-	5.9

[a]RILEM TC 274-TCE (2022)
[b]Czajkowski et al. (2022)

ods. Another important future work is cross-referencing of a number of available databases for improved robustness. Further, it should be noted that different parameters were collected in the different case studies based on their respective purpose. The overlap of the case studies is the evaluation of name or ID and density as mapping parameters.

Since the only parameters mapped were the name or ID and the density, there is a great risk of misidentifying some materials, leading to incorrect data mappings. One potential way of overcoming this would be comparing further data attached to each resource, such as the material description commonly provided. To investigate such qualitative parameters, natural language processing (NLP) methods could be applied to compare great amounts of textual data.

We acknowledge that the proposed method only resolves the issue of selecting consistent material data for an existing geometric model for use with a variety of simulation tools. Further work is needed to streamline the workflows which generate each necessary analysis model from this single geometric model with added material data.

Conclusion

In the present study, the mapping of life cycle data from the Swedish Klimatdatabas and the Swiss KBOB database to databases for energy, daylight, and structural material properties was investigated. This was done by comparing the names or IDs and densities of materials to check whether the materials could be mapped.

A divergence of ±15 % was detected for structural materials, whereas the divergence of up to 64 % was detected for insulation materials. This indicates that while the mapping using the proposed methodology is possible, it introduces a great uncertainty in terms of the material properties.

We propose the development of robust, nationally and internationally relevant material databases including life cycle information along with energy, daylight, and structural material properties. The mapping of materials used when developing such databases could include the material name or ID, the density, and other parameters. These databases need to be periodically updated and revised as the life cycle impact of materials is improved through the sustainable transition of the production industry. The benefit of the use of such a holistic database is reducing modelling and analysis time, avoiding user errors, and allowing direct comparison between analyses, by only selecting material information once as shown in Figure 1.

In future work, we propose that the applied methodology of matching materials through resource name or ID and density is further tested by comparing several available databases, and by testing it in different national contexts.

Acknowledgements

The research presented in this paper was partly funded by the Swedish Energy Agency, project number 51715-1.

References

Basic, S., Hollberg, A., Galimshina, A., and Habert, G. (2019). A design integrated parametric tool for realtime life cycle assessment – bombyx project. IOP conference series: earth and environmental science, 323:012112. Publisher: IOP Publishing.

Boverket (2023). About Boverket. Online: https://www.boverket.se/en/start/about/about-bover ket/(accessed 2023-12-12).

Cavalliere, C., Habert, G., Dell'Osso, G. R., and Hollberg, A. (2019). Continuous BIM-based assessment of embodied environmental impacts throughout the design process. Journal of Cleaner Production, 211:941–952.

Czajkowski, L., Kocewicz, R., Weres, J., and Olek, W. (2022). Estimation of thermal properties of straw-based insulating panels. Materials, 15:1073. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8838588/.

Eco Merchant (n.d.). Steico flex 036 wood fibre insulation batts (100mm x 375mm wide). Online: https://www.ecomerchant.co.uk/interior/insulation/steico-flex/steico-flex-wood-fibre-insulation-375mm.html (accessed 2023-12-13).

Emami, N., Heinonen, J., Marteinsson, B., Säynäjoki, A., Junnonen, J.-M., Laine, J., and Junnila, S. (2019). A life cycle assessment of two residential buildings using two different LCA database-software combinations: Recognizing uniformities and inconsistencies. Buildings, 9(1).

Estève-Bourrel, P., Beckett, C., Bosché, F., and Habert, G. (2023). How load-bearing earth construction can meet today's energy demands? In International Conference on Biobased Building Materials 2023, Vienna. https://www.researchgate.net/publication/374000156_How_load-bearing_earth_construction_can_meet_today%27s_energy_demands.

European Committee for Standardisation (2014). EN-1992-1-1:2004+A1:2014 - Eurocode 2: Design of concrete structures. Part 1-1: General rules - Rules for buildings, bridges and civil engineering structures. https://eurocodes.jrc.ec.europa.eu/EN-Eurocodes/eurocode-2-design-concrete-structures.

European Committee for Standardisation (2016). EN 338:2016 - Structural timber. Strength classes. https://www.en-standard.eu/csn-en-338-structural-timber-strength-classes/.

Fenz, S., Bergmayr, J., Plattner, N., Chávez-Feria, S., Poveda-Villalón, M., and Giannakis, G. (2021). Integration of building material databases for IFC-based building performance analysis. In Proceedings of the International Symposium on Automation and Robotics in Construction, volume 2021-November, pages 182–189. ISSN: 2413-5844.

Hollberg, A. (2016). Parametric Life Cycle Assessment: Introducing a time-efficient method for environmental building design optimization. PhD thesis, Bauhaus-Universität, Weimar.

Hollberg, A., Lichtenheld, T., Klüber, N., and Ruth, J. (2018). Parametric real-time energy analysis in early design stages: a method for residential buildings in Germany. Energy, Ecology and Environment, 3(1):13–23.

Hong, S.-H., Lee, S.-K., and Yu, J.-H. (2019). Automated management of green building material information using web crawling and ontology. Automation in Construction, 102:230–244.

Jusselme, T., Rey, E., and Andersen, M. (2020). Surveying the environmental life-cycle performance assessments: Practice and context at early building design stages. Sustainable Cities and Society, 52:101879.

KBOB (2023). Données écobilans dans la construction. Online: https://www.kbob.admin.ch/kbob/fr/home/themen-leistungen/nachhaltiges-bauen/oekobilanzdaten_baubereich.html (accessed 2023-12-13).

Khoshnava, S. M., Rostami, R., Mohamad Zin, R., Štreimikienė, D., Mardani, A., and Ismail, M. (2020). The Role of Green Building Materials in Reducing Environmental and Human Health Impacts. International Journal of Environmental Research and Public Health, 17(7):2589. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

Li, H., Armiento, R., and Lambrix, P. (2020). An ontology for the materials design domain. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12507 LNCS:212–227. ISBN: 9783030624651.

Meex, E., Hollberg, A., Knapen, E., Hildebrand, L., and Verbeeck, G. (2018). Requirements for applying LCA-based environmental impact assessment tools in the early stages of building design. Building and Environment, 133:228–236.

Pinterić, M. (2021). Building Physics: From physical principles to international standards. Springer International Publishing.

Preisinger, C. (2013). Linking structure and parametric geometry. Architectural Design, 83:110–113.

Purup, P. B. and Petersen, S. (2020). Requirement analysis for building performance simulation tools conformed to fit design practice. Automation in Construction, 116:103226.

RILEM TC 274-TCE (2022). Testing and Characterisation of Earth-based Building Materials and Elements: State-of-the-Art Report. https://link.springer.com/10.1007/978-3-030-83297-1.

Robert McNeel & Associates (2022). Rhinoceros 3D. Online: https://www.rhino3d.com/ (accessed 2023-12-12).

Röck, M., Saade, M. R. M., Balouktsi, M., Rasmussen, F. N., Birgisdottir, H., Frischknecht, R., Habert, G., Lützkendorf, T., and Passer, A. (2020). Embodied GHG emissions of buildings – the hidden challenge for effective climate change mitigation. Applied Energy, 258:114107.

Sadeghipour Roudsari, M. and Pak, M. (2013). Ladybug: A parametric environmental plugin for grasshopper to help designers create an environmentally-conscious design. In Proceedings of BS 2013: 13th Conference of the International Building Performance Simulation Association, pages 3128–3135.

Société suisse des Ingénieurs et des Architectes (SIA) (2020). SIA 2032:2020 - L'énergie grise – Établissement du bilan écologique pour la construction de bâtiments. http://shop.sia.ch/collection%20des%20normes/architecte/sia%202032/f/2020/F/Product.

Svenskt trä (2003). Värmeegenskaper. Online: https://www.traguiden.se/om-tra/materialet-tra/traets-egenskaper-och-kvalitet/termiska-egenskaper1/varmeegenskaper/ (accessed 2023-12-13).

Säwén, T. (2023). Early stage architectural design practice perspectives on life cycle building performance assessment. Licentiate thesis, Chalmers University of Technology, Gothenburg, Sweden.

Säwén, T., Magnusson, E., Sasic Kalagasidis, A., and Hollberg, A. (2022a). A characterisation framework for parametric building performance simulation tools. In Proceedings of BuildSim Nordic 2022, volume 362. E3S Web of Conferences.

Säwén, T., Magnusson, E., Sasic Kalagasidis, A., and Hollberg, A. (2022b). Tool characterisation framework for parametric building LCA. In IOP Conference Series: Earth and Environmental Science, volume 1078, page 012090.

The Engineering ToolBox (2009). Solids - densities. Online: https://www.engineeringtoolbox.com/density-solids-d_1265.html (accessed 2023-12-13).

Theißen, S., Höper, J., Wimmer, R., Zibell, M., Meins-Becker, A., Rössig, S., Goitowski, S., and Lambertz, M. (2020). BIM integrated automation of whole building life cycle assessment using german LCA data base ÖKOBAUDAT and industry foundation classes. In IOP Conference Series: Earth and Environmental Science, volume 588. ISSN: 1755-1307 Issue: 3.

US DOE BTO (2023). Energyplus. Online: https://www.energyplus.net/ (accessed 2023-12-13).

Vitro Glazings (2020). Clear technical product data. Online: https://www.vitroglazings.com/media/rnff0xlj/vitro-clear-technical-pds.pdf (accessed 2023-12-13).

AN ONTOLOGY FOR SIGNAL STRENGTH ESTIMATION OF NOMADIC 5G NETWORKS ON CONSTRUCTION SITES

Lukas Kirner[1], Jyrki Oraskari[1], Marit Zöcklein[2],
Andreas Möller[3] and Sigrid Brell-Cokcan[1]
[1]Individualized Production, RWTH Aachen University, Aachen, Germany
[2]Center Construction Robotics, Aachen, Germany
[3]UNIBERG GmbH, Bebensee, Germany
kirner@ip.rwth-aachen.de

Abstract

Network availability and quality is vital for using robotics, real-time digital twins or augmented reality in construction. This paper explores using semantic web models to perform radio propagation calculations for construction sites. We have created an ontology describing and linking data to calculate how building element installation affects attenuation losses. Our findings enable to predict network quality in the context of construction progress to ensure network performance and coverage for mobile construction robotics. Hereby, network optimisation tasks can be automated or planned in time for re-adjusting antenna positions and orientation.

Introduction

The construction sector faces major challenges in creating affordable housing and modernising infrastructure despite rising material prices and a shortage of skilled workers. To meet these challenges, it is necessary not only to optimise the design of buildings but also to dramatically increase the productivity of construction processes. In the methods commonly used in practice, team productivity is strongly influenced by effective communication in coordinating construction activities and fast access to information (Ahsan et al., 2009). Nowadays, this means that a stable network connection is of great benefit. Network availability and quality are crucial to further digitalising the sector and enabling the use of digital twins and robotics, which have increased productivity in other sectors.

Construction activities can be monitored by various sensors, including laser scanners, depth cameras, ground penetrating cameras, RGB cameras, radio frequency identification, inertial measurement units, global navigation satellite systems, and wireless sensor networks (Rao et al., 2022). These sensors are also needed for mobile robots to operate safely on construction sites - especially to provide situational awareness in human-machine collaboration. The success of construction robotics will depend not only on the capabilities of each machine but also on the ability to share observations and, thus, context. This is highly dependent on a robust and high-performance network infrastructure.

Common, widely used wireless network technologies, such as WLAN (e.g. IEEE 802.11), have range, throughput, jitter, latency and reliability limitations that significantly hinder construction automation. Additionally, a lot of devices do not back up Quality of Service (QoS) functionalities. A promising approach to overcome this is the introduction of 5G technologies in the form of campus networks for construction sites. For example, the reliability and low latency promised by 5G networks play an important role in autonomous and remote-controlled construction machinery (Lee et al., 2022). In general, high latency is detrimental to real-time robot control.

While the capabilities of 5G networks can potentially improve network quality and coverage on construction sites, these environments present specific challenges. Unlike fixed factory floor deployments, environmental factors on a construction site are permanently changing. Therefore, networks must be able to adapt and move as construction progresses. Otherwise, limited network performance and connectivity results (Din and Bernold, 2017). As with other wireless radio networks, the quality of 5G networks depends on the characteristics and design of the antenna and the propagation of radio waves. Factors such as propagation losses, including path loss, reflection, diffraction, refraction and scattering, can limit the effectiveness of radio networks (Erunkulu et al., 2020).

Predicting network quality at specific locations on a construction site by estimating the loss of signal strength, considering construction progress, building information, and antenna configuration could significantly contribute to digitalization and automation in construction. The knowledge gained could, for example, influence the navigation preferences of mobile robots to avoid areas with poor network quality.

More promisingly, it could be used to optimise a network setup, where antenna settings, position and orientation can be adjusted based on construction progress. Recent antenna technologies such as Massive Multiple Input Multiple Output (MIMO) allow software-optimised radio, such as beamforming optimizations. Concepts like O-RAN and SD-RAN enable applications to steer RAN so that fewer physical positioning adjustments are required. Still, such applications require exact information about site topology and progress.

To enable this, the proposed network quality prediction requires harmonising and linking data that do not originate from the same source system and cannot currently be put into a machine-readable context. Linked Data and Semantic Web technologies are ideal for linking such heterogeneous data from different knowledge domains. In addition, there are already approaches for Building Information Modeling (BIM) and construction process modelling in the form of Linked Building Data (LBD) ontologies, ifcOWL (Pauwels and Terkaj, 2016) and the results of the Internet of Construction (IoC) project (Brell-Cokcan and Schmitt, 2024). Therefore, the approach described in this paper investigates the development and use of a semantic data model using an ontology to represent the data required for the calculation of radio wave propagation.

State of the Art

Fundamentals in Radio Wave Propagation

Radio waves are electromagnetic waves. As such, they are affected by reflection, refraction and absorption phenomena when propagating through matter. Commonly, the propagation of an electromagnetic wave is described with the wave equation:

$$\begin{aligned} E(x,t) &= E_0 \cdot e^{-i\left(\omega t - \frac{\omega}{c} x n\right)} \mid \text{with } n = n' + i \cdot n'' \\ &= E_0 \, e^{-\frac{\omega n''}{c} \cdot x} \cdot e^{-i\left(\omega t - \frac{\omega n'}{c} \cdot x\right)} \end{aligned} \tag{1}$$

where the first term gives the exponential decay of the wave's amplitude with propagation for $n' > 0$. For the intensity $I(x)$ holds

$$\begin{aligned} I(x) &= \frac{1}{2} \varepsilon_0 c \cdot |n| \cdot |E(x)|^2 \\ &= \frac{1}{2} \varepsilon_0 c \cdot |n| \cdot E_0{}^2 \cdot e^{-\frac{2\omega n''}{c} \cdot x} = I_0 \cdot e^{-\alpha \cdot x} \end{aligned} \tag{2}$$

where α is the attenuation coefficient, and the last identity is known as Beer-Lambert's law. The intensity of the transmitted radio wave is related to its power by multiplication with the unit area $I \cdot 4\pi d^2 = P$. Typically, the transmit power of an antenna is given in the unit dBm. This is equivalent to the power ratio of the signal power in reference to $P_0 = 1\,\text{mW}$ on a $\log_{10}$-scale as in equation 3.

$$P\,[\text{dBm}] = 10 \cdot \log_{10}\left(\frac{P_{RX}\,[\text{mW}]}{1\,\text{mW}}\right) \tag{3}$$

On the same logarithmic scale, the received signal strength P_{RX} at the distance d can be expressed as the sum of the transmit power P_t of the transmitting antenna, its respective antenna gain G_{TX}, the propagation losses along the path between the transmitter and the receiver TL, and the gain of the receiving antenna G_{RX} on the logarithmic scale.

$$P_{RX}\,[\text{dBm}] = P_t\,[\text{dBm}] + G_{TX}\,[\text{dBi}] - TL\,[\text{dB}] + G_{RX}\,[\text{dBi}] \tag{4}$$

The antennas' gain and transmit power are adjustable properties, whereas, for propagation losses, the environment plays a crucial role. Propagation loss describes the reduction in signal strength when an electromagnetic wave traverses a medium capable of absorbing or dispersing a portion of the wave energy. In a simplified model, the total propagation losses can be considered as the sum of the free space path loss (FSPL), reflection loss (RL), and attenuation by penetrating building elements (PL).

$$\text{TL}\,[\text{dB}] = \text{FSPL}\,[\text{dB}] + \text{RL}\,[\text{dB}] + \text{PL}\,[\text{dB}] \tag{5}$$

FSPL refers to "the loss between two isotropic radiators in free space, expressed as a power ratio". The Friis transmission equation (Friis, 1946) can be written as

$$\text{FSPL}\,[\text{dB}] = 20 \log_{10}\left(\frac{4\pi d f}{c}\right) \tag{6}$$

where d denotes the distance between the transmitter and receiver and f the frequency.

Reflection losses describe the signal lost when an object's surface reflects a fraction of incoming radio waves. The reflectance is a material property and depends strongly on the frequency of the incident radio wave and the angle under which the wave hits the surface.

$$P_r = R(f) \cdot P_{in} \tag{7}$$

Lastly, and especially in the construction environment, penetration loss or attenuation occurs when radio waves penetrate building elements since most real-world materials are dielectric materials characterised by permittivity, permeability, and conductivity. To quantify the impact of material properties on wave propagation, we follow the calculations in ITU (2023) and consider the attenuation distance Δ where the amplitude of the electromagnetic field vector has fallen by a factor $\frac{1}{e}$:

$$\Delta = \frac{1}{k_0 \sqrt{\eta'}} \sqrt{\frac{2\cos(\delta)}{1-\cos(\delta)}} \tag{8}$$

In equation 8 η' denotes the real part of the complex relative permittivity $\eta = \eta' + i \cdot \eta''$. The permittivity of the material is then $\varepsilon = \eta \varepsilon_0$. Since the angle of the loss tangent $\delta = \arctan\left(\frac{\sigma}{\varepsilon\omega}\right)$ relates the conductivity σ with the permittivity, the attenuation distance also depends on these properties of the material and the frequency of the wave. In the limit of a good conductor, the attenuation length simplifies to

$$\Delta_{conductor} = \lim_{\sigma \to \infty} \Delta = \frac{1}{k_0 \sqrt{\eta'}} \sqrt{\frac{2}{tan(\delta)}} \tag{9}$$

whereas for a nearly perfect dielectric, it becomes

$$\Delta_{dielectric} = \lim_{\sigma \to 0} \Delta = \frac{1}{k_0 \sqrt{\eta'}} \frac{2}{tan(\delta)} \tag{10}$$

The attenuation distance is related to the attenuation rate as shown in equation 11.

$$A\left[\text{dB}\,\text{m}^{-1}\right] = \frac{20 \log_{10}(e)}{\Delta} \tag{11}$$

As suggested by ITU (2023), we use equations 9 and 10, to express the attenuation rates for conductors and approximately ideal dielectric materials as functions of η' and σ resulting in

$$\begin{aligned} A_{conductor} &= 20 \log_{10}(e) \sqrt{\frac{\pi}{\varepsilon_0 c_0{}^2}} \cdot 10^9 \cdot \sqrt{\sigma f\,[\text{GHz}]} \\ A_{dielectric} &= \frac{20 \log_{10}(e)}{2\varepsilon_0 c_0} \cdot \frac{\sigma}{\sqrt{\eta'}} \end{aligned} \tag{12}$$

Both σ and η' depend on the frequency and can be characterised as follows:

$$\eta' = af^b \quad \text{and} \quad \sigma = cf^d \tag{13}$$

To calculate the total amount of propagation losses, the derived attenuation rates have to be multiplied by the distance the wave travels within the material. Since building elements often consist of several layers of materials, these calculations need to be done for all layers. The thus found value for the expected signal strength can be compared to the RSSI value that can be found in the network statistics of each connected device that fulfils the IEEE standards 802.11 or 802.15.4.

Research on building materials affecting radio wave propagation

Current research on the electromagnetic behaviour of building materials is mainly conducted at the city scale or focuses on the finished building and how the choice of materials affects the signal quality. Exemplary approaches focus on the use of blueprints (Seidel and Rappaport, 1994) and the 3D geometry of the building (Ullah et al., 2020) for ray tracing simulations. Similarly, empirical studies of building penetration losses indicate the effect of façade material, structure and thickness, radio parameters such as frequency, signal strength, incidence angle and the incident field's polarisation. A study by García Sánchez et al. (2022) focused on the impact of building penetration losses at 3.5 GHz across various facade types, examining factors such as the incidence angle, polarization, and material effects within 5G systems. Their findings included an attenuation variation of up to 8 dB based on the incident angle. An example of a compilation of such studies can be found in (ITU, 2021). Like Yamamoto et al. (2019), we relate the received radio signal strength at a dedicated measurement point to the distance between the transmitting antenna and the receiver. Yet, in contrast, in construction, considerations about the propagation losses that affect the radio signal along its path through physical objects need to be included. While only applied studies are available for many phenomena such as scattering, and therefore no values exist that can be meaningfully transferred to individual components and their materials, the necessary characteristic values for permittivity and conductivity are collected in ITU-R P.2040 (ITU, 2023). Therefore, in this paper, we concentrate primarily on mapping these values.

Linked Building Data

Linked Building Data (LBD) leverages Linked Data principles (Heath and Bizer, 2011), connecting information relevant to applications in the Architecture, Engineering and Construction sectors. World Wide Web Consortium (W3C) standards such as the Resource Description Format (RDF) (Lassila et al., 1998) are used. LBD benefits interdisciplinary collaboration (Pauwels et al., 2017) and interoperability (Beetz, 2009). To promote data accessibility, the W3C Linked Building Data (LBD) Community Group has focused on developing modular approaches in recent years(Rasmussen et al., 2020). For material properties, the Digital Construction Building Material (DICBM) ontology, developed by Valluru et al. (2020), is a modular framework to improve the management of building material information within the BIM collaboration process. This ontology encompasses components such as MaterialDefinition, LayerSet, Layer and MaterialProperty. However, it does not include properties for describing the electromagnetic behaviour of construction materials. In contrast to materials, radio antennas and their properties have been completely absent and, therefore, cannot currently be modelled in the LBD context.

Methodology

Scope and competency questions

Calculating values to dynamically predict network quality and coverage on construction sites is mainly related to antennas, their properties, the site topology, and the building plan, including geometry and used materials. Moreover, the construction process, which involves scheduled and performed progress and auxiliary elements such as formwork or scaffolds, needs to be included as changing spatial conditions. As there is not yet the possibility to model and link all the needed data in RDF, an ontology is created to introduce the necessary classes and properties while connecting them to existing approaches from the linked building data domain, as the ontology complements and extends these when possible. The methodology deployed is described by Noy and McGuinness (2001). According to their guide, the first step is to clarify the focus and scope of the new ontology. Three questions regarding the scope (SQ) are defined and answered for that purpose.

SQ1: What domain is covered by the ontology?

The ontology covers radio network modelling and signal strength predictions for construction sites.

SQ2: What is the purpose of the ontology

The ontology should enable the calculation of the Received Signal Strength Indicator (RSSI) calculation to predict radio networks' relative signal strength at any spot on a construction site. Therefore, it is necessary to be able to save and link data relevant to radio propagation of the frequency used connected with the building information modelling and construction process data. It should enable the simulation of wireless communication networks with respect to changing spatial conditions of the construction site and propose possible adaptations to the nomadic network infrastructure.

SQ3: What questions should the ontology answer?

The ontology should be able to answer questions about the topology-related setup of the wireless network infrastructure and its current configuration. It should help assign and query values for building materials describing their type permittivity and conductivity to query and calculate propagation losses dynamically and independently of the deployed radio frequency.

Competency questions (CQ) were developed according to the scope specification. These competency questions are technical-functional, describing what queries should be able to be answered when data is instantiated using the ontology. These queries are not intended to directly answer questions about the expected network quality, as this requires complex calculations. Our goal is to prove that context-enriched values can be retrieved to calculate or simulate approximate values incorporating dynamic construction process information. The CQs are shown in Table 1.

Table 1: Defined competency questions (CQ).

No.	Questions
CQ1	What are the characteristics and position of the transmitting antenna?
CQ2	What are the characteristics and position of the receiving antenna?
CQ3	What building elements that interfere with radio wave propagation are expected to be realized at a given time?
CQ4	What is the material composition of the building elements in question?
CQ5	What are the materials' type, permittivity and conductivity properties concerning radio waves?

Reusing existing ontologies is one of the core principles of Linked Data. Our investigation into existing ontologies that predict radio propagation within construction sites reveals a lack of developed approaches. Therefore, the competency questions cannot be answered with any existing solution known to us. However, for concepts that describe the building elements and the construction process, there are ontologies that we consider to be applicable and mature. These can be built upon and incorporated to achieve greater interoperability. Table 2 gives an overview of the reused or linked ontologies.

Table 2: Overview of the connected ontologies.

Name	Main classes and purpose
bot	*bot:Zone*, *bot:Element*, *bot:Interface* Ontology describing topological concepts of a building (Rasmussen et al., 2018)
ifc	*Ifc:Root* Web Ontology Language (OWL) representation of the IFC schema (Pauwels and Terkaj, 2016)
ioc	*ioc:process* Ontology developed within IoC to describe processes (Kirner et al., 2024)
opm	*opm:PropertyState* An ontology for describing properties that change over time (Rasmussen et al., 2020)
qudt	*qudt:Quantity*, *qudt:Unit* A unified model of measurable quantities and units (QUDT.org, 2015)

The Construction Site Network Ontology (CNO)

The **CNO** ontology is designed as an application-level ontology. It is intended to be used in combination with the LBD ontologies (Oraskari et al., 2021) and *ifcOWL* (Pauwels and Terkaj, 2016) to add construction context, as well as the Internet of Construction Process Ontology (ioc) (Kirner et al., 2024), which enables dynamic scheduling and status information of construction processes. The main classes of the **CNO** ontology describe the missing dynamic data needed to calculate the expected signal strength according to the radio propagation fundamentals. Specifically, these are *cno:Antenna*,*cno:Obstacle* and *cno:Material*. Most proposed additions are object and data properties that extend or link existing classes like *ifc:IfcMaterial* to add metadata relevant to radio propagation.

cno:Antenna

The major new concept to add is the *cno:Antenna*. Simplified, the antennas are subdivided into *cno:Transmitter*, which describes the main antenna and *cno:Receiver*, in this case, the antennas of the devices in the network. Antennas must have object properties that describe the relevant properties. In this first draft, these are limited to position and direction, power, gain, and frequency. As antennas for wireless communication are quite complex, these properties will be extended in future iterations.

cno:Obstacle

An object that leads to reflection and propagation losses caused by its dielectric properties is described with the class *cno:Obstacle*. Although the focus is clearly on the building elements, obstacles, such as formwork, can also be auxiliary or temporary objects of the construction process. In addition, cars or construction machinery are relevant. For building elements, position and rotation can be queried via LBD geometry descriptions or data from the *ifcOWL* instances. Geometry information in conjunction with antenna position and properties can be used to calculate FSPL or coarse reflection effects (see equation 6).

cno:Material

For the object propagation loss, the material composition of the building elements is decisive. Real-life building elements rarely consist of a single material. The *ifc:IfcMaterialLayer* instances can provide material, layerthickness, and direction. Therefore, we propose extending *ifc:IfcMaterial* via *rdfs:subclassOf* with *cno:Material* to add metadata regarding the relevant material type (*cno:Conductor* or *cno:Dielectric*). The values needed to calculate the attenuation rate according to equation 11 are linked as *cno:hasRelativePermittivityFrequencyCoefficient*, and *Exponenet* as well as *cno:hasConductivityFrequencyCoefficient* and *Exponent*. They can be retrieved from Table 3 of ITU-R P.2040-3 (ITU, 2023). The model from which these equations were deduced is only valid in a certain frequency spectrum, so a *cno:FrequencyRange* must also be attached to each material.

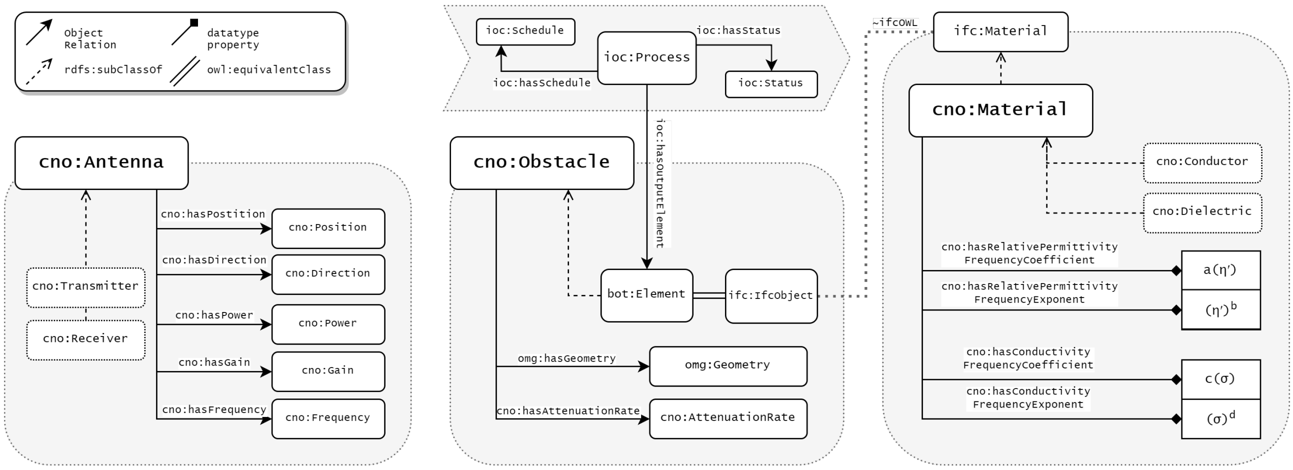


*Figure 1: Main concepts and properties of the **CNO** ontology.*

All classes and properties presented aim to enable the calculation of attenuation rates in a construction site context. For object propagation losses, they can be added to a *cno:Obstacle* with the help of *cno:hasAttenuationRate*. This should be a dynamic, time-stamped value enabled using the OPM ontology's *opm:CurrentDataState* class. This is considered a main improvement over static material bound values found throughout the state-of-the-art. On the one hand, depending on the composite material, the number of layers is due to changes in the course of the construction process. On the other hand, as shown in equations 8, attenuation rates also depend on frequency, which is not necessarily static, depending on the network setup deployed.

Properties of the ontology

The ontology described here represents a first draft of an OWL-based ontology describing radio networks on construction sites. In its current state, the ontology consists of twelve classes, six object properties, and six data-type properties. Figure 1 shows the basic concepts, as well as object and datatype properties proposed.

Evaluation and use-case

A dataset derived from the operational 5G network on the reference site Aachen Melaten was created to evaluate the ontology presented. This does not aim to validate whether the calculated values are precise but to prove that the established competency questions can be answered. Therefore, a dataset created with the help of the CNO ontology provides the data required to predict signal strengths in a dynamic construction site environment.

Due to scientific comparability, a known BIM model, the duplex[1], is used as a source of building information. It is converted to RDF and extended by *ioc:processes*, which describes, using *ioc:Schedule* that a certain building element is expected to be erected at a given time.

The modelled scenario consists of the transmitting antenna, an AW3374-T0-F (1) mounted on a Liebherr L1-24 tower crane, and a Milesight UR75 5G Router (2). The Router is mounted on a Heros 224 mobile AGV from Innok Robotics. A multilayer outside wall of the BIM model is chosen as the element of interest (3). Figure 2 shows the scenario and its components.

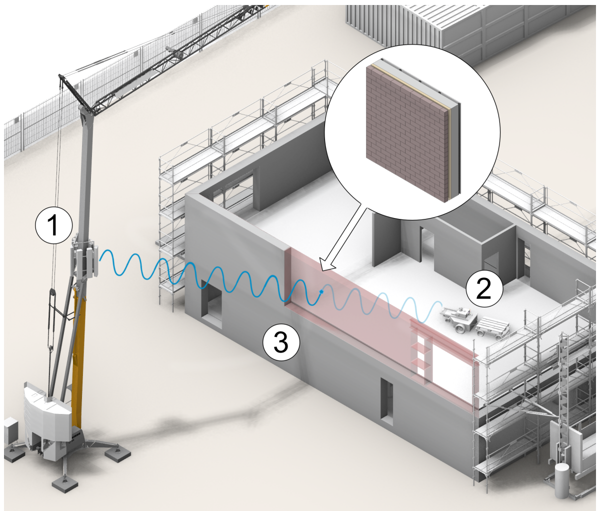


Figure 2: Scenario on the reference construction site in Aachen.

The Antenna and its properties were modelled manually. While values like gain or direction are among the characteristics provided by the antenna's manufacturer, power and frequency are subject to the network setup. It is to be noted that these values are linked to the antenna to simplify the data model and queries. Practically, the radio in use determines these values. The antenna position is designed as a dynamic feature to enable nomadic network setups in reaction to progressing construction processes. The resulting triples added to the graph are shown in figure 3.

[1]https://portal.nibs.org/files/wl/?id=O9inoVWDytV4P3UTS7ildqhVKcTyVlcs

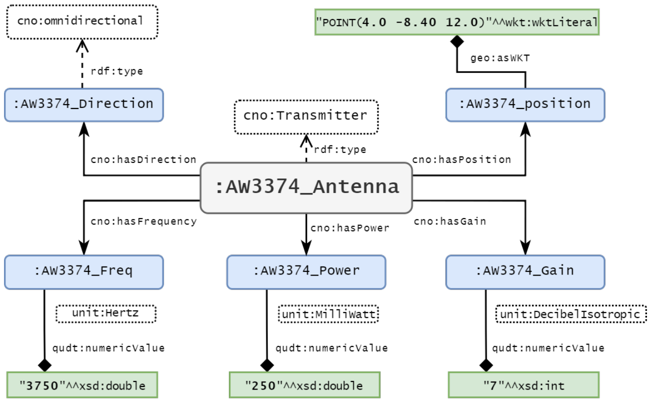


Figure 3: Example for an instantiated cno:Antenna including metadata.

CQ1-2 can be answered by querying the depicted instances in addition to a receiver that is modelled accordingly. Listing 1 presents a query retrieving the transmitter, frequency, and receiver positions. Querying for *opm:CurrentDataState* ensures that the currently valid value is retrieved, which implies the dynamic characteristics of the proposed approach.

Listing 1: SPARQL Query for CQ1

```
PREFIX cno: <http://w3id.org/cno#>
PREFIX qudt: <http://qudt.org/schema/qudt#>
PREFIX unit: <http://qudt.org/vocab/unit#>
PREFIX geo: <http://www.opengis.net/ont/geosparql#>
PREFIX opm: <http://www.w3id.org/opm#>

SELECT ?f ?posTX ?posRX WHERE
   {?Transmitter a cno:Transmitter;
      cno:hasPosition ?posTXState;
      cno:hasFrequency ?fstate.
   ?fstate a opm:CurrentDataState;
      qudt:numericValue ?f;
      qudt:unit unit:Hertz.
   ?posTXState a opm:CurrentDataState;
      geo:asWKT ?posTX.
   ?Receiver a cno:Receiver;
      cno:hasPosition ?posRXState.
   ?posRXState a opm:CurrentDataState;
      geo:asWKT ?posRX.}
```

Retrieving the cartesian coordinates of the transmitter and receiver and the transmitter frequency enables us to calculate the free space path loss as in equation 6.

$$\mathrm{FSPL} = 20\log_{10}\left(\frac{4\pi \cdot 21\,\mathrm{m} \cdot 3.75\,\mathrm{GHz}}{c_{air}}\right) = 70.37\,\mathrm{dB} \qquad (14)$$

To answer CQ3, the process model linked to the building model must be queried. The structure of the query depends on how detailed the process model was created. In its most simplistic form, a *ioc:Process* is created for each *bot:Element*, which describes the "finalization" of said element. The object property *ioc:hasOutputElement* describes that the instantiated Process creates the linked element. Finally, a schedule is added to that process via *ioc:hasSchedule*. To keep it simple, the schedule is only connected to *xsd:datetime*, which expresses the scheduled end time of the process with the help of *prov:endedAtTime*. Querying elements that are the output of processes with this schedule metadata and using a SPARQL Filter for the timestamp will retrieve the necessary building element IRIs.

All values from Table 3 of ITU-R P.2040-3 are collected in a JSON file to include the relevant material properties for radio wave propagation to the data model. Hereafter, material aliases are added to map the materials from the Table to materials in the IFC. Depending on which authoring software is used, these can differ slightly. The Duplex was modelled with Revit, so exemplarily, "Brick" needs to be mapped to "Masonry - Brick". Finally, a Python script queries the data model for IFC materials, locates the materials via the aliases in the JSON file, and updates the graph with the needed material properties. An excerpt of the resulting database is depicted in figure 4. It shows the process logic on the top, LBD element and *ifcOWL* structure, and two of the material layers on the bottom. For enhanced readability, we refrained from labelling *ifcOWL* object properties.

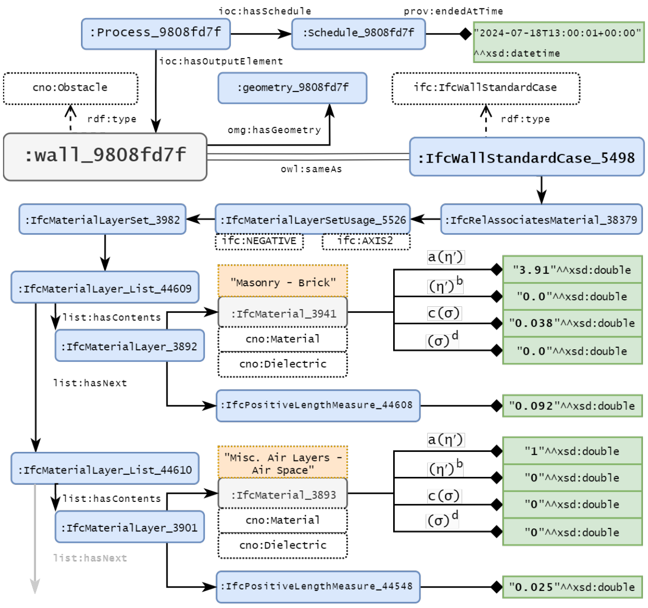


Figure 4: Excerpt of the data model extended with the help of the cno classes and properties.

Querying the database to answer QC4-5 mainly involves traversing the *ifcOWL* structure, which is complex. While for many use cases, the LBD ontologies and, most specifically, *props* in combination with geometry is more than sufficient, we need the full expressiveness of the IFC data model here. RDF lists need to be processed when navigating via *relAssociates* to the material layers. This is best done by using the SPARQL 1.1 property path features. Listing 2 shows a query realising this. The asterisk in *list:hasNext** shows that we follow this property to the end of the list via a path of "zero or more occurrences".

Listing 2: SPARQL Query for CQ5-7

```
PREFIX cno: <http://w3id.org/cno#>
PREFIX qudt: <http://qudt.org/schema/qudt#>
PREFIX unit: <http://qudt.org/vocab/unit#>
PREFIX list: <https://w3id.org/list#>
PREFIX express: <https://w3id.org/express#>
PREFIX ifc: <https://standards.buildingsmart.org/IFC/DEV/
    ↪ IFC2x3/TC1/OWL#>

SELECT ?Axis ?Dir ?dLayer ?MatType ?a ?b ?c ?d WHERE
  { BIND (<Element-IRI> as ?Element)
  ?IfcRelAdMat ifc:relatedObjects_IfcRelAssociates ?Element;
  ifc:relatingMaterial_IfcRelAssociatesMaterial ?IfcMat.
  ?IfcMat a ifc:IfcMaterialLayerSetUsage;
       ifc:forLayerSet_IfcMaterialLayerSetUsage ?
           ↪ IfcMatLayerSet;
       ifc:layerSetDirection_IfcMaterialLayerSetUsage ?Axis;
       ifc:directionSense_IfcMaterialLayerSetUsage ?Dir.

  ?IfcMatLayerSet ifc:materialLayers_IfcMaterialLayerSet ?
      ↪ IfcMatLayer_List.
  ?IfcMatLayer_List list:hasNext* ?IfcMatLayer_List_Elem.
  ?IfcMatLayer_List_Elem list:hasContents ?IfcMatLayer.
  ?IfcMatLayer ifc:material_IfcMaterialLayer ?IfcMatLayered;
      ifc:layerThickness_IfcMaterialLayer ?IfcMatLayered_d.
  ?IfcMatLayered_d express:hasDouble ?dLayer.

  ?IfcMatLayer ifc:material_IfcMaterialLayer ?Mat.
  ?Mat rdf:type ?MatType;
      cno:hasRelativePermittivityFrequencyCoefficient ?a;
      cno:hasRelativePermittivityFrequencyExponenet ?b;
      cno:hasConductivityFrequencyCoefficient ?c;
      cno:hasConductivityFrequencyExponent ?d.}
```

This query returns the sequence of materials in the layered element. The direction, axis and thickness of each layer describe the geometric properties. The material properties introduced in the cno ontology contain the values to be inserted into the required propagation loss equations. These depend on the type of material, either a *cno:Conductor* or a *cno:Dielectric*. For the first layer queried, "Masonry - Brick" (see figure 4), we calculate:

$$\begin{aligned} \eta' &= 3.91 \cdot 3.75\,\text{GHz}^{0.0} = 3.91 \\ \sigma &= 0.0238 \cdot 3.75\,\text{GHz}^{0.16} = 0.81\,\text{S}\,\text{m}^{-1} \\ A_{dielectric} &= 1636 \cdot \frac{0.03}{\sqrt{3.91}} = 24.33\,\text{dB}\,\text{m}^{-1} \end{aligned} \tag{15}$$

Multiplying the calculated attenuation rate by the layer thickness of 0.092 m gives an attenuation of about 2.23 dB. This process must be repeated for all 6 layers of the component, resulting in a calculated propagation loss of 20.61 dB and an average attenuation rate of 49.4 dB m^{-1}. While this value is consistent with the approximate values found in the literature, it should be noted that this calculation is highly simplified. On the one hand, the layered wall has a stud metal layer, which can significantly affect propagation due to its conductive behaviour. On the other hand, we do not consider other effects, such as Snell's law or additional reflections that are relevant here.
To complete this proof of concept, the transmitting and receiving antenna gain is taken from listing 1. Focusing on the FSPL and the element propagation loss, combining the values to get the total propagation loss and inserting it into equation 4, we can calculate the received power in dBm and thus the expected RSSI (see equation 16).

$$36.98\,(\text{dBm}) + 9\,(\text{dBi}) - 90.97\,(\text{dB}) + 5\,(\text{dBi}) = -39.99\,(\text{dBm}) \tag{16}$$

Conclusions and Outlook

Construction sites differ from other production environments due to the ever-changing site topology, leading to restricted radio network connectivity and performance. To address this, an ontology is designed and developed that introduces and links concepts relevant to radio wave propagation with existing building and construction process modelling approaches. These concepts have been developed with respect to both the scientific fundamentals and the material properties available in official reports. To test the usability of the approach, a sample data set describing the 5G installations of the reference construction site in Aachen and a BIM model enriched with scheduling was generated and queried. The values obtained could be used in simplified, basic equations to determine the network quality, thus demonstrating that the approach makes it possible to link the necessary variables with the construction process context.
The described use case can prove that the necessary data can be retrieved to perform basic estimations. For this, the developed ontology as open and extensible data standard the facilitates access to information independent of software vendors, ensuring accessibility and interoperability. However, the ultimate goal is to feed more complex simulation environments to enable accurate network quality estimation for network optimization. To achieve the required accuracy, the ontology will need to be extended to include material properties such as albedo, roughness, or refractive index to allow for the calculation of effects such as reflection, refraction, and scattering, which were beyond the scope of this work. Future research will aim to validate more complex, construction context-dependent simulations with precise measurements in real-world scenarios. In this way, further insights will be gathered to create a network infrastructure that can adapt to the evolving construction environment by predicting network coverage and quality, thus driving innovation in the construction industry, such as digital twins and construction robotics.

Acknowledgments

This work is part of the research project EConoM, funded by the Federal Ministry for Digital and Transport of Germany within the initiative InnoNT (funding number 19O122009F). It was supported within the TARGET-X framework, a project funded by the Smart Networks and Services Joint Undertaking (SNS JU) under Horizon Europe (funding number 101096614). The authors are responsible for the content.

References

Ahsan, S., El-Hamalawi, A., Bouchlaghem, D., and Ahmad, S. (2009). Applications of converged networks in construction. International Journal of Product Development, 7(3/4):281–300.

Beetz, J. (2009). Facilitating distributed collaboration in the AEC/FM sector using Semantic Web Technologies. PhD thesis, Built Environment. https://doi.org/10.6100/IR652808.

Brell-Cokcan, S. and Schmitt, R. H., editors (2024). IoC - Internet of Construction; 1. Springer Fachmedien Wiesbaden, Wiesbaden.

Din, Z. U. and Bernold, L. E. (2017). Experimental study of signal behavior for wireless communication in construction. Construction Innovation, 17(4):475–491.

Erunkulu, O. O., Zungeru, A. M., Lebekwe, C. K., and Chuma, J. M. (2020). Cellular communications coverage prediction techniques: A survey and comparison. IEEE Access, 8:113052–113077.

Friis, H. T. (1946). A note on a simple transmission formula. Proceedings of the IRE, 34(5):254–256.

García Sánchez, M., Iglesias, C., Cuiñas, I., and Expósito, I. (2022). Building penetration losses at 3.5 ghz: dependence on polarization and incidence angle. Electronics, 12(1):106.

Heath, T. and Bizer, C. (2011). Linked data: Evolving the web into a global data space. Synthesis lectures on the semantic web. theory and technology, Springer Nature Switzerland Ag - Cham 6330, Gewerbestrasse 11 , HANDEL. Switzerland., 1(1):1–136.

ITU (2021). Compilation of measurement data relating to building entry loss, report itu-r p.2346-4.

ITU (2023). Effects of building materials and structures on radiowave propagation above about 100mhz. recommendation p.2040-3. itu-r.

Kirner, L., Oraskari, J., Wildemann, P., and Brell-Cokcan, S. (2024). Internet of Construction Process Ontology (ioc) v 0.5.

Lassila, O., Swick, R. R., et al. (1998). Resource description framework (rdf) model and syntax specification. W3C recommendation.

Lee, H. J., Krishnan, A., Brell-Cokcan, S., Knußmann, J., Brochhaus, M., Schmitt, R. H., Emontsbotz, J. J., and Sieger, J. (2022). Importance of a 5g network for construction sites: Limitation of wlan in 3d sensing applications. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction, volume 39, pages 391–398. IAARC Publications.

Noy, N. F. and McGuinness, D. L. (2001). Ontology development 101: A guide to creating your first ontology. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880.

Oraskari, J., Senthilvel, M., and Beetz, J. (2021). Shacl is for lbd what mvdxml is for ifc. In Proc. of the Conference CIB W78, volume 2021, pages 11–15.

Pauwels, P. and Terkaj, W. (2016). EXPRESS to OWL for Construction Industry: Towards a Recommendable and Usable ifcOWL Ontology. Automation in Construction, 63:100—133. 10.1016/j.autcon.2015.12.003.

Pauwels, P., Zhang, S., and Lee, Y.-C. (2017). Semantic web technologies in AEC industry: A literature overview. Automation in Construction, 73:145–165. http://doi.org/10.1016/j.autcon.2016.10.003.

QUDT.org (2015). Quantities, units, dimensions and types (qudt). doi:10.25504/FAIRSHARING.D3PQW7.

Rao, A. S., Radanovic, M., Liu, Y., Hu, S., Fang, Y., Khoshelham, K., Palaniswami, M., and Ngo, T. (2022). Real-time monitoring of construction sites: Sensors, methods, and applications. Automation in Construction, 136:104099.

Rasmussen, M., Lefrançois, M., Bonduel, M., Hviid, C., and Karlshøj, J. (2018). OPM: An ontology for describing properties that evolve over time. In Proceedings of the 6th Linked Data in Architecture and Construction Workshop, volume 2159, pages 24–33.

Rasmussen, M. H., Lefrançois, M., Schneider, G., and Pauwels, P. (2020). BOT: the Building Topology Ontology of the W3C Linked Building Data Group. Semantic Web.

Seidel, S. Y. and Rappaport, T. S. (1994). Site-specific propagation prediction for wireless in-building personal communication system design. IEEE transactions on Vehicular Technology, 43(4):879–891.

Ullah, U., Kamboh, U. R., Hossain, F., and Danish, M. (2020). Outdoor-to-indoor and indoor-to-indoor propagation path loss modeling using smart 3d ray tracing algorithm at 28 ghz mmwave. Arabian Journal for Science and Engineering, 45(12):10223–10232.

Valluru, P., Karlapudi, J., Menzel, K., Mätäsniemi, T., and Shemeikka, J. (2020). A semantic data model to represent building material data in aec collaborative workflows. In Boosting Collaborative Networks 4.0: 21st IFIP WG 5.5 Working Conference on Virtual Enterprises, PRO-VE 2020, Valencia, Spain, November 23–25, 2020, Proceedings 21, pages 133–142. Springer.

Yamamoto, B., Wong, A., Agcanas, P. J., Jones, K., Gaspar, D., Andrade, R., and Trimble, A. Z. (2019). Received signal strength indication (rssi) of 2.4 ghz and 5 ghz wireless local area network systems projected over land and sea for near-shore maritime robot operations. Journal of Marine Science and Engineering, 7(9):290.

REVOLUTIONISING INDIAN HIGHWAY PROJECTS: UNLEASHING THE POWER OF DATA INTEGRATION WITH CUTTING-EDGE DATA LAKE

Chenchu Murali Krishna[1], Kirti Ruikar[2], and Kumar Neeraj Jha[3]
[1]Indian Institute of Technology Delhi, New Delhi, India
[2]Loughborough University, Loughborough, United Kingdom
[3]Indian Institute of Technology Delhi, New Delhi, India

Abstract

Indian highway projects grapple with multifaceted delays, from contractual hurdles to legal complexities. Addressing this, the National Highway Authority of India (NHAI) embraces a cloud-based Data Lake, ushering in a "Fully Digital" paradigm. This advanced tool, validated through a literature review and NHAI stakeholder interviews, forecasts and manages delays systematically. The study reveals NHAI's adept integration of Data Lake technology across project phases, offering tangible solutions and underlining the transformative potential for data-driven decision-making in highway infrastructure projects. Recommendations advocate wider adoption, continual development, and prioritising education and research in delay management, heralding a new era of operational efficacy.

Keywords: Construction Delays, Data Lake, Highway Projects, Delay Management.

Introduction

In the realm of constructing India's expansive road network, measuring approximately 5.89 million kilometers, the continuous challenge of ensuring the timely completion of highway projects stands as a significant obstacle (Ministry of Road Transport & Highways Year End Review, 2022). This herculean task encounters numerous hurdles ranging from unpredictable weather conditions and complex land acquisition processes to bureaucratic permissions and the looming threat of supply chain disruptions. A study by the National Cooperative Highway Research Program (NCHRP) reveals that a staggering 40 per cent of projects experience delays, leading to escalated costs, compromised quality, and a noticeable decline in public satisfaction. These delays not only inflate project budgets and diminish quality but also erode trust among the public. Traditionally, project managers have relied heavily on their expertise and intuition to anticipate and manage potential delays. Yet, this conventional approach often falls short of identifying hidden delays until they materialise (Tripathi *et al.*, 2023). The imperative arises for a paradigm shift toward a more data-centric approach to delay management.

In this era of data innovation, data lakes have emerged as critical components for managing enormous quantities of structured and unstructured data efficiently. Functioning as a centralised repository, a data lake allows for the storage of data in its original format. It facilitates detailed analysis through an array of data mining and machine-learning techniques (Mathis, 2017). This seismic shift towards data lakes provides a transformative solution for highway project delay management, moving away from the traditional, manual forecasting techniques that are time-consuming data gathering and inherent to inaccuracies.

Data lakes are preferred over traditional approaches like data warehouses and data marts due to their unmatched flexibility, scalability, democratised access to data, and cost-effectiveness (Hai *et al.*, 2023). Unlike rigid data warehouses that require predefined schemas, data lakes store raw data in its original form, accommodating diverse data types seamlessly (Singh *et al.*, 2022). Scalability is another key advantage, enabling data lakes to handle the growing data volumes of highway projects seamlessly. This scalability, coupled with the lower total cost of ownership, makes data lakes an attractive option for organisations and innovation by providing a centralised repository for organisational data, unlike data marts tailored to specific domains (El Haddadi *et al.*, 2020). Additionally, data lakes facilitate advanced analytics techniques like AI and Machine learning, empowering organisations to extract actionable insights from their data. The decision to prioritise data lakes over well-defined data structures and explicit information requirements reflects a strategic choice to embrace technologies such as UML, graph databases, and ontologies (Sawadogo, 2019; Singh and Ahmad, 2019; Nambiar and Mundra, 2022). A comprehensive comparison of these technological options is shown in Table 1, highlighting the unique strengths and advantages that data lakes offer in meeting the evolving demands of construction big data challenges in the digital age.

In the dynamic landscape of highway construction, where the precision of timing and meticulous management of resources is paramount, the integration of data lakes emerges as a pivotal element for advancing project efficiency (Nargesian *et al.*, 2018). This advanced system empowers project managers to aggregate and scrutinise data from a myriad of sources, encompassing historical project data, predictive weather analytics, and real-time insights from sensors strategically deployed at construction sites. The advent of machine learning algorithms within these data lake frameworks acts as a formidable guard against potential delays. These algorithms are adept at identifying complex patterns, recognizing trends, and discovering connections within the vast pool of integrated data, offering not only the ability to forecast delays but also the means to proactively prevent them (Gudivada *et al*., 2017). Imagine a scenario in highway construction where delays are not merely managed but actively prevented (Durdyev and Hosseini, 2020). This scenario is made possible through the seamless integration of varied data sources, all housed within the secure confines of a data lake, orchestrated by the strategic application of machine learning algorithms. This not only underscores the role of data lakes as mere storage facilities but elevates them to the role of architects in a data-centric revolution in highway project management. This paradigm shift underscores the narrative that progress in the construction sector transcends the traditional reliance on physical materials, pivoting towards the strategic leverage of data. As the industry leans into this transformative approach, it stands to redefine delays not as mere hurdles but as opportunities to refine resource allocation, boost quality, and uplift public satisfaction.

Table 1: Comparison of the different technologies with Data Lake

S. No.	Database	Stores Raw Data	Structured Data Support	Schema Flexibility	Horizontal Scaling	Querying & Analytics	Why Data Lake is Preferred
1	Data Lake	✓	✓	✓	✓	✓	Flexibility in storing raw data in its original format
2	Data warehouse		✓		✓	✓	Structured storage for consistent querying and reporting
3	NoSQL Databases	✓	✓	✓	✓	✓	Flexibility in schema and scalability for diverse data types
4	Object Storage	✓			✓		Scalability and cost-effectiveness for storing large objects
5	Unified Modeling Language					✓	Conceptual modelling and communication in software development
6	Graph Databases				✓	✓	Ability to model relationships between entities effectively
7	Ontologies					✓	Explicit representation of domain knowledge and relationships

Data Lake Implementation at NHAI

The NHAI has introduced a data lake, marking a significant step forward in managing highway projects. This innovation brings enhanced efficiency, transparency, and the ability to make decisions based on data. The development involved careful planning and implementation, beginning with the initial phase of recognising the essential need for a centralised data repository. This necessity arose from the complexities involved in managing highway projects. Consequently, NHAI decided to establish a cloud-based data lake, laying the foundation for a revolutionary digital infrastructure. In collaboration with technology partners, the NHAI initiated the design of a data lake's architecture, focusing on ensuring scalability, reliability, and the capacity to handle various data types (Khine and Wang, 2018).

The selection of cloud infrastructure from premier providers like AWS or Azure was made to establish a strong foundation for the platform's operations (Sharma, 2018). The development phase involved dedicated teams working diligently to realise the data lake's vision, integrating it effectively with existing systems such as project management and GIS for a smooth transition. Stringent testing measures were put in place to maintain the utmost standards of data integrity, security, and efficiency. This included thorough user acceptance testing with participation from stakeholders across different functions, ensuring it meets the changing requirements of NHAI. Additionally, detailed training programs were organised for end users such as contractors, engineers, project directors, and regional officers, enabling them to utilise the data lake fully in their daily activities.

The data lake was smoothly incorporated into the NHAI's pre-existing workflows, replacing manual procedures with automated systems. This enhancement significantly improved efficiency and accountability across various operations. The integration encompassed project documentation, communications, workflow monitoring, timelines, notifications, and financial transactions related to projects, and all facilitated through the data lake interface. This allowed stakeholders to make well-informed decisions based on up-to-the-minute data insights. Furthermore, the compulsory inclusion of drone surveys for every project enabled thorough monitoring and analysis, substantially strengthening NHAI's capability in managing projects. The data lake serves a broad array of end users, including contractors, engineers, project directors, regional officers, and headquarter staff. Each group of stakeholders enjoys access to customised features and functionalities designed to meet their unique requirements, thereby empowering them to make decisions rooted in data, which in turn promotes project success and organisational development.

NHAI data lake architecture

The data lake architecture employed by the NHAI is designed to efficiently handle large volumes of diverse data, facilitating data-driven decision-making and operational efficiency within the organisation. It consists of key components such as the ingestion layer, storage layer, processing layer, catalogue and metadata management, governance and security, analytics and visualisation, and finally sharing and collaboration components. The simplified breakdown of the key components is as follows:

Data Ingestion Layer: This layer focuses on collecting data from various sources, including internal databases, APIs, and external sources like documents and images. It ensures the seamless ingestion of structured and semi-structured raw data types relevant to NHAI's operations (Singh and Ahmad, 2019).

Data Storage Layer: Once data is ingested, it is stored in a distributed file system, such as Hadoop Distributed File System (HDFS) or cloud-based storage solutions like Amazon S3 or Azure Blob storage. This ensures secure and efficient storage of data, enabling accessibility for further processing and analysis (Inmon, 2016).

Data Processing Layer: In this layer, tools and frameworks like Apache Spark, Apache Flink, or Apache Beam are utilised to process and transform raw data into usable formats. Batch and stream processing techniques enable real-time or near-real-time data processing, enhancing NHAI's analytical capabilities (Sharma, 2018).

Data Catalog and Metadata Management: NHAI's data lake architecture incorporates a metadata management system to catalogue and organise the vast amount of stored data. This metadata provides valuable information about the structure, format, and lineage of the data, facilitating easy discovery and understanding of available datasets (Madera and Laurent, 2016).

Data Governance and Security: Data governance policies and security measures are integral components of NHAI's data lake architecture. Access control mechanisms, encryption, data masking, and compliance with regulatory requirements ensure the privacy and security of sensitive information stored within the data lake (Abraham *et al*., 2019).

Data Analytics and Visualisation: NHAI's data lake enables data analytics and decision-makers to perform advanced analytics and derive insights from the stored data. Tools like Apache Hadoop, Apache Spark, or specialised analytics platforms facilitate data analysis and visualisation through dashboards or reports, empowering informed decision-making (Nagel *et al.,* 2021).

Data Sharing and Collaboration: NHAI's data lake architecture includes features for sharing data and collaborating with internal and external stakeholders. This may involve providing APIs for accessing data, implementing data-sharing agreements, enabling secure data exchange protocols, and fostering collaboration and knowledge sharing within the organisation (Couto and Ruiz, 2022).

Data lake for delay management

In the domain of highway construction, the data lake emerges as a powerful tool for effective delay management. Acting as a centralised repository, it adeptly accommodates structured and unstructured data of any magnitude, enabling the integration and analysis of information from various sources, such as construction sites, traffic sensors, and weather reports. This versatility makes the data lake a crucial component in effective delay management, which is evident in its diverse applications (Edison and Singla, 2020).

Primarily, the real-time monitoring capabilities of data lakes empower project managers to gather instantaneous data from varied sources ranging from sensor metrics and weather projections to traffic dynamics and social media updates (Gorelik, 2019). This comprehensive data snapshot enables proactive decision-making, allowing project managers to discern potential delays and take corrective measures promptly. Secondly, through the lens of predictive analytics, data lakes craft models that discern trends in historical data, translating them into actionable insights for anticipating and mitigating future delays. For instance, a meticulous analysis of past weather patterns facilitates pre-emptive measures against weather-induced schedule disruptions (Nambiar and Mundra, 2022).

Furthermore, data lakes facilitate a detailed optimisation of resource allocation by identifying areas of underutilisation or excess usage. This optimisation strategy not only streamlines project timelines but also acts as a strong deterrent against delays. Additionally, these reservoirs of information play a pivotal role in root cause analysis by storing and examining historical data on previous delays. Armed with this knowledge, organisations can effectively implement measures to minimise and prevent similar delays in future projects. Additionally, data lakes transcend their role as mere repositories by fostering collaboration among project stakeholders. By providing a centralised hub for project-related data, they facilitate seamless communication and collaboration across diverse teams, ensuring alignment toward common project objectives (Hagstroem *et al.*, 2017; Chomo, 2019).

In summary, data lakes serve as invaluable assets in the ongoing effort to minimise delays in highway construction projects. Through the integration and analysis of data from diverse sources, project managers gain deep insights, enabling them to make informed decisions and proactively guide projects toward successful completion. This fusion of big data within the data lake seamlessly addresses the complexities of highway construction delay factors, showcasing a nuanced alignment between the features of the data lake and the diverse challenges encountered in construction projects, as illustrated in Table 2.

Methodology

This study was conducted within the framework of the NHAI and involved a diverse group of stakeholders. The study aimed to understand the subjective experiences and interpretations of these stakeholders' concerning delays in highway projects. By adopting a collaborative approach, the research sought to generate knowledge that reflects the multifaceted perspectives of those involved, thus acknowledging their crucial role in shaping the findings. The methodology of the study was structured into four stages, designed to align with the research objectives. The investigation utilised both secondary and primary data sources. Initially, the research identified the problem of project delays through an extensive literature review, the author's direct involvement with highway projects, and informed discussions with experts.

Table 2: Matrix that maps delay management features for highway projects against common causes of delays

Cause of delay	Delay management features			
	Real-time monitoring	***Predictive analytics***	***Resource Allocation & optimisation***	***Collaboration***
Poor planning and scheduling of the project	✓			✓
Site clearances	✓			✓
Land acquisition and rehabilitation issues	✓	✓		✓
Weather and force majeure-related delays	✓	✓	✓	✓
Project changes and redesigns	✓	✓	✓	✓
Permit and approval delays	✓	✓		✓
Stakeholder conflicts	✓	✓	✓	✓
Poor resource management	✓		✓	✓
Inadequate project monitoring	✓	✓		✓
Poor communication between construction parties	✓			✓

This phase of the research identified 182 publications related to project delays and data lakes, with 56 unique articles remaining after duplicates were removed. These articles, dating from 2006, were sourced from databases such as Scopus, Elsevier, and Web of Science. Following the identification of delays, theoretical propositions were formulated to address these delays, based on analysis of the initial data. Qualitative data was collected through semi-structured interviews with six stakeholders, including deputy managers experienced in data lake-integrated projects, two project managers and two senior research followers actively involved in data lake-related projects. Table 3 shows the details of the interviewers. These interviews aimed to gather insights on the identified causes of delays, the theoretical propositions, and the impact of data lake processes and technologies in mitigating these issues. Interviewees are chosen for their efficacy in eliciting information about non-observable phenomena. The final stage involves formulating recommendations and conclusions derived from the content analysis of the collected interview data. This comprehensive approach, integrating data, information, and knowledge through the prism of a data lake, has the potential to revolutionise the landscape of effective data utilisation in highway construction projects, leveraging the vast reservoir of big data available, paving the way for a revolution in the field.

Table 3: Details of the Interviewees

Parameters	Stakeholders					
	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Designation	Deputy Manager	Deputy Manager	Project Manager	Project Manager	Senior Researcher	Senior Researcher
Experience (Years)	18	20	15	15	4	3

Propositions

The theoretical propositions of this study are drawn from a thorough review of the literature, focusing on the attributes of a data lake anticipated to contribute significantly to the mitigation of project delays. The study primarily explores the implementation of a data lake in the context of Indian highways, aiming to quantify its advantages vis-à-vis the prominent causes of delays. However, the study's scope is limited to examining a data lake's potential to improve communication, information flow, and coordination among project stakeholders.

The study proposes further investigation into how a data lake can enhance communication and collaboration in highway projects (Nargesian *et al.*, 2018; Giebler *et al.*, 2019). The existing literature research lays a robust foundation for the formulation of key propositions, integral to guiding interview questions design and subsequent data collection:

1. *Proposition 1*: Implementation of a Data Lake in highway projects enhances data visualisation and construction process comprehension, thereby mitigating delays in progress payments and expediting client decision-making.
2. *Proposition 2:* Adoption of a data lake in NHAI projects facilitates effective delay management while fostering improved communication and coordination between the client and other stakeholders.

These propositions serve as the compass for the study, directing the formulation of insightful interview questions and the meticulous collection of data from pertinent stakeholders. This approach allows for a detailed exploration of the benefits and challenges inherent in implementing a data lake within the context of Indian highways, unravelling the untapped potential of data integration for practical usage in highway construction projects.

Interview Analysis

Proposition 1: Implementation of a Data Lake in highway projects enhances data visualisation and construction process comprehension, thereby

mitigating delays in progress payments and expediting client decision-making.

The efficacy of a data lake in highway construction projects is evident in its ability in visualising diverse datasets and operational aspects. Consolidating all data into a singular repository provides a comprehensive overview of project status, progress, and performance, enabling stakeholders to make data-driven decisions. Addressing concerns outlined in prior research such as Keane and Caletka (2015) and Du *et al.* (2018) pertaining to insufficient information about construction progress and communication gaps among stakeholders causing delays in progress payments and decision-making, the content analysis of interviews underscores the substantial benefits that implementing a data lake brings to the fore in managing delays and enhancing decision-making processes.

The content analysis of interviews highlights that implementing a data lake in highway projects brings significant advantages for managing delays and improving decision-making. Data visualisation enables quick identification of concerns and prompt corrective actions. Access to centralised data facilitates decision-making based on real-time insights. However, successful implementation requires understanding the data and selecting meaningful metrics. Transparent processes and workflows within the data lake enhance utilisation and support process improvement. It is important to note that implementing a data lake alone is not sufficient; it requires a comprehensive approach to leverage its benefits effectively.

The interviews emphasise that the implementation of a data lake in highway projects significantly bolsters delay management and decision-making processes. The key insights derived from the content analysis illuminate critical considerations:

1. *Strategic Metric Selection for Visualization*: Based on the input from deputy manager-level stakeholders and project managers who have experience with data lake-integrated projects, it is unanimously agreed that selecting meaningful metrics for data visualisation is critical. This strategic metric selection ensures that the visualised data is relevant and resonates with its audience, which is crucial for effective decision-making. Strategic metric selection in data lake is key for enhancing business performance through a data-driven approach. Metrics related to operational efficiency, data integration, and business analytics are essential for effective visualisation (Laurent *et al.*, 2020; Barbierato *et al.*, 2021; Kumar and Chundi, 2023).
2. *Expedited Decision-Making Process:* The consensus among interviewees highlights that the implementation of data lakes significantly accelerates the decision-making process by offering streamlined access to comprehensive and centralised data, fostering a culture of prompt and well-informed decisions. Beyond speeding up decision-making, data lakes bring technological advantages by improving flexibility and scalability, democratising data access. They also transform the business paradigm by addressing big data challenges and driving digital transformation, ultimately enhancing business intelligence. By providing a centralised repository of virtually inexhaustible raw data for analytical activities, data lakes enable enterprises to profoundly improve their decision-making processes and business intelligence, thereby transforming their overall business paradigms (Terrizzano *et al.*, 2015; Johny and Pillai, 2022).
3. *Visualisation of Project Status and Impact of Delays:* The interview findings highlight the significance of visualising data pertaining to schedule overruns and activity delays to enhance decision-makers' understanding of project status, with projection of the completion schedule offering insights into the impacts of delays and enabling proactive mitigation efforts. Data lakes at NHAI play a vital role in democratising data access and supporting divers' analytics tasks within enterprises. The data lake structure comprises five layers, enabling effective visualisation of project status by integrating multiple time series data sets. The adaptability of data lakes is exemplified by the NHAI dashboard, which positions stakeholders of data lakes as invaluable tools for the visualisation and analysis of project statuses (Fang, 2015; Mathis, 2017; Kumar and Chundi, 2023; Schneider *et al.*, 2023).
4. *Process Establishment and Improvement:* Interview data affirm that data lakes play a pivotal role in facilitating the establishment and documentation of processes and workflows within NHAI. This analysis helps in identifying areas that need improvement, resulting in greater efficiency in overall processes. Additionally, data lakes enable the use of advanced data-driven analysis techniques, which significantly aid enterprises in optimising the NHAI business operations. This not only captures key process parameters but also serves as a foundation for rigorous analysis. The insights derived from this analysis become instrumental in identifying areas for improvement, ultimately leading to enhanced overall process efficiency (Nagel *et al.*, 2021; Schneider *et al.*, 2023).

In conjunction with the literature review propositions, the synthesis of interview findings underscores that the implementation of data lakes in highway projects has the potential to be a game-changer in decision-making through data visualisation. However, it is imperative to acknowledge that successful implementation necessitates not only adequate resources for maintenance and development but also the addressing of a discernible skill gap in data lake expertise. The collaboration and knowledge-sharing among industry professionals, as underscored by senior research follows actively engaged in data-lake-related projects, emerge as critical components for establishing best practices and standards in data implementation and data lake utilisation. The call to allocate resources for ongoing support, encompassing infrastructure, data integration, quality management, and governance, resonates as a strategic imperative for organisations venturing into the realm of data lake implementation in highway construction projects.

Proposition 2: Adoption of a data lake in NHAI projects facilitates effective delay management while fostering improved communication and coordination between the client and other stakeholders.

In delving into the realm of effective data utilisation within highway construction projects, insights obtained from interviews with key stakeholders underscore the crucial role of a data lake in enhancing communication dynamics. The literature, supported by expert opinions, emphasises the importance of seamless communication to prevent delays, with inadequate exchange leading to suboptimal designs and planning (Saini, 2015).

Interview analysis highlights the significant impact of a data lake on communication between clients and stakeholders, serving as a keystone for real-time updates, meticulous tracking of communication flows, and seamless access to relevant information within a centralised platform (Malacarne *et al.*, 2018). This data lake design intricately reflects all communications within stakeholders' dashboards, ensuring immediate action based on document flow, approval status, and remarks of respective officers. This creates a harmonised and transparent communication channel that fosters enhanced coordination, mitigating delays attributed to internal or coordination factors. The implementation of a data lake in highway projects not only improves communication but also streamlines decision-making processes, leading to more efficient project management and ultimately, successful project completion.

In essence, the interview analysis manifests in phrases that unveil the data lake's diverse role:

1. *Improving Communication between Clients and Stakeholders*: The interviews with key stakeholders, including deputy manager-level experts in data lake-integrated projects, project managers, and senior research fellows deeply entrenched in data lake-related endeavors, yielded profound insights into the enhancement of communication dynamics between clients and stakeholders. A pivotal observation emerged: the strategic design of the data lake serves as a communication nexus, seamlessly reflecting all relevant interactions within stakeholders' dashboards. This design ensures not only visibility but catalyses immediate action on requests, leveraging the dynamic flow of documents, approval statuses, and remarks from respective officials.

   The real-time updates from the data lake proved instrumental in transforming project dynamics. Stakeholders, armed with instantaneous information on project status, documentation details, and approvals, navigated a landscape of enhanced coordination. This, in turn, emerged as a powerful countermeasure to delays induced by internal factors or coordination challenges. The interviews underscored a paradigm shift in communication dynamics, where the data lake became a pipeline for real-time collaboration and coordination, moving projects towards streamlined efficiency and reduced delays.
2. *Central Repository for Documents and Information:* In illuminating insights drawn from interviews with key stakeholders, the data lake emerges as a central repository for approved requests, project documents, drawings, notices, and more. According to the perspectives gleaned from two deputy manager-level stakeholders intimately acquainted with data lake integration, two project managers, and two senior researchers who are deeply entrenched in data lake-related projects, the data lake's role as a document repository is pivotal. This repository function, as articulated by the interviewees, not only ensures the organised storage of essential documents but also stands as a safeguard against human errors. The availability of information within this consolidated data lake reduces the likelihood of oversights and inefficiencies, contributing to smoother functioning. Importantly, the improved accessibility to this repository acts as a catalyst, streamlining processes and, consequently, playing a significant role in the reduction of delays within the context of highway construction projects. The consensus among these key stakeholders underscores the transformative impact of a data lake as more than just a repository—it is a dynamic facilitator that not only safeguards against errors but actively contributes to the efficiency and agility of project processes, aligning seamlessly with the overarching goal of minimising delays in highway construction projects.
3. *Facilitating Prompt Resolution of Requests:* In the interview analysis, stakeholders with expertise in data lake-integrated projects unanimously highlighted the data lake's instrumental role in expediting request resolutions and curbing delays. The real-time updates afforded by the data lake emerge as a cornerstone for well-informed decision-making, providing up-to-date insights that prove pivotal in navigating project intricacies. The synergy of improved coordination and communication, facilitated by the data lake, contributes substantially to the efficiency of project management. This integrated approach ensures not only the prompt resolution of requests but also an overarching enhancement of project efficacy through streamlined communication channels and data-driven decision-making.

Conclusions and Recommendations

The research, centered on leveraging data lakes to streamline Indian highway projects, crystallises into two pivotal propositions. The first proposition underscores the transformative impact of visualising data within construction endeavors through data lake implementation. This strategic utilisation addresses delays in progress payments and decision-making with unprecedented efficacy. The second proposition underscores the instrumental role of data lakes in elevating communication and coordination among project stakeholders, thereby enhancing overall operational efficiency in delay mitigation. The study unequivocally concludes that data lakes represent a potent solution, offering substantial potential to rectify delays by optimising data management, communication, and coordination.

In extrapolating these findings, the imperative for the NHAI becomes evident. Prioritising two key domains data governance and skilled professionals emerges as the strategic pathway to maximise data lake utilisation.

1. ***Data Governance***: Effective data governance is essential for optimising the use of data lakes. To achieve this, organisations should first establish a clear data governance framework by setting clear goals, such as ensuring data quality, securing sensitive data, complying with regulations, and forming a cross-functional team to oversee policy implementation (Gillan, 2021). It is also important to maintain detailed documentation and manage metadata to keep data easily accessible and understandable. Implementing role-based access controls and maintaining audit trails helps protect sensitive information and monitor data usage. Organisations should classify data based on sensitivity and handle it, accordingly, applying necessary security measures for highly sensitive data. Managing the data lifecycle by setting policies for data retention, archiving, and deletion, and conducting regular audits ensures data remains compliant with legal and business requirements (Brous *et al.*, 2016). Continuous monitoring, measuring effectiveness, and adapting governance practices as necessary will maintain the data lake as a secure, compliant, and efficiently managed resource. This approach empowers organisations to fully leverage their data lakes while ensuring data security and compliance (Duzha *et al.*, 2023).
2. ***Skilled Professionals***: To effectively manage data lake systems in the construction industry, professionals require a combination of technical skills and soft skills. Technical skills encompass knowledge of data platforms like data warehouses and data lakes, proficiency in data management tools such as Delta Lake and Snowflake and understanding of data processing strategies like those in Hadoop (Kaur *et al.*, 2023). Additionally, soft skills like effective communication and presentation are vital for managing both data and human resources within construction teams. Professionals in the construction industry must possess skills in sensor data analytics and data science, including machine learning. Understanding data analytics applications across various construction phases is crucial. A methodology integrating Building Information Modelling (BIM) and Business Intelligence (BI) tools enables collaborative data management. The adoption of information communication technology (ICT) facilitates decision-making and project management (Huang *et al.*, 2012). Managing data lakes requires proficiency in data platforms, management tools, and processing strategies. Investing in skilled professionals is crucial for deriving insights and optimising operations. The NHAI recognises the importance of this and focuses on recruitment, retention, and training programs. Collaborative efforts with academia and industry experts further contribute to bridging the skills gap and fostering innovation within the construction industry. This holistic approach ensures that data lake systems are managed efficiently, driving progress, and enhancing the construction sector's competitiveness.

By focusing on these strategic imperatives of data governance and skilled professionals, NHAI is poised to unlock the full potential of their data lake. This approach not only facilitates effective analysis of highway data but also enhances operational efficiency and safety measures across the expansive highway network.

References

Abraham, R., Schneider, J. and Vom Brocke, J. (2019) 'Data governance: A conceptual framework, structured review, and research agenda', International Journal of Information Management, 49, pp. 424–438.

Barbierato, E., Gribaudo, M., Serazzi, G. and Tanca, L. (2021) 'Performance evaluation of a data lake architecture via modeling techniques', in European Workshop on Performance Engineering. Springer, pp. 115–130.

Brous, P., Herder, P. and Janssen, M. (2016) 'Governing Asset Management Data Infrastructures', in Procedia Computer Science. Elsevier B.V., pp. 303–310.

Chomo, T. (2019) Deploying Data Lake for Big Data Management. Masaryk University.

Couto, J.C. and Ruiz, D.D. (2022) 'An overview about data integration in data lakes', in 2022 17th Iberian Conference on Information Systems and Technologies (CISTI). IEEE, pp. 1–7.

Durdyev, S. and Hosseini, M.R. (2020) 'Causes of delays on construction projects: a comprehensive list', International Journal of Managing Projects in Business, 13(1), pp. 20–46.

Duzha, A., Alexakis, E., Kyriazis, D., Sahi, L.F. and Kandi, MA (2023) 'From Data Governance by design to Data Governance as a Service: A transformative human-centric data governance framework', in Proceedings of the 2023 7th International Conference on Cloud and Big Data Computing, pp. 10–20.

Edison, J.C. and Singla, HK (2020) 'Development of a scale for factors causing delays in infrastructure projects in India', Construction Economics and Building, 20(1), pp. 36–55.

Fang, H. (2015) 'Managing data lakes in big data era: What's a data lake and why has it became popular in data management ecosystem', in 2015 IEEE International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER). IEEE, pp. 820–824.

Giebler, C., Gröger, C., Hoos, E., Schwarz, H. and Mitschang, B. (2019) 'Leveraging the Data Lake: Current State and Challenges', Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11708 LNCS,

Gillan, A. (2021) 'Governance: the key driver for data-driven innovation', Computer Fraud & Security, 2021(4), pp. 10–13.

Gorelik, A. (2019) 'The enterprise big data lake: Delivering the promise of big data and data science.' O'Reilly Media.

Gudivada, V., Apon, A. and Ding, J. (2017) 'Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations', International Journal on Advances in Software, 10(1), pp. 1–20.

El Haddadi, O., El Hamlaoui, M., Dkaki, T. and Nassar, M. (2020) 'Data Lake and Digital Enterprise.', in ENASE, pp. 423–429.

Hagstroem, M., Roggendorf, M., Saleh, T. and Sharma, J. (2017) A smarter way to jump into data lakes.

Hai, R., Koutras, C., Quix, C. and Jarke, M. (2023) 'Data lakes: A survey of functions and systems', IEEE Transactions on Knowledge and Data Engineering [Preprint].

Inmon, B. (2016) Data Lake Architecture: Designing the Data Lake and avoiding the garbage dump. Technics publications.

Johny, M.G. and Pillai, S. (2022) 'Analysing the vital role of enterprise data lake in the era of digital transformation', in AIP Conference Proceedings. AIP Publishing.

Kaur, P., Kaushik, A. and Kapoor, A. (2023) 'Skills and Responsibilities of Data Wrangler', Data Wrangling: Concepts, Applications and Tools, p. 19.

Keane, P.J. and Caletka, A.F. (2015) Delay analysis in construction contracts. Wiley Blackwell Oxford, UK.

Khine, P. P. and Wang, Z.S. (2018) 'Data Lake: a new ideology in big data era', ITM Web of Conferences, 17, p. 03025.

Kumar, A. and Chundi, P. (2023) 'Data Lakes', in Encyclopedia of Data Science and Machine Learning. IGI Global, pp. 410–424.

Laurent, A., Libourel, T., Madera, C. and Miralles, A. (2020) 'The Gravity Principle in Data Lakes', Data Lakes, 2, pp. 187–199.

Mathis, C. (2017) 'Data Lakes', Datenbank-Spektrum, 17(3), pp. 289–293.

Ministry of Road Transport & Highways Year End Review-2022: Ministry of Road Transport and Highways.

Nagel, S., Corea, C. and Delfmann, P. (2021) 'Cognitive effects of visualisation techniques for inconsistency metrics on monitoring data-intensive processes', Information Systems Management, 38(4), pp. 342–357.

Nambiar, A. and Mundra, D. (2022) 'An Overview of Data Warehouse and Data Lake in Modern Enterprise Data Management', Big Data and Cognitive Computing. MDPI.

Nargesian, F., Zhu, E., Miller, R.J., Pu, K.Q. and Arocena, P.C. (2018) 'Data Lake management: Challenges and opportunities', in Proceedings of the VLDB Endowment. VLDB Endowment, pp. 1986–1989.

Saini, M. (2015) A framework for transferring and sharing tacit knowledge in construction supply chains within lean and agile processes. University of Salford (United Kingdom).

Sawadogo, P.N. (2019) 'Textual Data Analysis from Data Lakes', in New Trends in Databases and Information Systems: ADBIS 2019 Short Papers, Workshops BBIGAP, QAUCA, SemBDM, SIMPDA, M2P, MADEISD, and Doctoral Consortium, Bled, Slovenia, September 8–11, 2019, Proceedings 23. Springer, pp. 558–563.

Schneider, J., Gröger, C., Lutsch, A., Schwarz, H. and Mitschang, B. (2023) 'Assessing the Lakehouse: Analysis, Requirements and Definition.', in ICEIS (1), pp. 44–56.

Sharma, B. (2018) Architecting data lakes: data management architectures for advanced business use cases. O'Reilly Media.

Singh, A. and Ahmad, S. (2019) 'Architecture of data lake', International Journal of Scientific Research in Computer Science, Engineering, and Information Technology, 5(2), p. 4.

Singh, J., Singh, G. and Bhati, B.S. (2022) 'The implication of data lake in enterprises: A deeper analytics', in 2022 8th International Conference on Advanced Computing and Communication Systems (ICACCS). IEEE, pp. 530–534.

Terrizzano, I.G., Schwarz, P.M., Roth, M. and Colino, J.E. (2015) 'Data Wrangling: The Challenging Journey from the Wild to the Lake.', in CIDR. Asilomar.

Tripathi, O.P., Hasan, A., Jha, K.N. and Jain, A.K. (2023) 'Evaluating Government Contracts for Delays, Delay Damages, and Levy of Compensation Provisions', Journal of Legal Affairs and Dispute Resolution in Engineering and Construction, 15(1).

REAL-TIME ASSESSMENT OF REGULATORY COMPLIANCE OF CONSTRUCTION SITES

Nicholas Nisbet[1 2], Zijing Zhang[1], Selçuk Çidik[1]
[1]University College London, London, UK,
[2]AEC3 UK Ltd

Abstract

Previous work on automatic compliance checking has targeted static descriptions of the built environment represented in Building Information Models. In contrast, this work examines the potential of semantic mark-up to capture and then apply health and safety regulations to live construction sites. There are significant differences in what constitutes a testable metric and what constitutes a fact when considering process rather than product compliance. Extensions to the RASE approach are defined and demonstrated to accommodate these differences. This leads to extending the concept of 'decidability' with considerations around the relative 'timeliness' of information.

Introduction

Previous work has applied the RASE method to building regulations and other static assessments of building proposals. The paper examines how automated compliance checking of construction site processes differ from the compliance checking of facility models. The attention on wellbeing, health and safety in construction (WHS) is widening from retrospective consideration of reconstructions of past accidents and near-misses scrutinized in enquiries and hearings (Pirzadeh et al., 2017). Besides safety, compliance requirements provide rules that may be best assessed live, rather than retrospectively. Thus, WHS presents a suitable case to develop theory and practice on automated compliance checking for processes in construction work.

There has been an increasing use of contemporaneous and real-time monitoring of construction activity on site (Cheng et al., 2013). This can support the immediate triggering of alarms and warnings about safety due to the proximity of incompatible entities, such as heavy plant and operatives. One source of the rules for such triggering is the legislation and regulations defining safety compliance. These rules can be checked against real-time construction activity data and static information about construction sites, such as project planning and resourcing (PERT) plans, 3D and 4D models and methods of work statements. This paper examines examples of real-time information sources and safety regulations to propose how semantic mark-up based on the RASE methodology (Nisbet et al. 2008) can be adapted to provide machine-operable rules that can be evaluated continuously for automated compliance checking for construction processes. RASE is a mark-up approach that highlights the logical structure and metrics in documents, as illustrated below.

Previous work

There has been previous work on rules about construction sites, dynamic assessment in other sectors, rule capture and the representation of time in predicate logic.

Many recent works have focused on 'prevention through design' looking for design features in static BIM models which may indicate specific hazards that may need consideration during construction planning (Johansen et al., 2023; Schwabe et al., 2019; Tekbas et al.,2020; Yuan et al., 2019). Shen et al. (2022) used the term 'dynamic' to refer to the importance of the sequence of processes anticipated when following work procedures. The target of the rule checking was a '4D' BIM description and narrative. Rules were obtained by transcribing knowledge onto a product/activity/location template along with the hazard and solutions. Queries were accepted in free text with NLP being used to map terms to the template. Bao et al. (2022) used a static BIM model and additional safety measures inserted automatically by third-party software. Scripts detected topological adjacency information. Sequencing was obtained from the project PERT (program evaluation and review) plan. Rules were translated into SQL (structured query language) expressions. The outcomes were alerts within the BIM authoring platform with recommendations for improvements prior to construction commencing.

Continuous real-time assessment of dynamic construction sites is being considered as an aspect of digital twins (Li et al., 2021). Teizer et al. (2022) proposed a combined view seeking to ensure safety through (a) design and construction planning, (b) risk detection and (c) learning and feedback. The choice of rules to be applied was left open.

Pradhananga (2015) explored the tracking of heavy plant using GPS and the visualization of those tracks around work, hazard, material, travel, loading and dumping zones. Zones could be taken from GIS (geospatial information system) or BIM (building information modelling) information or deduced from the behaviour of vehicles. Productivity was measured by repetitive cycle times. Speed and proximity were detected as leading indicators for hazards. Xu (2023) developed a fuzzy-logic tree combining live monitoring weather, location and worker IoT feeds to generate a continuous feed assessing worker well-being and risk. Rules were acquired by agreeing decision tables and fuzzy logic tables of risk and association. Other sectors have similar needs but within a static built environment and with static operational patterns. For example, Yang et al., (2023) examined vehicle conflicts at airports. The work deduced rules from safety critical correlations between vehicle speeds, accelerations and convergence. Li et al. (2022) combined engineering and soil monitoring values to alert possible construction risks, using SPARQL (RDF query language) applied to the combined data sets. Nisbet et al (2023) has raised concerns concerning the use of SPARQL.

RASE is a mark-up approach to identify the logical structure and semantic metrics in normative text, by identifying sections and phrases as Requirements, Applicability, Selections or Exceptions. RASE has been applied to static rule situations such as comparing building proposals against planning (zoning) and building control (technical) regulations (Beach et al., 2020 and 2023). RASE has been shown to be able to generate propositional and predicate logic statements (Nisbet et al., 2022). RASE was selected for its 'no-code' approach offering transparency to both inspectors and constructors.: mark-up can be added, reviewed and, if necessary, improved, by domain experts. This 'white-box' approach is in contrast to conventional translation and programming which remains opaque to domain experts.

Whereas most discussions of propositional and predicate logic assume a static context, built around the present tense, rules about live environments may need to consider 'temporal logic' capturing the sequence of the processes of interest. This has extensive literature, given its importance in legal and contractual and computational analysis, summarized in Lamport (1994).

There are two candidate modes for temporal logic, one based on representing events, and the other based on representing intervals. Temporal logic can be expressed in predicate logic by considering the role of four verb tenses. Ploug et al. (2012) summarised previous work including Prior and Kripke who proposed four tenses P "It has at some time been the case that …" F: "It will at some time be the case that …" H: "It has always been the case that …" and G: "It will always be the case that …" which can be combined to generate more complex tenses. In order to integrate event-based view into predicate logic the operators 'before' and 'after' can be used to relate two disjointed predicate statements.

In summary, literature has focused on identifying invariant rules on facility designs and construction plans that must be true continuously but has not addressed the dynamic narrative around events on construction sites and the means to capture and apply such rules. There is a gap in the literature related this real-time assessment of construction process compliance.

Method

Given the uncertainties and lack of previous work, a design science research approach has been taken to explore and iterate towards a plausible approach. This allows the solution space to be explored without necessarily discovering all the limitations of the solution. Part of that solution space includes the use of existing ontologies around safety and built environment.

This approach was supported by a workshop invitation issued by the Health and Safety Executive (HSE), a UK regulatory body, as part of their 'Discovering Safety' programme. The HSE encouraged the engagement of three commercial solution providers each addressing live construction site information, and the lead author as an information integrator.

The following section examines some relevant theory and ontology resources. The next section then looks at the preparatory work applying RASE semantic mark-up to Clause 22 of the UK Construction (Design and Management) Regulations 2015 (C(D&M) 2015), prior to considering specific information sources. This paper then reports on experiments to examine information sources around some specific scenarios, how these map to target ontologies and how these information resources can be continuously tested. Each experiment iterated progressively towards an improved rule representation and a rule-engine to test its utility.

Theory

This section examines theory and ontology for safety compliance checking of construction processes including for the target model, for dictionary resources and for rules.

Ontology for construction processes including health and safety

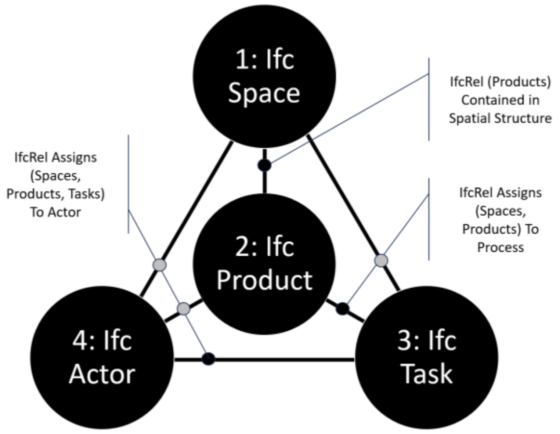
Both IFC (ISO 16739-1, 2024) and classification standards (ISO 12006-2, 2015) use a four-layer model for the descriptive and narrative representation of the built environment. The top-most layer describes the built environment in terms of named locations and spaces. These spaces are defined and supported by physical objects. These physical objects are affected by processes, activities and events. The processes are supported by resources including actors and construction aids.

This suggests that a tabular representation (Figure 1) can be used to track all the activities and timings (3) happening on the construction site with associated locations (1), products (2) and actors (4).

1: Location		2: Product		3: Activity			4: Actor	
Name	Description	Name	Description	Name	Description	Date and Time	Name	Description
S1	Site	B1	Facility	45246	Shift	16/11/2023 08:00	-	-
EX1	Excavation	-	-	IN1	Introduction	16/11/2023 08:10	CW1	Crew
EX1	Excavation	-	-	IN2	Introduction	16/11/2023 08:15	EX1	Excavator
EX1	Excavation	-	-	VW1	Inspection	16/11/2023 08:25	SE1	Engineer
EX1	Excavation	-	-	AP1	Approval	16/11/2023 08:35	SE1	Engineer
EX1	Excavation	-	-	EV1	Landslip	16/11/2023 08:47	CW1	Crew

Figure 1: Tabular information model and sample

In situations where there are multiple information sources available, a unifying ontology may be needed. IFC offers a comprehensive descriptive and narrative schema. Typical BIM models provide the physical product and some spatial location entities. To consider the dynamics of a construction site activity, tasks, events and actors are included in addition. Figure 2 shows the four IFC entity types in a tetrahedral arrangement. All six possible inter-relationships are provided by three IFC objectified relationship entities. The ability to sequence activity is also illustrated in the sample.




```
#11 = IFCSPACE(...  , 'EX1', 'Excavation 1', $, #91, $, $, .ELEMENT., .EXTERNAL., $);
#12 = IFCMEMBER(... , 'SP2', 'Excavation support', $, #92, #93, $);
#13 = IFCTASK(...   , 'SH3', 'Shift 3', 'Work shift', 'SH3', '', 'Shift start', .F., $);
#14 = IFCACTOR(...  , 'SE4', 'Site engineer 4', 'Engineer', #94);

#21 = IFCRELCONTAINEDINSPATIALSTRUCTURE(... , (#12),    #11);
#22 = IFCRELASSIGNSTOPROCESS(...            , (#11), $, #13, $);
#23 = IFCRELASSIGNSTOPROCESS(...            , (#12), $, #13, $);
#24 = IFCRELASSIGNSTOACTOR(...              , (#11), $, #14, $);
#25 = IFCRELASSIGNSTOACTOR(...              , (#12), $, #14, $);
#26 = IFCRELASSIGNSTOACTOR(...              , (#13), $, #14, $);

#26 = IFCRELSEQUENCE(...                    , #13,      #15, 0, .FINISH_START.);
```

Figure 2: IFC ontology and sample

Some information sources may lack structuring so consideration can be given to automated or manual RASE mark-up of diary and monitoring entries (Figure 3) either to read directly, obtain tabular information or generate IFC data. Figures 3, 4, 5 and 6 show Requirements in blue with underlining, Applicability in green with dashed underlining, Selection in purple with dotted underlining and Exceptions in orange with double underlining.

Figure 3: Narrative text with RASE as an information model

UK BSI PAS 1192-6 (2018) offers a model of safety risks. The comparability of risks in a shared risk register is supported by the consideration of likelihood and consequence. Likelihood is considered as the outcome of the effects of physical product (P), process activity (A) and spatial location (L). This 'PAL' model omits the consideration of the actors responsible for or engaging in activities so as to shift the emphasis away from operator error and towards systemic factors. This is intended to make safety risk information more shareable, generalisable and reviewable. Regulatory considerations often additionally address actor roles so that responsibility and accountability can be assigned when compliance is assessed. Uniclass (2023) tables can be used to classify the locations (SL), products (Pr), activities (Ac), roles (Ro) and also the resulting risk (RK).

Ontology for dictionaries

Previous work has shown that dictionaries can be used to relate concepts from different domains, including relating names, descriptions and classifications and properties, or concepts from natural language and specific data schemas (ISO 12006-3, 2022). Nisbet et al. (2024) showed how ordinary text and tables can be marked-up to serve as dictionary resources.

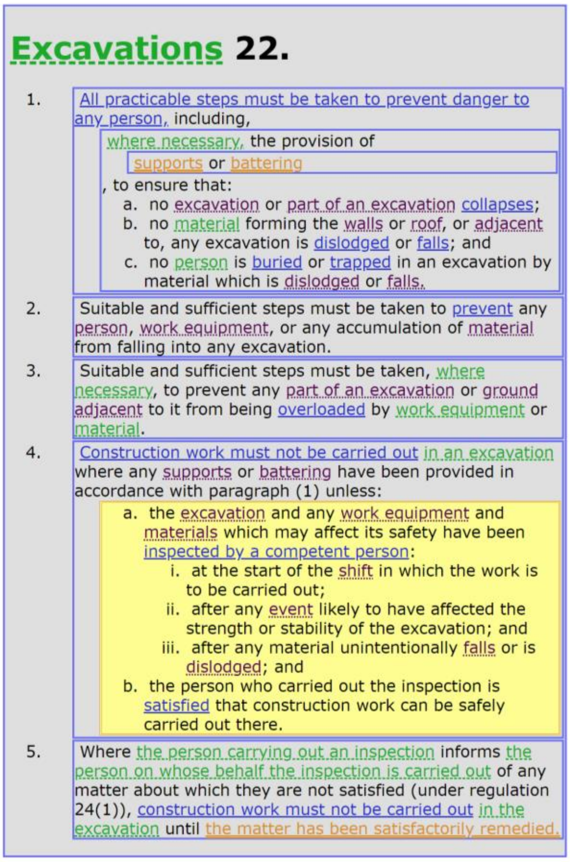
Ontology for normative knowledge

RASE captures the ontology of knowledge embodied in First Order Logic. This can be summarized as an executable tree hierarchy of logical operators applied to objectives and testable metrics. Normative knowledge as a logical statement of how the world is required to be, may be wrapped in deontic logic around duties and obligations which are usually handled externally by administrative or operational action. In the current example the action may include 'enforcement' by the HSE authorities.

Preparations

EU UK legislation and secondary legislation and regulations are publicly accessible. An example is EU Directive 92/57/EEC (EP 1992) which is implemented in the UK as the C(D&M) 2015 regulations that are applicable to most construction activity. These regulations comprise clauses 1-15 concerning the allocation of responsibilities and roles, and clauses 16-39 which define 'general requirements for all construction sites'. Clause 22 considers the steps expected for the safe execution of excavations. Sub-clause 22(1) is problematic in that it refers to 'all practicable steps' and the ill-defined qualification 'if necessary'. This sub-clause will be considered in the Discussion section. The text of clause 22 is available as HTML and has been marked up with RASE (Figure 4) using 'AEC3 Require1' (AEC3, 2024).

Sub-clause 22(4) is a particularly significant clause in that it explicitly bans construction work within excavations unless certain conditions are met. The Requirement is that no construction activity is carried out. This has Application to excavation locations but only if a Selection from supports and battering (sloping sides) is present. A substantial Exception then identifies the events which should trigger an inspection and satisfaction. Whilst most of the events are discrete and independent, the inspection and satisfaction are related and dependant to those events. This relationship is not captured by the visible mark-up. However, each RASE mark-up can contain metadata which can be made visible when a mouse hovers over the mark-up ('mouse-over event'). Previously this metadata has given each mark-up an identifier 'id' and the naming of the RASE class 'Type'. Optionally metrics can have 'Property', 'Comparator', 'Target' and 'Unit' added to deconstruct any numeric metric. Optionally sections can have a 'Outcome' to record where any intermediate outcome is of interest. In HTML each piece of metadata is stored as an attribute on the tags, prefixed with 'data-rase' which combines the convention for extensions to HTML markup with a RASE specific identifier.



Excavations 22.

1. All practicable steps must be taken to prevent danger to any person, including,
where necessary, the provision of
supports or battering
, to ensure that:
a. no excavation or part of an excavation collapses;
b. no material forming the walls or roof, or adjacent to, any excavation is dislodged or falls; and
c. no person is buried or trapped in an excavation by material which is dislodged or falls.

2. Suitable and sufficient steps must be taken to prevent any person, work equipment, or any accumulation of material from falling into any excavation.

3. Suitable and sufficient steps must be taken, where necessary, to prevent any part of an excavation or ground adjacent to it from being overloaded by work equipment or material.

4. Construction work must not be carried out in an excavation where any supports or battering have been provided in accordance with paragraph (1) unless:
a. the excavation and any work equipment and materials which may affect its safety have been inspected by a competent person:
i. at the start of the shift in which the work is to be carried out;
ii. after any event likely to have affected the strength or stability of the excavation; and
iii. after any material unintentionally falls or is dislodged; and
b. the person who carried out the inspection is satisfied that construction work can be safely carried out there.

5. Where the person carrying out an inspection informs the person on whose behalf the inspection is carried out of any matter about which they are not satisfied (under regulation 24(1)), construction work must not be carried out in the excavation until the matter has been satisfactorily remedied.

Figure 4: RASE markup of Clause 22

One way of understanding this complexity of related events is to ask what it is that potentially fails. It is not the individual events, nor the excavation but the sequence of events and inspections – the timeline – that passes or fails. Some metrics depend on the state of the excavation, in terms of the presence or absence of 'work', 'supports' and so on. Based on the work on 'temporal logic' discussed in the literature review section, it is necessary to record that certain events such as 'inspections 'and 'satisfactions' are not only required but must be in a specific relationship to the events and states, specifically coming later in a timeline.

This means that the requirement for inspection is actually for a 'subsequent inspection' and the requirement for satisfaction is actually for 'subsequent satisfaction'. The 'subsequent' test is in each case relative to a previous entry in the timeline. This can be achieved by adding an additional item of metadata 'Reference' which can ensure that the test for an inspection is applied, not to the whole timeline (has there been an inspection ever?) but to the timeline since the triggering event (has there been an inspection since that triggering event?), and similarly can ensure that the test for 'satisfaction' is applied, not to the whole timeline (has there been a satisfaction event ever?) but to the timeline since the inspection event (has there been satisfaction since that inspection?).

With this addition, the marked-up regulation can be reported back, for example as a conceptual graph (Solihin et al., 2015) of distinct metrics as seen in figure 5.

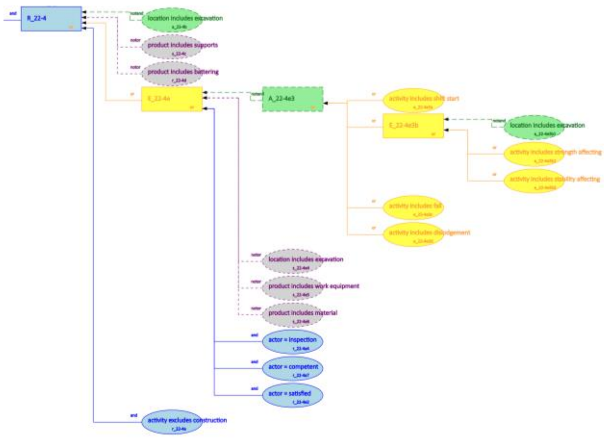


Figure 5: Reflection of Clause 22(4) as a conceptual graph

Reviewing the list of metrics, it is clear that some terms need synonyms. These can be plain language synonyms, or can relate the terms to the table headings (Figure 1) or IFC concepts (Figure 3). Such synonyms can be provided for locations, products, activities and roles. Within a specific project, there may further synonyms, for example 'SE1' may be synonymous with 'Simon Engineer' or 'se1@enginering.com'. Both the general and the project dictionary can be held as RASE documents (figure 6), or in a database service.

eng-GB: Site engineer	**eng-GB: Site engineer**
Uniclass: Ro_50_20_11	**Uniclass: Ro_50_20_11**
P101: SE1	**P101: SE1**
IFC: IfcActorRole.Engineer	**IFC: IfcActorRole.Engineer**

Figure 6: Example dictionary entry

Experiments

Two scenarios were investigated. In each, the site diary and site monitoring data streams were available as separate timelines. These were provided by the participants as anonymised examples based on actual site data. The site diary is maintained by the site supervisor, logging the significant events. Many site-based ('field') and construction management solutions offer applications to support this activity, including AI-supported voice-and video recording. Site monitoring is dependent on the use of electronic tagging of individuals via their construction hard-hats and vehicles via secure attachments. The specific application used also supported the representation of polygonal safety zones defined from 4D BIM or other inputs.

Some safety issues were detectable by considering a data stream alone. In experiment 1, site monitoring can create a timeline where checking Clause 22(4) can detect if work is being conducted in an excavation. There is a vehicle incursion into the excavation zone, which is detected by the site monitoring. This is event is defined as potentially affecting the stability of the excavation, but work continues, so immediately the rule-engine reports that the site is non-compliant because although a visit by the site engineer has been detected, no acceptance can be sensed (Figure 7).

In experiment 2, the number of aspects that can be checked for compliance expands considerably by combining data streams. In this second scenario, a shift starts, there is a fall of earth and so supports are installed. Supplying the rule-engine with a combined data stream from the site diary and the site monitoring highlights that the excavation was compliant up until the supports were installed but before the excavation had been inspected. Once inspected and approved, the excavation reverts to being compliant (Figure 8).

Overall, two experiments using the live collation of information were used to assess the compliance of a construction site. Alerts have been generated when the excavation is in a non-compliant state. Any alert can be supplied to the site supervisor, site engineer and/or the operatives concerned. These alerts are exactly defined by the sense of the original regulatory text, and potential remedies are suggested.

1: Location		2: Product		3: Activity			4: Actor		CDM	
Name	Description	Name	Description	Name	Description	Date and Time	Name	Description	Compliant	Description
EX1	Excavation	-	-	IN1	Introduction	16/11/2023 09:10	CW1	Crew	TRUE	work in progress
EX1	Excavation	-	-	IN4	Introduction	16/11/2023 10:05	LR1	Lorry	FALSE	stability risk not inspected or satisifed
EX1	Excavation	-	-	VW2	Inspection	16/11/2023 11:00	SE1	Engineer	FALSE	stability risk inspected but satisifed

Figure 7: Non-compliance detected but unresolved from site monitoring alone (experiment 1)

1: Location		2: Product		3: Activity			4: Actor		CDM	
Name	Description	Name	Description	Name	Description	Date and Time	Name	Description	Compliant	Description
S1	Site	B1	Facility	45246	Shift	16/11/2023 08:00	-	-	TRUE	
EX1	Excavation	-	-	IN1	Introduction	16/11/2023 08:10	CW1	Crew	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	IN2	Introduction	16/11/2023 08:15	EX1	Excavator	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	VW1	Inspection	16/11/2023 08:25	SE1	Engineer	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	AP1	Approval	16/11/2023 08:35	SE1	Engineer	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	EV1	Landslip	16/11/2023 08:47	CW1	Crew	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	NT1	Notification	16/11/2023 08:50	SE1	Engineer	TRUE	excavation but no supports or battering in use
EX1	Excavation	-	-	VW2	Inspection	16/11/2023 09:00	SE1	Engineer	TRUE	excavation but no supports or battering in use
EX1	Excavation	BA1	Battering	IN3	Introduction	16/11/2023 09:05	CR1	Crew	FALSE	material fall but not inspected or satisfied
EX1	Excavation	-	-	AP2	Approval	16/11/2023 09:20	SE1	Engineer	TRUE	excavation inspected and satisifed
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Figure 8: Non-compliance detected and resolved from combined sources (experiment 2)

Temporary non-compliance (in the examples a matter of minutes) may be felt to be too strict. There is no guidance or common knowledge as to what period of time is reasonable before compliance is to be re-established. Reasonableness will depend on what action is attached to the lapse. If an alert sounds on a mobile device, then 5 minutes might be appropriate. If a siren or alarm sounds, then a longer period may be more appropriate.

Discussion

This paper has addressed the question of how rules for the continuous assessment of safety compliance within a site context can be obtained and made operable explicitly and without specific coding. The experiments have shown that some clauses of the C(D&M) 2015 regulations can be assessed in 'real-time', but others may be problematic as they lack 'decidability' due to the temporality of the evaluated processes.

Decidability

Zhang et al. (2022) identified four kinds of ambiguity in normative text which pose obstacles for decidability in automated compliance checking. These are (1) intentional, (2) grammatical, (3) tacit knowledge and (4) domain-specific. A dictionary can hold the collective tacit knowledge, domain-specific mappings and agreed analytic tools to address these obstacles (Beach et al., 2023). In addition, some C(D&M) clauses are ambiguous if taken as metrics defined by examination of the information sources, but if read as objectives, defined by satisfying subsequent clauses, these can be evaluated. We can distinguish testable metrics from objectives, even if their appearance in source regulations appear similar. Previous assumptions about the exclusion of clauses by consideration of the internal characteristics have been replaced with criteria relating to this external context.

Table 1: Obstacles to decidability

Obstacle	Description (and resolution)	*C(D&M) 2015 examples*
Grammar	Two or more possible parsings (resolved by mark-up).	*'work equipment or material'* *(see 22(3))*
Ambiguity	Two or more assessment methods (resolved in dictionary).	*'adjacent'* *(see 22(1))*
Subjectivity	No common assessment method.	*'practicable'* *(see 22(1))*
Objective	Dependant on following sections and metrics (resolved by mark-up or omission).	*'danger'*
Temporality	No information or authority available.	*(see 22(1))*

Thus, decidability is dependent on managing grammar, ambiguity, subjectivity and additionally 'objectives' and 'temporality' (Table 1). Grammatical ambiguity can be resolved by domain experts when performing the RASE markup. In the example it could be that either all material or only work material is meant. Other ambiguities need to be resolved by developing the intermediate dictionary so that there is a common and singular matching of regulatory and descriptive/narrative models, for example 'adjacent' is untestable unless the dictionary provides a plausible distance. Subjectivity, for example 'practicable', arises where perceptions can differ completely between applicant and inspector. This identification of sections as objectives can be resolved by mark-up or its omission. On the other hand, 'temporality' depends on the distinction between leading and live indicators of the state of the construction site as opposed to trailing outcomes such as 'buried or trapped' in clause 22(1c) (Table 2).

Table 2: Safety and compliance indicators

Category	Description	*Safety examples*	*C(D&M) 2015 examples*
Leading indicators	Predictive metrics	*Safety management underway*	*Clauses 1-15*
Live indicators	Direct metrics	*Inspections and satisfactions*	*Clauses 16-35*
Trailing indicators	Outcome metrics	*Safety records and evidence*	*Clause 22(1c)*

Leading and live indicators may be available from documentation or from contemporaneous recordings by sensors and participants. These sources can be combined in a unified information model accessible to a rule engine.

Trailing outcomes are frequently referred to in the first sub-clause in each C(D&M) clauses 16-35. The construction of such sentences gives them a similar appearance to 'actual' requirements, the difference being that the value of the individual metrics is knowable only in retrospect. These can be handled either by deliberately not marking them up or by developing a 'Future World Assumption' (FWA) that requirements and exceptions may be taken as positive but applicability and selections may be taken as negative, at least until they become knowable.

Within in the DRSM (design science research method) framework, the evaluation of the user acceptance of the outcomes is postponed until the artefact has been developed. Such evaluation would need to take into account a number of different organisation contexts into which a live compliance checking system could fit, such as private to the constructor, private to the inspectorate or accessible to both the constructor and inspectorate.

Conclusions

Automated compliance checking has previously been focused on static descriptions of a facility. This work has shown that the RASE methodology can be extended to capture the expectations for dynamic contexts such as construction sites, especially the relative timing of separate requirements. Using a suitable ontology to track real-time events can structure the increasing availability of real-time information on construction sites so as to support the continuous assessment of compliance, such as safety compliance. Definitions of compliance may mix leading, live and trailing metrics which can affect the 'decidability' of those rules. The decidability of individual metrics within regulations depends on addressing ambiguity, subjectivity and especially the 'temporality' of information availability. Whilst the other barriers can be resolved (as listed in Table 1), the use of metrics having trailing 'temporality' can only be resolved by treating these as 'objectives'. This represents a contribution to the real-team process compliance checking in construction.

A limitation of this work has been the limited choice of clauses as examples, and the limited number of information sources used. Further research could be less selective about the clauses considered so as to confirm the completeness of the approach used here. Other leading and real-time information sources, such progress charts, could be integrated.

Further work is being proposed to broaden the number of use-cases supported by the mark-up of construction regulations. This could include decision support tools and educational and training experiences.

Acknowledgments

This work has been conducted with the support of AEC3 UK Ltd who provided the 'AEC3 Require1' application, ASITE Ltd who provided sample anonymised site diary information and Wearable Link Ltd (trading as PLINX) who provided sample anonymised site monitoring logs, whilst collaborating with the UK HSE.

References

AEC3 Require1 regulatory workbench, http://www.aec3.eu/require1/AEC3_Require1.html. (accessed 2023-12-03)

Bao, Q., Zhou, J., Zhao, Y., Li, X., Tao, S. and Duan, P., 2022. *Developing a rule-based dynamic safety checking method for enhancing construction safety.* Sustainability, 14(21), p.14130.

Beach, T.H., Hippolyte, J.-L., Rezgui, Y. *Towards the adoption of automated regulatory compliance checking in the built environment* (2020) Automation in Construction, 118, art. no. 103285. doi: https://doi.org/10.1016/j.autcon.2020.103285

Beach, T., Yeung, J., Nisbet, N. and Rezgui, Y., 2024. *Digital approaches to construction compliance checking: Validating the suitability of an ecosystem approach to compliance checking.* Advanced Engineering Informatics, 59, p.102288. https://doi.org/10.1016/j.aei.2023.102288

BSI PAS 1192 Part 6 2018 https://knowledge.bsigroup.com/products/specification-for-collaborative-sharing-and-use-of-structured-health-and-safety-information-using-bim. BSI 2018

C(D&M) 2015: *The Construction (Design and Management) Regulations 2015* https://www.legislation.gov.uk/uksi/2015/51/contents/made (accessed 2023-12-03) .

Cheng, T. and Teizer, J., 2013. *Real-time resource location data collection and visualization technology for construction safety and activity monitoring applications.* Automation in construction, 34, pp.3-15.

EP 1992: EUROPEAN PARLIAMENT and COUNCIL OF THE EUROPEAN UNION. *Directive 92/57/EEC* 24 June 1992. https://www.legislation.gov.uk/eudr/1992/57/contents (accessed 2023-12-03) .

ISO 12006-2:2015 *Building construction. Organization of information about construction works. Part 2: Framework for classification* https://www.iso.org/standard/61753.html . (accessed 2023-12-03)

ISO 12006-3:2022 *Building construction. Organization of information about construction works. Part 3: Framework for object-oriented information*. https://www.iso.org/standard/74932.html . (accessed 2023-12-03)

ISO 16739-1:*2024 Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema* https://www.iso.org/standard/84123.html . (accessed 2024-03-31)

Johansen, K. W. (2024). *Automated performance assessment of prevention through design and planning (PtD/P) strategies in construction.* Automation in Construction, 157. https://doi.org/10.1016/j.autcon.2023.105159

Lamport, L., 1994. *The temporal logic of actions.* ACM Transactions on Programming Languages and Systems (TOPLAS), 16(3), pp.872-923.

Li, B., Nielsen, R.O., Johansen, K.W., Teizer, J., Larsen, P.G. and Schultz, C., 2021. *Towards digital twins for knowledge-driven construction progress and predictive safety analysis on a construction site*. In Leveraging Applications of Formal Methods, Verification and Validation: Tools and Trends: 9th International Symposium on Leveraging Applications of Formal Methods, ISoLA 2020, Rhodes, Greece, October 20–30, 2020, Proceedings, Part IV 9 (pp. 153-174). Springer International Publishing.

Li, X., Yang, D., Yuan, J., Donkers, A. and Liu, X. 2022. *BIM-enabled semantic web for automated safety checks in subway construction.* Automation in Construction, https://doi.org/10.1016/j.autcon.2022.104454.

Pradhananga, N. and Teizer, J., 2013. *Automatic spatio-temporal analysis of construction site equipment operations using GPS data.* Automation in construction, 29, pp.107-122.

Nisbet N, Wix J and Conover D. 2008. *The future of virtual construction and regulation checking*, in Brandon, P., Kocaturk, T. (Eds),Virtual Futures for Design, Construction and Procurement, Blackwell, Oxfordshire. doi: 10.1002/9781444302349.ch17

Nisbet N., Ma L., Aksenova G. *Presentations of RASE knowledge mark-up.* (2022) Proceedings of the European Conference on Computing in Construction, pp. 184 - 190, DOI: 10.35490/EC3.2022.162

Nisbet, N., Zhang, Z. and Ma, L., 2023, July. *Automated generation of SPARQL queries from semantic mark-up.* In EC3 Conference 2023 (Vol. 4, pp. 0-0). European Council on Computing in Construction.

Nisbet, N., Zhang, Z., Ma, L., Chen, W. and Çıdık, M.S., 2024. *Semantic correction, enrichment and* *enhancement of social and transport infrastructure BIM models.* Advanced Engineering Informatics, 59, p.102290. https://doi.org/10.1016/j.aei.2023.102290

Pirzadeh, P. and Lingard, H., 2017. *Understanding the dynamics of construction decision making and the impact on work health and safety.* Journal of Management in Engineering, 33(5), p.05017003.

Ploug, T., and P. Øhrstrøm, 2012, *Branching Time, Indeterminism, and Tense Logic: Unveiling the Prior-Kripke Letters,* Synthese, 188(3): 367–379.

Schwabe, K., Teizer, J. and König, M., 2019. *Applying rule-based model-checking to construction site layout planning tasks.* Automation in construction, 97, pp.205-219.

Shen, Q., Wu, S., Deng, Y., Deng, H. and Cheng, J.C.P., 2022. *BIM-Based Dynamic Construction Safety Rule Checking Using Ontology and Natural Language Processing.* Buildings 2022, 12, 564.

Solihan et al., 2015 *A Knowledge Representation Approach to Capturing BIM Based Rule Checking Requirements Using Conceptual Graph*, CIB W78 2015 proceedings.

Teizer, J., Johansen, K.W. and Schultz, C., 2022. *The concept of digital twin for construction safety.* In Construction Research Congress 2022 (pp. 1156-1165).

Tekbas, G. and Guven, G., 2020. *BIM-based automated safety review for fall prevention.* In Advances in Building Information Modeling: First Eurasian BIM Forum, EBF 2019, Istanbul, Turkey, May 31, 2019, Revised Selected Papers 1 (pp. 80-90). Springer International Publishing.

Uniclass 2023. https://uniclass.thenbs.com/#browse (accessed 2023-12-03)

Xu, R. (2023). *Predictive worker safety assessment through on-site correspondence using multi-layer fuzzy logic in outdoor construction environments.* Heliyon, 9(9). https://doi.org/10.1016/j.heliyon.2023.e19408

Yang K., Yang H., Zhang J., Kang R. *Safety and Efficiency Evaluation Model for Converging Operation of Aircraft and Vehicles* (2023) Aerospace, 10 (4), art. no. 343, DOI: 10.3390/aerospace10040343Ref

Yuan, J., Li, X., Xiahou, X., Tymvios, N., Zhou, Z. and Li, Q., 2019. *Accident prevention through design (PtD): Integration of building information modeling*

SEMANTIC MAPPING ANALYSIS OF DIGITAL BULDING LOGBOOK/ PASSPORT MODELS

Sara Karami[1,4,5], José-Paulo de Almeida[2,5], Claire Ellul[3], Alberto Cardoso[1,4,5] [1]Department of Informatics Engineering, University of Coimbra, Coimbra, Portugal [2]Department of Mathematics, University of Coimbra, Coimbra Portugal
[3]Department of Civil, Environmental & Geomatic Engineering, University of College London, London, United Kingdom
[4]Centre for Informatics and Systems of the University of Coimbra (CISUC)
[5]Institute for Systems Engineering and Computers at Coimbra (INESCC)
sarakarami@dei.uc.pt, zepaulo@mat.uc.pt, c.ellul@ucl.ac.uk, alberto@dei.uc.pt

Abstract

A Digital Building LogBook/Passport (DBL/DBP) is a repository for static ”as built” and dynamic ”Key Performance Indicators (KPIs)” throughout the life cycle of the buildings. This ongoing research uses the semantic mapping approach to analyse and compare the current recommended DBL/DBP models to identify the similarities/differences between them. The findings of this investigation show that due to the lack of comparison studies on DBL/DBP models in academic literature, the similarity percentage between proposed categories and elements is very low, and a knowledge-based method would be needed that makes the comparison not straightforward. In particular, the lack of metadata caused an increasing uncertainty percentage of elements. The focus of this paper is on assessing the DBL/DBP models (comparison of categories and elements, metadata evaluation), further work is needed to develop an integrated model.

Introduction

Buildings account for 40% of total energy consumption in the EU and are a major contributor to greenhouse gases (GHG) and energy consumers globally (Koltsios et al., 2022). Current European regulations aim to achieve the Paris Agreement as a nearly zero operating energy target in buildings (Ahmed et al., 2022). Net-Zero energy buildings are defined as a building or construction that has a zero-net consumption of energy or zero carbon emissions over a set period (Sartori et al., 2012) and are classified into four types: zero-energy sites, zero-emissions buildings, zero-energy sources, and zero-cost energy buildings (Ahmed et al., 2022). Space heating, hot water production, lighting, and the operation of various electric appliances, all use operational energy and account for the majority of total life cycle energy use (Vourdoubas, 2017).

Collecting Net-Zero emission building data in a single repository facilitates monitoring of greenhouse gas emissions across the entire life cycle of constructed assets within global, regional, and national frameworks (Society, 2020). Implementing a unified repository enables the thorough tracking and analysis of emissions data, hence facilitating the identification of patterns, trends, and areas that require improvement. Decision-makers can then use this information to guide the creation and application of strategies and policies to help the built environment reach Net-Zero emissions targets.

Building LogBooks and Passports

Töpfer (1997) introduced the first definition of a Building LogBook (BL) as a tool to improve the transparency of technical properties, standards of building services, quality of use, and operation costs of buildings by communicating comprehensive information to clients and buyers of the new property, developers, and real estate agents. The term ”Building Passport (BP)”, has now replaced the Building LogBook. A BP is a simple format for presenting both the ”birth certificate” and ”health certificate” of a building. The birth certificate specifies the key performance elements of the building to operate sustainably during the design process or after the first year of operation (Virta et al., 2012). The ”health certificate” compares the building's operation to previous years' operations and provides a short and long-term repair and retrofitting plan (Virta et al., 2012). The Global Alliance for Buildings and Construction (GlobalABC) defined the BP as a repository for all construction types and real estate activities throughout the life cycle. It will be able to create and update itself, feed on data from various data sources (e.g., Internet of Things (IoT), and Digital Twins (DT)), and enable effective virtual representations of physical assets UNEP (2020).

Semantic mapping analysis is the process of calculating the relationships between texts based on the semantic similarities methods. These methods combine automatic and knowledge-based measurements, to determine similarities and differences between phrases and sentences based on their meaning (Chandrasekaran and Mago, 2021).

This paper provides an overview of current BP/BL models to identify their main differences and similarities with semantic methods (see Experiment section), which are the fundamental basis to assess the potential for creating a single integrated model. Understanding the variation between models helps in fully comprehending the primary (minimum, lowest common denominator) BL/BP data elements, by examining element commonality and similarity. It also enables the development of a fully comprehensive model that covers all elements of all BPs. Both models help to organize data in a way that reflects the relationships among different pieces of information that can be understood by both technical and non-technical users, which is crucial for effective decision-making and system development.

This paper only explores approaches currently available from the literature, and hence the effort required to integrate these data sources. The resulting integrated model is aimed to be used by various use cases, who will be able to develop a sub-set (view) on the model that contains only the data that they require, and hence identify how to source the required data in the different passport contexts and countries, greatly reducing the time for data discovery and integration.

Experiment

Current DBL/DBP initiatives

Numerous research initiatives received funding through the European Horizon 2020 program to encourage increased energy renovations across the European Union. EU Commission study (EU) proposed a DBL data model to enhance energy efficiency and sustainability (Dourlens-Quaranta et al., 2020). As part of the iBRoad project, Libório et al. (2018) proposed a BL data model to support the implementation of a building's renovation road map for single-family houses based on occupant needs and preferences. As part of the ALDREN project, Sesana et al. (2020) proposed a tailored conceptual building Renovation Passport model to understand the state of the non-residential buildings and inform building owners about the technical energy performance status. Within the X-tendo project, a BL model that can be attached to the EPCs to inspire the next EPC generation was proposed (Toth et al., 2021). The EUB SuperHub project proposed a DBL model to achieve the EU's sustainable buildings (Malinovec Puček et al., 2023). The DBL report used the existing European INSPIRE initiative and proposed the DBL semantic data model (van der Ende et al., 2023). A green Building Renovation Passport (UKGFI) was proposed to enable data for energy performance estimates and retrofit assessments in (Small-Warner and Sinclair, 2022). THE BIM4EBB[1] developed an interoperable BIM-based toolkit for efficient building renovations. The BIM4Ren[2] proposed workflow for the renovation process and information requirements.
The GlobalABC proposed the main contents of BP based on the integration of Woningpas[3], the Building Renovation Logbook initiative, Ukraine[4], and the Building data Collection initiative (UNEP, 2020). Miller (2016) developed a BP to capture building energy efficiency data on residential buildings. Within the chronicle[5] project, a DBL framework is developed to achieve sustainability targets and long-term maintenance and renovation plans. National initiatives such as Madaster[6], Passeport Efficacité Énergétique[7], PAS-E[8], BASTA Logbook[9], Property Logbook[10],and Building Renovation Logbook[11], do not provide an official data model. The initiatives that provided the data model are listed in Table 1.

Previous Attempts at Integration

Although the approach of integrating BL/BP models was not addressed before, various Energy Performance Certificate (EPC) integration initiatives exist ((EPC and BP are the sources of a building's operation consumption and energy performance)). Serna-GonzÁLez et al. (2021) proposed a harmonization EPC data model by combining models from Italy and Spain. The Hale Studio tool was used to map the target data model to the source data model(Serna-GonzÁLez et al., 2021). Popa et al. (2022) created an integrated EPC model from England, France, Scotland, and Ireland to determine a dwelling's energy performance rating. Pouliot et al. (2018) used the Open II (Open Information Integration) tool to compare the geospatial standards from syntactic, structural, and semantic points of view. The Hale Studio [12] and OpenII [13] are powerful data transformation and harmonization tools, which don't include functionality for comparing and identifying similarities between texts.

Semantic similarity

Semantic similarity is a technique for assessing the degree of semantic equivalence between two texts, providing a quantified measure of similarity instead of a binary determination of whether they are similar or dissimilar (Chandrasekaran and Mago, 2021). Existing methods for short text similarity calculation can be roughly divided into two categories: word-level semantic-based and semantic modelling-based (Yang et al., 2021).
word-level semantic-based: The similarity of two short texts is calculated by aggregating the similar words in both texts, including knowledge-based and corpus-based methods.

- **Knowledge-based methods:** This method uses the shortest path method (Yang et al., 2021), Lexical (Farias et al., 2016), and syntactical similarity (Costin and Eastman, 2019), with the knowledge-based approach to computing the semantic similarity between two short texts (Chandrasekaran and Mago, 2021). This approach allows for a representation of human common sense and demonstrates improved performance compared to other methods (Yang et al., 2021). The main drawback of shortest path and lexical, respectively, is that the texts must be synonymous, and the lexical technique only considers zero and 100% similarity. The syntactically similar measure quantifies the similarity of two text strings for approximate string matching or comparison. For example, the

[1]https://www.bim4eeb-project.eu/
[2]https://bim4ren.eu/
[3]https://woningpas.vlaanderen.be/
[4]https://eeplatform.org.ua/
[5]https://www.chronicle-project.eu
[6]https://madaster.com
[7]https://www.experience-p2e.org/
[8]http://pas-e.es//en
[9]https://www.bastaonline.se
[10]https://www.propertylogbook.co.uk/
[11]https://eeplatform.org.ua/
[12]https://wetransform.to/halestudio/
[13]https://sourceforge.net/projects/openii/

strings ”Sam” and ”Samuel” are similar and have the same character sequence (Pradhan et al., 2015).

- **Corpus-based methods:** This method extracts the context information of words from large corpora and then induces the distributional properties of words (Yang et al., 2021). It does not capture the semantic relationships between words. It treats each term in isolation, without considering synonyms, related terms, or context.

Semantic modelling-based: Semantic modelling-based methods model a text as a whole to obtain the semantic vector and calculate the similarity of two short texts by performing algebraic operations on the obtained vectors (Yang et al., 2021).

- **Explicit Semantic Representation:** ESR models focus solely on word-level representations, overlooking the valuable information conveyed by syntax (Yang et al., 2021).

Methodology

To evaluate the potential of combining existing BP towards creating a comprehensive passport model, their similarities and differences in elements (attributes) classification, and metadata availability need to be investigated. The overall approach is organized as follows:

Step 1: Identify LogBooks/Passports. This included a comprehensive literature review, using the following search terms in the Google Scholar engine in all fields of the publications, such as titles, abstracts, keywords, and the entire text of the papers: ”LogBook”, ”Building LogBook”, ”Building Passport”, ”Material Passport”, ”District LogBook”, ”Building Renovation Passport”, ”Green Building”, ”Building Certificate”. Publications from 2000 to 2023 in the form of journal and, conference papers were searched, with language limited to English to allow for semantic matching.

Step 2: High level comparison. The papers that provided information about the BP structure and classification were selected and reviewed to identify the detailed data model presence that could be used for comparison, and the presence of metadata describing the model. Following this stage, BPs with sufficient detail for semantic comparison were selected. To enable the comparison process, a single BP was selected, to be the model to which all other models will be compared.

Step 3: Detailed data model comparison for the selected BP

- **Format, type, and level of classification investigation:** This step involved understanding how different elements in the models are classified (grouped), the types of classifications used, and the hierarchical classification level.

- **Manually Transfer Models into Excel:** The models were initially in PDF format, and for the initial comparison, each model was manually entered into a distinct Excel sheet. This involved taking each element in the model and manually typing it out into the corresponding sheet. A sheet was created for each classification identified above, with a column created for the elements from each BP.

- **Exclude the dynamic data:** To evaluate the infrastructural elements of BP and due to the different natures of static and dynamic data, the dynamic elements were excluded.

- **Import into a Database:** SQL (Structured Query Language) is commonly used for querying relational entities (Date, 2011). Each sheet model in CSV format was imported into a database; i.e. one entity (table) was created for each classification for each BP. This enables semantic comparisons to be performed using Transact-SQL queries.

Step 4: Semantic comparison. The similarities and differences between the classifications and elements of each model by the other models were examined. The knowledge-based and syntactical methods were used to conduct a more detailed semantic comparison of the category and element at the static information level. The LIKE query was executed to model the syntactical method, which allows finding all textual elements similar to the source element. This method is acknowledged as a semi-automatic method in this paper. The following is the comparison process:

- **Identifying the elements with similar terminology:** A semi-automatic method was used to find textually similar element names.

- **Identifying the elements with similar meaning but different terminology:** A knowledge-based approach was used to find the elements that did not share the same terms but were similar in meaning. Where available, metadata was used where this was not available a comprehensive exploration of the documentation was undertaken to confirm any interconnection between the data element names.

- **Identifying the elements with similar terminology potentially different meaning:** As above, after using the semi-automatic method, a knowledge-based approach was used to review and clarify the elements, using metadata and BP documentation where possible.

- **Identifying the elements with uncertain meaning:** A final category of elements are those where the meaning is uncertain and where the documentation and metadata do not provide sufficient clarity to match the elements.

Step 6: Key information comparison. The main BP content proposed by the GlobalABC UNEP (2020) was compared to categories and elements of identified models to assess their similarity.

Step 7: Similarity methods comparison. The average similarity percentage was examined to determine whether semi-automatic or knowledge-based approaches performed better.

Step 8: Statistical analysis. Aggregating all of the outcomes above, the similarity percentage between categories and elements was calculated to assess how much information could be collected from each model. Elements with a high level of certainty were chosen for analysis, while those with unclear meanings were excluded. It is important to note that elements lacking consistent terminology; i.e. where meaning was uncertain, have been removed from the analysis.

Results

Analyzing BL/BP models

A total of 7 BL/BP data models were identified and summarised in Table 1. They have a different number of categories that provide different information. The first column of the Table 4, Table 5, Table 6, Table 7, and Table 8 provides the category description.

Table 1: Summary of Identified BP

Source	Summary	Metadata Available	Number of Categories	Total Number of elements
EU	Proposing a DBL model to achieve energy efficiency and sustainability	Partial	7	76
X-tendo	Proposing a DBL model to inspire the next generation of EPCs information	No	7	173
EUB	Proposing a data structure for DBL to achieve energy efficiency and sustainability	Partial	7	167
ALDREN	Proposing a Building Renovation Passport (BRP) model to understand the technical energy performance status of non-residential buildings	No	5	66
iBRoad	Proposing a framework for managing energy performance, executed maintenance activities, and building plans	Partial	4	73
UKGFI	Proposing a green BRP model to improve energy efficiency	No	8	75
DBL report	Proposing a semantic DBL data model	Partial	7	86

Performing semantic comparison

A preliminary comparison was performed to select a more complete model regarding the building's fundamental elements. EU and X-tendo models provided more information on the buildings, and EU was selected as a starting point. In the semi-automatic part, a script with the LIKE operator was executed per category, and the identified elements with the same input terms were displayed in the results. Similar results were then imported into the repository, which was used for the next comparisons. This process was repeated for each model until it reached the final one, allowing all models to be compared. It is important to note that some steps were iterative since previously analyzed information had to be reanalyzed several times along the line, as the need to understand the problem better increased in subsequent rounds. The following are some illustrative examples of the comparison process.

- **elements with similar terminology:** The ”Floor area” element in the EU model's *Building descriptions and characteristics* category, was identified as a match by the automated checking process to the ”Total floor area (unheated and heated area)” in the EUB model's *General Building Information* category, ”Floor area” in the X-tendo model's *Building descriptions and characteristics* category, ”Net floor area” in the iBRoad model's *General and administrative information * category and ”Floor area and building volume” in the UKGFI model's *Building type* category.

- **elements with dissimilar terminology but similar meaning:** The ”Building information model/Design and plans of the building” element in the X-tendo, EU model's *Building descriptions and characteristics* category, could be assumed to be similar to the ”Layout of the whole building for multi-unit properties (i.e. block of flats, terrace housing) ” in the UKGFI's *Building type* category, and ”3D model/Architectural plans” in the EUB model's *Building Documentation BIM* category. One of the main limitations of an automated approach is identifying similarities, where there are no identical terms between models.

- **elements with similar terminology but uncertain meaning:** The ”Address” has been considered in the EU and X-tendo models's*Administrative* category, which is undefined and is related to the building owner's address, occupier address or the building's address. The meaning of the ”Storage” element in the X-tendo model's *Building descriptions and characteristics* category, and the ”Energy supply and storage” in the UKGFI's *Building Services* category are unclear.

- **elements with uncertainty meaning:** The meaning of ”Renewable energy systems” in the EU model's *Building descriptions and characteristics* category, ”Any existing and planned local energy schemes” in the UKGFI model's *Building Service* category, ”Optimizing self-consumption of locally generated energy” in the ALDREN model's *Technical system* category, and the ”isDescribedByNativeGIS” in the DBL report's *Location* category, are unclear. In the EU and X-tendo models, the ”U value of different components” is proposed as an element in the ”Building Performance” category, although its meaning and definition are unclear.

- **elements with uncertainty metadata:** The description of ”Ventilation systems” or ”Cooling systems” in iBRoad model's *Building construction information* category is "Description of the building ventilation systems" and "Description of the building cooling systems", which do not encompass the essential requirements that should be incorporated into it.

duplicated elements: The "Safety Manual" proposed in the EU model includes building operations and maintenance, and security procedures, whereas the " Maintenance records and information" element, also has been proposed by EU, X-tendo, UKGFI, and iBRoad models. The "Building pictures" in the EUB report's *Building Documentation BIM* category was repeated in the EUB report's *General Building Information* category.

Our results demonstrated that significant progress in defining a comprehensive BL/BP model has been made, though there is still a lack of interoperability between elements, data consistency, metadata availability, and information exchange.

Comparing some key information

The percentage of similar elements proposed by more than 3 models is 6 %, and the similarity percentage between elements is low (See section Statistical analysis), making it challenging to identify the key elements. Table 2 compares the main content proposed by GlobalABC and identified BP models (categories and elements). As can be seen, in some cases, the contents are similar to the group of model elements (e.g., Building description, Technical features, Use and operation). In these cases, similar elements were found using the knowledge-based method in addition to the semi-automatic method. The semi-automatic method was used to find similar elements to the *Contracts*, *Material inventory*, *Building description*, *Operation*, *Energy Performance Certificate*, and *Maintenance*. The knowledge-based was performed to find similar elements to the *Identification of buildings*, *Identification of the plot*, *Material inventory* and *Technical features*. The models did not include the *Environmental performance and carbon footprint*, *Surfaces*, *Dismantling and recycling strategy*, *Results of user satisfaction survey* and *Certificate* elements. The EUB model only incorporated the *Proof of maintenance* (as the maintenance report element), *Design documents* (as the Architectural plans and As-built plans), and *Operational cost*.

Comparing similarity methods

Table 3 compares the semi-automatic and knowledge-based methods. The similarity was calculated by the average number of similar certain and uncertain elements (elements were unclear or did not have similar words but were similar in their meaning) by the total number of its elements. The semi-automatic method performed better for the X-tendo and EUB models, while the others were processed using the knowledge-based approach.

Table 2: Comparison of some key information

Content	EU	X-tendo	EUB	ALDREN	iBRoad	UKGFI	DBL report
Identification of plot and plot-related characteristic	**Element:** Building owner	**Element:** Buiding owner	**Element:** Building owner	**Element:** Owner name		**Element:** Owner's details	**Element:** Owner
Identification of building	**Element:** Unique building identifier	**Element:** Unique building identifier	**Element:** Unique building identifier			**Element:** Unique building identifier	
Contracts	**Element:** Utility,Service contract,Insurance	**Element:** Utility,Service contract,Insurance	**Element:** Utility,Service contract,Insurance				
Energy Performance Certificate / sustainability label	**Element:** EPC rating	**Element:** EPC rating	**Category:** EPC certificate		**Element:** EPC number,Issue date		
Material inventory	**Category:** Building material inventory	**Category:** Building material inventory	**Element:** Wall,Floor,Roof			**Element:** Materials Used	**Category:** Structure & Material
Building description	**Category:** Building descriptions and characteristics	**Category:** Building descriptions and characteristics					
Technical features and characteristics	**Category:** Building operation and use	**Category:** Building operation and use	**Category:** Building Element Information	**Category** :Building operation and use	**Category:** Building operation and use		
Use and operation data / consumption	**Category:** Building performance	**Category:** Building performance	**Category:** Building performance		**Category:** Building energy performance	**Category** :Energy consumption & use behaviour	**Category** :Prformance
Maintenance manuals	**Element:** Maintenance log	**Element:** Maintenance log	**Element:** Maintenance log		**Element:** Maintenance information	**Element:** Maintenance records	

Table 3: Comparison of similarity methods

	X-tendo [%]		EUB [%]		ALDREN [%]		iBRoad [%]		UKGFI [%]		DBL report [%]	
Source	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based
EU	37	4	20	6	0	4	1	4	0	10	6	3
X-tendo	-	-	16	4	1	2	0	4	1	5	2	1
EUB	-	-	-	-	2	4	1	7	0	6	2	3
ALDREN	-	-	-	-	-	-	1	5	1	6	3	11
iBRoad	-	-	-	-	-	-	-	-	0	6	2	4
UKGFI	-	-	-	-	-	-	-	-	-	-	0	2

Performing statistical analysis

The similarity percentage of classifications for each model to the other models was calculated by the average number of similar categories divided by the total number of categories in the model. The EU model with 57%, X-tendo with 57%, and EUB with 51 % were the most identical to the other models. ALDREN with 40%, iBRoad with 35%, DBL report with 25%, and UKGFI with 10 % were in the subsequent ranks. As can be seen in Table 4, Table 5, Table 6, Table 7, and Table 8, the "Building Documentation BIM" category is not included in the other models, or the "Envelope" category of the ALDREN model is relevant to the "Building construction" category of iBRoad. The location element has been considered in the other models but not in a separate category; ALDREN's model is the only one considered a category for location.

Table 4 demonstrates the similarity percentage of EU elements to the X-tendo, EUB, ALDREN, iBRoad, UKGFI, and DBL report models using the semi-automatic and knowledge-based methods. It was calculated by the number of similar elements with clear meaning in the EU model by the total number of its elements. Table 5, Table 6, Table 7, and Table 8 shows the similarity comparison for the other models. As can be seen, the overall matching percentage in terms of category and elements, with X-tendo and then EUB, is significantly higher than with the other models. The low percentage of similarities arises from diverse elements incorporated within the same categories across different models. For example, "Building age" is part of the iBRoad "General and administrative information" category, while the "Year Built" is included in the EU "General Building Information" category.

Table 4: EU similarity comparison with X-tendo, EUB, ALDREN, iBRoad, UKGFI, DBL report

EU	X-tendo [%]		EUB [%]		ALDREN [%]		iBRoad [%]		UKGFI [%]		DBL report [%]	
	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based
Administrative	33	0	17	17	0	0	0	0	0	17	0	0
Building descriptions and characteristics	14	0	5	0	0	0	0	10	0	5	4	0
General Building Information	29	0	14	14	0	14	0	14	0	14	0	14
Building operation and use	0	0	0	0	0	0	0	0	0	0	0	0
Building performance	33	0	0	0	0	33	16	0	0	17	0	0
Finance	46	0	7	0	0	0	0	0	0	0	0	0
Building material inventory	0	0	0	0	0	0	0	0	0	0	0	0

Table 5: X-tendo similarity comparison with EUB, ALDREN, iBRoad, UKGFI, DBL report

X-tendo	EUB [%]		ALDREN [%]		iBRoad [%]		UKGFI [%]		DBL report [%]	
	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based
Administrative	38	0	0	0	0	0	0	8	0	0
Building descriptions and characteristics	3	0	3	0	1	0	1	0	1	0
General Building Information	11	0	0	0	0	0	0	0	0	0
Building operation and use	0	0	0	0	0	0	0	0	0	0
Building performance	20	0	0	0	0	0	20	0	0	0
Finance	23	0	0	0	0	0	0	0	0	0
Building material inventory	0	0	0	0	0	0	0	0	0	0

Discussion

The BL/BP is a summary of all the key information about the building, including the original design, commissioning, and handover details, and information on its management and energy performance, the material used, sustainability performance, indoor environmental quality, and potential energy throughout the building life cycle (Small-Warner and Sinclair, 2022).

Table 6: EUB similarity comparison with ALDREN, iBRoad, UKGFI, and DBL report

EUB	ALDREN [%]		iBRoad [%]		UKGFI [%]		DBL report [%]	
	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based
Administrative	19	0	0	0	0	0	10	4
General Building Information	8	4	8	8	8	8	4	0
Building Element Information	0	7	0	13	0	0	0	0
Building operation and use	0	0	0	0	0	0	0	0
Building performance	0	0	10	20	0	0	0	0
Finance	0	0	0	0	0	0	0	0
Building Documentation BIM	0	0	0	0	0	0	0	0

Table 7: iBRoad similarity comparison with UKGFI, DBL report & UKGFI to DBL report

iBRoad	UKGFI [%]		DBL report [%]		UKGFI	DBL report [%]	
	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based		Semi-automatic	Knowledge-based
General and Administrative	7	18	7	14	Ownership & governance	0	0
					Building Service	0	0
Building construction	0	0	0	0	Energy consumption & use behaviour	0	0
					Building Type	0	0
Building energy performance	0	0	0	0	Building Fabric	0	0
					Climate resiliency	0	0
Building operation and use	0	0	0	0	Circular economy considerations & enhanced climate information	0	0
					Building information	0	0

Table 8: ALDREN similarity comparison with iBRoad, UKGFI, DBL report

ALDREN	iBRoad [%]		UKGFI [%]		DBL report	
	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based	Semi-automatic	Knowledge-based
Building features	0	20	0	20	10	30
Location data	0	0	0	0	12	37
Building geometry	16	0	16	0	8	0
Envelope	0	0	0	0	0	12
Technical system	0	0	0	0	0	0

The main purpose of our approach was to evaluate the differences and similarities between the data models that underpin each BP and assess whether semi-automated or knowledge-based approaches would be useful to create an integrated data model of the various BP. This work showed that the currently proposed BL/BP models, firstly have different categories (groupings or classifications of elements) which, when examined in detail, overlap in some cases and are not aligned with the proposed categories by GlobalABC (UNEP, 2020). This adds a layer of complexity to the potential for automated integration, which is further complicated by the fact that elements (attributes) within each classification sometimes overlapped or were included in different groupings in the different BP. Underpinning the automation challenges are gaps in the lack of an agreed and standardized conceptual BL model accompanied by detailed metadata, which would provide clear element definitions, including the data type, and sufficient data descriptions. This could be further complicated when it is considered that not all these BP originals in English-speaking countries - and local terminology - and translations- could cause additional issues. Methodologically, the use of text-based comparison (LIKE) for matching could also introduce false positives (See section Step 4 for examples) and manual checking was required to validate the results. The approach of selecting a single model (EU model) towards full integration of models available from the literature, was however very beneficial as it was avoided.

The comparison results showed that although the most common categories are Administrative information, Building Identification, Building performance, and Building operation (Table 2), the similarity percentages between elements are low (Table 4, Table 5, Table 6). There is a lack of clear consensus regarding the Building geometry category, suggesting the need for further discussion. The similarity percentage between models on Building equipment information (e.g., Heating, Cooling, DHW, etc) is low, which is substantial in proposing the integrated model. The models made no distinction between Building and Building unit elements, which are required for creating an integrated mode. The EUB model incorporated the EPBD indicators (Energy Performance Building Direction) [14], however further discussion is required to distinguish between EPBD and integrated BP elements. The categories in the integrated BP model must be harmonized according to the proposed GlobalABC report's categories. Although the static information was examined, the EUB model only included the KPI indicators. This indicates that identifying real-time data and considering the Building Automation System (BAS) in the integrated model requires significant effort.

Overall, therefore, it can be concluded that to develop an integrated model an extensively manual process will be required, starting from creating digital versions of PDF models and moving from there. Indeed, the automated SQL comparisons did not prove reliable enough to even be a first pass for the approach, given that in 32 cases similar terminology was used with different meanings, so false positives were observed (Section Methodology- Step 4). An extensive effort to identify key information in an integrated model is required due to the 42 elements shared by more than 3 models.

[14] https://eur-lex.europa.eu/resource.html?uri=cellar:c51fe6d1-5da2-11ec-9c6c-01aa75ed71a1.0001.02/DOC_1

Conclusions and further work

The BL/BP is a repository for the comprehensive collection of the most important performance characteristics and technical data feed in from various data sources (e.g., Internet of Things(IoT)) and should enable effective virtual representations of physical assets (Tharma et al., 2018). The overall purpose of this specific task of our ongoing research was to examine both the parallels and distinctions of the current recommended BL/BP models in terms of their element's contribution to the physical, energy performance characteristics, and technical data of a building using semantic similarity methods. The results showed that, while great progress has been made, the similarities between proposed models are quite low, and a knowledge-based strategy would be required. The outcome of the process above will be an integrated UML and physical data model that will allow users to trace back each element to its source or sources (BP or EPC) and make a connection between different entities. Users of this integrated model will thus be able to select the required subset of elements from the full model (as a view) and easily understand whether such data is available from data currently collected by BP or EPC. Future work will also entail the integration of the BP identified above with EPC models. These provide the necessary energy consumption information and may help achieve Net-Zero carbon objectives.

Acknowledgments

This research is being partly supported by the Portuguese Foundation for Science and Technology (FCT) I.P./MCTES through national funds (PIDDAC), within the scope of INESC-Coimbra R&D Unit (Ref. UIDB/00308/2020 Ref. UI0308/geoBIM.1/2022) and of CISUC R&D Unit (Ref. UIDB/00326/2020).

References

Ahmed, A., Ge, T., Peng, J., Yan, W.-C., Tee, B. T., and You, S. (2022). Assessment of the renewable energy generation towards net-zero energy buildings: a review. Energy and Buildings, 256:111755.

Chandrasekaran, D. and Mago, V. (2021). Evolution of semantic similarity—a survey. ACM Computing Surveys (CSUR), 54(2):1–37.

Costin, A. and Eastman, C. (2019). Need for interoperability to enable seamless information exchanges in smart and sustainable urban systems. Journal of Computing in Civil Engineering, 33(3):04019008.

Date, C. J. (2011). SQL and Relational Theory: How to Write Accurate SQL Code. O'Reilly Media, Inc.

Dourlens-Quaranta, S., Carbonari, G., De Groote, M., Borragán, G., De Regel, S., Toth, Z., Volt, J., and Glicker, J. (2020). Building logbook state of play. report 1 of the study on the development of a european union framework for buildings' digital logbook. Report 1.

Farias, T. M. d., Roxin, A., and Nicolle, C. (2016). A brief overview of semantic interoperability for enterprise information systems. In International Conference on Informatics in Economy, pages 38–52. Springer.

Koltsios, S., Fokaides, P., Georgali, P.-Z., Tsolakis, A. C., Chatzipanagiotidou, P., Klumbytė, E., Jurelionis, A., Šeduikytė, L., Kontopoulos, C., Malavazos, C., et al. (2022). An enhanced framework for next-generation operational buildings energy performance certificates. International Journal of Energy Research, 46(14):20079–20095.

Libório, P., Fragoso, R., Monteiro, C. S., Fernandes, J., Silva, E., and Casteele, T. V. (2018). The logbook data quest: Setting up indicators and other requirements for a renovation passport.

Malinovec Puček, M., Khoja, A., Bazzan, E., and Gyuris, P. (2023). A data structure for digital building logbooks: Achieving energy efficiency, sustainability, and smartness in buildings across the eu. Buildings, 13(4):1082.

Miller, W. (2016). Capturing sustainable housing characteristics through electronic building files: the australian experience. In Sustainable Built Environment Conference 2016 Proceedings, pages 190–199. Karlsruher Institut fur Technologie (KIT).

Popa, A., Ramallo González, A. P., Jaglan, G., and Fensel, A. (2022). A semantically data-driven classification framework for energy consumption in buildings. Energies, 15(9):3155.

Pouliot, J., Larrivee, S., Ellul, C., and Boudhaim, A. (2018). Exploring schema matching to compare geospatial standards: application to underground utility networks. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences-ISPRS Archives, 42(4/W10):157–164.

Pradhan, N., Gyanchandani, M., and Wadhvani, R. (2015). A review on text similarity technique used in ir and its application. International Journal of Computer Applications, 120(9):29–34.

Sartori, I., Napolitano, A., and Voss, K. (2012). Net zero energy buildings: A consistent definition framework. Energy and buildings, 48:220–232.

Serna-GonzÁLez, V., HernÁNdez, M. G., Miguel-Herrero, F. J., Valmaseda, C., Martirano, G., Pignatelli, F., and Vinci, F. (2021). Harmonisation of datasets of energy performance certificates of buildings across europe.

Sesana, M. M., Rivallain, M., and Salvalai, G. (2020). Overview of the available knowledge for the data model definition of a building renovation passport for non-residential buildings: The aldren project experience. Sustainability, 12(2):642.

Small-Warner, K. and Sinclair, C. (2022). Green building passports: a review for scotland.

Society, R. (2020). Digital technology and the planet: Harnessing computing to achieve net zero uk.

Tharma, R., Winter, R., Eigner, M., et al. (2018). An approach for the implementation of the digital twin in the automotive wiring harness field. In DS 92: Proceedings of the DESIGN 2018 15th International Design Conference, pages 3023–3032.

Töpfer, K. (1997). Gebäudepaß–eine serviceleistung für bauherren und nutzer. Bundesbaublatt, 9(97):608.

Toth, Z., Maia, I., Rosa, N., and Volt, J. (2021). Technical specifications of energy performance certificates data handling: Understanding the value of data.

UNEP (2020). The building passport: A tool for capturing whole life data in construction and real estate – practical guidelines.

van der Ende, M., Flickenschild, M., Borst, T., Raes, N., Cai, A., Gankema, Y., and Schinkel, R. (2023). Dbl semantic data model, providing standard form and meaning to digital building logbook data.

Virta, M., Hovorka, F., and Lippo, A. (2012). Building passport as a tool to evaluate sustainability of building. World Green Building Council Publication: London, UK.

Vourdoubas, J. (2017). Creation of zero co2 emissions residential buildings due to operating and embodied energy use on the island of crete, greece. Open Journal of Energy Efficiency, 6(4):141–154.

Yang, J., Li, Y., Gao, C., and Zhang, Y. (2021). Measuring the short text similarity based on semantic and syntactic information. Future Generation Computer Systems, 114:169–180.

GRAPH-BASED DIGITAL DECISION SUPPORT SYSTEMS: INTRODUCING BTWIN, A TOOLKIT FOR PERFORMANCE-BASED BUILDING MANAGEMENT

Angelo Massafra[1], Ugo Maria Coraglia[1], Giorgia Predari[1], Riccardo Gulli[1]
[1]Alma Mater Studiorum University of Bologna, Bologna, Italy

Abstract

The digitization of the AECO industry is increasing the need for digital decision-support systems (DDSS) in building management. Despite technological progress, developing such systems still demands significant resources. In particular, integrating diverse building data remains a challenge. Addressing this, the paper introduces BTwin, a toolkit designed to simplify the prototype development of DDSSs supporting decision-making in performance-based building management. This software library facilitates linking data from multiple sources into graph networks, like building information models and time series databases, providing a low-code and open-source framework for creating user-friendly web applications using connector logic.

Introduction

The built environment strongly impacts global energy consumption and produces high environmental emissions. Although it sits at the crossroads of many EU policies, it is still affected by a persistent challenge of limited innovation, which is probably the primary impediment to its digital and ecological transformation.

To face these challenges, in recent years, the Architecture, Engineering, Construction and Operation (AECO) sector has started its digital transition, pushed by the emergence of cutting-edge technologies, including Artificial Intelligence (AI) (Zhang et al., 2022), Building Information Modeling (BIM) (Deng et al., 2021), Internet of Things (IoT) (Mannino et al., 2021), Building Performance Simulation (BPS) (de Wilde, 2023), and Digital Twins (DT) (Lu, 2022). Notably, DTs have risen prominently, serving as a catalyst for developing digital decision support systems (DDSS) to enhance decision-making in built asset management, leading to the spread of performance-oriented applications for multiple uses, such as space usage monitoring, analysis of energy needs, reporting of energy and water consumption, estimation of emissions due to building usage, as well as analysis of indoor environmental quality regarding thermal, visual, acoustic comfort, and air ventilation (Petri et al., 2023).

Despite their rapid growth, DTs are still in their infancy and require, on the one hand, prototyping and technical development and, on the other, further conceptualization (Boje et al., 2020). The lack of standardized definitions, the heterogeneous and vast amount of data and data sources, and the not scientific approach that usually characterizes built asset management (Abuimara et al., 2021) hinder the organic development of such technology and, more in general, of DDSSs, limiting the possibility of easily transforming building data in usable knowledge.

Developing such systems poses a resource-intensive challenge, with economic, financial, and cultural benefits that are sometimes difficult to comprehend during the conceptualization phase (Ma et al., 2022). These challenges are highlighted by the AECO sector's lack of data integration, marked by the presence of numerous data silos belonging to various stakeholders and created using distinct languages and frameworks, leading to significant data interoperability issues (Merino et al., 2022). These digital solutions require high expertise from many domains to enable secure and streamlined data flow, reliable and scientifically valid information generation, and user-friendly, accessible and usable knowledge to non-digital experts. The scientific community, therefore, is called first to theorize these systems and second to prototype solutions demonstrating the tangible benefits of such technologies.

In order to respond to these needs, this paper introduces BTwin, a software toolkit designed for simple and cost-effective prototyping of DDSSs for building management. The toolkit allows for semantically linking building data from diverse sources (like BIM, BPS, meters, and sensors) within graph networks and visualizing it organically through interactive dashboards, offering a low-code and open-source solution for prototyping web-based DDSSs. Implemented on JSON-LD (JavaScript Object Notation for Linked Data) principles, BTwin structures data within graphs by following the axioms and meanings defined in a federated ontology. After linking data, the toolkit enables the creation of web applications that extract data from graphs using user-executable queries through a graphical user interface (GUI). It then presents this data in various formats, informing the application user about the key performance indicators (KPIs) that characterize building performance. The software library performs these tasks in Python, a programming language known for its simple syntax, making it extremely popular for various software projects.

In this article, after discussing the state of the art in developing graph-based DDSSs in the AECO sector, the theoretical framework that supported the development of the tool is presented. Furthermore, the toolkit is applied to a simplified case study to demonstrate its applicability and provide some examples of its use.

Related work and background

Knowledge graphs

When people express their thoughts to other people in the places where they work, they, consciously or unconsciously, create graphs. It means identifying specific objects and labelling the relationships between them to explain a problem, share a concept, or explain a domain, in other words, to transfer knowledge.

Such graphs can be produced either by humans or computers. Those intelligible to computers typically comprise nodes (or vertices), denoting entities and subjects, alongside edges representing the connections between these nodes (relationships or links). Both nodes and edges can be characterized by semantic attributions and property descriptions (Paulheim, 2016).

Accordingly, organizing domain concepts and information in knowledge graphs (KGs) allows for, first, assigning semantic and ontological meaning to data, as well as enabling semantic query, extraction and analysis of data from complex data structures; second, proficient visualization and, therefore, comprehension of a knowledge domain and its segmentation into knowledge sub-domains.

Beyond these capabilities, KGs also excel in analyzing extensive datasets entrenched in diverse formats. Moreover, if coupled with semantic web technologies (Pauwels et al., 2017), KGs can allow fast accessibility to information on the web, an issue that has become a priority for modern DDSSs.

Ontologies and schemas for data representation

Understanding the terminology used in a KG is crucial for systematically organizing and sharing the knowledge it contains in a universally comprehensible manner. In KGs, organic information description can be achieved through ontologies, formal, shared, and explicit representations of a domain's conceptualization, crafted to ensure that all information is clearly defined for logical reasoning.

Recently, the surge in digital tools within the construction sector has led to the definition of various ontologies to describe the AECO world and its subdomains (Pritoni et al., 2021). Among the most used ontologies and data schemas, the following are considered relevant to this research: Industry Foundation Classes (IFC); Building Topology Ontology (BOT); Brick Schema; Semantic Sensor Network Ontology (SSN); the Sensor, Observation, Sample, and Actuator (SOSA) ontology, and the Energy Management Key Performance Indicator Ontology (EKO).

IFC, the open BIM standard supported by buildingSMART, is the most recognized schema for BIM data. Its integration with semantic web technologies is allowed by the IFC web ontology language (ifcOWL) (buildingSMART International, n.d.). The Building Topology Ontology (BOT) is a simplified ontology proposed by W3C that exclusively addresses the core building concepts revolving around the building's topology, including physical and conceptual components and their relationships (Rasmussen et al., 2017). Brick is an open-source schema that standardizes the semantic descriptions of buildings' physical, logical, and system assets and interrelationships. Its primary focus is on smart building applications and equipment representation (Balaji et al., 2018). SSN describes sensors and their observations, features of interest and samples, procedures, and actuators. It is often used to describe BAS data with semantic tags (Compton et al., 2012). SOSA, instead, redesigns SSN to provide a lightweight, general-purpose specification for modeling the interactions between entities involved in observations, actuation, and sampling (Janowicz et al., 2018). Finally, EKO enhances multi-level energy management and energy performance information exchange (Li et al., 2019).

Graph-based approaches for digital twins

In the contemporary landscape, graphs have assumed a pivotal role across many data-driven applications, also influencing the AECO sector (Lygerakis et al., 2022).

Various researchers have introduced and tested the use of KGs to develop digital systems for building operations, contributing to forming a solid approach to the subject, although relying on various technologies and data schemas. For instance, Merino et al. experimented with using KGs to deliver the DT of an educational building aimed at fault detection and diagnosis of HVAC systems for facility management purposes (Merino et al., 2022). In their study, these authors mainly leveraged IFC and Brick schemas for data representation and a data lake architecture for data storage. Chamari and colleagues, instead, explored the integration of BMS, IoT, and BIM via KGs, utilizing IFC and Brick within RDF graphs (Chamari et al., 2022). They proposed a hybrid approach for managing time series data, which was stored in a JSON-based MongoDB to overcome the limitations of graphs in handling time series with millions of data points. Differently, Hosamo et al. conducted interesting work on DT for fault detection and comfort optimization, integrating BIM, sensor time series, and maintenance records. Although the research refers to BOT, SSN, and Brick schemas, they used a custom ETL (Extract, Transform, Load) process to enable condition monitoring and predictive operation (Hosamo et al., 2023). Another significant example is the work of Ramonell and his team, who proposed a system for data integration within DTs, leveraging property graphs in Neo4j (Ramonell et al., 2023). This approach allowed for mapping semantic information according to different ontologies, keeping the graph as the interoperability backbone of the systems.

All these contributions, which propose hybrid strategies to enhance BIM-based KGs with time-series data, underscore the significance and utility of such technologies in modeling building knowledge, highlighting both potentials and limitations. However, an aspect not covered in the mentioned articles is the Return on Investment (ROI) for developing these applications. Indeed, it remains a challenge to ascertain whether the benefits of developing such complex systems can be financially justified for the administrations that adopt them, especially considering that substantial funding programs are usually behind these research initiatives. Consequently, leveraging the theoretical frameworks established in these studies and the insights derived from these pioneering explorations, this paper seeks to introduce a toolkit designed to streamline the initial prototyping process of such digital environments. The goal is to make the process more cost-effective and accessible to research and development groups of smaller scale or with limited resources, thus promoting wider adoption and innovation in this field.

Materials and methods

Ontology definition

BTwin assumes that a building can be represented as a graph composed of nodes (subjects) and edges (predicates). To define organically information, nodes are designated to classes (and subclasses), whereas edges are associated with types of relationships, following the ontology depicted in Figure 1. This federated ontology, designed by leveraging BOT, IfcOWL, Brick, SSN, SOSA, and EKO, integrates these already established ontologies within a unified and extensible ontological model.

At the foundational level, the federated ontology employs the BOT's zones (i.e., 'bot:Building', 'bot:Site', 'bot:Storey', and 'bot:Space'). These objects establish a hierarchy necessary for representing the building's spatial configuration. BOT serves as the backbone, allowing for the detailed specification of physical locations and the inclusion of elements (such as sensors) within spaces. BOT is also used to model 'bot:Interface' elements resembling interface elements bounding the spaces (e.g. walls, floors, roofs, windows and doors).

Interlaced with BOT is IfcOWL, which is used in the ontology to map the properties of spatial elements as 'ifc:Property' within 'ifc:PropertySet', instrumental in providing detailed descriptions of element properties and their attributes. Moreover, IfcOWL allows for the representation of sensor elements, considered as physical electronic devices, using the 'ifc:Sensor' class.

Enhancing the ontology is Brick. The Brick framework enables classifying sensor elements as 'brick:Sensor'. While there seems to be an overlap with 'ifc:Sensor', Brick sensors are denoted as 'brick:Point', which are input points symbolizing the value captured by a device or instrument engineered to detect and measure various variables. This indicates that a single IfcSensor device can embody multiple points (thanks to the 'brick: hasPoint' relationship). For instance, a 'brick:HumiditySensor' and a 'brick:TemperatureSensor' points may continuously provide data to an 'ifc:Sensor' device. The link between sensors and spaces is given by the relationship' brick:hasLocation'. Additionally, Brick facilitates the organization of spaces into zones utilizing the 'brick:Zone' class, which can encapsulate groupings of rooms such as energy zones or fire compartments. Furthermore, the relationship 'brick:hasLocation' establishes semantic interoperability with the BOT spatial hierarchy, thereby enriching the data model with spatial context and associations.

The ontology further incorporates the SSN and SOSA ontologies through the 'sosa:Observation' and 'sosa: ObservableProperty' classes. These are pivotal for modeling the data generated by sensors (thanks to the

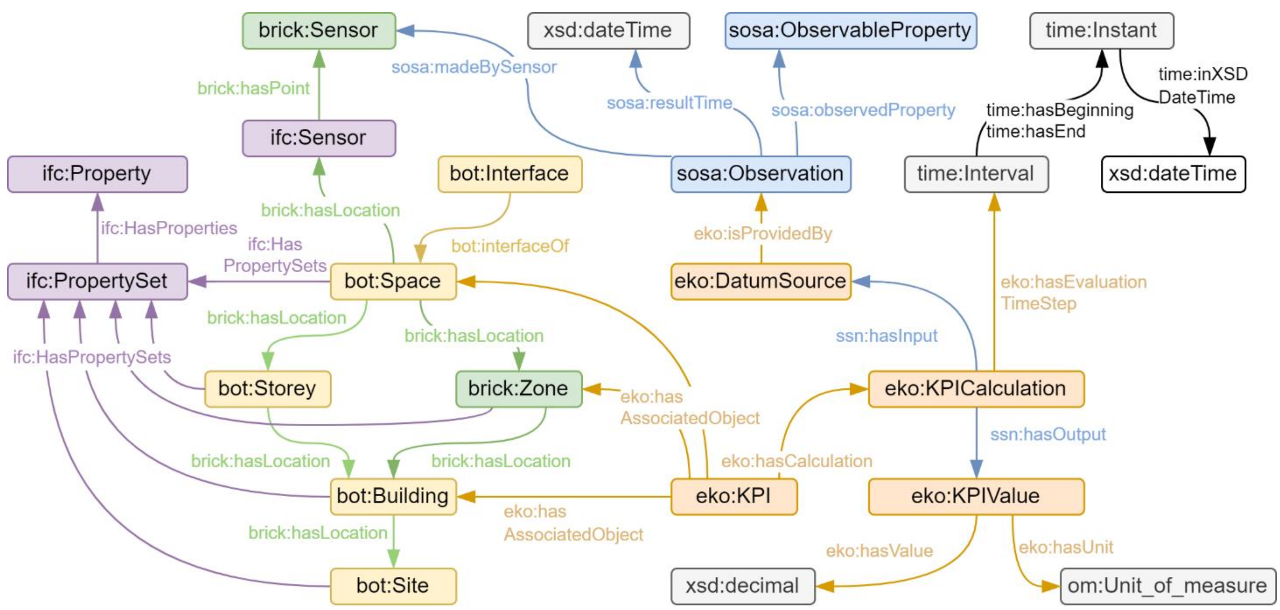


Figure 1: The federated ontology supported by BTwin.

'sosa: madeBySensor' relationship) and the observational processes in a semantically rich context.

EKO, which includes classes such as 'eko:KPI', 'eko:KPICalculation', and 'eko:KPIValue', is instead used to describe the performance aspects of spatial elements. These classes are used to define, calculate, and store the values of performance metrics, which are crucial for evaluating building performance. The 'eko: hasAssociatedObject' relationship provides the necessary linkages between performance-related and spatial concepts. Additionally, this part of the ontology integrates time-related classes, such as 'time:Instant' and 'time:Interval', to represent temporal aspects, which are essential for capturing the dynamics of sensor performance data and observations over time.

Data format

The federated ontology is enhanced with a distinct syntax employing JSON-LD to illustrate entities and their interconnections in graphs. JSON-LD is a streamlined format for Linked Data, leveraging the prevalent JSON structure to enable JSON data to function seamlessly globally. It significantly improves the ease with which data can be read and written by humans, offering a clear advantage over other formats like the EXPRESS format used in IFC. Due to its extensibility and adaptability, this data model allows the serialization of data and the creation of graph-based data models. Additionally, it is lightweight, rendering it suitable for data exchange within web environments and the development of web APIs and applications.

Specifically, BTwin adopts the JSON-LD format to represent static elements, whereas it sources dynamic data, like sensor observations, from databases tailored for time-series data (e.g., SQL or MongoDB), which are more adept at managing large data volumes than graph databases. Within the JSON-LD data format, each building element corresponds to a JSON object. This JSON object functions as a dictionary, possessing a unique identifier, the indication of its class within the ontology, and the relationships with the other objects within the KG. In BTwin, this data format can be read and transformed into a KG using specific functions that leverage the NetworkX library, a Python package that provides a comprehensive set of tools for creating, manipulating, and analyzing graph networks.

Figure 2 presents an example of a JSON-LD syntactic description in BTwin. The 'context' section enables the mapping of terms used in the document to Internationalized Resource Identifiers (IRIs) to provide precise meanings. In this part, reference ontologies (such as Brick and BOT) are mentioned, as well as the keys used in various JSON dictionaries to describe the objects. On the other hand, the 'graph' section describes the nodes that form the graph and maps the existing relationships between them ('relationships' keys).

The example in Figure 2 shows the topological structure of a sample building through BOT. This structure comprises <storeyX>, defined as a 'bot:Storey' and representing a level within the building, and by <space1>, which is a 'bot:Space' located on <storeyX>. The Brick Schema is employed to standardize descriptions of sensor data, with <sensor1> being an 'ifc:Sensor' situated in <space1>, and having a temperature 'brick:Point' connected to <temperaturePoint1>. Moreover, IFC is used to model the properties of spatial entities, i.e., <PSet_SpaceThermalRequirements1> is depicted as a set of thermal requirements for <space1>, including a property about the temperature setpoint for winter. The interconnections between the elements, given by the 'relationships' key, are essential for decoding the spatial hierarchy and placing sensors within the building.

```
{
"@context": {
    "bot":"https://w3id.org/bot#",
    "brick":"https://brickschema.org/schema/Brick#",
    "ifc":"https://standards.buildingsmart.org/.../OWL#",
    "name":"http://schema.org/name"
    },
"@graph": [
      {
        "@id": "storeyX",
        "@type": "bot:Storey"
      },
      {
        "@id": "space1",
        "@type": "bot:Space",
        "relationships": {
            "brick:hasLocation": [
                  {
                    "@id": "storeyX", "@type": "bot:Storey"
                  }]}
      },
      {
        "@id": "sensor1",
        "@type": "ifc:Sensor",
        "relationships": {
            "brick:hasLocation": [
                  {
                    "@id": "space1","@type": "bot:Space"
                  }]}
      },
      {
        "@id": "temperaturePoint1",
        "@type": "brick:Sensor",
        "relationships": {
            "brick:isPointOf": [
                  {
                    "@id": "sensor1","@type": "brick:Sensor"
                  }]}},{
        "@id": "PSet_SpaceThermalRequirements1",
        "@type": "ifc:IfcPropertySet",
        "name": "PSet_SpaceThermalRequirements",
        "relationships": {
            "ifc:IfcPropertySetDefinition": [
                  {
                    "@id": "space1",
                    "@type": "bot:Space"
                  }]},
        "ifc:hasProperties": [
              {
                "@type": "ifc:IfcPropertySingleValue",
                "name": "SpaceTemperatureWinterMin",
                "nominalValue": {
         "type": "ifc:IfcThermodynamicTemperatureMeasure",
                    "value": 20
                    }
        }
      ]
    }
  ]
}
```

Figure 2. Example JSON-LD representation.

Toolkit architecture

BTwin consists of a Python library dependent on other reference libraries widely supported by the programming community. Leveraging the principles of linked data and using a hybrid approach for integrating KGs with time-series data allows for fast prototyping of DDSS dashboards. A

s stated, this software library employs a data format based on JSON-LD. This format can be parsed and converted into network graphs by relying on specific functions integrated with the NetworkX library.

Using the so-called 'connector' logic, the tool organically integrates with external libraries and BIM formats, promoting scalability and modularity in the data modeling steps. Alternatively, for simpler cases, static graph data can be modeled directly within BTwin without the use of third-party software.

Table 1 below shows the alignment of the federated ontology adopted by BTwin with some of the classes of the BIM formats supported by the toolkit: IFC, Autodesk Revit, Topologic, and EnergyPlus' Input Data Format (IDF). The representation of such formats in KGs is obtained through the use of connectors that rely on IfcOpenShell, Autodesk Revit APIs, Topologicpy, and Eppy, respectively.

For example, the connector with Topologic allows the representation of Topologic elements directly within graphs (e.g., Topologic cells are transformed into nodes labelled as 'bot:Space'). Similarly, the connector with EnergyPlus, developed in BTwin by integrating the APIs of the Eppy library, enables the linking of simulation results to the spatial elements whose behavior is simulated (provided that the zones in the IDF have the same name as the zones in the KG), as well as the calculation of energy-related KPIs. Additional connectors allow for interfacing with Autodesk Revit, using the APIs provided by Autodesk, and IFC, using IfcOpenShell. These connectors are useful for extracting the properties of spatial elements within BIM models and associating them with the graph nodes. Furthermore, compatibility with Arduino for modeling sensor data has been tested, and preliminary experiments have been conducted with ChatGPT APIs to enable the extraction of information from graphs using textual prompts and interpreting it through OpenAI's Large Language Models (LLMs).

For better computational efficiency, dynamic data, such as sensor readings, are not stored in graphs but in more streamlined formats within dedicated databases, such as SQL. These databases employ a straightforward schema that includes only the sensor's unique identifier, the timestamp, and the recorded value. The connection to the time-series data is established via a link to the brick:Sensor's UID. An illustrative example of tabular time series data is depicted in Table 2. This approach enables time-series databases to handle and store many records efficiently. Meanwhile, the graph database maintains relationships among static elements separately, allowing for the processing of time-series data solely when computing KPIs.

Once the static and dynamic data are interconnected using this hybrid approach, the toolkit enables various functions for querying, aggregating, processing, and visualizing the data stored within the KG. In particular, the visualization module relies on Plotly and Dash libraries for the creation of web dashboards. The results are web micro-applications that can be easily published on the Internet using solutions such as Docker and AWS, following, for instance, methods similar to the one demonstrated by (Betti et al., 2022).

Table 1: BIM formats supported by BTwin and class alignment.

BTwin	IFC	Autodesk Revit	Topologic	Energy Plus
bot: Building	IfcBuilding	Building	Cell Complex	Building
bot: Storey	IfcBuilding Storey	Level	-	-
bot:Space	IfcSpace	Space	Cell	Space
brick: Zone	IfcZone	Zone	Cluster	Zone
bot: Interface	IfcWall, IfcRoof, IfcSlab, IfcDoor, IfcWindow	Wall, Roof, Floor, Door, Window	Face, Aperture	Surface

Table 2: Example of a time-series data

Timestamp	Value	brick:Point UID	Unit
2023-11-15 10:00	18.3	tempPointX	°C
2023-11-15 10:15	18.6	tempPointX	°C
2023-11-15 10:30	18.8	tempPointX	°C

Demonstration

This section applies the toolkit to develop a sample application for monitoring indoor environmental quality parameters, namely temperature, humidity, and illuminance. For illustrative purposes, a fictional case study is employed to demonstrate a potential workflow. This hypothetical building comprises four spaces: three square offices, each with dimensions of 5x5 meters and a height of 3 meters, and an adjacent corridor measuring 15x2 meters with the same height. This setup creates a data-sharing environment that, through a user-friendly interface, enables the evaluation of the number of hours during which the setpoints for temperature, humidity, and

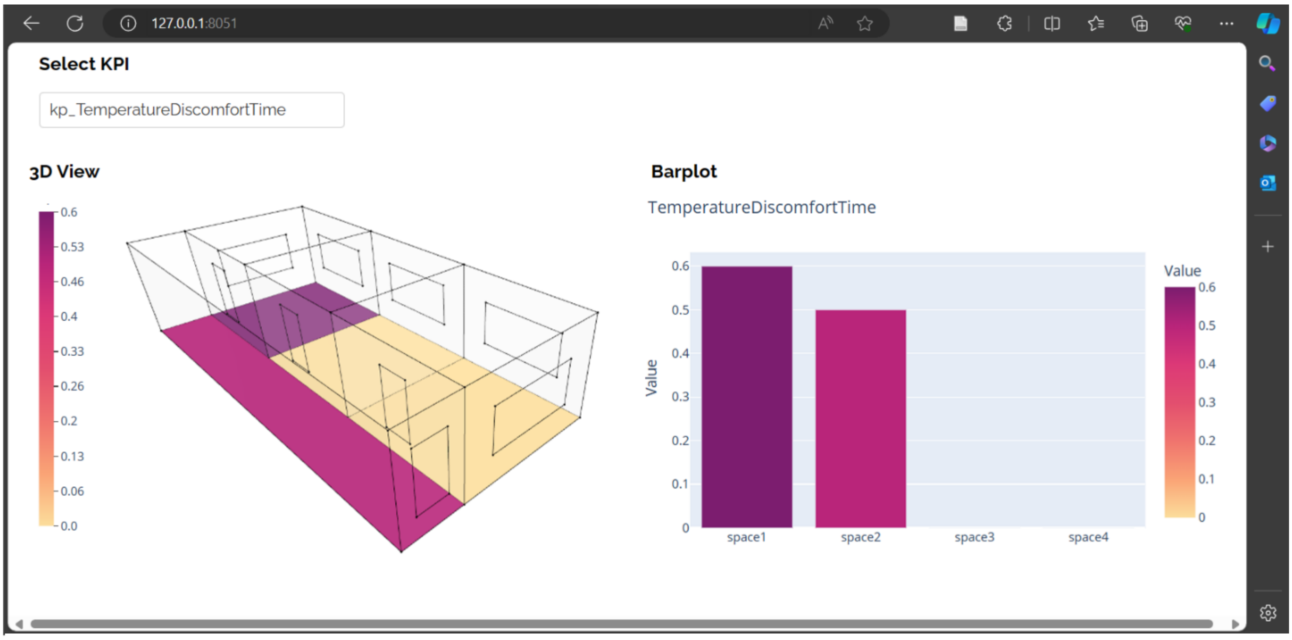


Figure 3: A dashboard developed for the sample application.

lighting are not met in these four spaces. The assessment is based on fictional measurements recorded on a winter day, as shown in Figure 3.

The architecture of the prototype DDSS is shown in Figure 4, while the main steps followed for the application's implementation are listed below.

1. First, the spatial hierarchy of the building was modeled. In particular, Topologic was used to create the case study's BIM model and the connector between Topologic and BTwin was used to transform it into JSON-LD format. The Topologic model also served for 3D visualization within Plotly.
2. Subsequently, sensor elements were added to the JSON-LD as 'ifc:Sensor' and 'brick:Sensor' elements. Moreover, relationships between these and the spatial elements were established.
3. Once all the spatial and sensor elements were identified, the properties of the spatial elements were added to the JSON-LD document. For each space, attributes related to target values of environmental comfort, defined in the KG as properties of the spaces grouped within property sets, were added. The values of the added properties were set, for example, at a minimum value of 20°C for temperature, a range between 30% and 60% for humidity, and a minimum value of 250 lux for lighting.
4. Thanks to NetworkX's connector, a KG containing the spatial elements, sensors, and property sets was created starting from the JSON-LD data. It is depicted in Figure 5.
5. Then, the data related to the measurements taken by the sensors was formatted in a MySQL database according to the notation in Table 2. For each space, temperature, humidity, and illuminance observations were recorded once per hour over a 10-hour working day in winter, resulting in 120 records.
6. Information was then extracted from the KG and MySQL to calculate the internal environmental comfort KPIs, space by space. These KPIs indicate the number of hours in which the target comfort values were not met. After being calculated, such KPIs were connected to the node representing the evaluated space in the KG for each space.
7. Thanks to Plotly's connector, a dashboard was created through Dash to display the KPIs on the geometries of the building's 3D model, previously created in Topologic, within graphs and tables (Figure 3).

From a methodological point of view, the following process can be repeated for different information uses, such as analyzing the energy needs of spaces, detecting indoor air quality, monitoring space usage, and so on.

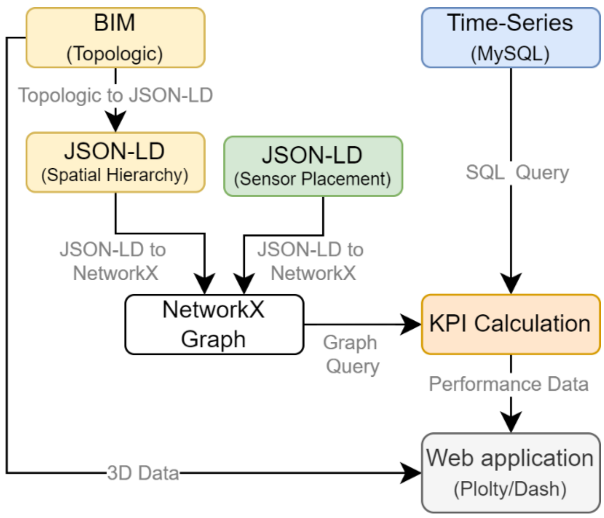


Figure 4. The system architecture of the prototype DDSS.

Figure 5: Knowledge graph developed for the sample application.

Conclusion

The AECO sector is currently experiencing a digital revolution, resulting in an increased need for instruments that facilitate decision-making in the realm of building management. Despite technological strides, a scarcity of these tools persists, primarily due to the substantial technical and economic investments required for their creation.

To provide a solution for streamlining the prototyping process of web-based DDSSs and address the gap in data integration in the AECO sector, this article presented BTwin, a software library designed for low-code development of DDSSs aimed at building performance-based management. The article, serving as a whitepaper for BTwin, provided the theoretical background of the toolkit, describing the axioms, concepts, and principles on which it is based. Additionally, it offered demonstrations using a simple fictional case study building to illustrate an example of its application.

This open-source toolkit, developed using a JSON-LD format and several Python functions, features a range of methods designed for semantically integrating building data from various sources, such as BIM, BPS, meters, and sensors, within KGs. It also enables data visualization through interactive dashboards, facilitating the monitoring and analysis of building performance-related aspects in a user-friendly manner. This approach could enable developers to easily create DDSS prototypes, allowing stakeholders to quickly preview dashboards during the development processes of these digital systems and provide feedback from a user-centered design perspective, thus ensuring tools tailored to actual needs. Moreover, its human-readable design makes it valuable for students and learners in the fields of BIM, DT, and KGs.

The toolkit is currently in development, with plans to expand its capabilities. Future enhancements will involve the integration of additional connectors to align the JSON-LD format with more complex data structures, such as gbXML (Green Building XML), to ensure compatibility with widely used BIM and web data exchange standards. Experiments will be conducted to utilize BTwin's KGs in more robust databases than NetworkX, such as Neo4j or MongoDB, and to incorporate real-time time-series databases. Moreover, special focus will be placed on refining the interfaces to simplify user access to data within the KG, enabling data queries through textual prompts. This feature will be developed utilizing Large LLMs like ChatGPT, which are experiencing rapid growth in the industry. Ultimately, future versions will demonstrate the tool's application in more complex real-world case studies, guiding the presented approach towards a DT perspective.

Acknowledgements

The publication was realized by a researcher with a research contract co-funded by the European Union - PON Research and Innovation 2014-2020 pursuant to Art. 24, paragraph 3, lett. a), of Law no. 240 of 30 December 2010 and subsequent amendments and Ministerial Decree no. 1062 of 10 August 2021.

References

Abuimara, T., Hobson, B.W., Gunay, B., O'Brien, W., Kane, M., 2021. Current state and future challenges in building management: Practitioner interviews and a literature review. Journal of Building Engineering 41, 102803. https://doi.org/10.1016/j.jobe.2021.102803

Balaji, B., Bhattacharya, A., Fierro, G., Gao, J., Gluck, J., Hong, D., Johansen, A., Koh, J., Ploennigs, J., Agarwal, Y., Bergés, M., Culler, D., Gupta, R.K., Kjærgaard, M.B., Srivastava, M., Whitehouse, K., 2018. Brick : Metadata schema for portable smart building applications. Applied Energy 226, 1273–1292. https://doi.org/10.1016/j.apenergy.2018.02.091

Betti, G., Tartarini, F., Nguyen, C., Schiavon, S., 2022. CBE Clima Tool: a free and open-source web application for climate analysis tailored to sustainable building design. https://doi.org/10.48550/ARXIV.2212.04609

Boje, C., Guerriero, A., Kubicki, S., Rezgui, Y., 2020. Towards a semantic Construction Digital Twin: Directions for future research. Automation in Construction 114, 103179. https://doi.org/10.1016/j.autcon.2020.103179

buildingSMART International, n.d. International Foundation Classes (IFC). Available online at: https://technical.buildingsmart.org/standards/ifc (Accessed 29/03/2024)

Chamari, L., Petrova, E., Pauwels, P., 2022. A web-based approach to BMS, BIM and IoT integration. CLIMA 2022 Conf. 2022: CLIMA 2022 The 14th REHVA HVAC World Congress. https://doi.org/10.34641/CLIMA.2022.228

Compton, M., Barnaghi, P., Bermudez, L., García-Castro, R., Corcho, O., Cox, S., Graybeal, J., Hauswirth, M., Henson, C., Herzog, A., Huang, V., Janowicz, K., Kelsey, W.D., Le Phuoc, D., Lefort, L., Leggieri, M., Neuhaus, H., Nikolov, A., Page, K., Passant, A., Sheth, A., Taylor, K., 2012. The SSN ontology of the W3C semantic sensor network incubator group. Journal of Web Semantics 17, 25–32. https://doi.org/10.1016/j.websem.2012.05.003

de Wilde, P., 2023. Building performance simulation in the brave new world of artificial intelligence and digital twins: A systematic review. Energy and Buildings 292, 113171. https://doi.org/10.1016/j.enbuild.2023.113171

Hosamo, H.H., Nielsen, H.K., Kraniotis, D., Svennevig, P.R., Svidt, K., 2023. Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method. Energy and Buildings 288, 112992. https://doi.org/10.1016/j.enbuild.2023.112992

Janowicz, K., Haller, A., Cox, S.J.D., Phuoc, D.L., Lefrancois, M., 2018. SOSA: A Lightweight Ontology for Sensors, Observations, Samples, and Actuators. https://doi.org/10.48550/ARXIV.1805.09979

Li, Y., García-Castro, R., Mihindukulasooriya, N., O'Donnell, J., Vega-Sánchez, S., 2019. Enhancing energy management at district and building levels via an EM-KPI ontology. Automation in Construction 99, 152–167. https://doi.org/10.1016/j.autcon.2018.12.010

Lu, Q., 2022. Digital twins in the built environment: fundamentals, principles and applications. ICE Publishing, London.

Lygerakis, F., Kampelis, N., Kolokotsa, D., 2022. Knowledge Graphs' Ontologies and Applications for Energy Efficiency in Buildings: A Review. Energies 15, 7520. https://doi.org/10.3390/en15207520

Ma, X., Li, X., Yuan, H., Huang, Z., Zhang, T., 2022. Justifying the Effective Use of Building Information Modeling (BIM) with Business Intelligence. Buildings 13, 87. https://doi.org/10.3390/buildings13010087

Mannino, A., Dejaco, M.C., Re Cecconi, F., 2021. Building Information Modeling and Internet of Things Integration for Facility Management. Literature Review and Future Needs. Applied Science 11, 3062. https://doi.org/10.3390/app11073062

Merino, J., Xie, X., Moretti, N., Chang, J.Y., Parlikad, A.K., 2022. Data integration for digital twins in the built environment. Presented at the 2022 European Conference on Computing in Construction. https://doi.org/10.35490/EC3.2022.172

Paulheim, H., 2016. Knowledge graph refinement: A survey of approaches and evaluation methods. Semantic Web 8, 489–508. https://doi.org/10.3233/SW-160218

Pauwels, P., Zhang, S., Lee, Y.-C., 2017. Semantic web technologies in AEC industry: A literature overview. Automation in Construction 73, 145–165. https://doi.org/10.1016/j.autcon.2016.10.003

Petri, I., Rezgui, Y., Ghoroghi, A., Alzahrani, A., 2023. Digital twins for performance management in the built environment. Journal of Industrial Information Integration. 33, 100445. https://doi.org/10.1016/j.jii.2023.100445

Pritoni, M., Paine, D., Fierro, G., Mosiman, C., Poplawski, M., Saha, A., Bender, J., Granderson, J., 2021. Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis. Energies 14, 2024. https://doi.org/10.3390/en14072024

Ramonell, C., Chacón, R., Posada, H., 2023. Knowledge graph-based data integration system for digital twins of built assets. Automation in Construction 156, 105109. https://doi.org/10.1016/j.autcon.2023.105109

Rasmussen, M.H., Pauwels, P., Hviid, C.A., Karlshøj, J., 2017. Proposing a Central AEC Ontology That Allows for Domain Specific Extensions, in: Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction., pp. 237–244. https://doi.org/10.24928/JC3-2017/0153

Zhang, F., Chan, A.P.C., Darko, A., Chen, Z., Li, D., 2022. Integrated applications of building information modeling and artificial intelligence techniques in the AEC/FM industry. Automation in Construction 139, 104289. https://doi.org/10.1016/j.autcon.2022.104289

DIGITAL BUILDING PERMITS AND DIGITAL BUILDING LOGBOOKS – CLUSTERING THE CHALLENGES AND REQUIREMENTS FOR ALIGNMENT

Pedro Mêda[1], Judith Fauth[2], Christian Schranz[3], Harald Urban[3], Hipólito Sousa[1]
[1]CONSTRUCT – Gequaltec, Porto University Faculty of Engineering, Porto, Portugal
[2]University of Cambridge, Cambridge, United Kingdom
[3]Research Unit Digital Building Process, TU Wien, Vienna, Austria

Abstract

Digitalisation supporting a more sustainable, efficient, and waste-friendly built environment is key. The requirements for this go far beyond the development of digital tools. The ability to collect, track, exchange, and trace data throughout processes is decisive. Digital Building Permit (DBP) and Digital Building Logbook (DBL) share these assumptions. However, a mutual realisation is lacking, and intuitions point to attaining added value. This vision motivates the research. Action research was elected to diagnose and discuss the viability of the vision and the associated challenges. Contributions demonstrate soundness and provide relevant insights into twin paths involving both concepts.

Introduction

Before becoming real and usable, buildings and civil engineering works need to be strategically defined, technically detailed, and constructed. This sequence of events is how, iteratively, construction "ripens and materialises" the idea and the object to be built. The underlying processes involve the production of large amounts of information. Digitalisation defies the transition from paper documentation to digital documents, metadata, and models (International Standards Organization, 2018). However, the EU Industrial strategy for construction (European Commission, 2021b) and the Green Deal requirements (European Commission, 2019) entail improvements at efficiency and sustainability levels, scoping the processes themselves as well as their outcomes (Papadaki *et al.*, 2023). To attain the proposed objectives, data-driven concepts with interfaces steeped by digital sobriety assumptions (Péréa, et al., 2023) must be positioned at the centre of the concerns.

Building permitting constitutes the authority of the local administration based on applying the measures laid down by law, aiming to secure the safety, sustainability, and compliance of buildings with the regulations (Fauth *et al.*, 2024). Digital Building Permitting (DBP) includes using digital data as input and output and digital tools to support or automate checking to tackle current limitations derived from analogue documents and processes (Malsane *et al.*, 2015). Several stakeholders structure building and infrastructure cadastral data for different purposes (Buchholz and Lützkendorf, 2023). Some data derives from national or local regulations and are kept by public authorities or private agents, from promoters to owners and users. Several gaps and inefficiencies are observed due to the wide range of processes and the often practised siloed and analogue approaches (Gómez-Gil, *et al.*, 2022). DBLs aim to collect, store, and link buildings' relevant data, fostering data transparency and increasing data availability on buildings-related properties to a broad range of market players (Dourlens S., et al., 2021). Several studies and projects are being developed to define and structure reference architectures and features to deploy DBPs and DBLs. However, little research exists seeking overlaps, potential relationships, challenges and added value from mutual developments and implementation.

This study builds upon previous research on DBP (such as Noardo *et al.*, 2022) and DBL (such as Mêda *et al.*, 2022) to focus on the abovementioned aspects. Stanford Center for Integrated Facility Engineering (CIFE) horseshoe and Action Research constitute the framework, anchored in a focus group. The focus group gathered experts on DBP and DBL to assess the relationship further and discuss the common elements. The first outcome is the answer to the research question: To what extent should DBP and DBL be considered related concepts? A second outcome involves identifying mutual challenges and processes with outlined associated requirements.

The article is organised as follows: The present Introduction section portrays the topic, the motivations for the work, the scientific approach, and the organisation. Second, the Research Design and Methods are presented, detailing the characteristics of the focus group from which most results stem. Follows the Diagnosis, where a brief overview of DBP and DBL is performed supported by scientific research and grey literature. This part includes a summary of the concerns, setting the bridge to the Action section, where the focus group contents are presented, namely, quantitative results from surveys and qualitative thoughts shared. Then, the Analysis and Discussion section proposes a framework with links, assessing and clustering the practical challenges and requirements of a common approach. Examples of processes and impacts on stakeholders' practices are debated. Finally, the Conclusions present the main remarks, elaborating on limitations and future directions.

Research Design and Methods

This research aims to demonstrate if DBP and DBL are related concepts. As part of the process, commonalities and requirements are evidenced, as well as clustering challenges at the implementation level. Given its potential to structure a transdisciplinary research process, the CIFE

Horseshoe Framework (see Figure 1) was selected as the guiding framework (Kunz and Fischer, 2007).~

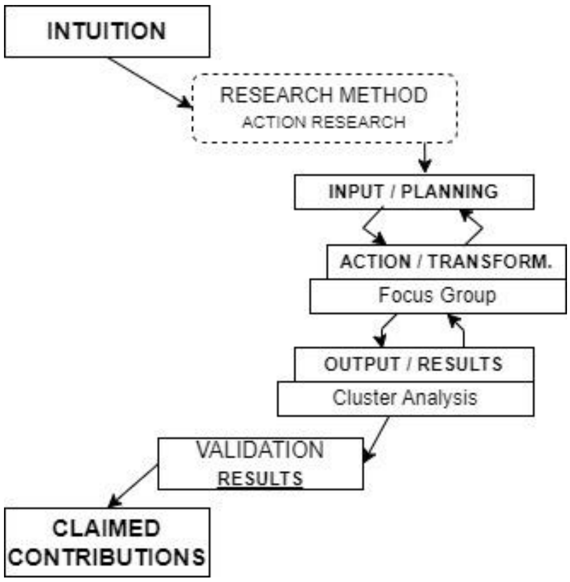


Figure 1: Research design using action research as the main method and The Stanford CIFE Horseshoe framework

The author's intuition establishes that DBP and DBL are related concepts. This relationship, which might seem obvious, does not have expression in the literature, nor is it a consensus among experts and industry stakeholders working on the associated topics and processes. Action research adopting a critical and pragmatic paradigm to investigate the intuition and propose improvements is used (Fellows and Liu, 2022).

The planning stage comprises DBP and DBL diagnosis anchored by relevant scientific and grey literature. The action stage involves a focus group in the form of an online workshop to discuss the alignment of the concepts and collect and explore qualitative and quantitative data on the requirements and challenges. The option for a focus group with selected experts was because it is the most suitable method to encounter the topic as this state. Following (Eeuwijk and Angehrn, 2017), it helps to identify and clarify shared knowledge among groups and communities, which would otherwise be difficult to obtain with a series of individual interviews.

Eighteen professionals were engaged, and the participation criteria were based on the sound knowledge/expertise on at least one of the concepts under discussion. In terms of background, the group had different origins from engineering (22,2%) and architecture (27,7%) practice to academia (27,7%) and geospatial sciences (11,1%). Regarding professional experience, 22,2% work between 15 and 20 years and 77,8% for over 20 years. The structure follows the characteristics addressed by (Jain, 2023), pursuing several iterative steps.

The outputs and research contributions comprise the discussion of all impressions and the analysis of the two surveys (see Figure 2) performed during the workshop. Cluster analysis (Romesburg, 1984) of all elements leads to the primary challenges and requirements and a proposal for a framework linking DBP and DBL.

Diagnosis

Digital Building Permit (DBP)

The process of issuing a permit for a building, independently of the phase, is a crucial milestone for any construction process. Despite being a small part of the mentioned life cycle, all relevant laws and regulations that ensure construction quality, user safety, environmental safety, etc., are enforced (Bloch and Fauth, 2023). The dependency of the legal framework and governmental processes makes the building permit process overly complex, prone to errors, non-transparent, and unpredictably lengthy (Ataide, *et al.*, 2023). Due to these characteristics, building permit issuance can be somewhat different in terms of process, information, and moment(s) during the construction process life cycle, among others. Some researchers have devoted their efforts to the topic of automated code compliance checking, for example (Bloch, *et al.*, 2023; Fischer *et al.*, 2023), which is one of the essential steps within a permit, mainly exploring the potential of using the BIM methodology (Amor and Dimyadi, 2021). Complementary to and following the international push for digitalisation, others have been looking to the building permitting process to identify ways to become more efficient and transparent through the use of digital tools (Bloch and Fauth, 2023; Fauth and Seiß, 2023; Krischmann, *et al.*, 2015).

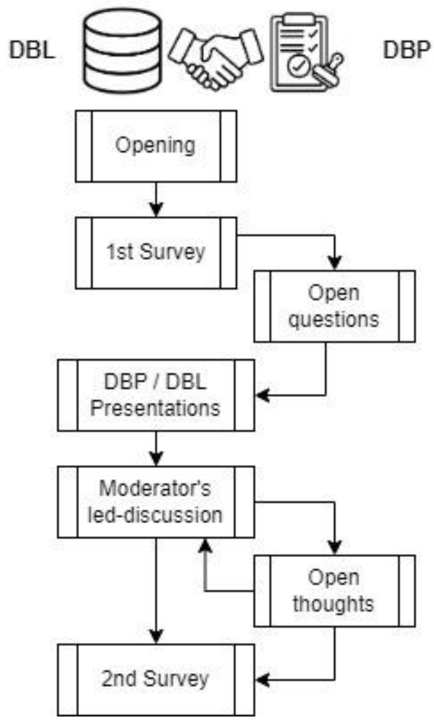


Figure 2: "DBL meets DBP" workshop – action structure

To support a digital transformation in the building permit domain, a basis of knowledge was needed, including not only the technological aspects of supporting industry in the development of the systems infrastructure but also the involved procedures, legislation on different levels and the involvement of various organisations (Fauth *et al.*, 2024). The DBP gained a specific space and attention under the EUnet4DBP initiative (Noardo *et al.*, 2020). Requirements and reviews were developed to clarify, provide awareness, identify common data and processes, and forecast common frameworks for DBPs (Noardo *et al.*, 2022; Fauth *et al.*, 2024).

Presently, there is the notion that DBP can represent a wide variety of building permitting processes running from early design phases to the handover and the commissioning of the building or infrastructure. Notwithstanding, almost all actions are driven by regulatory obligations and led by public authorities or designated bodies, meaning that similarities exist in how the process should run despite the differences in information requirements and compliance checks. The processes tend to be narrow regarding time frame, and DBP efficiency and accomplishment are linked to this objective. Finally, the issuance of a permit relies on information that is collected and structured for this specific purpose. Despite the issuance of the building permit, the storage of the information used for the process and the building permit report itself should be stored and kept during the life cycle of the construction. For example, a building usually changes during its life and most likely needs other building permits. It would be highly beneficial to have the necessary information directly by hand.

Digital Building Logbook (DBL)

An enormous amount of information is generated and used during a construction life cycle (Espinoza-Zambrano, *et al.*, 2023). However, this information is often collected in isolation, meaning that it is searched and managed to serve the purpose of a defined stakeholder in a given temporal moment of that life cycle. Other stakeholders must do the same for their specific objectives, searching for and managing the same or different information. According to (Miller *et al.*, 2014), this way of doing things leads to several gaps. Additionally, the process is highly error-prone and inefficient, diluting most of the benefits of information use. The proposal for the new Energy Performance of Buildings Directive (EPBD) (European Commission, 2021a) calls attention to the importance of effectively collecting and managing this data. The DBL concept was introduced as the common repository for all building relevant data, setting links to relevant existing databases, documentation, records and processes (Dourlens S. *et al.*, 2021).

Several recent studies have contributed to improving and clarifying the understanding regarding the DBL concept (Mêda *et al.*, 2021; Gómez-Gil, Espinosa-Fernández and López-Mesa, 2022; Alonso *et al.*, 2023). Most research works individually focus on specific aspects of the DBL, from the processes level to the data sources and functionalities, digital technologies or event contributions to progress indicators associated with the environment. Despite the wide range of discussed aspects, the DBL is broader and, as a result, also its challenges and added value. This wide range of aspects makes it very difficult for many stakeholders to understand the benefits when they are out of their scope of action or boundaries. Developing a business process modelling notation (BPMN) associated with the DBL life cycle was vital as a kick-off to detail the associated main processes (Mêda *et al.*, 2022). The objective was to set a background that enabled further granularity at process, data, and relationship levels.

The vision of the processes and their evolution throughout the construction process life cycle feeds the intuition of the potential relationships with DBP. Complementary, the outcomes from the most recent study on DBL technical guidelines (Grow, 2023) presented a proposal for DBL components where data from regulatory steps is to be collected and managed.

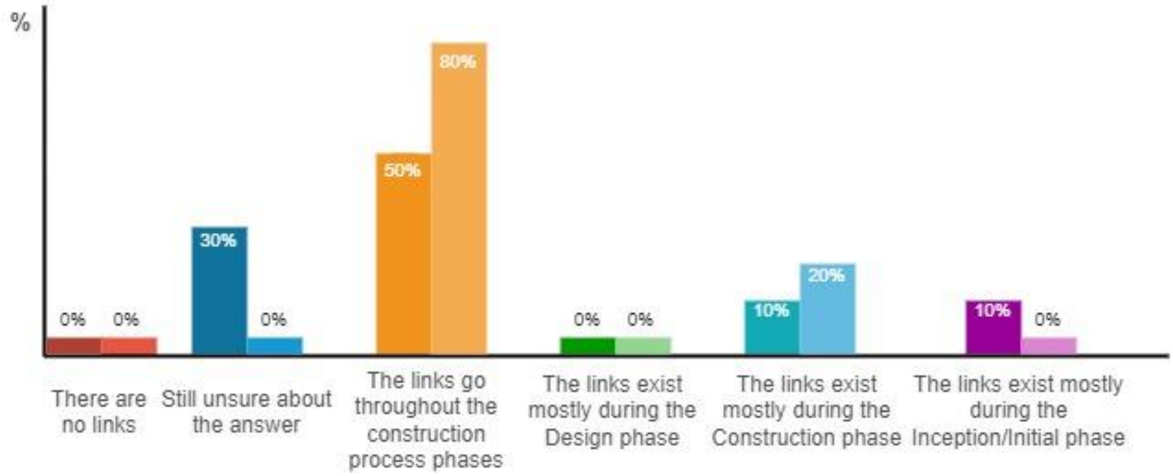


Figure 3: Focus group perception on the existing links between DBP and DBL

Action – Focus Group Workshop

A ninety-minute session was structured to accommodate several activities: two surveys, open questions, presentations, and moderators-led discussions and thoughts on challenges and priorities for future actions, all contributing to the research objectives. The two surveys were placed at the beginning and end of the session to collect quantitative data on the perception of concepts, their relationship, and how the visions evolved (see Figure 2). Figure 3 provides an overview of the respondents' visions and how they shifted during the session regarding the relationships between DBP and DBL. Darker colours (the left ones) represent the results from the first survey, and the lighter ones are the results from the second.

Interestingly, the answers point to evident links between the two concepts despite the fuzziness of the first survey on how they materialise during construction. In the second survey, most identify links throughout the construction process phases. It is one of the aspects worth further discussion.

The presentations mainly used contents already detailed in the diagnosis section. The objective was to harmonise the knowledge level regarding DBP and DBL. Several topics were raised during the discussion. Notwithstanding and for the purpose of the research, it is relevant to stress the shared thoughts associated with "Information Requirements", "Common Processes", and "Stakeholders".

When asked to share thoughts regarding the information requirements of DBP and DBL, many stated that DBP will have most of the requirements defined by regulations. At the same time, DBL seems to be more open regarding requirements. Depending on the evolutions and business models, a more risk-driven approach might be applicable. Given the wide range of situations these concepts can cover, defining a minimum set of information requirements was mentioned as needed, followed by mechanisms to assess the reliability of the information presented. The discussion also approached the existing knowledge and standards developed for Building Information Modelling (BIM). In this respect, it was found relevant to try "not to reinvent the wheel" and, in opposition, make all efforts to use the existing standards related to information requirements as they are or with updates.

A general consensus was observed regarding the existing links with Geographic Information Systems (GIS) when debating the common processes. A large majority of the participants shared that despite the wide range of processes surrounding permits, "DBP itself can be an integrated part of the DBL process". This declaration was a surprising insight, considering the objectives. To conclude, it is worth highlighting that it is assumed that DBP and DBL hold and deal with General Data Protection Regulation (GDPR)-sensitive data. Due to that, both concepts need to consider the processes to ensure GDPR compliance as part of their reference architectures.

To finalise, some thoughts were shared on the stakeholders' engagement: the awareness needs, the training, and the competencies to understand, proficiently apply and operate the technologies and processes surrounding DBP and DBL. Similarly, the previous knowledge and experiences with BIM implementation

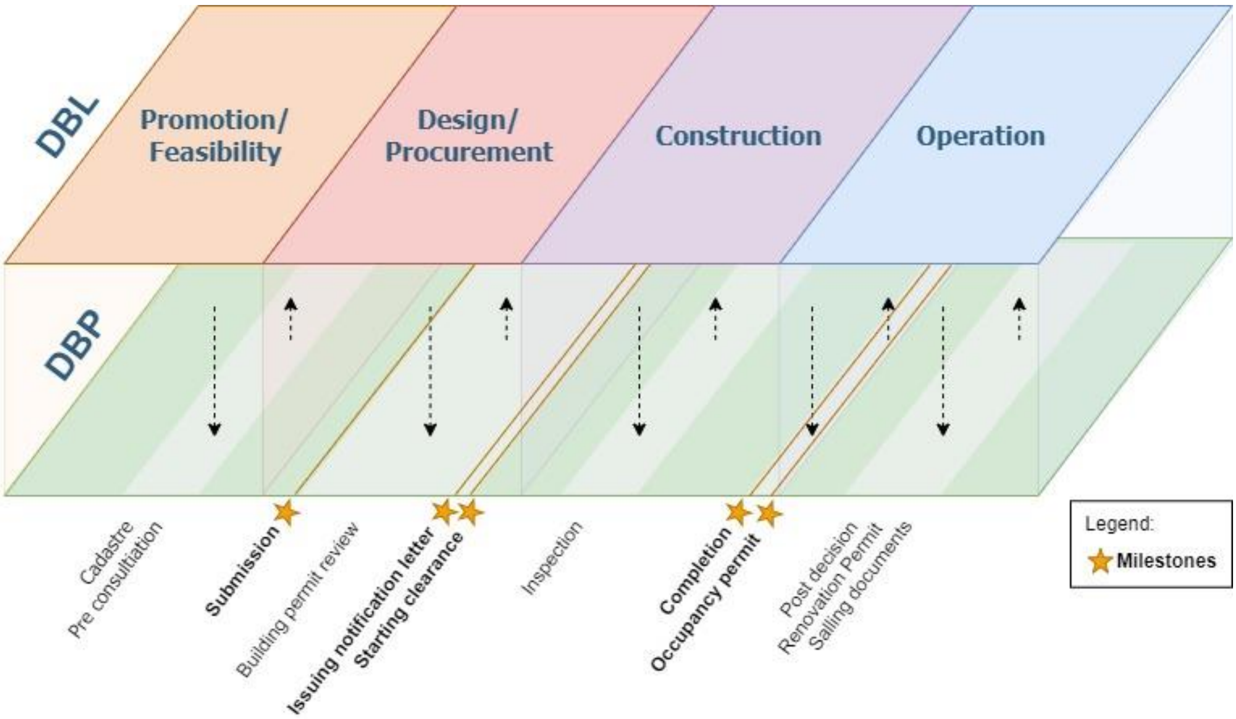


Figure 4: Process-related links between DBP and DBL

were highlighted as relevant to allow shortcuts and anticipate barriers.

Analysis and Discussion

Processes and Phases

It is possible to state that building permit processes are highly country- or even municipality-dependent. Despite this fact, there is the notion that building permitting is associated with verification, meaning that it occurs at a specific moment and often as a milestone belonging to the construction life cycle. In contrast, the processes involving the materialisation of the building logbook are more continuous. Assuming that DBL can work as a Digital Twin enabler, the concept will surpass the construction process life cycle itself, serving the purposes of the construction life cycle. Considering all the visions and statements, it is clear that DBP and DBL are related concepts, and several links throughout the construction life cycle may exist depending on the range of the building permitting processes.

Figure 4 envisages the links between DBP and DBL, using background knowledge from targeted research performed by (Fauth *et al.*, 2023) and (Mêda *et al.*, 2022) and the concerns expressed during the focus group.

Figure 4 presents a high-level framework proposal on how to merge DBL and DBP processes, establishing the data flows between both. Considering previous research, a broad vision of this permit process was considered from the building permit side, where these are required at several moments throughout the life cycle. In this respect, it is worth highlighting the cadastre evaluation and pre-consultations with the authorities to prepare a building application in the early phases and before design. The application adjustments during the review, including compliance checking and authorisation for construction, occupancy permit, and permit prior to a renovation intervention (similar to cadastre pre-consultation when there is a pre-existence and when the permit process starts again).

Looking in detail at one of the possible examples and seeking increased granularity, it is possible to pinpoint a situation during the design phase where the information requirements for the building permit (for example, building permit review performed by the Municipality) will need to be considered for the design development. All elements should be developed or deployed as part of the DBL. Depending on the type of automation, the elements can be sent via DBL or the communication can be enabled with the DBP system, which is managed by the authority. After the analysis, the building permit or the correction notice is sent using the established communication mechanism (as previously and in the opposite direction). The records are kept, and all relevant stakeholders are informed of the results (depending on the individual authorisation). The elements must be centralised or stored in a distributed ledger, depending on the system's architecture.

Requirements and Challenges

As described, the thoughts shared during the focus group were prosperous. They allowed the identification and clustering of common requirements and challenges. Eleven aspects were considered the most relevant to prioritise in terms of what concerns the challenges. Relations to BIM, awareness, training and knowledge, and related services, meaning additional services linked with DBL, and services that must be organised to feed DBP and/or DBL, were identified as the aspects where requirements exist at process, stakeholder and data levels. A second group of challenges was identified, comprising process and stakeholder requirements. These involve governance (existing standards, regulations, use of open standards and BIM, challenges in terms of revising and setting new legal diplomas), the identification and establishment of mechanisms linking existing databases, implementation costs, and all the environment associated with data security and accessibility. One challenge was identified with data and stakeholder requirements related to establishing common terminology for DBP and DBL. Data framework and Information reliability challenges were identified in association with the data level. Finally, the challenge associated with the maintenance of data and curation of the systems throughout the building life cycle was identified as bearing requirements at the process level. Figure 5 overviews the challenges and associated requirements clustered from the focus group inputs.

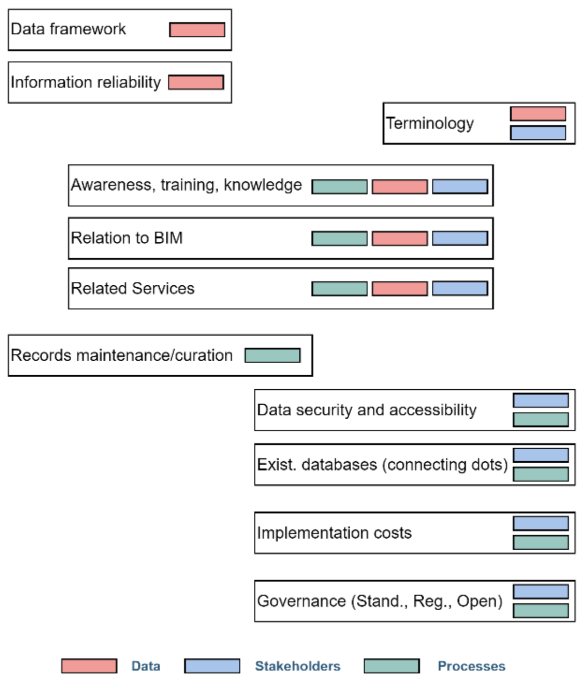


Figure 5: Requirements and Challenges deriving from the cluster analysis of the focus group inputs

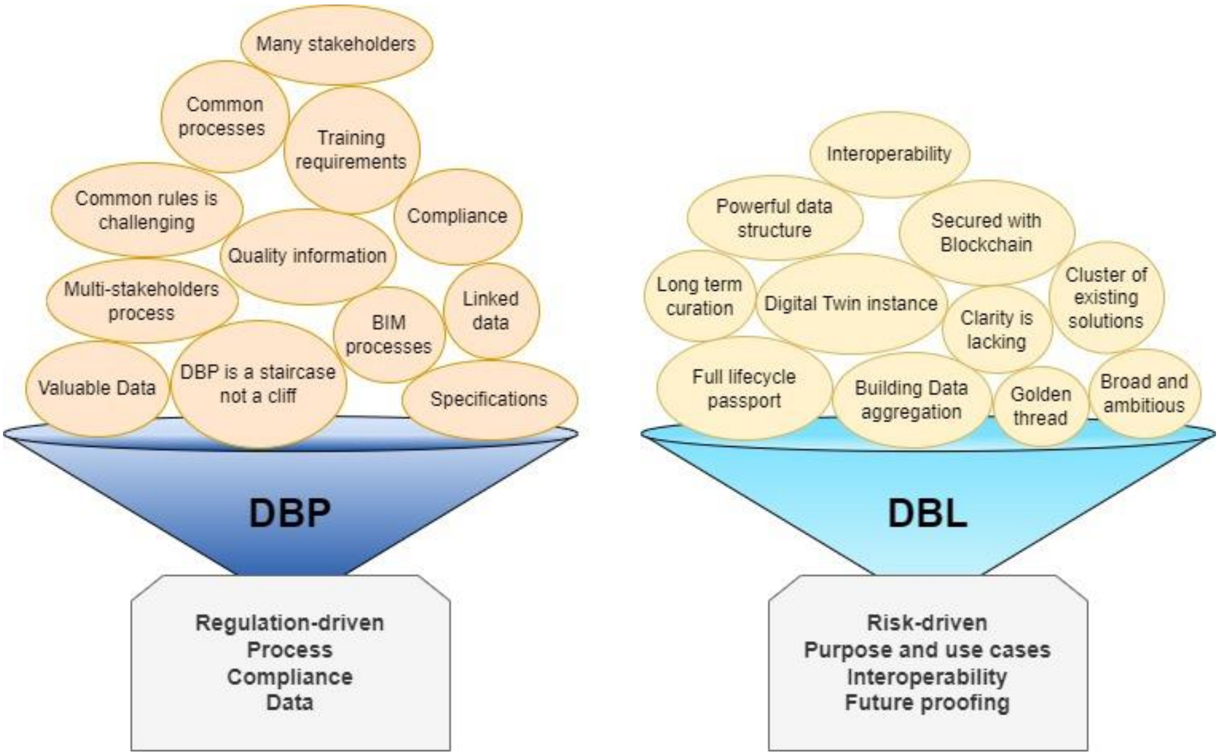


Figure 6: DBP and DBL associated concepts and main clustered characteristics

Conclusions

This research proves that DBP and DBL are related concepts and have several potential interconnections throughout the construction life cycle. Nevertheless, it is paramount to stress that DBP is more time-framed (specific processes). In contrast, DBL is more continuous through the construction life cycle. Combining the findings with previous outcomes, it is worth mentioning that there are no time lags between DBP and DBL as this is deployed together with the idea, and DBP will interact, depending on the range of processes, through the life cycle. This vision is supported by the evidence that the concepts partially share information requirements, relationships with databases, data security and reliability concerns, governance, and curation care (Figure 6).

DBP and DBL can be set as independent systems. However, given the overlaps in their scope and for the sake of digital sobriety, it is relevant to consider a shared landscape for terminology and datasets. This aspect should encompass other systems and concepts. Additionally, it is worth highlighting that the efficiency and complete accomplishment of the objectives of both concepts might be disturbed if there is the need to feed each one manually with data from the other.

From a strategic viewpoint, there can be several benefits from the synergies and proper links and data sharing between DBP and DBL. The implementation challenges will be higher due to the coordination needs from a systems architecture perspective. However, several savings can be achieved in awareness, training, and getting the confidence of stakeholders for use. As discussed, several stakeholders' activities might overlap, meaning adoption can become more straightforward by working correctly with the synergies.

This study's limitations are related to the defined scope and level of granularity. Nevertheless, this was set intentionally to set the background for future and more in-depth research activities. This research focused on specific dimensions, such as processes, data, technologies, solutions, and stakeholders' needs, setting assumptions to work further on the articulation between DBP processes and DBL.

As studies point to the integration between BIM and GIS, DBP and DBL, which are still in their infancy, should benefit from a joint endeavour, contributing to the accomplishment of digital twin transition goals for the built environment and effectively raising most relevant stakeholders into the Construction 4.0 dimension. If building permitting is a critical process in the construction life cycle, using and producing building-related data. Logbooks aim to be golden thread enablers, and complete data traceability implies embedding the digital processes for building permitting as part of the digital logbooks' framework.

Acknowledgements

Pedro Mêda and Hipólito Sousa would like to thank: Base Funding - UIDB/04708/2020 with DOI 10.54499/UIDB/04708/2020 (https://doi.org/10.54499/UIDB/04708/2020) of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC).

Judith Fauth has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034337.

References

Alonso, R. *et al.* (2023) 'Towards the definition of a European Digital Building Logbook: A survey', *Heliyon*, 9(9). Available at: https://doi.org/10.1016/j.heliyon.2023.e19285.

Amor, R. and Dimyadi, J. (2021) 'The promise of automated compliance checking', *Developments in the Built Environment*, 5(December 2020), p. 100039. Available at: https://doi.org/10.1016/j.dibe.2020.100039.

Ataide, M., Braholli, O. and Siegele, D. (2023) 'Digital Transformation of Building Permits: Current Status, Maturity, and Future Prospects', *Buildings*, 13(10). Available at: https://doi.org/10.3390/buildings13102554.

Bloch, T., Borrmann, A. and Pauwels, P. (2023) 'Graph-based learning for automated code checking – Exploring the application of graph neural networks for design review', *Advanced Engineering Informatics*, 58(August). Available at: https://doi.org/10.1016/j.aei.2023.102137.

Bloch, T. and Fauth, J. (2023) 'The unbalanced research on digitalization and automation of the building permitting process', *Advanced Engineering Informatics*, 58(June), p. 102188. Available at: https://doi.org/10.1016/j.aei.2023.102188.

Buchholz, M. and Lützkendorf, T. (2023) 'European building passports: developments, challenges and future roles', *Buildings and Cities*, 4(1), pp. 902–919. Available at: https://doi.org/10.5334/bc.355.

Dourlens S. & Carbonari G., De Groote M., Borragán G. & De Regel S., Toth Z., V.J.& G.J. (2021) *Study on the Development of a European Union Framework for Digital Building Logbook - Final Report*. Brussels. Available at: https://doi.org/10.2826/659006.

Eeuwijk, P. Van and Angehrn, Z. (2017) *How to.... conduct a focus group discussion (FGD): Methodological manual by Peter van Eeuwijk and Zuzanna Angehrn*, *Swiss TPH*. Basel.

Espinoza-Zambrano, P., Marmolejo-Duarte, C. and García-Hooghuis, A. (2023) 'Libro del Edificio Electrónico (LdE-e): Advancing towards a Comprehensive Tool for the Management and Renovation of Multifamily Buildings in Spain', *Sustainability (Switzerland)*, 15(4). Available at: https://doi.org/10.3390/su15042957.

European Commission (2019) *The European Green Deal*, *COM(2019) 640 final*. Brussels.

European Commission (2021a) *Proposal for a Directive of the European Parliament and of the Council on the energy performance of buildings (recast). COM(2021) 802.*, *Official Journal of the European Union*. Brussels. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX %3A52021PC0802&qid= 1641802763889.

European Commission (2021b) *The New Industrial Strategy for Europe*, *COM(2020) 102 final*. Available at: https://doi.org/10.1007/s10272-021-0967-8.

Fauth, J. *et al.* (2023) 'Requirements and framework for Gaia-x-based building permit processes', *Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference*, 4. Available at: https://doi.org/10.35490/ec3.2023.216.

Fauth, J. *et al.* (2024) 'Taxonomy for building permit system - organizing knowledge for building permit digitalization', *Advanced Engineering Informatics*, 59 (May 2023), p. 21. Available at: https://doi.org/10.1016/j.aei.2023.102312.

Fauth, J. and Seiß, S. (2023) 'Ontology for building permit authorities (OBPA) for advanced building permit processes', *Advanced Engineering Informatics*, 58(October). Available at: https://doi.org/10.1016/j.aei.2023.102216.

Fellows, R. and Liu, A. (2022) *Research methods for construction*. Wiley-Blac. Oxford.

Fischer, S. *et al.* (2023) 'Automation of escape route analysis for BIM-based building code checking', *Automation in Construction*, 156(October 2022), p. 105092. Available at: https://doi.org/10.1016/j.autcon.2023.105092.

Gómez-Gil, M., Espinosa-Fernández, A. and López-Mesa, B. (2022) 'Review and Analysis of Models for a European Digital Building Logbook', *energies*, 15, p. 24. Available at: https://doi.org/https://doi.org/10.3390/en15061994.

Grow, D. (2023) *Technical guidelines for digital building logbooks - Guidelines to the Member States on setting up and operationalising digital building logbooks under a common EU framework*. Brussels.

International Standards Organization (2018) *ISO 19650-1 - Organization and digitization of information about buildings and civil engineering works , including BIM -Information management using building information modelling*. Switzerland.

Jain, N. (2023) *What is Focus Group Research? Definition, Types, Methods, and Examples*, *Ideascale*. Available at: https://ideascale.com/blog/what-is-focus-groupresearch/#toc_What_is_Focus_Group_Research.

Krischmann, T., Urban, H. and Schranz, C. (2015) ‘Entwicklung eines openBIM Bewilligungsverfahrens’, *Bauingenieur*, 90(5), pp. 335–344. Available at: https://doi.org/https://doi.org/10.37544/0005-6650-2020-09-61.

Kunz, J. and Fischer, M. (2007) *CIFE Research Questions and Methods*, *Stanford Digital Repository*. Stanford.

Malsane, S. *et al.* (2015) 'Development of an object model for automated compliance checking', *Automation in Construction*, 49(PA), pp. 51–58. Available at: https://doi.org/10.1016/j.autcon.2014.10.004.

Mêda, P. *et al.* (2021) 'Incremental Digital Twin Conceptualisations Targeting Data-Driven Circular Construction', *Buildings*, 11(11), p. 554. Available at: https://doi.org/10.3390/buildings11110554.

Mêda, P. *et al.* (2022) 'A Process-Based Framework for Digital Building Logbooks', in EC3 (ed.) *2022 European Conference on Computing in Construction*. Ixia: EC3, p. 8. Available at: https://doi.org/10.35490/EC3.2022.183.

Miller, W. *et al.* (2014) *Strategies and solutions for housing sustainability: building information files and performance certificates*. Queensland. Available at: https://eprints.qut.edu.au/67271/.

Noardo, F. *et al.* (2020) 'Integrating Eexpertises and Ambitions for Data-Driven Digital Building Permits – The EUNET4DBP', *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIV-4/W1-, pp. 103–110. Available at: https://doi.org/10.5194/isprs-archives-XLIV-4-W1-2020-103-2020.

Noardo, F. *et al.* (2022) 'Unveiling the actual progress of Digital Building Permit: Getting awareness through a critical state of the art review', *Building and Environment*, 213(January), p. 108854. Available at: https://doi.org/10.1016/j.buildenv.2022.108854.

Papadaki, I. *et al.* (2023) *Transition pathway for Construction*, *Directorate H – Ecosystems III: Construction, machinery & standardisation*. Brussels. Available at: https://ec.europa.eu/docsroom/documents/53854.

Romesburg, H.C. (1984) *Cluster Analysis for Researchers*. Belmont: Lifetime Learning Publications United States.

LEVERAGING GRAPH BASED SEMANTIC ENRICHMENT FOR ENHANCED AUTOMATED CODE-CHECKING

Tanya Bloch[1], Jiabin Wu[2], Maen Alnuzha[1] and André Borrmann[2]
[1]Technion Israel Institute of Technology, Haifa, Israel
[2]Technical University of Munich, Munich, Germany

Abstract

Despite advances in BIM technology, achieving a comprehensive machine-readable design representation for Automated Code Checking (ACC) remains a challenge. We propose decomposing regulatory checks into sets of semantic enrichment tasks, each tailored with a suitable solution. We demonstrate this approach with a case study on Israeli regulations for security rooms, focusing on two distinct semantic enrichment tasks. We use the Girvan-Newman algorithm to isolate single apartments in a floor plan. A GNN then categorizes security rooms based on the ratio of supporting walls, according to the ranges defined in the regulations. The approach paves the way for a highly automated ACC methodology, with potential implications for diverse regulatory contexts.

Background

The conventional approach to ACC involves two pivotal tasks: the translation of human-written text into machine-readable rules and the establishment of a machine-readable representation of the building design. The ability to translate regulatory documents into hard-coded rule sets is limited due to ambiguous, subjective, or vague regulatory statements, requiring manual interpretation and contextual knowledge (Nawari, 2019; Zhang and El-Gohary, 2017). Despite the prevalence of BIM technology and the adoption of Industry Foundation Classes schema (IFC) as an industry standard, achieving a machine-readable representation of the design that contains all semantic information required for ACC, remains a challenging task that often leads to manual preprocessing of the BIM models.

Moreover, matching regulatory concepts to those represented in the computer-readable design remains highly reliant on human understanding. Recent research proposed implementing ML approaches for ACC, eliminating the need for compiling explicit rule sets (Bloch et al., 2023). However, the efficacy of ML-based code checking is contingent on comprehensive input information. As BIM models often lack the necessary highly rich semantic information, semantic enrichment offers an automated solution for supplementing it.

Complex code clauses typically require querying the topological relationships among building elements, for example, checks of wall continuity across security rooms in Israeli Home Front Command regulations, rules of accessible route and line of visibility in the New Zealand Building Code, checks on support of plaster skins in the International Residential Code, rules related to the travel path and distance in the ADA Standards for Accessible Design and the International Building Code.

Verifying building regulations necessitates a thorough understanding of the complex relationships among various elements throughout the building's structure. Consequently, code checking requires extended data structures or a proof-of-solution describing how the design proves compliance rather than merely fulfilling prescribed criteria. For example, graph processors can be integrated to address implicit spatial properties (Solihin and Eastman, 2015). In this work, we decompose the ACC task into a series of semantic enrichment tasks aiming to automatically supplement needed information for the checking process, focusing on complex requirements that involve relational aspects of the design. We turn to graph-based semantic enrichment techniques to address these specifications, leveraging the inherent relational structure within the design data.

Semantic enrichment

Semantic enrichment of BIM models emerged as a solution for lifting the need for multiple domain-specific Model View Definitions (MVDs) by reasoning over explicitly represented information to derive new facts about the model (Belsky et al., 2016). While some efforts previously focused on querying BIM to extract implicit information stored in models (Borrmann et al., 2006; Mazairac and Beetz, 2013; Wülfing et al., 2014), they often fell short in providing explicit representations of inferred information as part of the building information, limiting downstream applications. The SEEBIM system, proposed by Belsky et al. (2016), leveraged domain expert knowledge, represented as logical statements, to infer new facts and explicitly add them to the model. Wu and Zhang (2019) also implemented a rule-based, iterative method for classifying BIM objects in IFC, leveraging geometric features to identify objects with similar geometric representations. They also recognized that pure geometry, without contextual information is limited in the ability to distinguish between elements with similar geometry. In parallel, semantic enrichment solutions that are based on semantic web technology have been suggested and illustrated as well. In fact, one of the motivations for using the Web ontology, as presented by Pauwels et al. (2017), is utilizing logical inference and proofs, which involves the use of First Order Logic (FOL) to derive new insights from the initial building model. Combining inferred information with the original data represents the core objective of semantic enrichment.

Departing from the FOL methodology, Bloch & Sacks (2018) explored different methods for semantic enrichment. Their work compares a rule-based and a ML-based approach for semantic enrichment by demonstrating both for classifying room types in residential apartments. They found machine learning worked better than rules for this specific task. With the recent developments in the ML domain, and the introduction of Graph Neural Networks (GNNs) (W. L. Hamilton et al., 2017), the applicability and benefit of such models has been demonstrated for the same room type classification problem (Wang et al., 2022). The major benefit of such models is the ability to leverage not only geometric but also contextual information about the building elements. Motivated by limitations in existing BIM software regarding semantic representation and interoperability, the authors proposed leveraging graph-based data structures coupled with enhanced GNN architectures that incorporate both node and edge features. They develop the SAGE-E algorithm and generate the Room Graph dataset containing apartment layouts. Experiments demonstrate the superior performance of SAGE-E over conventional ML methods for room type classification, validating the promise of GNNs for BIM semantic enrichment. This groundbreaking effort establishes a research foundation for applying GNNs to augment BIM semantics. As suggested in Bloch and Sacks (2020), different semantic enrichment tasks can benefit from different approaches to solution. Until now, the more classic graph-based methods remain mostly unexplored in the context of semantic enrichment.

Graph representation of building design information

Building models encapsulate a wealth of architectural and engineering design knowledge, along with specific native design intentions tailored to meet the predefined requirements. In the realm of representing the embedded design properties and the complex topological relationships among building elements, graph structures have emerged as a prevalent tool (Vilgertshofer and Borrmann, 2017).

The adoption of graph-based methodologies facilitates the transformation of building models of structures into networks. These networks comprise nodes and edges, which represent building objects and interrelationships. The graph representation hinges on the specific nature of the building's design structure and the intended application scenarios. Furthermore, the choice of the graph structures depends on the varying objectives of the query. For example, dependency graphs are utilized to predict the clash change components (Hu et al., 2023), and parametric building graphs are employed to match detailing patterns (Abualdenien and Borrmann, 2021).

Method

In this work, we propose a comprehensive, hybrid approach for addressing the challenges of ACC. We demonstrate the approach through a case study, focusing on the effectiveness of graph-based semantic enrichment in addressing the requirements defined by the Israeli Home Front Command regulations for security rooms (Home Front Command, 2010). The experiment described below is structured to fulfill two essential tasks integral to the checking process. The first task is to correctly identify the security rooms in the BIM model. As previous work on the subject demonstrated effective solutions under the assumption that we can isolate individual apartments in the building, we implement community detection to automatically isolate spatial groupings representing individual apartments. The second task addressed in the described experiment is automatically recognizing the ratio of vertical connectivity between the walls of the security rooms. Recognizing the complexity of calculating this ratio, we propose a ML-based classification approach using GNNs. The GNN model was trained to categorize the ratio of vertical continuity into predefined ranges specified by the regulations, streamlining the checking process without the need for precise numeric calculations. The checking process is then complete using a set of logical statements. By implementing diverse computational techniques, we demonstrate an enhanced ACC process with no additional information requirements for the modeler.

Experiment and results

Israeli regulations for security rooms

Security rooms (bomb shelters) are designated spaces within a building specifically designed and constructed to provide protection during emergency situations related to military actions. These spaces are fully constructed from reinforced concrete and equipped with specialized windows and doors that are designed to withstand shrapnel from a blast. As per Israeli Home Front Commands regulations (2010), there are several restrictions for the components of the individual security rooms, such as room dimensions, thickness of slabs and walls, dimensions of openings, etc. However, collections of security rooms are also examined together. The security rooms must be vertically stacked so that the reinforced concrete walls are continuous and reach the foundations. Due to functional design requirements, parking, storage, or other non-protective spaces (security space) often reside above or below the security rooms, where reinforced concrete walls may interfere with the functionality of these spaces. Alterations like wall removals, added openings, or weaker materials in those spaces can disconnect the security room from its base.

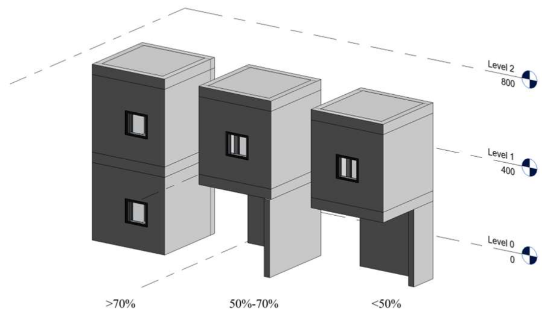


Figure 1: Ratio of vertical continuity of security room walls.

Regulations thus allow for some vertical discontinuity contingent upon certain conditions. One key determinant that dictates additional measures required is the percentage of the security room walls that are continuous and reach the foundation. We will refer to this as the "ratio of vertical continuity."

The regulations delineate three ranges, as illustrated in Figure 1, for the ratio of vertical continuity: spaces with over 70% continuous walls require no further enhancements, effectively extending the requirements for individual security rooms through all associated levels. Spaces with 50-70% vertical continuity and spaces showing less than 50% vertical continuity must incorporate thicker structural walls and slabs and are restricted in the accepted sizes of openings. Note that the ratio of vertical continuity can be calculated per wall individually, thereby dictating additional constraints on the permissible horizontal displacement between walls. And can be calculated for the entire space as the sum of the inner boundaries defined by the walls that compose the space. The ratio of vertical continuity for spaces determines the requirements for the concrete thickness of slabs and walls in the security shaft. The ability to check the compliance of a design to the concrete thickness requirements is the focus of this work. The calculation of the ratio of vertical continuity for walls and spaces involves a large set of guidelines, conditions, and exceptions. The topology of the walls supporting the security rooms, their orientation, sizes of openings and distances between them, and their proximity to existing wall corners, etc., can impact the calculation of the vertical continuity ratio since, under specific conditions, openings must be subtracted from it.

To automate this complex calculation based on BIM data, parametric and relational details of building components that are not explicitly present in the models must be considered. While some of the needed information is intrinsic to individual components, some rely on relationships between model elements, which may be more difficult to retrieve.

In this work, we demonstrate a workflow towards a highly automated compliance checking of these regulations using a case study of a residential building that includes overall five security shafts.

Community detection and room classification

As explained in the previous sections, any kind of ACC routine is contingent upon information requirements that are not always given in the BIM model in their explicit form. Spaces are the most primitive building elements that are crucial for many checks. Given that there are no standards or enforced naming conventions, the room names provided by designers become unreliable, which poses challenges for downstream applications. Given that our goal is to avoid introducing information specifications to the design process, the implementation of semantic enrichment to support ACC emerges as a promising solution. In the context of automated space function classification, it has been demonstrated that semantic enrichment for classifying abstract elements, such as spaces, is better addressed with ML than rules (Bloch and Sacks, 2018). Later, a graph-based approach for room type classification has also been demonstrated using GNNs (Wang et al., 2022). In both cases, however, the basic assumption was that we operate within the boundaries of an apartment, and the question of how we can isolate the apartments without additional information requirements for the designers remained open. We propose an automatic solution by implementing a community detection algorithm.

Community detection algorithms are useful for analyzing complex networks to identify subgroups within them. In the context of buildings, where spatial relationships and connectivity are crucial, algorithms for community detection can help uncover distinct spatial groupings indicative of separate apartments. In this test case, we implement the Girvan-Newman algorithm (Girvan and Newman, 2002) for community detection, which was originally designed for social networks. The principle of this algorithm is isolating communities in a graph by

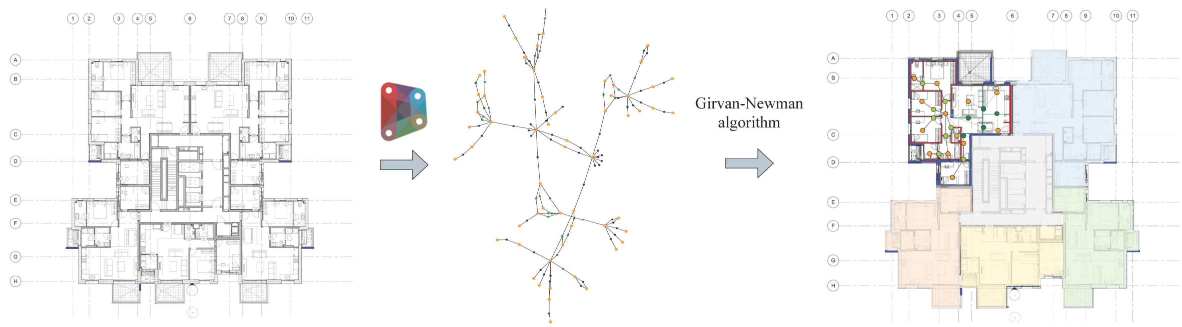


Figure 2: Graph representation of individual floors and community detection for isolating apartments.

iteratively removing edges with high *betweenness centrality,* a measure of how often an edge lies on the shortest path between two nodes. It is important to note that while the Girvan-Newman algorithm is considered relatively slow, it has been determined that it can be effectively used on graphs with less than 1000 edges (Fortunato, 2010). As we are concerned with floorplan layouts that translate into small graphs (in our case, no more than 200 edges), the algorithm provided accurate results within seconds. Other community detection algorithms were not examined.

To implement the suggested technique, we followed the workflow described in Figure 2. Utilizing the API of the BIM authoring tool (Autodesk Revit) and Dynamo (Autodesk, 2022), we collect the building elements that engage in any connection relationship within the spatial layouts, including spaces and doors. Additionally, designers typically use separation lines to divide distinct functional areas within shared spaces. Thus, both the connecting doors, typically situated in boundary walls, and separation lines are recognized as accessible or connecting elements. Furthermore, we perform topological queries on the BIM model to determine the relationships between spaces and accessible elements, forming the undirected edges within the graph structure. For each selected element, the spatial data (floor level) and the associated semantic data (element type) are queried and attached to the related graph node. Within the graph structure representing the whole building model, the elements that belong to each floor are determined by grouping the unconnected subgraphs. As a result, each floor plan of the model is individually processed, leading to the automated generation of accessibility graph structures for our test case.

Once graph extraction was complete, the Girvan-Newman algorithm for community detection was applied on the typical floors, and all individual apartments were correctly identified, after which the previously developed methods for room type classification can be applied. At the end of this stage, we assume all the security rooms in the model are identified and correctly tagged.

Graph Neural Networks for semantic enrichment

Calculating the ratio of vertical connectivity is a complex task. However, the absolute numerical value itself holds minimal significance. What is important is discerning the category of further requirements associated with the ratio—whether it falls within the 70%, 50-70%, or below 50% range. In essence, complex calculation can be avoided by implementing a ML model tailored specifically for this classification task. Given that the calculation heavily relies on the relationships between walls, GNNs emerge as a valuable tool for this classification task. Unlike traditional models, GNNs possess the unique ability to account for and leverage the complex interconnections among building elements.

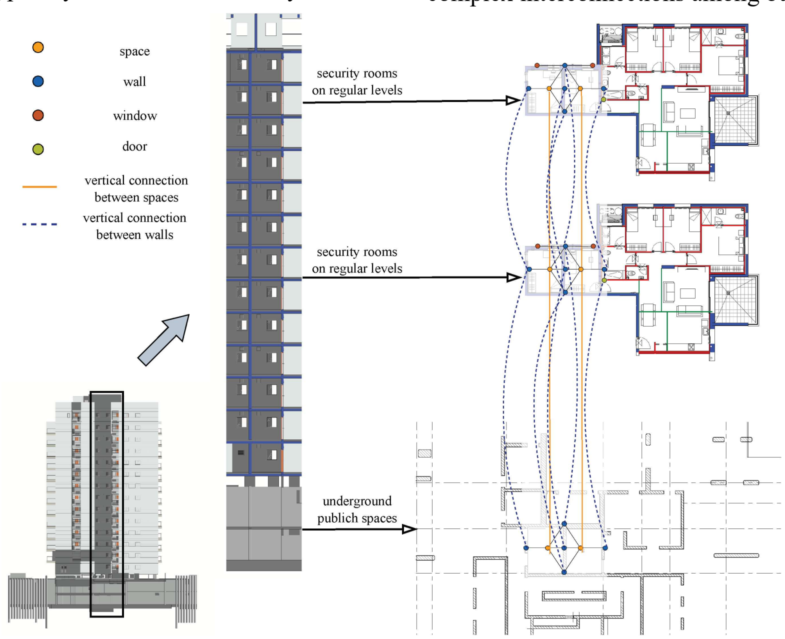


Figure 3: Graph representation of security shafts - extracting the vertical graphs.

To enable the application of GNNs, two steps are implemented. First, we aggregate all security rooms within a single shaft and link them to any continuous walls that support the security rooms. Given the high complexity of the interconnections among security-room-related elements, we interpret the design model to generate a shaft-centered graph structure utilizing the authoring tool's API via Dynamo (Autodesk, 2022). The topological query analyzes all shaft-related building elements across all floors.

By taking the bounding box coordinates of all shafts as horizontal references, we identify all security spaces that align horizontally with the shafts on all floors. We then filter out the irrelevant spaces and perform a bounding box geometric analysis at the space level to calculate the connecting relationships in the vertical direction. To achieve this, we vertically extend the space bounding boxes with a pad dimension threshold, which equals to the identified slab thickness. We also analyze the bounding relationships, including the boundary interactions between spaces and walls and the hosting relationships between walls and openings like windows and doors. This analysis translates the topological relationships of building elements and their hosts into the graph structure. Similarly, we apply the bounding box analysis to the selected walls to provide edge information about whether the structural walls of security rooms are continuously stacked until the foundations are reached. The abovementioned topological queries and geometric calculations are accomplished in an automated manner by taking the shaft-related building elements extracted from the design authoring tool as a basis.

The generated shaft-centered graph structure includes nodes that correspond to the individual components from the building models - walls, spaces, doors, and windows (Figure 3). Each node contained continuous geometric features like minimum and maximum x, y, z coordinates, area, length, height, and width, as well as intrinsic categorical features indicating the specific component type. The graph edges represent connection relationships between components based on the BIM layouts. Within the security shaft, each space is connected to the spaces above and below. Moreover, these spaces link with the associated walls, while the doors and windows are connected to the host wall. The walls of each space are linked with the aligned walls above and below that belong to security rooms on a level above and level below.

and the resulting graph representation was used to train a GNN model for node classification to categorize shaft components by vertical connection length ratio. To train the model, a set of 710 BIM models of shafts representing security rooms and other components that may be found in these shafts were generated and labeled manually using the Revit platform. To introduce realistic design variations into the models, the researchers received guidance from one of Israel's building control companies, in addition to reviewing several building control reports. One example of such a design variation is interconnected security shafts (where two security rooms share a wall) versus shafts comprising individual security rooms. Overall, the data set consists of 51,630 nodes and 78,811 edges, capturing the connections and relationships between the various components in the shafts.

The use of synthetic data comes in response to the great challenge of collecting relevant data, as industry stakeholders are usually reluctant to share BIM models. Hence, we follow the workflow suggested by (Bloch et al., 2023), where fully synthetic data was used for training a GNN model for ACC. It could be shown that the trained model had similar accuracy rates when applied for making predictions for real design.

In this case, the labels correspond to the three distinct classes characterized by the proportions of vertical connection ratios, as defined in the regulations. Namely, 70%, 50-70% range, and less than 50%. As illustrated in Figure 4, there are four types of elements represented as nodes in the graphs: spaces, walls, doors, and openings. Hence, we introduce an additional class that is related to "not applicable" (NA) nodes representing windows and doors.

The graph representations of these models were obtained using an automated workflow based on data extracted from the models to a spreadsheet. Since the synthetic data includes only individual security room shafts, the graph extraction process was different, but the resulting graph representation structure is identical to the structure obtained by the process depicted in Figure 4.

Using the described synthetic data set, a two-layer Graph Attention Network (GAT) model (Veličković et al., 2017) was trained to perform a node classification task. The training set was split into 70% for training, 15% for validation and 15% for testing. The testing phase provided a comprehensive assessment of the model's capabilities, even though the test phase was performed using a portion of the synthetic data set. The testing accuracy achieved 94.58% reflecting the model's proficiency in correctly classifying nodes within the graph. Precision, recall, and F1 score further validate the model's effectiveness, with values exceeding 94%. The described results are the outcome of iteratively evaluating multiple configurations of GNN models (such as GCN (Kipf and Welling, 2016) and SAGE (W. Hamilton et al., 2017)), model architectures and graph structures. Within the variations in graph structure, we evaluated graphs with different building elements to node mappings and with different feature vectors assigned to each node.

Table 1 illustrates the confusion matrix providing the results of predictions vs. actual labels of the test set. It provides insights into the model's classification behavior, with high counts along the diagonal indicating accurate predictions with a notable emphasis on correctly identifying all the 2025 instances in the "not applicable" class (class 4). For class 1 (>70%), the total count of nodes was 3,778, among these 3,650 instances were correctly predicted. However, 62 instances were misclassified as class 2 (50-70%) and additional 66 instances were misclassified as class 3 (<50%).

Table 1 Confusion matrix for synthetic test set

		Actual			
		>70%	**50-70%**	**<50%**	**NA**
Predicted	**>70%**	3650	62	66	0
	50-70%	196	449	19	0
	<50%	74	3	1201	0
	NA	0	0	0	2025

Class 2 (50-70%) showed the lowest accuracy, with 664 instances, out of which 449 instances correctly predicted, but 196 misclassified as class 1, and 19 misclassified as class 3. These misclassifications can be attributed to the imbalanced data set and insufficient number of examples covering these cases. Class 3 (<50%) predictions reached a better classification accuracy in comparison to class 2.

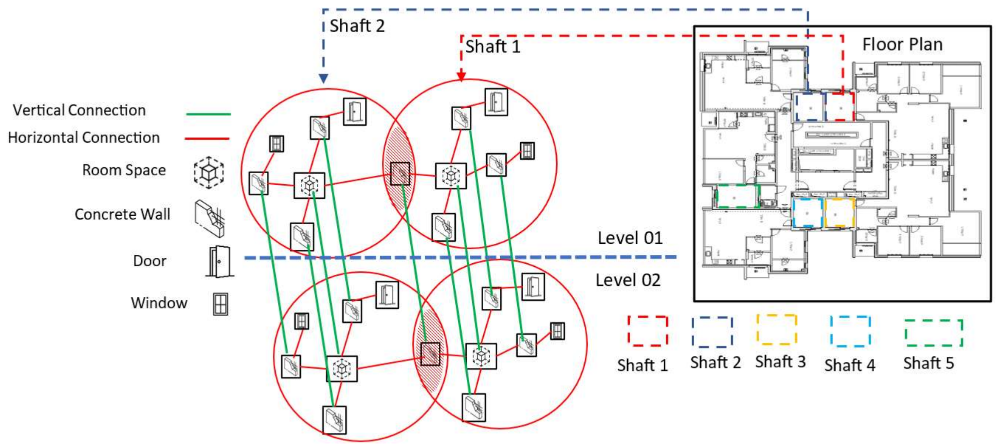


Figure 4: Graph representation of security shafts.

The total count of true class was 1278, out of which 1201 instances were correctly predicted. However, 74 instances were misclassified as Class 1, and an additional three instances were misclassified as Class 2. This also indicates the difficulty in identifying class 2 as it has high similarity to both class 1 and class 3. A better representation of this class can improve the results. In case the model is still not accurate enough in identifying class 2, we might think of a slightly different workflow that relies on distinguishing between the common case (>70% continuity) and the other cases.

After testing the model using a portion of the synthetic data, it was utilized to generate classifications on a real test case model. The real model is a residential building comprising 11 floors. Each floor accommodates five apartments, while the ground and first floors contain extra common and technical service areas. Additionally, the building includes two underground levels dedicated to parking, storage, and technical rooms. Each apartment in the building features a dedicated security room.

Notably, the model has five security shafts, where the first and second shafts are interconnected, as are the third and fourth shafts, via a shared separating wall (see Figure 4). The fifth shaft stands independently. Applying the classifier to the described real case, the model predictions achieved an accuracy of 86.84%. The confusion matrix for predictions is described in Table 2.

Table 2 Confusion matrix for the real case

		Actual			
		>70%	**50-70%**	**<50%**	**NA**
Predicted	**>70%**	208	6	6	0
	50-70%	9	21	9	0
	<50%	23	2	32	0
	NA	0	0	0	102

The complexity of the chosen test case is a result of aiming to maximize the available parking spaces, making the configuration of the walls on the lower levels of the building is not a standard case. While most of the elements follow the preferred guidelines of at least 70% vertical connection ratio, we can see representation of the other cases as well. It's noteworthy that the real-world instances include unique configurations that are not encountered and were not present in the training data.

Results of classification were injected back into the BIM model to be used for further checking requirements for reinforced concrete elements' thickness. Based on the predicted class for the spaces, a set of logical rules was inactivated to check the model for compliance with requirements for reinforced concrete element thickness. The overall checking process thus becomes "hybrid" by relying partially on graph-based and rule-based techniques. The supporting wall ratio for each space determines what requirements are applicable for the walls and slabs in the safety areas. Obviously, a misclassification of the elements by the GNN can lead to incorrect checking results. In this case, 90% of the security rooms in the model (nodes representing spaces) were classified correctly by the GNN model. The overall checking results based on this classification showed that out of a total of 265 walls and slabs checked, only 27 elements were incorrectly identified as "not compliant" while they truly are compliant with the code. Namely, 27 false positive cases. Zero false negative cases were identified. In the context of ACC, this distinction is very important since the non-compliant cases will have to be reevaluated by the designer\checker, but the false negatives may be approved despite not adhering to the code. ML models are not deterministic and may provide unreliable results. This limitation can be mitigated by further development of data sets, refinement of the models, and by constant validation against real cases. Additionally, by incorporating human oversight into the process, a more robust and reliable compliance-checking framework can be established.

Discussion and Conclusions

The extensively researched ACC domain may benefit from a significant shift in the overall approach. Hard-coded rules, being explicitly predefined and manually encoded, are often too restrictive. Furthermore, there is a tendency to confine the entire process to one single method, predominantly relying on rule-based systems. As a result, we miss opportunities for leveraging mutually beneficial methods that may lead to higher levels of automation in the overall process. Such approaches may incorporate learning techniques in addition to the human coded predetermined rules. The flexibility provided by the learning techniques is crucial in the field of ACC, where building regulations can be ambiguous or complex. One of the main limitations of such models is that they are not deterministic therefore, may be unreliable in the process of ACC. However, with large enough data sets and continuous exposure to diverse examples, these models can enhance their performance, making them well-suited for handling tasks from the code compliance checking domain. This work introduces a novel approach for ACC, leveraging semantic enrichment to enhance the level of automation that can be achieved. The approach is based on the capacity to decompose regulatory clauses into distinct semantic enrichment tasks, each addressed by a suitable solution. We demonstrate this approach with a test case, addressing specific requirements of Israeli regulations for security rooms. As described in this paper, most of the methods implemented in this case are based on graph theory, including community detection algorithms, and Graph Neural Networks. The introduced workflow is highly automated, with very little information requirements for the designers, showcasing the ability to reach a higher level of automation in the process by implementing diverse methods for every code clause. Although the training process includes manual work, it is not part of the checking as the model can be pretrained. The checking process itself relies on minimal data requirements and minimal manual BIM preprocessing, mainly making sure that the spaces have been created and that the security rooms are properly tagged.

Limitation

Several limitations are acknowledged in this study. Firstly, the transformation of the accessibility and connectivity relationships into graphs requires a well-defined BIM model, where space elements are comprehensively placed and correctly separated in the building layouts. While we aim not to add information requirements or specifications, good quality modeling practice is expected. The utilized building graphs, which are only parts of the various graph structures, are selected according to the topological features of the envisaged regulatory requirement. Thus, the applicability of the proposed graph-based enrichment for code checking on other regulations is to be justified.

Second, while we demonstrate the potential of GNNs for ACC, we acknowledge the limitation of the developed classifier, which is tailored to the context of the Israeli regulations for security rooms. Implementing a similar technique for different regulations would necessitate the collection or creation of new, context-specific synthetic datasets for training the models. Furthermore, the automated routines for data extraction must be modified to generate graphs with needed structures. Another limitation of using a synthetic data set for training is that real-world data may follow assumptions that are not well represented in the synthetic set. While we tried to limit this problem by working under the guidance of a building control company, our access to the design documents was limited. This underscores the importance of future efforts to address the scalability of the approach to accommodate a broader spectrum of regulations, including generating graph datasets tailored to support various requirements. The demonstrated promising model performance across various evaluation metrics is encouraging to continue developing this research direction.

Contribution

This work contributes to the ACC domain, addressing several existing challenges and paving the way for a more automated and adaptable approach to code compliance checking. The proposed approach for code checking, which leverages graph representation, introduces a novel way of thinking about ACC processes, where we do not have to be confined solely to rules. This approach facilitates a more comprehensive investigation of design knowledge. It has the capacity to elucidate not only the instances of non-compliance but also to subsequently link other elements that may be interconnected with these issues. Such enriched results empower designers with detailed insights, enabling more effective resolution strategies (Wu et al., 2023). This approach lays the groundwork for the advanced automated design adaptation method known as "Design Healing." By providing a more nuanced understanding of the design and its compliance with building codes, designers can leverage this foundational data to make informed adjustments, enhancing the overall efficacy and compliance of the building model.

This work lays the groundwork for future research on automated code checking, encouraging the exploration of hybrid models that harness the strengths of diverse computational techniques for enhanced regulatory compliance. The proposed workflow is fully automated, with minimal information requirements to be supplemented by users. The success of the approach underscores the potential of semantic enrichment to enhance the level of automation achieved in ACC.

Acknowledgments

The research was funded by the Israeli Ministry of Construction and Housing. The presented research also received funding from the TUM International Graduate School of Science and Engineering (IGSSE).

References

Abualdenien, J., Borrmann, A., 2021. PBG: A parametric building graph capturing and transferring detailing patterns of building models, in: Proc. of the CIB W78 Conference 2021.

Autodesk, 2022. Dynamo.

Belsky, M., Sacks, R., Brilakis, I., 2016. Semantic enrichment for building information modeling. Comput.-Aided Civ. Infrastruct. Eng. 31, 261–274. https://doi.org/10.1111/mice.12128

Bloch, T., Borrmann, A., Pauwels, P., 2023. Graph-based learning for automated code checking – Exploring the application of graph neural networks for design review. Adv. Eng. Inform. 58, 102137. https://doi.org/10.1016/j.aei.2023.102137

Bloch, T., Sacks, R., 2020. Clustering Information Types for Semantic Enrichment of Building Information Models to Support Automated Code Compliance Checking. J. Comput. Civ. Eng. 34, 04020040.

Bloch, T., Sacks, R., 2018. Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models. Autom. Constr. 91, 256–272. https://doi.org/10.1016/j.autcon.2018.03.018

Borrmann, A., Van Treeck, C., Rank, E., 2006. Towards a 3D spatial query language for building information models, in: Proc. Joint Int. Conf. of Computing and Decision Making in Civil and Building Engineering (ICCCBE-XI).

Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486, 75–174. https://doi.org/10.1016/j.physrep.2009.11.002

Girvan, M., Newman, M.E.J., 2002. Community structure in social and biological networks. Proc. Natl. Acad. Sci. 99, 7821–7826. https://doi.org/10.1073/pnas.122653799

Hamilton, W., Ying, Z., Leskovec, J., 2017. Inductive representation learning on large graphs. Adv. Neural Inf. Process. Syst. 30.

Hamilton, W.L., Ying, R., Leskovec, J., 2017. Representation Learning on Graphs: Methods and Applications. CoRR abs/1709.05584.

Home Front Command, 2010. Specifications for Building Shelters. Protective structures department, Home Front Command, Ramle, Israel.

Hu, Y., Xia, C., Chen, J., Gao, X., 2023. Clash context representation and change component prediction based on graph convolutional network in MEP disciplines. Adv. Eng. Inform. 55, 101896. https://doi.org/10.1016/j.aei.2023.101896

Kipf, T.N., Welling, M., 2016. Semi-supervised classification with graph convolutional networks. ArXiv Prepr. ArXiv160902907.

Mazairac, W., Beetz, J., 2013. BIMQL An open query language for building information models. Adv. Eng. Inform. 27, 444–456. https://doi.org/10.1016/j.aei.2013.06.001

Nawari, 2019. A Generalized Adaptive Framework (GAF) for Automating Code Compliance Checking. Buildings 9, 86. https://doi.org/10.3390/buildings9040086

Solihin, W., Eastman, C., 2015. Classification of rules for automated BIM rule checking development. Autom. Constr. 53, 69–82. https://doi.org/10.1016/j.autcon.2015.03.003

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P., Bengio, Y., 2017. Graph attention networks. ArXiv Prepr. ArXiv171010903.

Vilgertshofer, S., Borrmann, A., 2017. Using graph rewriting methods for the semi-automatic generation of parametric infrastructure models. Adv. Eng. Inform. 33, 502–515. https://doi.org/10.1016/j.aei.2017.07.003

Wang, Z., Sacks, R., Yeung, T., 2022. Exploring graph neural networks for semantic enrichment: Room type classification. Autom. Constr. 134, 104039. https://doi.org/10.1016/j.autcon.2021.104039

Wu, J., Dubey, R.K., Abualdenien, J., Borrmann, A., 2023. Model Healing: Toward a framework for building designs to achieve code compliance, in: ECPPM 2022-eWork and eBusiness in Architecture, Engineering and Construction 2022. CRC Press, pp. 450–457.

Wu, J., Zhang, J., 2019. New automated BIM object classification method to support BIM interoperability. J. Comput. Civ. Eng. 33, 04019033.

Wülfing, A., Windisch, R., Scherer, R., 2014. A visual BIM query language, in: eWork and eBusiness in Architecture, Engineering and Construction: ECPPM 2014. p. 157.

Zhang, J., El-Gohary, N.M., 2017. Integrating semantic NLP and logic reasoning into a unified system for fully-automated code checking. Autom. Constr. 73, 45–57. https://doi.org/10.1016/j.autcon.2016.08.027

NLP-BASED DATA-ENRICHMENT FOR BUILDING MANAGEMENT

Hamada Elshaboury[1], Fulvio Re Cecconi[2], Enrico De Angelis[3], Luciano Baresi[4], Vincenzo Scotti[5]
[1,2,3]ABC Department, Politecnico di Milano, Milan, Italy
[4,5]DEIB Department, Politecnico di Milano, Milan, Italy

Abstract

One of the main challenges in building management is dealing with the large amount of unstructured data produced throughout the asset's life cycle. At handover, Building Information Models often provide low-quality, incomplete data, necessitating extensive manual rework to elicit information from many sources. To minimize this manual rework, this study proposes an automated *Information Extraction* (IE) procedure, applied to the design and construction documents to extract information, enrich a model (COBie format), and maximize the transfer of structured data to the client. The proposed approach is based on *Natural Language Processing (NLP)* and adopts *Transformer-based Named Entity Recognition* (NER) and *Relation Extraction* (RE) methods for IE. It was evaluated and achieved good performance, with average F1 scores of 0.73 for NER and 0.91 for RE, representing a step toward a reliable tool for an enhanced data handover process.

Introduction

During the design and construction phase, a significant number of text documents are produced, not only by the design team but also by the contractor, suppliers, client, and consultants. Design documents and specifications are followed by building performance reports (structural safety, energy, acoustics...), contractor's method statements, risk assessments, health, safety, and maintenance plans, inspection and commissioning reports, product sheets and declarations, etc. However, all these documents contain poorly accessible, unstructured information that can only be retrieved through manual searches (Marzouk and Enaba, 2019).

Building Information Modelling (BIM) has emerged as the leading digital solution in the Architecture, Engineering, and Construction (AEC) industry to support the design, construction, and subsequent management phases, as facility managers can use the same physical and functional information defined in the early phases to plan and manage maintenance operations. However, the actual situation is not ideal. On the one hand, digital models are often unable to easily incorporate all the data produced (also because of the low digital readiness of stakeholders and their models). On the other hand, the information created is often incomplete, inaccurate, inconsistent, complex, large, and poorly standardized. As a result, BIM, as an enabler for providing reliable information for building management, faces numerous challenges (Tsay *et al.*, 2022; Ullah *et al.*, 2019).

In response to these issues, US agencies, with the help of the National Institute of Building Sciences, have developed a standard digital data schema known as Construction Operations Building Information Exchange (COBie) to streamline access and manipulation of BIM data. A COBie file can be exported as an Excel spreadsheet or IFC STEP file from any BIM model. However, a COBie file reflects the quality of the BIM model from which it was generated. In addition, some of the information in the BIM model dataset is not exported in the COBie file (Kumar and Teo, 2020), increasing the manual effort required to have complete information to support the management of maintenance and operations planning.

Recognizing these challenges, this study uses NLP to enhance the handover process, as shown in Figure 1, by developing a methodology in three steps: data preparation (1), information extraction (2), and integration into COBie (3).

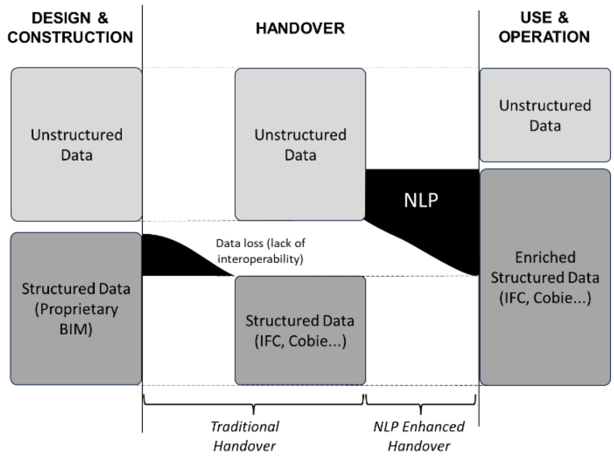


Figure 1: Handover problem statement and proposed solution

This paper is organized as follows: an introductory *State of the Art*, a *Methodology* section detailing the proposed approach, a third section for *Experiments*, and, eventually, a *Conclusion* and *References* section.

State of the Art

NLP and Data Management in the AEC Industry

NLP joins linguistics with computer science, empowering our lives with tools and techniques to perform machine reading, comprehending, and analyzing human-written documents (Salama and El-Gohary, 2013). As a result of these capabilities, NLP has been widely applied in a highly documented sector as the AEC industry (Marzouk and Enaba, 2019) to improve the management of many processes such as safety (Zhang *et al.*, 2019), risk assessment (Zou *et al.*, 2017), quality control (Jeon *et al.*, 2021), document management (Qady and Kandil, 2014), and code compliance checking (Zhang and El-Gohary, 2013; Zhang and El-Gohary, 2015).

The cornerstone of improving construction management is built on the efficiency and effectiveness of data management through transforming unstructured texts into structured information, enabling organizing and analyzing information, and identifying patterns supporting the enhancement of planning and decision-making. This transformation from unstructured data into structured information, as shown in Figure 2, is executed through a suite of NLP tools, such as tokenization, part-of-speech (POS) tagging, phrase structure analysis (PSG), stemming and integrating with machine learning (ML) and deep learning (DL) techniques to perform advanced tasks such as Text Classification (TC), Document Clustering (DC), Information Extraction (IE), and Information Retrieval (IR) to support data management.

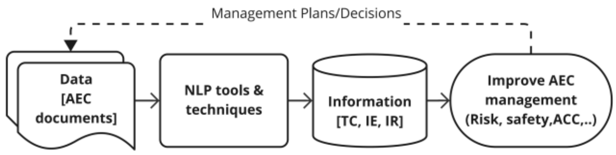


Figure 2: Methodology of NLP application in the AEC industry

Information Extraction (IE)

IE is a process that aims to automatically extract specific information from unstructured data and represent such information in a structured format (Salama and El-Gohary, 2013). IE includes sub-tasks such as *Named Entity Recognition* (NER), which recognizes and classifies entities into predefined categories, and *Relation Extraction* (RE), which extracts the semantic relations between these entities (Zhang and El-Gohary, 2013). IE methods include rule-based, machine learning (ML)-based, and deep learning (DL)-based approaches.

In rule-based approaches, the target information is extracted based on pattern-matching rules coded by experts. These rules utilize NLP tools such as POS and PSG for sentence analysis, term matching, and semantic analysis (Zhang and El-Gohary, 2015). For example, (Zhang and El-Gohary, 2013) proposed a semantic and rule-based approach to extract information from building codes to support compliance checking. Also, (Li *et al.*, 2016) developed a rule-based approach to extract spatial rules and check compliance of utility spatial specifications with requirements. (Liu and El-Gohary, 2021) proposed a dependency parsing model to extract dependency relations between the semantic information elements from bridge inspection reports and represent them in semantic information sets. Although rule-based approaches have low scalability and flexibility and require manual efforts for pattern formalization, they have high accuracy.

In ML-based approaches, models rely on learning from data (Russell and Norvig). For example, (Zhang and El-Gohary, 2016) developed an approach to classify relations between BIM and regulations concepts using ML classifiers (SMV, NB, DT, K-NN). ML-based approaches are more flexible for modifications, less costly, and require fewer manual efforts than rules-based approaches. However, manual work is still necessary for preprocessing and preparing datasets for training and testing.

DL-based models comprise multiple neural network layers from various algorithms, allowing them to deeply understand and represent complex and unstructured data for NLP tasks (Lecun *et al.*, 2015). Recently, sequence-to-sequence models such as Recurrent Neural Networks (RNNs), LSTM and RGU, along with Convolutional Neural Networks (CNNs), have been used extensively for IE. For example, (Zhong *et al.*, 2020) proposed an approach to support quality management checking using a Bi-LSTM-CRF-based NER model to identify and label entities in the clauses and an LSTM-MLP-based RE model to classify relations between entities into five predefined groups. (Zhang and El-Gohary, 2022) used a Bi-LSTM-MLP-based RE model to extract the semantic relations of building code sentences. (Wang and El-Gohary, 2023) proposed a CNN and RNN-based method to extract relations between entities that describe fall protection requirements from safety regulations and represent entities with their relations in query graphs.

Most recently, transformer-based models have gradually become a trend, utilizing multi-headed self-attention mechanisms to deeply understand text context and representation (Vaswani *et al.,* 2017) and enabling the development of pre-trained models (e.g., BERT) (Devlin *et al.*, 2018). For example, (Zhang and El-Gohary, 2023) used a BERT–based model to determine the semantic relation probability between regulatory and IFC concepts. DL-based approaches are more effective in handling various data types that exhibit high levels of complexity, achieving a high semantic analysis and word representation.

Data Handover for Building Management

There is a paramount interest in using project data for the life cycle with the evolution of Building Information Modelling, which promotes the incremental collection of data (Lindkvist and Whyte, 2013). Nevertheless, limited studies have focused on effective and automated building information sharing during the commissioning and close-out stage for efficient building handover, operations, and maintenance through its life cycle (Singh and Anumba, 2023). As a result of this lack of research in the field, the automated digital practices to perform the commissioning and handover process are still largely unknown.

Many studies (Cavka *et al.*, 2017; Tan *et al.*, 2018) collectively underscore the importance of knowledge trans-

fer, information flow, and the use of technology in enhancing information management during handover for better asset and facility management. COBie, a specification for the Construction Operations Building Information Exchange, aims to streamline data handover from design and construction to facility management (Schwabe *et al.*, 2018). It is a Model View Definition (MVD) of the Industry Foundation Classes (IFC) as defined in ISO (ISO 16739-1:2018) and was first introduced by the US Army Corps of Engineers to enable organizations to electronically capture and record important project data at the point of origin. Because the fundamental purpose of COBie is to efficiently obtain information (relevant to facility management) generated in the design and construction phases, an Excel-based spreadsheet that anyone can easily access and handle is often used (Shin *et al.*, 2022). A COBie datasheet consists of 20 workbooks in which data about the facility is stored systematically. The columns and their location in the workbooks are fixed, and changing the location of columns is restricted. Different colors are used to define data inside workbooks (Kumar and Teo, 2020). Although COBie has been widely used (Maltese *et al.*, 2017)(Kumar and Teo, 2020) (Asare *et al.*, 2023) in practice, creating COBie deliverables can be problematic due to misunderstandings among end users and insufficient software implementation. This can ultimately lower acceptance among practitioners.
COBie data capturing and compiling is a complex process; after each project phase ends, a COBie sheet deliverable is required, which needs to be verified with the project stage requirements known as COBie "data drops" (Love *et al.*, 2014). A typical data drop process requires, among other activities, extracting IFC models from native BIM models from different disciplines, merging and verifying for consistency, updating missing information through documents, and verifying that the dependency and links are maintained. This implies that data needs to be entered manually inside the COBie datasheet. The transfer of information from as-built documents is one of the biggest problems, as much of the information is not transferred in COBie format or is transferred with errors. Natural Language Processing (NLP) tools can improve the information management process by increasing the amount of structured data transferred to the client via COBie.

Methodology

In this section, we outline the proposed data enrichment methodology, which comprises three steps, as depicted in Figure 3: data preparation, information extraction, and integration into COBie. Recently, NLP has offered many DL-based tools to automate text analysis and process large batches of documents, so the proposed methodology exploits NLP techniques for language analysis to automate IE and integration.

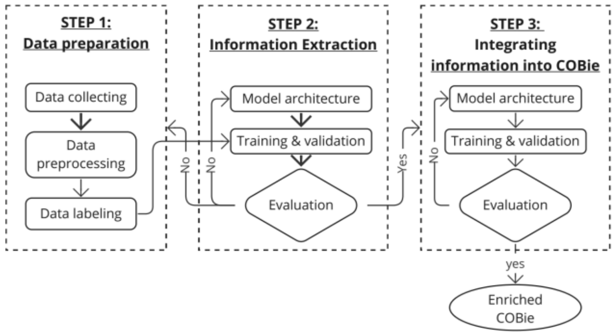


Figure 3: Methodology framework

Step 1: Data Preparation

Given the use of deep learning, a data-driven technique, ground truth data is required to train and evaluate the proposed models. This step, which includes data collection, preprocessing, and labeling, is crucial for the model's development.
For data collection, a corpus of unstructured documents for design and construction was collected from case studies provided by one of Italy's largest engineering companies in Milan and converted into plain text format for easier processing.

Table 1: Categories of the Information Entities

Entity Category	Definition
Object ID	Identify the ID code of an entity "Wall 01"
Object	Identify an entity "internal wall"
Description	Clarify a description for an entity, action, attribute
Section	Identify the geographical location of an object "north-facing windows"
Action	Identify an activity for an object "demolishing", "building", "painting"
Attribute	Identify a character that distinguishes an object's "thickness," "airtightness," and "or strength."
Value	Identify value "15", a class "U", or category "GBK(A)" for an attribute.
Unit	Identify a unit measure for attribute value "cm", or "REI" for fire resistance.
Constraint	Identify a specific condition for an attribute or value or action "equal to", "in a range", "partially, completely, "fully".
Reference	Identify a regulation, standard, or code explaining how the attribute must be evaluated or measured and the object or action must be realized.

After collecting the raw documents, data preprocessing was conducted to remove non-textual parts like figures, headers, and footers, and sentence segmentation was applied to isolate the individual sentences. Labeling involves annotating a specific word or more with the corresponding label or the relation between two entities. This process begins with defining labels for the target information entities and relation types the proposed models intend to recognize and extract. In this research, the defined categories for the information entities and their relation types are presented in Tables 1 and 2 and illustrated in Figure 4.

Table 2: Relation types between entities

Relation Type	Definition
Identifies	Links the unique identifier code (*Object ID)* to the corresponding *Object*
PartOf	Links an *Object* to another *Object*, establishing a parent-child relationship and the hierarchical structure
DescribedBy	Links an *Object* or *Action* to a *Description* that provides more details
LocatedIn	Links an *Object* to a specific geographical location (*Section*) in a building
RequireAction	Links an *Object* to an *Action* that specifics the activities or tasks to be realized
HasAttribute	Links an *Object* to an *Attribute* that specifies its characteristics
HasValue	Links an *Attribute* with a specific *Value* for its characteristics
MeasuredIn	Links a *Value* to a *Unit*, specifying the unit of measurement for the value
HasConstraint	Links an *Action* or *Value* to a *Constraint*, specifying the conditions for this value or action.
Referenced	Links either an *Object*, *Action*, or *Value* to a *Reference* establish the connection between a regulation, standard, or code and its application to an object, action, or value.

An open-source data labeling platform, "Label Studio" was used to facilitate manual labeling, applying two distinct notations. For the NER task, the Beginning-Inside-Outside (BIO) labeling schema was used, where each word in a sentence was labeled manually to indicate whether it was the beginning or inside of a specific entity or not a part of any entity (outside). In the RE task, a single relation type was assigned between each input pair of entities using the tagging schema of the SemEval 2010-Task 8 datasets to mark the entity's positions with <e1> and </e2> XML-like tags, as shown in Figure 6.b.

Step 2: Information Extraction (IE)

In this study, we introduce a modular IE pipeline comprising two respective tasks: 1) *Named Entity Recognition* (NER) to process a plain text document or a specific segment and extract all entities that exist in the text according to Table 1; and 2) *Relation Extraction* (RE) to extract and classify the relation between each pair of entities following Table 2. By iteratively applying these two steps across the entire document corpus, we methodically construct a comprehensive and structured knowledge base that encapsulates atomic entities and delineates their interrelations, as shown in Figure 5. This step consists of three sub-steps: model development, training, and evaluation, as discussed below.

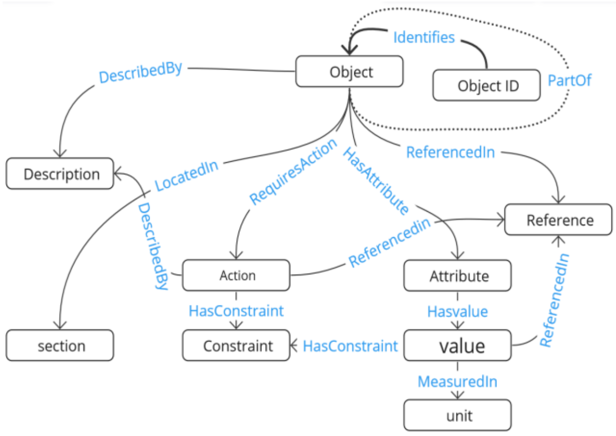


Figure 4: Entity categories and Relation types

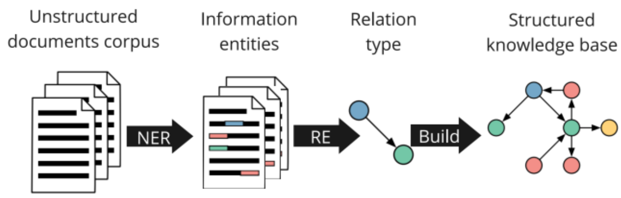


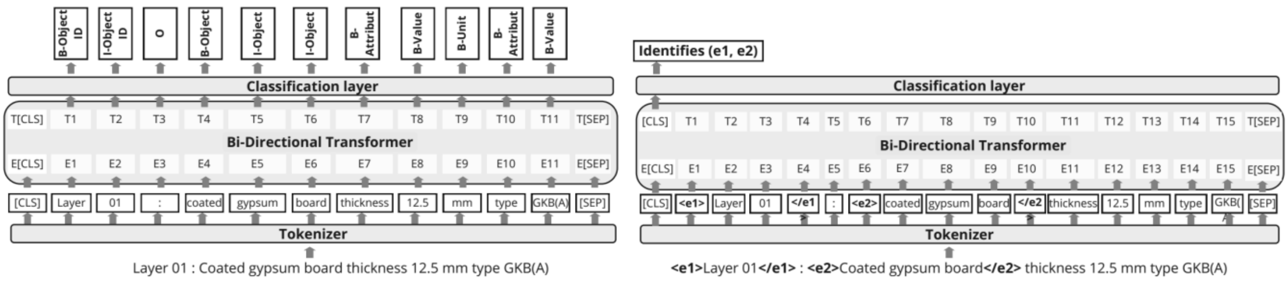
Figure 5: Information Extraction (IE) pipeline

Model development

Following current state-of-the-art NLP applications built using a transformer neural network, pre-trained language models were selected as the base for the NER and RE tasks composing the pipeline. As a classification task, a bi-encoder architecture is more suitable where the model architecture consists of an input, encoding, and output layer, as shown in Figure 6 and described below:
Input (embedding) Layer: converts each word in the input text to a vector representation by tokenizing sentences, padding, converting tokens to IDs, adding an attention mask and two special tokens [CLS] and [SEP] to determine the starting and ending of each sentence.
Encoding Layer: feeds the embedding vectors of tokens from the input layer into transformer encoder blocks passing with the multi-head attention followed by the feed-forward network, with add & normalize layers, and the output of each block is passed to the next one as input. As a result, the final output from the last block represents a highly contextual embedding vector (hidden state), considering both the left and right contexts, semantics, and synthetic relationships.
Output (classification) layer: takes the final hidden states from the last layer as input, applying a SoftMax function to calculate probabilities.
In the NER task, as depicted in Figure 6. a, given a piece of text passing through the input and encoding layers, the

language model computes a probability distribution over the class labels for each input token and selects the most probable class label associated with each token. The class label indicates whether the token corresponds to the beginning, inside, or outside of a specific entity. Once a text is tagged, named entities are extracted, searching for the beginning of each entity and taking that first token with all the inside tokens associated with that same entity.

In the RE task, as illustrated in Figure 6. b, the input piece of text has four additional tokens to mark the beginning and end of the first and second entities; passing through the input and encoding layers, the language model computes a probability distribution over the class labels and selects the most probable class label associated with the input text to tag the relation type between the two marked entities.



a. Transformer-based NER model

b. Transformer-based RE model

Figure 6: Architecture of the language model for the NER and RE tasks

Model training

As pre-trained language models on a large general domain corpus, there is a need to fine-tune these models for specific downstream tasks (i.e., NER and RE). The objective is to minimize the negative log-likelihood of the target class(es) associated with the input text. Cross-entropy, serving as a cost function, calculates the discrepancy between the actual and predicted labels/relations, assessed after the model's forward pass; during the backpropagation, the optimizer updates the model weights based on the gradient of the loss and learning rate scheduler.

Evaluation

In the realm of IE, there are three common metrics for performance evaluation: Precision (P), Recall (R), and F1-score (Zhai and Massung, 2016), as delineated in Equations 1 to 3, with their values ranging from 0 to 1, where 0 is the worst score, and 1 is the best. For instance, R = 1 signifies that the model correctly identified all actual positives, P = 1 indicates that all positive predictions were correct, and F1 = 1 indicates a perfect harmonic balance between R and P.

$$Recall\ (R)\ = \frac{TP}{TP+FN} \quad (1)$$

$$Precision\ (P)\ = \frac{TP}{TP+FP} \quad (2)$$

$$F1 = 2\ x\ \frac{P\ x\ R}{P+R} \quad (3)$$

Step 3: Information integration into COBie

In this step, we merge the extracted information with the COBie format, where the relations, as depicted in Table 2 , establish interactive connections and contextual associations between the entities, creating a semantic-structured knowledge base to support the integration process. To achieve this, two models will be developed: 1) a Transformer-based model to classify the extracted entities from Step 2 into COBie sheets, 2) a Rule-based model to match the entities with COBie sheets, and embed information into COBie data fields. Further details and experiments on this step will be pursued in future research directions.

An example to illustrate the application of the proposed methodology to enrich the COBie schema is shown in Figure 7.

Experiments

In this section, we describe the implementation of the IE pipeline, training settings, and evaluation results.

Data preparation

Wall specification documents from the collected corpus were used as a case study to prepare the dataset to train and test the models. These documents include detailed information regarding wall layers, materials, characteristics, and state-of-the-art construction procedures. Following the procedure outlined in Step 1, a dataset was developed encompassing approximately 2400 entities and 1000 relations representing all entity categories and relation types in Figure 4.

Implementation

The proposed IE pipeline was implemented using two prominent bidirectional language models, BERT (Devlin *et al.*, 2018) and RoBERTa (Conneau *et al.*, 2019), from the Transformers library in the Hugging Face model hub. Table 3 Provides the considered models' variants and reference names; some were pre-trained on multiple languages, while others were pre-trained on Italian. The implementation was conducted using PyTorch, built with Python 3, and executed on a T4 GPU provided by Google Colaboratory.

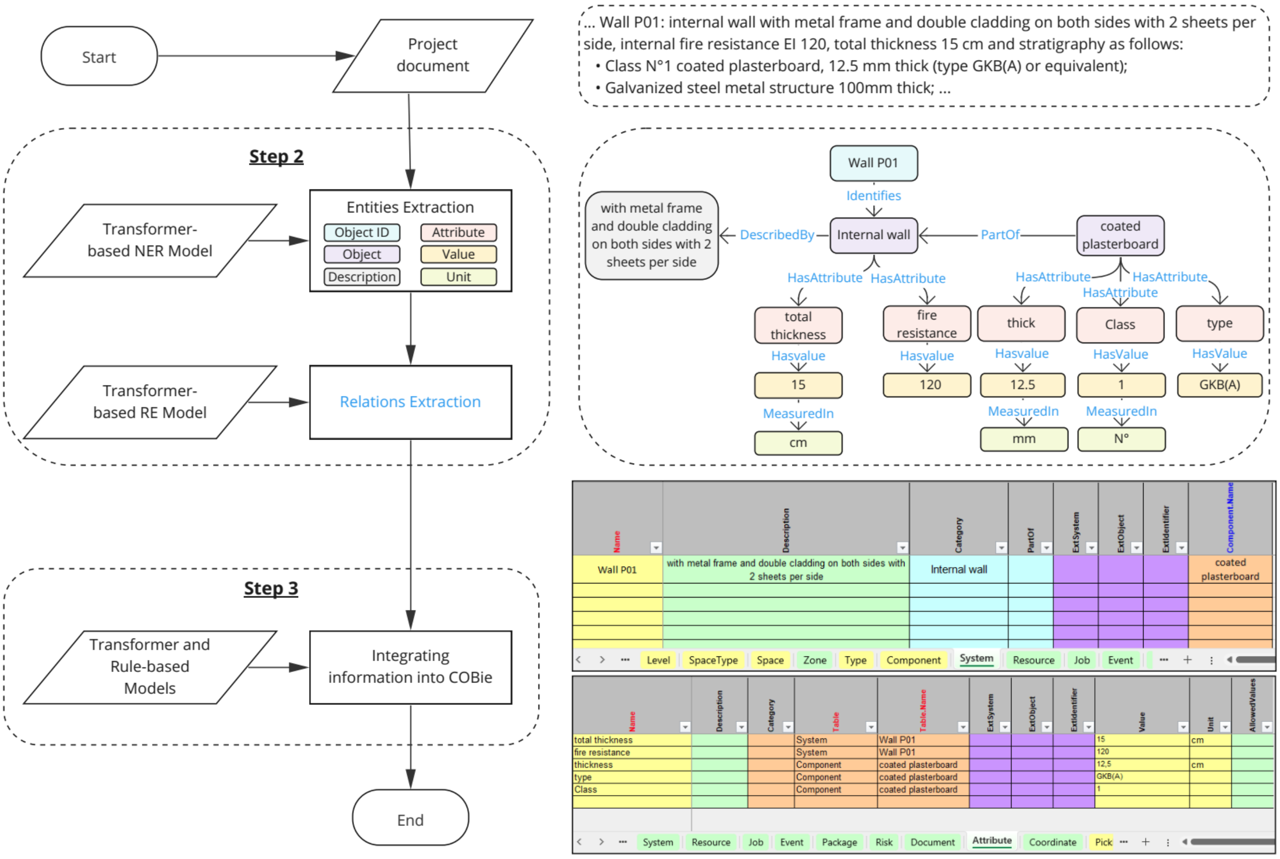


Figure 7: Enriche COBie schema using the proposed methodology

Table 3: Variants of BERT and RoBERTa models for the NER and RE tasks

Reference Name	No. of Layers	Parameters
Multilingual Bert (Cased - Uncased)	12 layers	110M
XLM-Roberta Large	24 layers	550M
Italian ALBERTO (Uncased) (Polignano *et al.*, 2019)	12 layers	110M
dbmdz-Italian Bert -XXL (Cased - Uncased)	12 layers	110M

Training

The language models were fine-tuned using the labeled dataset, which was split into training, validation, and testing sets with 80%,10%, and 10%, respectively, for the NER task and 64%, 12%, and 24%, respectively for the RE task. The configuration and hyperparameters stayed the same for the models, as provided in Table 4.

Evaluation and results

The evaluation was conducted using the testing dataset alongside selected transformer-based models. Table 5 Provides the results of the experiments where the "*xlm-roberta-large*" model emerges as the most effective in the IE pipeline, achieving the highest precision, recall, and F1 scores for both tasks. It achieved F1 scores of 0.73 for NER and 0.91 for RE, respectively, outperforming other models with an average increase of 8.8% for NER and 5.2% for RE tasks.

Table 4: Values of the hyperparameters of the language models

Hyperparameter	Value
Maximum length	128
Batch size of data loader	32
Adam learning rate	1e-5
Dropout Rate	0.1
Early stopping condition	5 epochs

Regarding the results of the RoBERTa model, the model indicated a good capability in identifying entities "Object", "Value", "Attribute", and "Unit", achieving relatively high F1-scores ranging from 0.75 to 0.90; conversely, it encountered challenges in recognizing entities "Description", "Action", and "Constraint" as reflected by lower F1-scores from 0.33 to 0.62. For the RE task, the model demonstrated a robust performance across various relation types achieving high F1-scores ranging from 0.88

to 0.97 for relations "HasValue","DescribedBy","HasAttribute","Partof","ReferencedIn", and "MeasuredIn" and moderate F1- scores from 0.67 to 0.78 for "LocatedIn", "RequiresAction", and "ConstrainedBy" ones. The limitations in the model's performance in identifying and recognizing these entities or relation types can be attributed to restricting their examples in the training dataset. Consequently, the model encountered interpretation complexities when distinguishing within the text. These results are satisfactory given the current volume of the dataset and underscore the potential for enhancing the performance once more data becomes available.

Table 5: Results of the proposed IE pipeline (NER and RE tasks) with the language models

Model	NER			RE		
	P	R	F1	P	R	F1
Bert-base-multilingual-cased	0.63	0.69	0.64	0.89	0.89	0.89
Bert-base-multilingual uncased	0.64	0.64	0.62	0.83	0.84	0.83
xlm-roberta-large	**0.74**	**0.74**	**0.73**	**0.91**	**0.91**	**0.91**
Bert_uncased_ italian_alb3rt0	0.69	0.68	0.67	0.86	0.86	0.86
dbmdz/bert-base-italian-xxl-cased	0.64	0.68	0.63	0.87	0.86	0.86
dbmdz/bert-base-italian-xxl-uncased	0.67	0.69	0.65	0.85	0.86	0.85

Conclusion

This study contributes to building management by enhancing information management during the handover process, enabling effective planning and management of maintenance operations. It proposed a methodology aiming to provide an enriched data schema (COBie) that integrates information from both the design and construction phases with less manual effort and time consumption, overcoming some of the prevalent challenges, such as data interoperability issues and the high level of documentation in the construction sector. The methodology comprised three main steps: data preparation, information extraction (IE), and integration into the COBie format. Entity categories and relation types were identified to prepare the labeled data, and two transformer-based models, BERT and RoBERTa, with their variants, were utilized to implement the proposed IE pipeline, automating *Named Entity Recognition* (NER) and *Relation Extraction* (RE) tasks. The proposed IE pipeline was tested using wall specification documents as a case study, and the "*xlm-roberta-large*" language model delivered the highest performance on both tasks, achieving an average F1-score of 0.74 in the NER task and 0.91 in the RE. These scores can be considered good results given the current data volume.

Future research directions include: 1) Implementing Step 3 to integrate the extracted entities and their relations into the COBie format; 2) Expanding the labeled dataset by considering a broader range of design and construction documents to improve the models' performance, scalability, and robustness; 3) Large Language Models (LLMs) will be used for the NER and RE tasks, and evaluated for their effectiveness in the given context, and 4) Performing a comparison to a COBie data extraction tool.

References

Asare, K.A.B., Liu, R. and Anumba, C.J. (2023). Building information modeling for airport facility management: the case of a US airport, Thomas Telford.

Cavka, H.B., Staub-French, S. and Poirier, E.A. (2017). Developing owner information requirements for BIM-enabled project delivery and asset management. *Automation in Construction*, 83:169–183.

Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., Grave, E., Ott, M., Zettlemoyer, L., & Stoyanov, V. (2019). Unsupervised Cross-lingual Representation Learning at Scale. Proceedings of the Annual Meeting of the Association for Computational Linguistics, 8440–8451.

Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *NAACL HLT 2019,* 1:4171–4186.

ISO 16739-1:2018 - Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema.

Jeon, J., Xu, X., Zhang, Y., Yang, L. and Cai, H. (2021). Extraction of Construction Quality Requirements from Textual Specifications via Natural Language Processing. SAGE Publications, 2675:222–237.

Kumar, V. and Teo, E.A.L. (2020). Perceived benefits and issues associated with COBie datasheet handling in the construction industry. *Facilities*, 39:321–349.

Lecun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. Nature Publishing, 521: 436–444.

Li, S., Cai, H., Asce, M. and Kamat, V.R. (2016). Integrating Natural Language Processing and Spatial Reasoning for Utility Compliance Checking. *Journal of Constr. Engineering and Management*, 142:04016074.

Lindkvist, C. and Whyte, J. (2013). Challenges and Opportunities Involving Facilities Management in Data Handover: London 2012 Case Study. *AEI 2013*, 670–679.

Liu, K. and El-Gohary, N. (2021). Semantic Neural Network Ensemble for Automated Dependency Relation Extraction from Bridge Inspection Reports. *Journal of Computing in Civil Engineering*, 35:04021007.

Love, P.E.D., Matthews, J., Simpson, I., Hill, A. and Olatunji, O.A. (2014). A benefits realization management building information modeling framework for asset owners. *Automation in Construction*, 37:1–10.

Maltese, S., Moretti, N., Re Cecconi, F., Ciribini, A.L.C. and Kamara, J.M. (2017). A Lean Approach to Enable Sustainability in the Built Environment through BIM. *Journal of Technology for Architecture and Environment*, 13:278–286.

Marzouk, M. and Enaba, M. (2019). Text analytics to analyze and monitor construction project contract and correspondence. *Auto. in Construction*, 98:265–274.

Polignano, M., Basile, P., de Gemmis, M., Semeraro, G., & Basile, V. (2019). AlBERTo: Italian BERT language understanding model for NLP challenging tasks based on tweets. *CEUR Workshop Proceedings*, 2481.

Qady, M. and Kandil, A. (2014). Automatic clustering of construction project documents based on textual similarity. *Automation in Construction*, 42:36–49.

Russell, S. and Norvig, P. (n.d.). Artificial Intelligence-A Modern Approach (3rd Edition).

Salama, D.M. and El-Gohary, N.M. (2013). Semantic Text Classification for Supporting Automated Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, 30: 04014106.

Schwabe, K., Dichtl, M., König, M. and Koch, C. (2018). COBie: A Specification for the Construction Operations Building Information Exchange. *Building Information Modeling,* 167–180.

Shin, S., Moon, H. and Shin, J. (2022). BIM-Based Maintenance Data Processing Mechanism through COBie Standard Development for Port Facility. *Applied Sciences,* 12:1304.

Singh, J. and Anumba, C.J. (2023). Building commissioning process and documentation: a literature review and directions for future research. *International Journal of Construction Management*.

Tan, A.Z.T., Zaman, A. and Sutrisna, M. (2018). Enabling an effective knowledge and information flow between the phases of building construction and facilities management. *Facilities*, 36:151–170.

Tsay, G.S., Staub-French, S. and Poirier, É. (2022). BIM for Facilities Management: An Investigation into the Asset Information Delivery Process and the Associated Challenges. *Applied Sciences,* 12:9542.

Ullah, K., Lill, I. and Witt, E. (2019). An overview of BIM adoption in the construction industry: Benefits and barriers. *Emerald Reach Proceedings Series*, 2:297–303.

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). Attention Is All You Need. Advances in Neural Information Processing Systems, 5999–6009.

Wang, X. and El-Gohary, N. (2023). Deep learning-based relation extraction and knowledge graph-based representation of construction safety requirements. *Automation in Construction*, 47:104696.

Zhai, C. and Massung, S. (2016). *Text Data Management and Analysis: A Practical Introduction to Information Retrieval and Text Mining*.

Zhang, F., Fleyeh, H., Wang, X. and Lu, M. (2019). Construction site accident analysis using text mining and natural language processing techniques. *Automation in Construction*, 99:238–248.

Zhang, J and El-Gohary, N.M. (2015). Automated Information Transformation for Automated Regulatory Compliance Checking in Construction. *Journal of Computing in Civil Engineering*, 29(4).

Zhang, J. and El-Gohary, N.M. (2016). Extending Building Information Models Semiautomatically Using Semantic Natural Language Processing Techniques. *Journal of Computing in Civil Engineering,* 30(5).

Zhang, J. and El-Gohary, N.M. (2013). Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking. *Journal of Computing in Civil Engineering*, 30(2).

Zhang, R. and El-Gohary, N.M. (2022). Hierarchical Representation and Deep Learning–Based Method for Automatically Transforming Textual Building Codes into Semantic Computable Requirements. *Journal of Computing in Civil Engineering*, 36(5).

Zhang, R. and El-Gohary, N. (2023). Transformer-based approach for automated context-aware IFC-regulation semantic information alignment. *Automation in Construction*, 145:104540.

Zhong, B., Xing, X., Luo, H., Zhou, Q., Li, H., Rose, T. and Fang, W. (2020). Deep learning-based extraction of construction procedural constraints from construction regulations. *Advanced Eng. Informatics*, 43:101003.

Zou, Y., Kiviniemi, A. and Jones, S.W. (2017). Retrieving similar cases for construction project risk management using Natural Language Processing techniques. *Auto. in Construction*, 80:66–76.

CONNECTING OBJECT DATA IN NATURAL STONE RENOVATION PROJECTS

Alex Donkers[1], Marijn Janssen Steenberg[1,2], and Pieter Pauwels[1]
[1]Eindhoven University of Technology, Eindhoven, The Netherlands
[2]Aannemingsbedrijf Nico de Bont B.V., Vught, The Netherlands

Abstract

Industry experts working on renovation projects rely on a wide range of information. Standards to connect this information are lacking, as well as tools to communicate the data with the industry expert. This study researches the application of semantic web technologies to connect heterogeneous data in renovation projects. Based on this connected data, we present a web-based linked building data viewer through which practitioners can ask relevant questions about a project. Close collaboration with industry experts led to time improvements and a positive attitude towards data-driven support systems.

Introduction

Historical buildings hold significant cultural and economic value and are key elements of a country's cultural heritage. Governments implement policies and regulations to safeguard and preserve those buildings. To do so, practitioners rely on a lot of information about the monument. This information is, however, often documented using various standards, data formats, and software systems, resulting in a situation of fragmented information that is difficult to manage.

The construction industry is increasingly adopting Building Information Modeling (BIM) as a method to digitize information in collaborative workflows. The Industry Foundation Classes (IFC) is a widely recognized standard to store information about objects in BIM models. Common Data Environments (CDE) provide ways to collaboratively use IFC-based BIM models and other information. However, the capabilities of those CDEs to provide users with an integrated overview of information are often limited, as the lack of data integration standards and the vendor lock-in of these systems make it difficult to link related information across different systems and stakeholders.

Semantic web technologies promise to alleviate some of those issues, as they enable connecting and structuring the currently fragmented datasets from various domains and stakeholders (Pauwels et al., 2017), and applying those technologies to integrate BIM models with built heritage information could help in structuring, validating, visualizing, and sharing information in monumental restoration and preservation projects. Efforts in this direction have already been made by other researchers. Hamdan et al. (2019) developed a modular ontology for defining damaged objects and their topology. Bonduel (2021) contributed to the development of a network of new and existing modular ontologies to integrate heterogeneous built heritage information. These works both conclude that more research is needed to apply linked data in real projects and that a better translation from real-world processes to semantic web technologies is necessary.

Next to the necessary development of more data integration standards, Hamdan et al. (2019) suggest that those standards should also be linked to practical actions so that integrated data can actually answer the questions of end users. Bonduel (2021) suggested the development of demo applications to translate the integrated data to end users. Various initiatives to build web-based linked building data viewers exist, such as LBDviz (Donkers et al., 2023), LD-BIM[1], RUB-IP (Hagedorn et al., 2023), BIM-SIM (Chamari et al., 2022), and ConSolid (Werbrouck et al., 2024). However, the current functionalities of those viewers are limited and use in a heritage context requires specific functionalities that are related to the requirements of end users in this specific domain (see Figure 2).

Other research has already made efforts that could help in structuring, validating, visualizing, and sharing data in the context of restoration projects. However, in practice, there still exists a research gap in structuring and visualizing such data in the construction phase of heritage projects. The created solution should use these new techniques while being integrated with existing processes and tools on site. This paper investigates the capabilities of applying semantic web technologies to integrate information from BIM models with other files and information used by practitioners. This work hereby focuses on the integration of data that is not exclusively restricted to RDF graphs. Instead, a software solution is relied upon that combines multiple datasets of multiple types, sometimes referred to here as 'connected data'. This setup and approach is not much different from several recent digital twin implementation initiatives.

This paper presents an ontology and a software tool to integrate data related to built heritage and make this data accessible to practitioners. Section 2 first presents state-of-the-art built heritage data integration, with a specific focus on natural stone renovation projects in the Netherlands. The processes in natural stone renovations, including the relevant information streams, are then

[1] https://ld-bim.web.app/

analyzed, including the relevant end users and their requirements. Those requirements are then translated into a software architecture. The information requirements are also translated into a network of ontologies, combining both existing ontologies and a newly developed extension. These methods are then finally used to develop a web-based tool that enables different end users to interact with the integrated information. The tool is validated in a real-world case study at the Dam Square in Amsterdam, The Netherlands.

Related Work

Conservation of Natural Stone Built Heritage

Historical buildings in the Netherlands hold cultural and economic value, attracting tourists and preserving the country's heritage. The conservation of these buildings is regulated through self-regulation and guidelines established by the Dutch Cultural Heritage Agency (RCE) and the Foundation ERM (Naldini & Hunen, 2019). The ERM guidelines provide technical specifications and quality standards for restoration work, ensuring the preservation of historical value. The process involves inspections, damage assessment, identification of cultural-historical value, determination of appropriate materials and techniques, and decision-making using the "Conservation ladder" (Lubelli et al., 2021). The whole conservation cycle should be documented to provide valuable information for future restoration projects. For natural stone restoration projects, aspects such as cultural-historical value, damage assessment, damage identification, material type, and shape need to be documented (ERM, 2020). Documenting these aspects is done using a wide range of standards and data formats, creating a situation that is difficult to manage.

Connected Data

Several initiatives have aimed to develop web-based building data. One of the first initiatives was the creation of ifcOWL, which translated the IFC schema into an RDF-based schema using semantic web technologies (Beetz et al., 2009; Pauwels & Terkaj, 2016). However, as it mirrors the IFC EXPRESS schema, it inherits the same complex structure and size. The Linked Building Data Community Group of the World Wide Web Consortium (W3C LBD CG) was created to identify and align ontology development initiatives for building data. The approach they took was to create a set of interoperable, flexible, and open standards covering different aspects and domains (Rasmussen et al., 2020). Rasmussen et al. (2017) took the first step in creating the Building Topology Ontology (BOT), which is a flexible and lightweight ontology to define the relationship between elements within a building. This is a basic ontology that describes building structures using the classes bot:Building, bot:Storey, bot:Space, and bot:Element. Implementing this ontology would result in a shift from file-based collaboration to data-centered collaboration using RDF, directly linking data instead of referring to other documents.

While the use of RDF graphs is immensely powerful to create a semantically rich cloud of linked data, several other types of data exist that typically do not fit this RDF graph principle. Several hybrid methods have therefore been proposed and deployed in the last few years that combine semantic web technologies with 3D models, pictures, point clouds, and so forth, e.g. Werbrouck et al. (2020). Another example is the use of semantic web technologies to describe information that does not fit into the BIM exchange schema IFC, for instance describing information about damages on objects while linking to the actual object in the IFC model (Hamdan et al., 2019). Also smart building implementations, e.g. based on Brick, adopt this approach to connect to telemetric data (Chamari et al., 2023). The adoption of semantic web technologies for the description of heritage buildings could help in integrating the different data sources into a full virtual representation. These hybrid methods, which integrate semantic web technologies with BIM models and other files, are considered a highly promising route for handling the data heterogeneity challenges in the built heritage domain.

Existing Ontologies Related to Natural Stone Restoration

Researchers proposed various ontologies that enable the representation, classification, and relationship modeling of historical buildings, their components, damages, inspections, interventions, geometry, and additional properties. Bonduel (2021) proposes the Construction Tasks Ontology (CTO), as an extension to BOT, providing terminology for construction tasks. The Damage Topology Ontology (DOT) contains classes and properties that define topological relations between the damage and the affected building components (Hamdan et al., 2019). Based on the same principles as BOT, DOT, and CTO also support the creation of extensions in a specific domain. Bonduel (2021) used this principle when converting the MDCS damage atlas into an ontology that extends DOT. Another example is the Natural Stone Damage Ontology (NSD), which is based on a different damage atlas (Hamdan, 2023; Seeaed & Hamdan, 2019). Seeaed & Hamdan (2019) have also developed the Stone Component Ontology (SCO), as an extension ontology of BOT, for semantically modeling natural stone facades. Furthermore, the Ontology for Managing Geometry (OMG) provides generic concepts for managing geometry descriptions in a semantic web context (Wagner et al., 2019). It focuses on general-purpose functionalities that allow the handling of single, multiple, and versioned geometry descriptions. OMG can be extended by the File Ontology for Geometry formats (FOG), which extends the OMG to describe file format, versioning, and object identifier information for geometry data (Bonduel et al., 2019). All these ontologies were developed using the principles of the W3C LBD CG, creating modular data structures that can be used following a plug-and-play principle to represent information in for example a natural stone restoration project.

These evolutions also happened in practice. The Dutch Cultural Heritage Agency (RCE) has published its registry of national monuments[2] as a large RDF graph or knowledge graph. This dataset is also linked with a registry of all buildings and addresses in the Netherlands[3] (BAG), which is one of the registries that is published and maintained as a knowledge graph by the Dutch cadaster. The RCE has also published thesauri terms as linked data using SKOS (CHT)[4]. This includes, for example, different types of damages and materials.

Methodology

This paper starts with natural stone renovation projects and investigates the use of the presented technologies and workflows (connecting data) for this data challenge. This paper first reviews the existing processes in natural stone renovation projects. Expert interviews with a project leader, planner, BIM coordinator, and BIM manager resulted in a list of competency questions which is translated to a use case diagram. Based on those competency questions and the processes in natural stone renovation projects, a network of ontologies is created that can support integrating the available data. The use case diagram is then translated into a system architecture that enables the end users to interact with the various types of information through a web application: Heritage LBDviz. A prototype of this application is tested using a real-world renovation project of the Dam Square in Amsterdam, The Netherlands, and validated against the competency questions in collaboration with the industry professionals. The development and validation of this tool was an iterative process with industry professionals where different functionalities were created and validated through three iterations: definition of the competency questions and use case diagram, development of the user interface, and validation.

Natural Stone Restoration Processes

The process of natural stone facade restoration consists of a series of tasks depending on the type of intervention that needs to be performed. To visualize this process, a BPMN schema was created, see Figure 1. The ERM guidelines (ERM, 2020) specify that information should be documented throughout the conservation cycle of a heritage property. The data column in Figure 1 shows when this data is collected. This gives an insight into when information is generated so that it can be associated with specific parts of the process. The schema does not show every possible workflow. However, it gives a clear overview of the main steps that are part of the process and the different types of workflows that are possible.

Firstly, an inspection of each object in the building is carried out in two steps. The first is aimed mainly at gathering information about the object. This information is often gathered in a spreadsheet. In the second step, it is decided what intervention is needed. This information is then used to produce a cost estimation, including the type of replacement material, which creates an additional spreadsheet of information. Once the interventions and costs are known, the actual work on the construction site begins. Three types of workflows are used in the execution of the interventions. Some interventions can be carried out in situ, creating a simple process with only one task: the repair. However, it is often necessary to carry out an intervention in the masonry workshop, which results in a series of additional tasks. In this case, the object must be dismantled from the building and then transported to the masonry workshop. Once the object has been repaired or reconstructed, it can be transported back to the site where it is reassembled in the building. Dismantling an object can require some of the surrounding objects to be dismantled as well. These are then transported to a location for temporary storage and, when the time is right, transported back to the site and used in a reassembly task. Progress is recorded at different steps by taking photos of the object. When all the work has been carried out and all the objects are back in place, a check is performed to ensure that the result meets the desired quality. Throughout the process, the multiple sources of data are passed on to different stakeholders to meet their information needs. This creates a situation where multiple sources of data exist in different places.

End User Requirements

During interviews with the industry experts about what information they need throughout the process, several competency questions were identified. Below are some of the most pressing ones.

CQ1 What is the location of the object in the building?

CQ2 What photos are available of a specific object?

CQ3 What is the volume, material, and shape of the object?

CQ4 What interventions are conducted on an object?

CQ5 How to give an overview of all relevant information on an object?

CQ6 What is the height of the object relative to ground level?

CQ7 What is the progress status of every object in the building?

The various competency questions were translated into use cases to be implemented as functionalities in the system (Fig. 2). They thus inform the normal user requirement definition stage of a regular software development process. These functionalities can then be developed during the different iterations of the prototype

[2] https://linkeddata.cultureelerfgoed.nl/

[3] https://labs.kadaster.nl/cases/bag-ld

[4] https://data.cultureelerfgoed.nl/term/id/cht.html

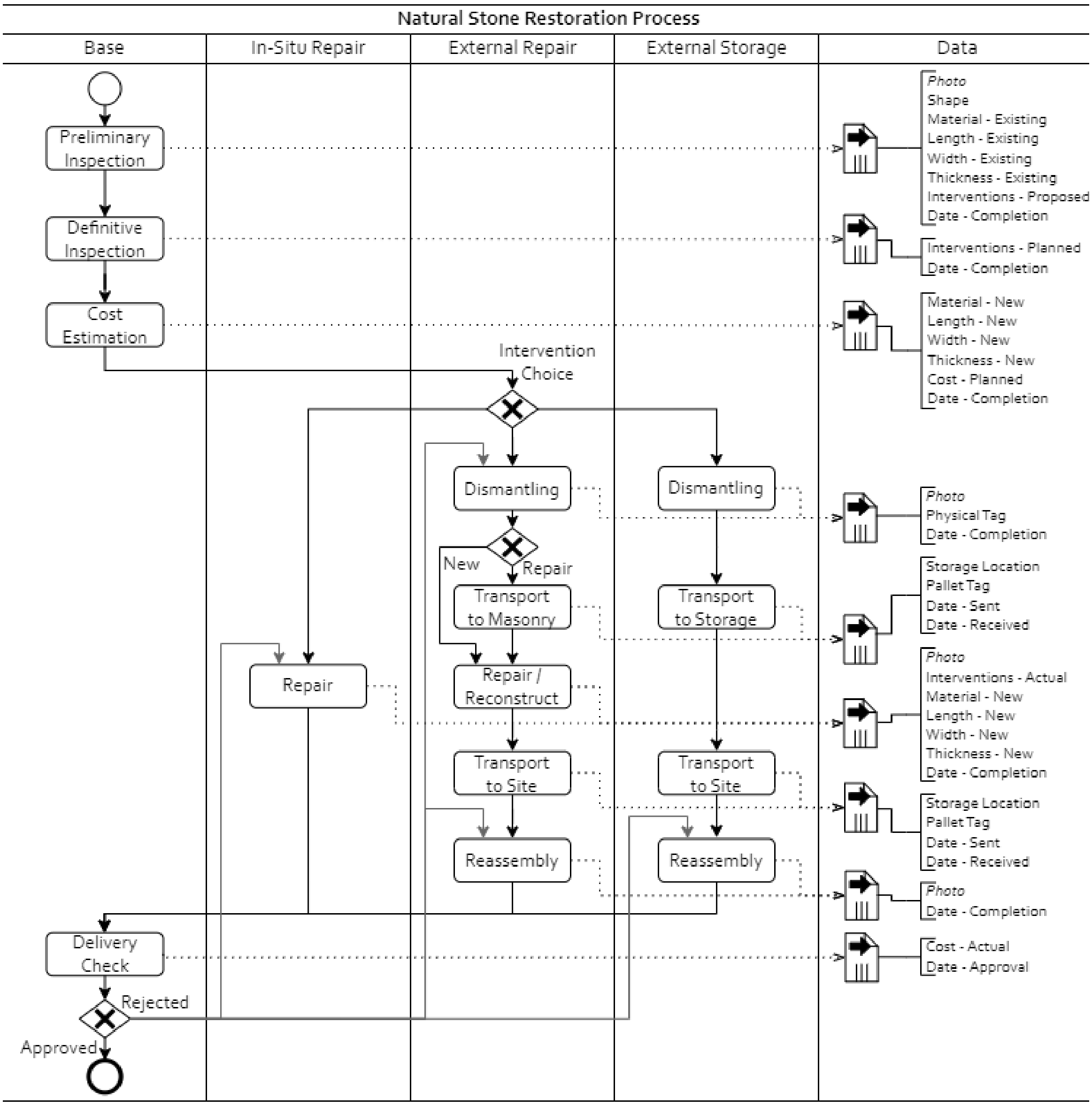


Figure 1: BPMN schema of a natural stone renovation process

development. Each use case either supports another use case or provides an answer to one of the competency questions. Multiple types of users are considered in the use cases being developed. The use cases for the beginner user focus on recognizable output, extended with new possibilities that introduce the user to the system. A more advanced user is provided with functionalities to find traditional information through the system. Finally, a pro user is also provided with the ability to create queries based on their extensive knowledge of the data structure.

System Architecture

Based on the existing processes and the end user requirements, a 5-layer conceptual system architecture was established for combining the RDF graphs with 3D models and other files, see Figure 3, following earlier found best practices (Donkers et al., 2023). The resources layer contains the heterogeneous data produced in the data layer in Figure 1, including BIM models, photos, and states. The data is structured using a network of ontologies. An ETL procedure then transforms some of the source information to RDF and loads the data into the RDF graph. The Connected Data layer combines an RDF graph in the backend with links to the 3D BIM model and other files. As said, this is similar to recurring Digital Twin development projects. A services layer provides capabilities to query data from the connected data layer and ask questions about this information. Users can interact with this information via user interfaces. This paper presents a 3D model viewer - Heritage LBDviz - and a spreadsheet exporter as user interfaces that implement functionalities based on the described use cases. The architecture defines the components independently of the selected solutions, focusing on the responsibilities of these different layers in the system. Implementation of the system architecture must include considerations for communicating between different layers.

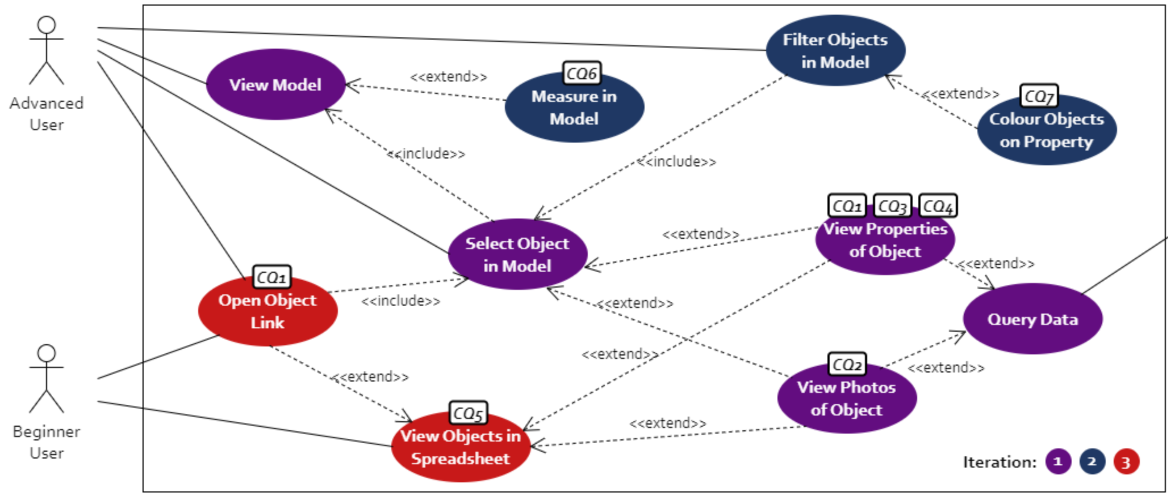


Figure 2: Use case diagram

Results

After development, the prototype was validated in a real-world setting. The restoration of the National Monument on Dam Square was chosen as a case study. The monument consists of a pylon with statues and a memorial wall with reliefs. In front of the monument, two lion figures stand as sentinels. The twenty-two-meter-high pylon, that is made of concrete, the memorial wall and the lions are cladded in travertine, a marble-like porous limestone from Tuscany, Italy. The last restoration took place in 1997. In 2022, 25 years later, the National Monument needed to be repaired or restored. During the restoration, the travertine elements were dismantled and transported to the stonemasonry workshop, where they were cleaned, repaired, and replaced where necessary, minimizing disruption to the Dam Square. The dataset provided does not include all the data described in the process schema, but the project offers a realistic representation of the current way in which practitioners operate and collect and manage data. The following data sources were available for this research: a Revit model and PDF elevations to give an insight into the geometry of the building objects; the Autodesk Build system used to carry out the inspection, which contains data on object properties and photographs; the actual set of photographs taken during the process; and finally some data documented in several spreadsheets, such as the interventions carried out on the elements and the dimensions of all the objects.

Network of Ontologies

Based on all the available data, a mapping was done to find all the classes for the case project data in the BOT ontology and all its extensions, such as CTO, DOT, and OMG. However, some of the data available in this case study could not be described using these existing ontologies. The Restoration Intervention Information (RII) ontology was created to define a set of classes and relationships to fill this gap. The ontology focuses on the interventions that are carried out on the heritage assets and the different properties of elements that are used as information in this process. The ontology describes classes of interventions based on the conservation ladder, defining the class rii:Intervention and its subclasses rii:Preservation, rii:Repair, rii:Reconstruction (Fig 4). These interventions are then linked to the cto:InspectionTask for which they were selected and the dot:Damage they are intended to cover. Finally, the cto:hasTaskContext is used to group the various cto:RepairTask instances that are performed to carry out the rii:Intervention. In addition to interventions, the ontology is also used to add external resources to the graph, specifically photos. The DOT ontology already defines a dot:ExternalResource, but this is specifically described as documentation that should only be damage-related. As not all information is damage-related, it was decided to use the class rii:ExternalResource with the subclass rii:ExternalImage, which is used in combination with the data property rii:filePath to link to externally stored images (Fig. 4).

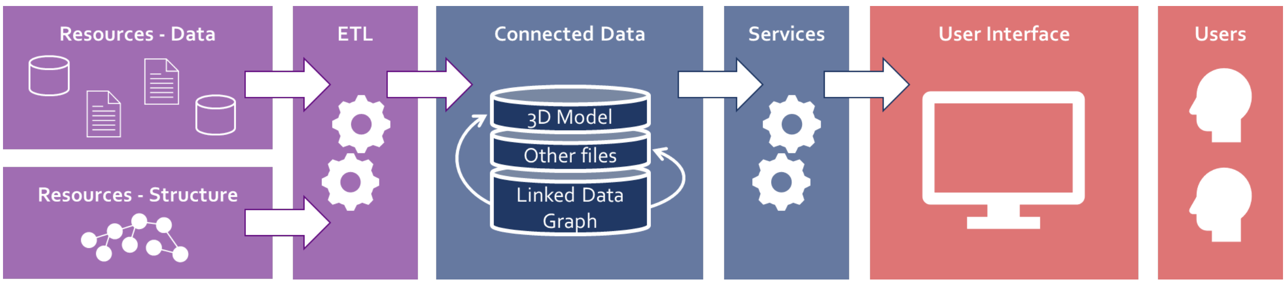


Figure 3: System architecture of Heritage LBDviz

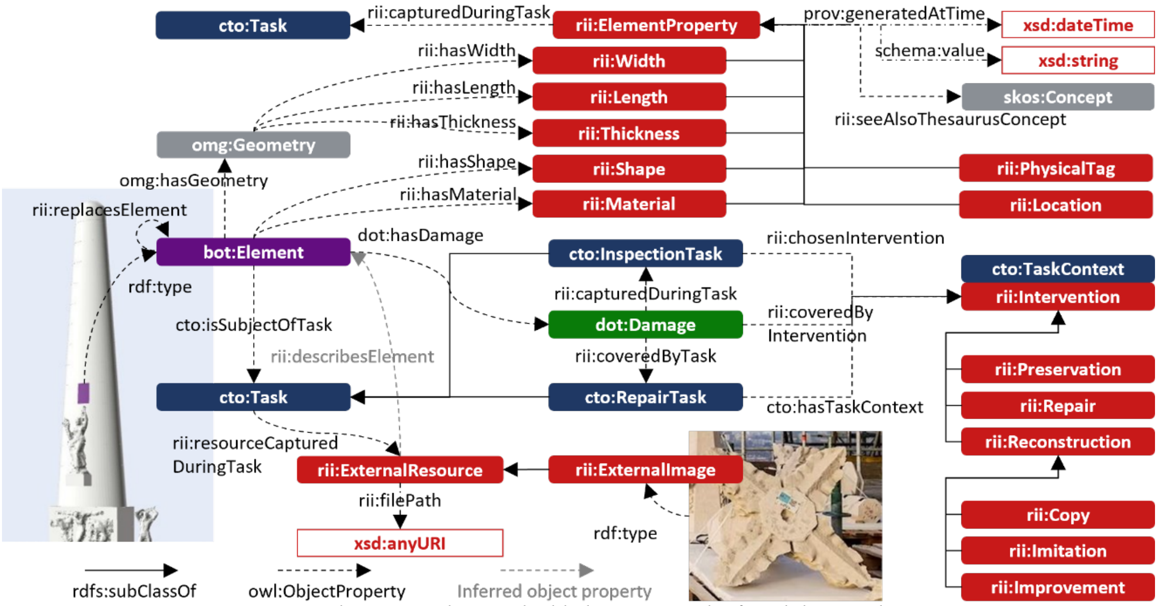


Figure 4: The RII ontology embedded in a network of modular ontologies

The ontology also includes an OWL property chain based on cto:isSubjectOfTask and rii:capturedDuringTaks, creating the relation rii:hasDescriptionInResource with the inverse rii:describesElement. The RII ontology is available at https://marijnjanssensteenberg.github.io/rii#.

Heritage LBDviz

To demonstrate the possibilities of our method in the context of restoration projects, a web application has been developed, serving as an easily accessible user interface that practitioners can visit through their browsers. This application - Heritage LBDviz - is an extension of the already existing LBDviz (Donkers et al. 2023). The tool (Figure 5) is a single-page web application combining an IFC viewer with several menu windows. IFC files are loaded through the IFC.js library, a JavaScript library that allows users to load, view, and edit IFC models in the browser. It also parses the IFC file to JSON so that one can interact with it through JavaScript. Behind this IFC file is a backend with the data, including RDF graphs and non-RDF data. This data can be accessed via Comunica.js, a meta query engine that can be used to execute SPARQL queries against multiple SPARQL endpoints or RDF files simultaneously. Heritage LBDviz[5] and a guide to recreate the tool[6] are available on GitHub.

Heritage LBDviz aims to fulfill the competency questions suggested by practitioners by adding the functionalities as shown in the use case diagram (Figure 2). First, beginner users can open the URI of a specific object, which will automatically open the viewer and highlight the specific object (CQ1). One can then use the various query functionalities to ask for information about this object, either by manually creating SPARQL queries or by using the predefined SPARQL queries hidden behind buttons in the user interface. This allows users to query information from the RDF graph, such as an object's material, size, and other properties (CQ2) and current and historical interventions on the object (CQ4), as well as non-RDF data such as photos (CQ3). All photos linked to an object can be queried and will be printed in the user interface (Figure 5). Next to printing results of queries in the user interface, Heritage LBDviz also allows users to export the results into spreadsheets. Those spreadsheets replace the traditionally manually created spreadsheets, saving time, while also allowing interaction between the spreadsheet and the viewer through hyperlinking cells with the URIs of objects (CQ5). Finally, users can interact with the IFC model in the viewer, for example by measuring distances in the 3D model (CQ6) and by highlighting objects. Using the ontology, Heritage LBDviz can also perform live reasoning and infer the current status of all elements in an IFC model after which all elements are highlighted using associated colors (CQ7).

Heritage LBDviz was validated using the Dam Square monument and the system was tested together with industry experts. Firstly, the tool seems to significantly save time in finding information on a specific object. Construction workers now don't need to go down to the workmen's shelter to find information on a block, and the normally manually created spreadsheets are now automatically generated. The fact that heterogeneous information can be accessed via one user interface is also expected to reduce errors. Inconsistent information regularly causes elements to contain errors after repair, delaying the process and increasing failure costs. The vendor-neutral approach makes projects less reliant on specific software, especially in multi-stakeholder projects. Finally, using the tool enhanced the awareness of the available data and its value, and industry experts could easily come up with suggestions on innovative functionalities of the tool.

[5] https://github.com/MarijnJanssenSteenberg/Heritage-LBDviz

[6] https://github.com/AlexDonkers/Frontends-and-LBD

Figure 5: The Dam Square in Heritage LBDviz

Discussion

This paper discusses an approach to integrating data related to renovation projects and making this data available to industry professionals. This paper is not the first paper to present data integration methods related to the topic of renovation projects. This project builds on these earlier projects and re-uses several existing ontologies. Furthermore, this paper presented an extension in the form of the RII ontology to define knowledge on interventions in restoration projects and specific properties that are commonly used in practice. The ontology follows best practices that suggest a modular approach and extends already existing ontologies. Although this ontology is tested in a natural stone renovation case study, its modularity should enable extensions toward other types of projects.

Secondly, this project confirms the feasibility and usefulness of making different types of data available through the appropriate use of semantic web technologies (RDF graphs) and state-of-the-art web development practices (service-oriented architecture). Moreover, it can be experienced that this approach to data integration reduces the gap between developers and data on the one hand, and practitioners on the other hand. As suggested in Donkers et al. (2023), co-creating the Heritage LDBviz environment together with practitioners significantly reduces barriers to using these data fittingly in practice. Validating the tool together with industry experts triggered ideas for more complex functionalities. This shows future opportunities for using such integrated data in restoration projects, and it shows that co-creating such tools in iterative cycles together with industry experts can create more awareness and engagement toward data-driven decision support systems.

Further to the above contributions, two points for further development were identified. First, Heritage LBDviz mainly focuses on visualizing data to answer the competency questions of users. However, validating the quality of the data is essential to make sure that the information is reliable. Future research in quality control methods, such as using SHACL, is needed to improve the reliability of the tool. Second, it is believed that access to a complete dataset of past restorations can create many opportunities for subsequent cycles of conservation. This may be better in reach using the proposed way of working, as it has a high level of modularity and scalability. More research is needed to understand how such long-term data preservation can be organized.

Conclusion

Conserving built heritage is a complex process that involves a lot of information. This research identifies two important challenges: we lack methods to integrate this information, especially beyond the RDF formatting, and we lack tools to make this information accessible and useful to practitioners working on renovation projects. This paper applies semantic web technologies and novel web development techniques to solve these challenges in close collaboration with the industry experts. After a thorough review of existing processes in natural stone renovations, using both literature and discussion sessions with industry experts, seven competency questions were defined. These led to the development of a network of ontologies, extended with the newly developed RII ontology. The knowledge from the review of renovation processes was translated to a use case diagram, based on which a system architecture for a web app was created. This web-based BIM viewer - Heritage LBDviz - aims to help the industry experts in answering their competency questions. The tool was validated together with those industry experts in a case study at the Dam Square. This validation showed that Heritage LBDviz was able to answer the competency questions and that it had the positive side effect of creating more engagement and awareness of such data-driven support systems.

The developed network of ontologies shows that it is possible to integrate heterogeneous data related to restoration projects. The web application shows that communicating this data via a simple user interface can help industry experts in their daily work and improve their understanding of the value of these data for future use cases. While the methods presented in this paper are specifically tested on natural stone renovation projects, we believe that the system architecture can be used in a much wider range of contexts, as long as there is a need to connect, structure, and communicate data.

References

Beetz, J., Van Leeuwen, J., & De Vries, B. (2009). IfcOWL: A case of transforming EXPRESS schemas into ontologies. *Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM*, *23*(1), 89–101. DOI: 10.1017/S0890060409000122

Bonduel, M. (2021). *A Framework for a Linked Data-based Heritage BIM* [KU Leuven]. https://lirias.kuleuven.be/retrieve/618662

Bonduel, M., Wagner, A., Pauwels, P., Vergauwen, M., & Klein, R. (2019). Including widespread geometry formats in semantic graphs using RDF literals. *Proceedings of the 2019 European Conference on Computing in Construction*, *1*, 341–350. DOI: 10.35490/ec3.2019.166

Chamari, L., Petrova, E., & Pauwels, P. (2022). A web-based approach to BMS, BIM and IoT integration: a case study. *CLIMA 2022 Conference*. DOI: 10.34641/clima.2022.228

Chamari, L., Petrova, E., & Pauwels, P. (2023). An End-to-End Implementation of a Service-Oriented Architecture for Data-Driven Smart Buildings. IEEE Access, 11, 117261-117281. Article 10287934. DOI: 10.1109/ACCESS.2023.3325767

Donkers, A., Yang, D., de Vries, B., & Baken, N. (2023). A Visual Support Tool for Decision-Making over Federated Building Information. *20th International Conference, CAAD Futures 2023*, 485–500. DOI: 10.1007/978-3-031-37189-9_32

ERM. (2020). *URL 2005: Inspecties van monumentale gebouwen*. Stichting Erkende Restauratiekwaliteit Monumentenzorg (ERM). https://www.stichtingerm.nl/kennis-richtlijnen/url2005

Hagedorn, P., Liu, L., König, M., Hajdin, R., Blumenfeld, T., Stöckner, M., Billmaier, M., Grossauer, K., & Gavin, K. (2023). BIM-Enabled Infrastructure Asset Management Using Information Containers and Semantic Web. *Journal of Computing in Civil Engineering*, *37*(1), 04022041. DOI: 10.1061/(ASCE)CP.1943-5487.0001051

Hamdan, A.-H. (2023). *Ein ontologiebasiertes Verfahren zur automatisierten Bewertung von Bauwerksschäden in einer digitalen Datenumgebung*. Technische Universität Dresden.

Hamdan, A.-H., Bonduel, M., & Scherer, R. J. (2019). An ontological model for the representation of damage to constructions. *CEUR Workshop Proceedings*, *2389*, 64–77. http://ceur-ws.org/Vol-2389/05paper.pdf

Lubelli, B., Pottgiesser, U., Quist, W. J., & Rextroth, S. (2021). *Dealing with Heritage: Assessment and Conservation* (S. Naldini, Ed.). BK Books. https://books.bk.tudelft.nl/press/catalog/book/796

Naldini, S., & Hunen, M. van. (2019). Guidelines for quality of interventions in built cultural heritage in The Netherlands. In *Professionalism in the Built Heritage Sector* (pp. 87–93). CRC Press. DOI: 10.1201/9780429397912-11

Pauwels, P., & Terkaj, W. (2016). EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in Construction*, *63*, 100–133. DOI: 10.1016/j.autcon.2015.12.003

Pauwels, P., Zhang, S., & Lee, Y. C. (2017). Semantic web technologies in AEC industry: A literature overview. *Automation in Construction*, *73*, 145–165. DOI: 10.1016/J.AUTCON.2016.10.003

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., & Pauwels, P. (2020). BOT: The building topology ontology of the W3C linked building data group. *Semantic Web*, *12*(1), 143–161. DOI: 10.3233/SW-200385

Rasmussen, M. H., Pauwels, P., Hviid, C. A., & Karlshøj, J. (2017). Proposing a Central AEC Ontology That Allows for Domain Specific Extensions. *Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction*, 237–244. DOI: 10.24928/JC3-2017/0153

Seeaed, M. K., & Hamdan, A.-H. (2019). BIMification of stone walls for maintenance management by utilizing Linked Data. *31st Forum Bauinformatik*, 165–172. DOI: 10.24086/ICACE2022/paper.879

Wagner, A., Bonduel, M., Pauwels, P., & Uwe, R. (2019). Relating geometry descriptions to its derivatives on the web. *Proceedings of the 2019 European Conference on Computing in Construction*, *1*, 304–313. DOI: 10.35490/EC3.2019.146

Werbrouck, J., Pauwels, P., Bonduel, M., Beetz, J., & Bekers, W. (2020). Scan-to-graph: Semantic enrichment of existing building geometry. *Automation in Construction*, *119*, 103286. DOI: 10.1016/j.autcon.2020.103286

Werbrouck, J., Pauwels, P., Beetz, J., Verborgh, R. & Mannens, E. (2024) ConSolid: a Federated Ecosystem for Heterogeneous Multi-Stakeholder Projects. *Semantic Web* (in press). DOI: 10.3233/SW-233396

KNOWLEDGE DRIVEN RULE-BASED BUILDING GEOMETRIC DIGITAL TWIN CONSTRUCTION – STATE OF ART REVIEW

Ya Wen[1], Mudan Wang[1], Mahendrini Ariyachandra[1], Ran Wei[1], Ioannis Brilakis[1] and Lizhao Xiao[2]
[1]University of Cambridge, Cambridge, United Kingdom
[2]University of Hong Kong, Hong Kong SAR

Abstract

Automatic construction of the building geometric digital twins (gDTs) is implemented based on the scanned data from the existing buildings. The highly detailed gDTs necessitates the top-down construction method informed by the predefined building knowledge. This paper reviews the building rules used for the model-driven building indoor environment gDTs construction. It is found that these knowledge is fragmented and lacks systematic analysis. The research categorise these rules into three groups and proposes an Ontology-based knowledge graphs (KGs) approach for organising this knowledge. This method aims to integrate rules for gDT construction to establish a knowledge system for the further research.

Introduction

The emergence of gDTs sheds light on both academic and practical perspectives in the digitalisation journey of the AECO industry (Dou et al., 2023). DTs are the virtual replicas of physical assets, processes, or systems that enable stakeholders to visualise, monitor, and analyse their projects with unprecedented precision. gDTs focus on the static information in the model, and these are meanly the geometric information (Pan et al., 2023). An accurate and semantically-rich gDT ensure a foundation decision-making system for the building condition real-time monitoring, simulation and optimisation.

The process of gDTs construction of existing buildings typically involves the steps of: 1) scanned data collection, 2) scanned data processing, and 3) gDTs recreation. There are numerous methods for surveying and data acquisition (Volk et al., 2014). Subsequently, the collected data should be detected, segmented, and classified efficiently. The first two steps involve most of the technically complex tasks to form the essential prerequisites for gDTs construction. Numerous studies focus on the data collection and processing workflow, while fewer research work assesses the gDT recreation process. The gDTs creation work is regarded to be the 'last mile' process and a simple procedure to convert the information extracted from collected data into a structured format. However, the efficient gDT recreation requires the prior knowledge of the buildings, which including the object shapes, the identifications of objects and their relationships. The rule-based construction method can help avoid the obvious irrationality in the gDT construction process (P. Tang et al., 2010). However, there is no research focuses on reviewing and analysing the building rules that frequently used in gDTs construction to the authors' knowledge.

Building grammar is utilised to present the common building design logic and the structural topology of the building objects (Hu et al., 2019). During the gDTs construction process, building grammars are mentioned in the literatures emphasising the rules that are related architectural design principle. These can be summarised as multiple rules being employed as pre-established knowledge to minimise the dependence on the accuracy of the scanned data. This knowledge serves as the crucial supports to gDTs construction and inference.

KG plays a crucial role in the knowledge management field. Ontology-based KG is a method of representing knowledge, which utilises ontological structures to categorise and interconnect the various entities and their relationships (Bassier et al., 2020). There are some existing ontology-based KGs that can present the building topology and its semantic information (Rasmussen et al., 2020). However, there is still a gap in how to present the knowledge transferring from the scanned data to create the gDTs and to integrate the gDTs with the specific building domain knowledge.

The research objective of this paper is to review the rules used in gDTs construction phase. In addition, the conceptual framework of an ontology-based KG is then introduced to not only present the model-driven gDTs construction rules but also allow the further domain knowledge integration. The proposed KG shall benefit the researchers with direct and intuitive access to the building knowledge as a top-down gDTs creation guidance. Based on the objective, the research questions are listed: a) what knowledge is frequently used as a guide of gDTs creation process, and b) how to organise them in an ontology-based KG as the knowledge base. The proposed KG will be further enriched in the following research with the detailed description of the entities, properties and the reasoning and querying rules, which will facilitate the construction of the semantically-rich gDTs and the inference of hidden objects that are occluded during the data collection process.

Methodology

State of art review is an important knowledge capture method to design a domain KG (Wang et al., 2022). This paper conducts a systematic review and concludes the knowledge utilised for the rule-based indoor environment gDT construction (see Figure 1). Leveraging the keywords that precisely encapsulate the target papers, the searches for 'digital twin generation' and 'scan to BIM' is implemented within the databases like Web of Science and Google Scholar, covering the past decade (2014 – early 2024). After excluding studies outside the relevant fields, a total of 825 papers were initially identified. Through further selection stages, which involved assessing content completeness and scrutinizing abstracts, only 18 papers that introduce the rules for model-driven gDTs construction of building indoor environment were chosen for further detailed analysis.

This knowledge is categorised into three perspectives, which are shape rules, topological rules and semantic rules. Different categories of rules work on the various aspects of information. The shape rules typically work on the geometric attribute of the one category of building element, including the building components and space. The topological rules contribute to the establishment of the spatial relationships between the building elements. While the semantic rules refer to the guidelines or constraints based on the underlying model's semantics. The semantic information can support to the realisation of various potential functions of gDTs, such as the hidden elements inference. Some of the rules are analysed and visualised through a case of the Civil Engineering Building in the University of Cambridge.

Based on that, we developed the conceptual framework of an ontology-driven KG. These rules shall be transformed into the key components of the gDTs construction KG, which are class, properties, constraints, and reasoning rules. The KG acts as a reusable knowledge base, offering a structured representation of the current state of knowledge during the gDT construction process.

Knowledge-driven rules of gDT construction

Knowledge-driven gDTs construction approach emphasises the top-down concep from the required output information. The quality of the scanned data is not a primary factor influencing the output, but the rules adopted during the creation process. Architectural grammar can be considered as the fundamental principle of the building design. In 1994, the knowledge-based computational method for architecture design was explored. It represents a complex procedure grounded in primitive elements and design rules (Carrara et al., 1994). Currently, the knowledge-driven rules based on the building grammars can be used for the building asset remodelling (Nikoohemat et al., 2021), prediction of the building componence configurations (Chen & Liu, 2023), building compliance checking and interior space navigation (Nikoohemat et al., 2021). The Table 1 concludes the key research about the knowledge-driven rule-based methods, and the corresponding target elements.

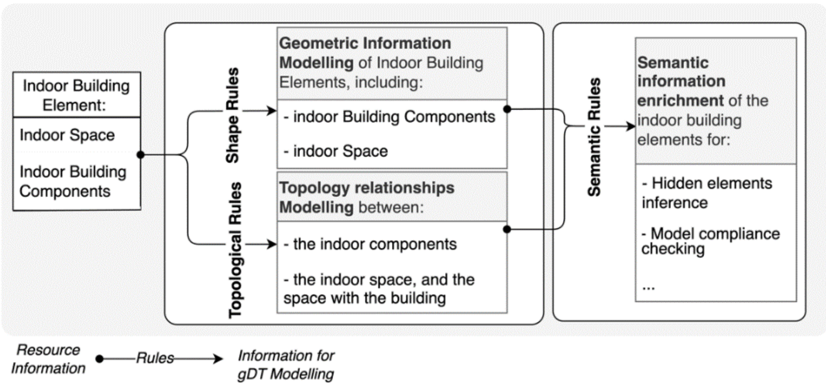


Table 1 Rules of building indoor environment gDTs construction

Research	Elements	Model-driven reconstruction method
(Khoshelham & Díaz-Vilariño, 2014)	interior walls	**shape-driven** rules for indoor wall layout reconstruction
(Becker et al., 2015)	interior hallway, rooms, staircase	three **topology-driven** rules for indoor layout and staircase reconstruction
(Ning et al., 2015)	Walls, windows, doors and roofs	**shape-driven and topology-driven rules** for the object-oriented in the scenes
(Abdollahi et al., 2023; Thomson & Boehm, 2015)	Interior walls	**shape-driven rules** for the geometric restrictions of interior walls
(Hu et al., 2019)	indoor room type inference	sixteen indoor room **topology-driven** rules to infer the room type from room layout
(Yang et al., 2019)	interior walls, roof, ceilings, doors, windows, stairs	**semantic-driven** representation for indoor environment reconstruction, especially the semantic representation of stair space

(Bassier & Vergauwen, 2020; Tran et al., 2019; Tran & Khoshelham, 2020)	interior walls	**topology-driven** rules for interior walls construction
(Nikoohemat et al., 2021)	indoor environment	eight **topology-driven** rules for building indoor component
(Cai & Fan, 2021)	interior rooms, doors	six **topology-driven** rules of the doors' position in the wall to infer spatial relationship of rooms
(Tang et al., 2022)	interior walls, roof, ceilings, doors, windows	**shape-driven** rules of planes and **semantic-driven** rules for 3D object-based indoor environment modelling
(Park et al., 2022)	indoor environment including main structural elements and furniture	**shape and topology-driven** rules of the boundary boxes of the indoor environment objects including floor, wall, column, door, window, table, chair, sofa and bookcase
(Chen & Liu, 2023)	indoor room layout inference	six **semantic-driven** rules to infer interior room layout from exterior building boundaries
(Kellner et al., 2023)	interior walls, windows and rooms	**shape-driven and topology-driven rules for the geometric features of the walls, windows and rooms;**
(Yang et al., 2023)	interior walls and indoor rooms	**topology-driven** rules of two walls, three walls and four walls
(Drobnyi et al., 2024)	interior components and space	**topology-driven rules** ('same-space', and 'same-object') for the indoor components and space

Shape rule-driven construction

The common building objects, such as the walls, floors, roofs, doors, windows, and space should be object-oriented constructed in the gDT creation process. The shape rules can be applied in both building elements modelling and interior spaces modelling.

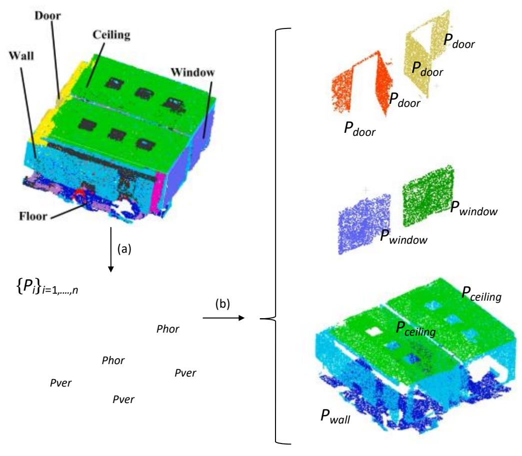


Figure 2 Shape rule-driven for building components construction (Tang et al., 2022) *(a) Decompose the data into planar primitives (b) semantic information enrichment to the planar primitives for further refinement.*

Shape rules for the building components

Wall modelling is an important task in the indoor gDT construction, typically requiring fitting with parallel and smooth surfaces rather than tessellated-as modelling. Bassier & Vergauwen (2020) defined a set of positional relationships as the foundational rules for the multiple cuboid walls. These relationships can serve as the prior knowledge for wall object creation, including: 1) intersection, 2) orthogonality, 3) blending, and 4) direct connection. Thomas and Boehm (Thomson & Boehm, 2015) proposed three geometric restrictions while the wall is generated, which are reject small wall, extend large wall and merge close planes with similar normal. Tang et al. (2022) utilised a grammar-enhanced method to decompose the entire scene into planar primitives using a two-level shape rule, based on the segmented point cloud clusters (refer to Figure 2). The first level of rules divides all planar primitives $\{P_i\}_{i=1,\ldots,n}$ into three categories according to their orientation: vertical planes (P_{ver}), horizontal planes (P_{hor}), and other planes (P_{oth}). The second level of rules further select the planar primitives that exhibit structural characteristics of the interior model and assign the semantic information to them. The characteristics are confirmed based on the geometric constrains and spatial orientation to identify doors (P_{door}), windows (P_{window}), ceilings ($P_{ceiling}$) and walls (P_{wall}).

The shape-based rules for identifying building components aim to discern the features of specific object categories. This approach is efficient as a reason of considering the ground truth shape of building elements. However, there is a lack of consistency in defining primitives and their geometric relationships across various building elements. Different researchers have adopted distinct methods for defining shape features. For instance, the positional relationships of interior walls are defined differently in the study by Bassier and Vergauwen (2020) and Khoshelham and DíazVilariño (2014). Despite these differences, both studies essentially describe the same real-world relationship between interior walls. In such scenarios, ontology-based KG holds the significant potential to facilitate the fusion of these diverse definition methods via unifying terminology.

Shape rules for the building internal space

Shape rules can be effectively applied to generate parametric indoor spaces (Abdollahi et al., 2023; Khoshelham & Díaz-Vilariño, 2014; Tran et al., 2019). This approach involves positioning cuboids within a zone defined by enclosing points to identify the space. Subsequently, topology relationship rules are typically employed in the next stage of space modelling.

The 4-tuple, $G = \{I,N,T,R\}$ is used to define a space in shape rules (Tran et al., 2019). Among that, I, represents the initial shape, providing the foundational form from the beginning of modelling. It serves as the baseline and starting point for further transformations. N denotes non-terminal shapes, which are intermediary forms during the modelling process. And T signifies terminal shapes, marking the end shape of space division. R refers to the production rules that indicate the transformation process. The relationships of these parameters can be presented as: $R : A \rightarrow B : cond., A \in N, B \in V, V = N \cup T$. The rules ($R$) can be categorised to different types, such as the initial shape place rule (R_{place}), the merging rule (R_{merge}) for two non-terminal shapes and the split rule (R_{split}) for space decomposition (Khoshelham & Díaz-Vilariño, 2014; Tran & Khoshelham, 2020). For instance, the R_{place} is presented through a transformation H, applied to the initial cuboid S, *where* $R_{placecuboid} : S \rightarrow HS$.

Shape rules can also define the relationship between building component and space as $G = \{N,T,R,S\}$, which represent the non-terminal space, terminal space, grammar axiom and starting symbol, respectively (Nikoohemat et al., 2021). With the basic elements defined, the relationship between two or more components can be further explored using topology rules. These two 4-tuples represent typical shape rules used for space construction, established on the similar conceptual basis but expressed in varying ways. Knowledge fusion is essential to merge similar concepts across different 4-tuples, aiming to realize a consistent workflow in the shape rule-driven space construction method.

Topology rule-driven construction

The logic applying topology rules for indoor 3D construction lies in establishing constraints on the interrelated topological connections of space and building structural elements. This method effectively addresses the challenges posed by incomplete and occluded elements in scanned point cloud data (PCD) (Ning et al., 2015; Tran et al., 2019; Yang et al., 2019).

Topology rules for indoor components

The building indoor environment comprises spaces and the building components, including structural elements, furniture and building service equipment. The allocation of walls is important for constructing the building indoor environment. The topological relationships between two walls (*collinear, perpendicular and intersect*), three walls (*T shape and pseudo T-shape*) and four walls (*cross shape*) are discussed (Yang et al., 2023). These relationships are important to overcome the occluded points at the corner in the collected PCD. Besides, the room boundary surface parameters can be adjusted based on the associated wall objects and the room object will be accurately rebuilt accordingly. Nikoohemat et al. (2021) highlighted the 8 potential relationships between 3D objects in building indoor elements, which are *disjoint, meet, equal, overlap, contain, cover, inside* and *coveredBy*. Furthermore, within the '*meet*' relationship, three sub- relationships are defined: support, supported and attachment.

The topological relationships between windows and walls are discussed by Kellner et al. (2023). These relationships can be utilised to not only generate the wall elements with attached door and windows, but also represented in the constructed gDT model. In addition, the integration of shape and topological relationships among the indoor building elements optimise the PCD semantic segmentation results to help with the high accuracy gDT construction. Park et al. (2022) define relationships of the indoor building element based on the boundary boxes of the objects, the spatial vectors of the boxes, and their referenced objects and referenced vectors. Except for the definition about the relationships between the objects, the topological relationships between planer surfaces are also explored (Drobnyi et al., 2024). The relationships of '*same-space*' and '*same-object*' are defined as a basis to form the topological relations between objects and spaces.

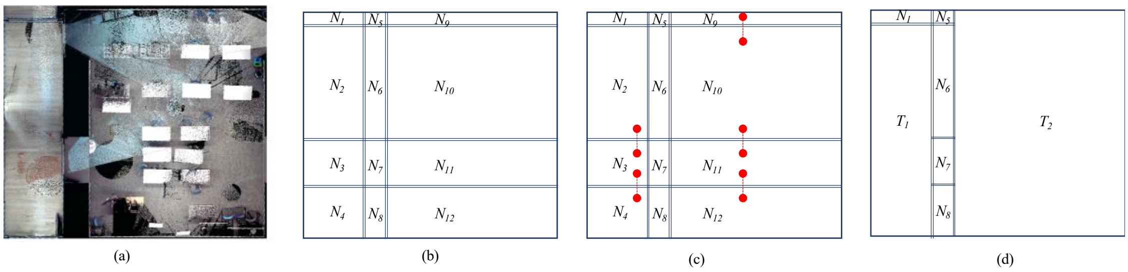


Figure 3 Merging non-terminal connected spaces (Civil Engineering Building in UOC as a case presenting the merge rules from (Tran et al., 2019): (a) top view of the selected scanned data, (b) top view of the cuboid shapes classified as the interior spaces, and defined as N (Non-terminal space), (c) find the adjacent spaces that without the common wall, and (d) the result with the merged cuboid defined as T (Terminal space).

Different research defines the various aspects of topological relations between the building indoor environment elements, which are laser scanning points, planer, object with the same or different semantic classification. The objectives of most of the rules are creating the semantic label to the disordered point cloud to construct the gDTs with not only the geometric, but also the topologic information. All the building elements with the same architecture style share the similar sets of features. However, the rules are fragmented. The systematic fusion of the rules can boost the knowledge-driven gDTs construction workflow.

Topology rules for indoor space

The indoor environment of a building not only consists of the building components but also the space. A room is defined as the terminal space enclosed by the walls, and it can be defined by the topological relationships of the spaces that compose them, and the building components, such as door or walls. These relationships are crucial for knowledge-driven gDT construction and inference. Tran et al. (2019) proposed a procedure for the room configuration rebuild including six rules: *placement, classification, adjacency, connectivity, merge and containment* rules. Three of the rules show the topological relationships of spaces including terminal and non-terminal spaces, which are: *adjacency, connectivity, and containment*. These relationships can be presented as: $R[t]: \{A1,A2\} \rightarrow \{A1[t],A2[t]\}: cond$, where, $R[t]$ is one of the constraints rules; $A1$ and $A2$ are terminal shapes or non-terminal shapes; and *cond* is the condition that identifies the topology relationship between shapes. The application of merge rule to combine two non-terminal spaces is exemplified in Figure 3, using UOC civil engineering building as an example.

The topological rules for the indoor space not only aid in room construction but can also facilitate the identification of further semantic enrichment (Cai & Fan, 2021; Tran et al., 2019). For instance, if two rooms are connected through an adjacency relationship, they share one common wall or a segment of it. Conversely, if a room partially shares a common wall with several rooms, it might indicate that this room is a corridor or hall space. Additionally, more semantic information can be inferred from the identification of room functions, such as the navigability of the space, and the existence of a door element. However, these relationships and the causation among these rules are not systematically organised. Integrating these into a KG could streamline the space construction and further support the reasoning of semantic information as proposed earlier.

In multi-story scenarios, the stair space is analysed by Yang et al. (2019). It is characterised as the composite space that connects two storeys. The stair space, denoted as S_{stair}, comprises two components: the stair-occupied space S_{sto} and the stair connection space S_{stconn}. The definitions of S_{sto} (green space) and the S_{stconn} (blue space) are illustrated in the Figure 4. Furthermore, the stair space connects with the slab free space S_{sf}, and the stair connection space, intersects with the associated slabs S_{slab}. This relationship can be expressed as: $S_{stair} = \{S_{sto}, S_{stconn}\}$, $S_{sf} = S_{slab} \cap S_{stconn}$.

Becker et al. (2015) emphasised the significance of hallway as critical elements within the building indoor environment. They proposed a strategy for constructing L-system hallway using spatial split rules and characteristics unique to hallway areas in buildings. Three functional control approaches are identified based on the topology relationships of segmented space, which are the *ActivationControl, LayoutSetting* and the *ConsistencyConstraints*.

The rooms, hallways or corridors and staircases are the typical indoor space. Tran et al. (2019) and Cai & Fan (2021) focused their research on individual rooms and corridors. Yang et al. (2019) examined the relationships of staircase spaces. However, current research lacks an integrated approach that encompasses the entire building interior space. The management of knowledge from these state- of-the-art studies could significantly enhance the capacity for indoor space construction in gDTs and provide robust evidence for subsequent semantic information retrieval.

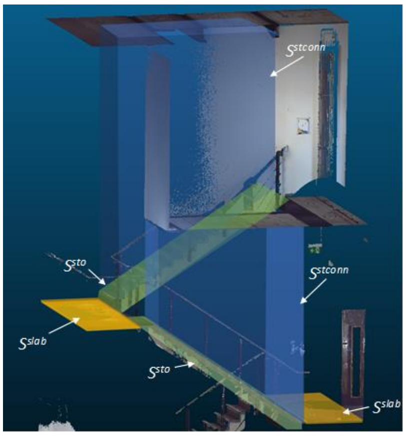


Figure 4 Definition of a stair connection space (Yang et al. 2019) *shown in the UOC case.*

Semantic rule-driven construction

The parametric and semantic properties of indoor spaces and components are pivotal information sources in current gDT construction practices (Jarząbek-Rychard & Maas, 2023; Liu et al., 2023). Semantic rules are commonly applied following the detection of building elements through shape rules, and the assessment of the topological relationships. For a comprehensive gDT, it is crucial to consider the unified and consistent characteristics of the entire indoor model, rather than focusing on isolated parts such as walls, doors or staircases (Yang et al., 2019). The importance of semantic rules is to generate the attributes for building elements beyond the parametric information that facilitate further utilisation. For example, the semantic rules are defined as classification and declassification to determine whether a room is an interior space or external space by Tran and Khoshelham (2020). This categorisation aids in supporting functions like indoor path planning. This research elaborates how semantic rules power the hidden element inference in the gDT construction process. Furthermore, this section reviews the key literature on the semantic rules applications and summarises the main semantic rules.

Hidden element inference

The knowledge-driven construction methods help with the prediction of space types (Hu et al., 2019; Xie et al., 2021). It provides a robust framework for building model construction, fostering a detailed and nuanced understanding of hidden objects and spaces within a 3D model. Chen and Liu (2023) underscored the importance of space syntax theory and design conventions, especially in the context of Singaporean dwelling units. This approach facilitates the inference of room types within a building.

Currently, most of the methods for inferencing building indoor environments focus on space or room predication. However, with semantic rules, further predication can be conducted. Hidden element inference, for instance, can be realised if default semantic information is available as prior knowledge. For example, the paths of the building pipeline could be predicted as guided by semantic rules by knowing the locations of the MEP (mechanical, electrical and plumping) terminals, and the spatial relationships between the equipment rooms and these terminals. However, there is less research contributing to this workflow.

Knowledge graph for gDT construction

A KG is a network of interconnected entities such as objects, events, concepts, represented in a graph format. In this graph, nodes typically symbolise entities or concepts, while edges represent the relationships between them (Bassier & Vergauwen, 2020). KGs utilize ontologies to define their structure, with Resource Description Framework (RDF) serving as a standard for describing the relationships (edges) within the graph.

The KGs can be utilised in building the knowledge of object identification and 3D model construction, and contributing to the knowledge fusion to the specific domains after the gDTs are constructed. Park and Cho (2022) develope the Entity-relationship diagram (ERD) to bridge the connections among point cloud, point, object, material, location. This semantic information is stored and converted in the point dimension to guide the information modelling of the building. The rules used to generate the gDTs can also be modelled in the KG, they can be regarded as the prior knowledge to facilitate the gDTs construction process. Hmida et al. (2013) propose the knowledge-based PCD detection and annotation strategy with an OWL ontology language, the Semantic Web Rule Language (SWRL) and the 3D processing built-ins. They divide the scene into various domain concepts depends on their characteristics and geometric relationships. This solution was tested in the railway objects. The knowledge from Geographic Information System (GIS) system is also integrated in the OWL structure for the further information enrichment to the 3D model. In addition, the approach of 'scan-to-graph' (STG) is proposed by Werbrouck et al. (2020). The STG method aims to integrate the Scan-to-BIM results with the sources of real-world asset via the semantic web technologies to overcome the uncertainties (occlusions, internal structure, etc.) of the collected data.

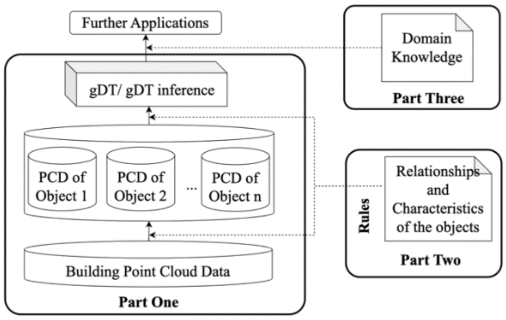


Figure 5 Conceptual framework of gDT construction KG

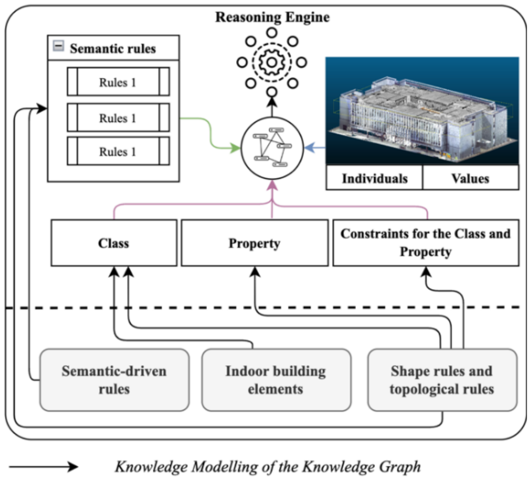


Figure 6 Key components of gDTs construction KG

Currently, there are less literatures focus on the KG development on both: 1) gDTs construction process and 2) the knowledge fusion for the further gDT applications, especially for the indoor building objects. Moreover, there are only limited rules are stored in the current KGs. In this paper, we developed the initial conceptual diagram of the knowledge graph to meet the two requirements with the considerations of shape, topological, and semantic rules. The structure of the ontology-based gDT knowledge graph is shown in Figure 5.

The first part of the framework focuses on transforming building PCD to object PCDs, and then the PCDs are converted to be the gDTs. The shape, topological, and semantic rules can be derived from the characteristics from the objects, and the relationships between the objects. Additionally, this step involves inferring parts of the building that might be occluded or hidden in the scan data based on the prior knowledge. The second part aims to establish the knowledge base (including the object characteristics and relationships) for the rules generations to facilitate the process of part one. This mainly involves the rules that discussed in the review parts of this paper. The establishment of the knowledge base of the rules can also be integrated with the existing ontologies that are commonly used, such as the Building Topology Ontology (BOT) (Rasmussen et al., 2020). In the third part, domain knowledge is utilised to enrich the gDTs for further development. This includes integrating the gDTs with other databases, such as energy consumption or environmental data, and refining the model for specific applications like simulation or analysis. The continuous improvement and updating of the gDTs are crucial to ensure they remain an accurate and useful representation of the physical structure.

Figure 6 presents the key components of the KG and the knowledge modelling mapping relationships from the building knowledge to the ontology-based KG. To enrich the ontology-based KG, building elements shall be converted to the classes, and the shape-driven and topology-driven rules will be presented as the classes, properties, constrains or the reasoning rules for the classes and properties of the KG. The semantic rules shall be expressed with the rules for the gDTs processing, semantic information enrichment and occluded gDTs inference. The objective of the proposed KG framework is to offer a comprehensive approach to enhance the gDTs construction workflow by incorporating the shape, topological, and semantic rules, ultimately paving the way for more sophisticated and accurate model-driven gDTs construction of building indoor environment.

Discussion

Understanding the indoor architectural rule and modelling it in a structured framework can support the recreation the building elements and the prediction of unreachable places in buildings. However, there are still some challenges in the current rule-based gDTs construction approaches. For example. the construction of the nonplanar surfaces and the non-Manhattan buildings are one of the big challenges (Tang et al., 2022; Tran et al., 2019). Moreover, the important connection building elements like doors and stairs are often manually added. Besides, some rules may only be applicable for specific types of buildings, such as churches or small-scale residential houses. The knowledge driven rule-based gDTs construction methods are important and should be further developed to compensate for the imperfections in the data collected from the physical world.

Conclusions

This paper has reviewed and analysed the rules in knowledge driven rule-based gDTs construction workflows from the perspectives of shape, topological, and semantic rules. It also proposes an initial conceptual KG framework for the structured representation of these rules and their interrelationships. Further development on the KG will encompass a broader range of building objects. Additionally, the continuous development of information reasoning and querying rules will also be conducted to facilitate the inference of hidden building objects and the enrichment of semantic gDTs.

Acknowledgments

The research leading to these results have received funding from the Horizon Europe UKRI Underwrite Innovate, under the grants G115919, G115676 and Digital Roads Prosperity Partnership (DR), supported by the Engineering and Physical Sciences Research Council (EPSRC) grant number [EP/V056441/1]. The authors gratefully acknowledge the collaboration of all academic and industrial project partners. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the institutes mentioned above.

Reference

Abdollahi, A., Arefi, H., Malihi, S., & Maboudi, M. (2023). Progressive Model-Driven Approach for 3D Modeling of Indoor Spaces. *Sensors*, *23*(13). https://doi.org/10.3390/s23135934

Bassier, M., Bonduel, M., Derdaele, J., & Vergauwen, M. (2020). Processing existing building geometry for reuse as Linked Data. *Automation in Construction*, *115*, 103180. https://doi.org/https://doi.org/10.1016/j.autcon.2020.103180

Bassier, M., & Vergauwen, M. (2020). Unsupervised reconstruction of Building Information Modeling wall objects from point cloud data. *Automation in Construction*, *120*. https://doi.org/10.1016/j.autcon.2020.103338

Becker, S., Peter, M., & Fritsch, D. (2015). Grammar-supported 3D indoor reconstruction from point clouds for "as-Built" BIM. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *2*(3W4), 17–24. https://doi.org/10.5194/isprsannals-II-3-W4-17-2015

Cai, Y., & Fan, L. (2021). An efficient approach to automatic construction of 3d watertight geometry of buildings using point clouds. *Remote Sensing*, *13*(10). https://doi.org/10.3390/rs13101947

Carrara, G., Kalay, Y. E., & Novembri, G. (1994). Knowledge-Based Computational Support for Architectural Design. *Proceedings of the 14th Annual Conference of the Association for Computer Aided Design in Architecture (ACADIA)*, *January 1994*, 5–12. https://doi.org/10.52842/conf.acadia.1994.005

Chen, C., & Liu, Z. (2023). A Fast Method for Identifying Room Configurations from Unit Boundaries in Existing Residential Buildings. *Buildings*, *13*(2). https://doi.org/10.3390/buildings13020357

Dou, Y., Li, T., Li, L., Zhang, Y., & Li, Z. (2023). Tracking the Research on Ten Emerging Digital Technologies in the AECO Industry. *Journal of Construction Engineering and Management*, *149*(3), 1–23. https://doi.org/10.1061/jcemd4.coeng-12290

Drobnyi, V., Li, S., & Brilakis, I. (2024). Connectivity detection for automatic construction of building geometric digital twins. *Automation in Construction*, *159*. https://doi.org/10.1016/j.autcon.2024.105281

Hmida, H. Ben, Cruz, C., Boochs, F., & Nicolle, C. (2013). *From 3D Point Clouds To Semantic Objects An Ontology-Based Detection Approach*. *June 2014*.

Hu, X., Fan, H., Noskov, A., Zipf, A., Wang, Z., & Shang, J. (2019). Feasibility of using grammars to

infer room semantics. *Remote Sensing*, *11*(13). https://doi.org/10.3390/rs11131535
Jarząbek-Rychard, M., & Maas, H. G. (2023). Modeling of 3D geometry uncertainty in Scan-to-BIM automatic indoor reconstruction. *Automation in Construction*, *154*(2023). https://doi.org/10.1016/j.autcon.2023.105002
Kellner, M., Stahl, B., & Reiterer, A. (2023). Reconstructing Geometrical Models of Indoor Environments Based on Point Clouds. *Remote Sensing*, *15*(18), 1–21. https://doi.org/10.3390/rs15184421
Khoshelham, K., & Díaz-Vilariño, L. (2014). 3D modelling of interior spaces: Learning the language of indoor architecture. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, *40*(5), 321–326. https://doi.org/10.5194/isprsarchives-XL-5-321-2014
Liu, W., Zang, Y., Xiong, Z., Bian, X., Wen, C., Lu, X., Wang, C., Marcato, J., Gonçalves, W. N., & Li, J. (2023). 3D building model generation from MLS point cloud and 3D mesh using multi-source data fusion. *International Journal of Applied Earth Observation and Geoinformation*, *116*(January), 103171. https://doi.org/10.1016/j.jag.2022.103171
Nikoohemat, S., Diakité, A. A., Lehtola, V., Zlatanova, S., & Vosselman, G. (2021). Consistency grammar for 3D indoor model checking. *Transactions in GIS*, *25*(1), 189–212. https://doi.org/10.1111/tgis.12686
Ning, X., Wang, Y., Hao, W., Zhao, M., Sui, L., & Shi, Z. (2015). Structure-Based Object Classification and Recognition for 3D Scenes in Point Clouds. *Proceedings - 2014 International Conference on Virtual Reality and Visualization, ICVRV 2014*, 166–173. https://doi.org/10.1109/ICVRV.2014.70
Pan, Y., Hu, Z., & Brilakis, I. (2023). Digital Twins and Their Roles in Building Deep Renovation Life Cycle. In *Palgrave Studies in Digital Business and Enabling Technologies: Vol. Part F1217* (pp. 83–96). Palgrave Macmillan. https://doi.org/10.1007/978-3-031-32309-6_6
Park, J., & Cho, Y. K. (2022). Point Cloud Information Modeling: Deep Learning–Based Automated Information Modeling Framework for Point Cloud Data. *Journal of Construction Engineering and Management*, *148*(2). https://doi.org/10.1061/(asce)co.1943-7862.0002227
Park, J., Kim, J., Lee, D., Jeong, K., Lee, J., Kim, H., & Hong, T. (2022). Deep Learning–Based Automation of Scan-to-BIM with Modeling Objects from Occluded Point Clouds. *Journal of Management in Engineering*, *38*(4). https://doi.org/10.1061/(asce)me.1943-5479.0001055
Rasmussen, M. H., Lefrançois, M., Schneider, G. F., & Pauwels, P. (2020). BOT: The building topology ontology of the W3C linked building data group. *Semantic Web*, *12*(1), 143–161. https://doi.org/10.3233/SW-200385
Tang, P., Huber, D., Akinci, B., Lipman, R., & Lytle, A. (2010). Automatic reconstruction of as-built building information models from laser-scanned point clouds: A review of related techniques. *Automation in Construction*, *19*(7), 829–843. https://doi.org/10.1016/j.autcon.2010.06.007
Tang, S., Li, X., Zheng, X., Wu, B., Wang, W., & Zhang, Y. (2022). BIM generation from 3D point clouds by combining 3D deep learning and improved morphological approach. *Automation in Construction*, *141*(December 2021). https://doi.org/10.1016/j.autcon.2022.104422
Thomson, C., & Boehm, J. (2015). Automatic geometry generation from point clouds for BIM. *Remote Sensing*, *7*(9), 11753–11775. https://doi.org/10.3390/rs70911753
Tran, H., & Khoshelham, K. (2020). Procedural reconstruction of 3D indoor models from lidar data using reversible jump Markov Chain Monte Carlo. *Remote Sensing*, *12*(5). https://doi.org/10.3390/rs12050838
Tran, H., Khoshelham, K., Kealy, A., & Díaz-Vilariño, L. (2019). Shape Grammar Approach to 3D Modeling of Indoor Environments Using Point Clouds. *Journal of Computing in Civil Engineering*, *33*(1). https://doi.org/10.1061/(asce)cp.1943-5487.0000800
Volk, R., Stengel, J., & Schultmann, F. (2014). Building Information Modeling (BIM) for existing buildings—Literature review and future needs. *Automation in Construction*, *38*, 109–127.
Wang, Q., Li, J., Tang, X., & Zhang, X. (2022). How data quality affects model quality in scan-to-BIM: A case study of MEP scenes. *Automation in Construction*, *144*. https://doi.org/10.1016/j.autcon.2022.104598
Werbrouck, J., Pauwels, P., Bonduel, M., Beetz, J., & Bekers, W. (2020). Scan-to-graph: Semantic enrichment of existing building geometry. *Automation in Construction*, *119*(January). https://doi.org/10.1016/j.autcon.2020.103286
Xie, L., Hu, H., Zhu, Q., Li, X., Tang, S., Li, Y., Guo, R., Zhang, Y., & Wang, W. (2021). Combined rule-based and hypothesis-based method for building model reconstruction from photogrammetric point clouds. *Remote Sensing*, *13*(6), 1–24. https://doi.org/10.3390/rs13061107
Yang, F., Li, L., Su, F., Li, D., Zhu, H., Ying, S., Zuo, X., & Tang, L. (2019). Semantic decomposition and recognition of indoor spaces with structural constraints for 3D indoor modelling. *Automation in Construction*, *106*. https://doi.org/10.1016/j.autcon.2019.102913
Yang, F., Pan, Y., Zhang, F., Feng, F., Liu, Z., Zhang, J., Liu, Y., & Li, L. (2023). Geometry and Topology Reconstruction of BIM Wall Objects from Photogrammetric Meshes and Laser Point Clouds. *Remote Sensing*, *15*(11). https://doi.org/10.3390/rs15112856

AUTOMATIC HVAC TOPOLOGY GENERATION USING BIM GEOMETRY CHECKING AND KNOWLEDGE GRAPH TECHNOLOGIES

Meng Wang[1], Georgios Lilis[1], Dimitris Mavrokapnidis[1], Kyriakos Katsigarakis[1], Ivan Korolija[1], Dimitrios Rovas[1]

[1]Institute for Environmental Design and Engineering, University College London, UK

Abstract

Developing digital twins of buildings can enhance the design process and streamline operational management. This study presents an automated framework designed to establish HVAC topologies of buildings using IFC-BIM data inputs. It achieves this by detecting geometric relationships among HVAC elements and generating a knowledge graph that captures their interconnections. A case study involving a recently constructed building in the UK was conducted to validate this framework. The resulting HVAC topology was compared against design drawings to identify any disparities, demonstrating the workflow's efficacy. This framework shows promising potential and generalization ability for application in various types of buildings.

Keywords

Building digital twins, Geometric relation check, Knowledge graph, HVAC topology

Introduction

A building digital twin is an advanced concept that can be used to improve building design, facilitate a building's operational management, and bridge the gap between building information and building simulation models. A seamless framework to establish Heating, Ventilation, and Air Conditioning (HVAC) topology automatically from BIM data exploits the latter potential and supports the automation of the generation of Building Energy Models, especially for the active components in building performance simulation. This HVAC topology generation involves multiple tasks such as the crafting of appropriate building-centric ontologies, the extraction of semantic relationships among HVAC elements, and finally the generation of semantic graphs conforming to the ontological schemes, capturing the HVAC topology.

Towards this direction, researchers established several building-oriented ontologies to enhance the building management system, e.g., BrickSchema developed by Brick Consortium (2019), to describe fluid flows through MEP distribution elements, e.g., Flow Systems Ontology (FSO) developed by Kukkonen et al. (2022), and to model connections between spaces and elements, e.g., Building Topology Ontology (BOT) Rasmussen et al. (2021). Seidenschnur et al. (2022) proposed a common data environment to represent a BIM-based model with the virtual commissioning, to establish a common data environment for HVAC design. They conducted two case studies to achieve the co-simulations between Modelica and EnergyPlus. Mavrokapnidis et al. (2023) used the Geometric Relation Checker (GRC) tool to identify pairs of IFC geometric instances and to generate an HVAC topology in knowledge graph form. They also briefly compared the applicability of BrickSchema, BOT, and FSO on data from the building domain. Many researchers tried to introduce ontologies as the bridge between building information and building energy modelling to achieve building simulation automatically. Wu et al. (2022) proposed an ontology-based framework for building performance simulation by integrating various data sources including building information modelling, building management systems, and weather stations. Their results showed a considerable reduction in modelling time but appropriate assumptions could not be avoided.

Different data sources of building information can be introduced into the HVAC-related Building Information Modelling to Building Energy Models (BIM2BEM) transformation. Not limited to IFC-based BIM models, CityGML can also serve as a data source and be integrated into the BIM2BEM workflow. Deng et al. (2023) developed a tool for district-scale urban building energy modelling in which the geometric information of buildings was extracted from CityGML data. To facilitate the BIM2BEM generation process, González et al. (2021) integrated BIM and BEM methodologies to conduct an energy performance analysis of a residential building. Their work mainly depended on the Revit platform considering many variables including lighting efficiency, plug-load efficiency, and HVAC systems. In their study, an architectural BIM model provided the majority of necessary data, enabling non-geometric information to be derived from experimental design. Ramaji et al. (2020) devised an extension for OpenStudio that converts IFC-based BIM models into building simulation models. Validation of their algorithm was conducted through a case study involving a three-story office building in the US, revealing several advantages and limitations. Their findings highlighted certain drawbacks of the current semi-automated BIM2BEM process, particularly when dealing with large and intricate structures.

As the number of BIM2BEM-related studies grew, researchers started to summarize the related developments investigating potential topics that require to be addressed in the future, e.g., Kamel and Memari (2019) and Farzaneh et al. (2019). They concluded that there was still work to be done, to reach a seamless workflow from BIM to BEM. In this context, two specific tasks were proposed to facilitate the development of an automatic BIM2BEM framework. The first task focuses on integrating models, while the second task revolves around creating a process map to seamlessly connect work and data flows. Researchers primarily concentrated on bridging the gap between BIM and BEM architectural data focused on the generation of the second-level space boundary topology Ying and Lee (2021), Rose and Bazjanac (2015), Lilis et al. (2017), rather than addressing data translation issues specific to MEP data, like those appearing in the HVAC data domain. A potential explanation for this decision might stem from the difficulties encountered when attempting to automatically generate an HVAC topology directly from BIM data—a crucial step in translating HVAC information within a BIM2BEM process. Achieving this automated HVAC topology generation requires employing multiple ontologies describing mechanical equipment (e.g., BrickSchema Brick Consortium (2019)) and material flows (e.g., FSO ontology Kukkonen et al. (2022)), within MEP networks. Additionally, various issues and errors in BIM models lead to challenges in achieving an accurate HVAC topology, often requiring labour-intensive manual corrections.

To tackle these challenges, this work develops a framework that automatically generates the HVAC topology of a building that can be used to mirror real-world conditions. To demonstrate its efficacy and feasibility, we apply this framework to a recently built structure with complex HVAC systems. Concurrently, the proposed framework seeks to enhance any automated BIM2BEM conversion process by providing a fast and efficient translation method for possible MEP data that could be part of this process.

To introduce this approach the structure of the paper is as follows: The methodology section outlines the framework and its pertinent stages. The case study section provides an in-depth overview of the chosen building and its HVAC setup. Following this, the results and discussions section delves into the generated HVAC topology, highlighting the strengths and limitations. Finally, the conclusion section summarizes the findings and proposes avenues for future research.

Methodology

The proposed HVAC topology generation process involves three major steps as illustrated in the block diagram of Figure 1. These steps are: (1) Extraction of information from BIM (IFC) files related to architectural (ARC) and mechanical, electrical, and plumbing (MEP) data, (2) Checking of the extracted BIM content to detect geometric (2a) and semantic (2b) relations among HVAC elements and space volumes, and (3) Establishment of the HVAC topology via knowledge graph generation and path-finding.

- In step 1 (Extraction), the IFC geometry exporter is used to extract the geometric representations (parametric and non-parametric) of the building space and MEP elements from the ARC-IFC and MEP-IFC files. These representations are then used to form XML files that are used as input to a geometric relation-checking (GRC) tool. These XML files contain the detailed geometric representations of each entity to be checked, together with their IFC global unique ID. Additionally, in this step, the IfcOpenShell is used for parsing the MEP IFC file to extract the semantic connectivity data.

- In step 2 (Checking), connectivity and containment entity pairs are extracted via geometric (step 2a) and semantic (step 2b) relation checks. In case semantic relations are absent from the BIM data, the GRC tool is applied to identify geometric relations among the solid representations of the MEP elements (containment, clash, and adjacency) and infer the missing semantics. Finally, the Oriented Bounding Box (OBB) method is introduced to address the HVAC elements that can not be handled by the GRC tool. Hence, a hierarchy of checks (BIM context - GRC - OBB) is formed to extract all connectivity and containment relations among the involved elements.

- In Step 3 (Establishment), an ontology-based knowledge graph is created to represent the HVAC elements and their connectivity, leveraging the outcomes of Step 2. Subsequently, a path-finding method is implemented to establish logical links from terminals to systems and from pumps to units, thereby simplifying the knowledge graph of the HVAC topology.

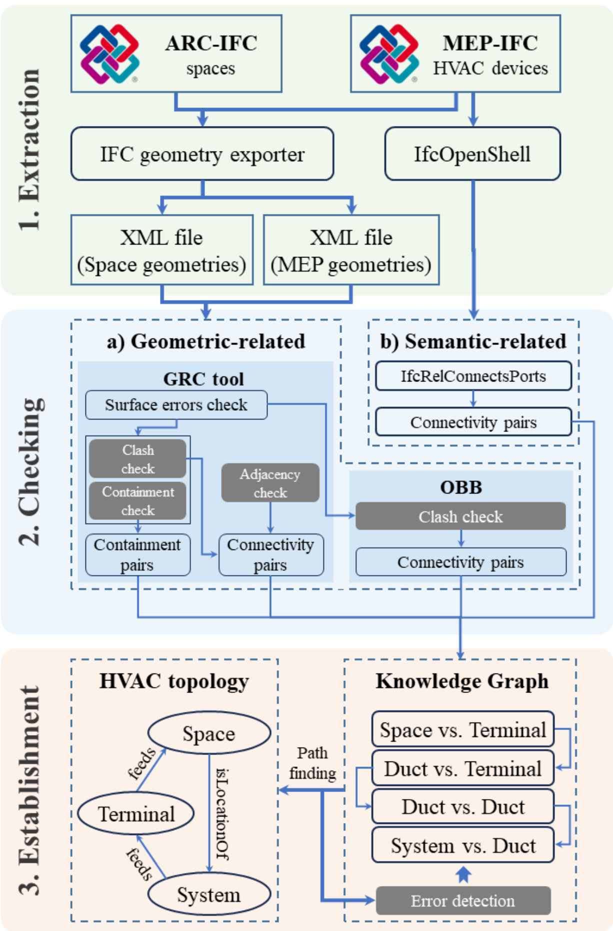


Figure 1: Introduced three-step framework

Information extraction from BIM - Step 1

The proposed workflow used two types of BIM models, i.e., ARC-IFC and MEP-IFC files. From the ARC-IFC file, only the class of IfcSpace was extracted and transferred to a set of entities in the knowledge graph that are detailed in the following sections. To extract the MEP information related to the duct network, data from multiple IFC classes were gathered, including (1) the classes of IfcAirTerminal, IfcDuctSegment, IfcDuctFitting, IfcAirToAirHeatRecovery, IfcBuildingElementProxy that are used to construct the air loop network of the HVAC system, and (2) the classes of IfcPump, IfcValve, IfcPipeFitting, IfcPipeSegment, and IfcFlowMeter that are used to construct the water loop network of the HVAC system. Instances of the previous classes that are contained in the IFC files were mapped, in a one-to-one manner, to entities in the knowledge graph.

Furthermore, the IFC Geometry Exporter Katsigarakis et al. (2021), was used to extract geometric descriptions of the elements from input IFC files, which is the prerequisite to implement the geometric relationship check that is conducted using low-level C++ geometric routines. The exporter extracted serialized geometric descriptions from the two types of IFC files, transformed them into a hierarchical tree-structured format, and created two XML files containing these transformed descriptions. The generated XML files were then used as input to the Geometric Relation Checker (GRC) tool, to identify the geometric relationship among the building's MEP and space elements.

Geometric relation checking - Step 2a

Three types of geometric relationships are defined: containment, clash, and adjacency to deduce semantic connections among the elements within the ARC and MEP IFC files. Detailed explanations of these relationships are provided in subsequent subsections.

Containment

A containment relationship between two solids exists when the surfaces enclosing one solid are entirely within the 3D boundaries of the other. Consequently, this relationship serves as a means to identify semantic links between spaces and terminals. It enables the detection of the specific space to which an HVAC terminal belongs. This association is particularly relevant, as numerous air terminal units, such as diffusers and grilles, are typically entirely enclosed within the volume of the building space, as depicted in Figure 2 a).

Clash

A clash relationship between two solids exists when surfaces or parts of the surfaces of the boundary representation of one solid are entirely within the 3D boundaries of the other. Figure 2 b) illustrates the clash relationship between two elements. This is the most common geometric relationship among many pairs of elements within the air loops, including spaces vs. terminals, terminals vs. ducts, ducts vs. ducts, and ducts vs. systems. Similarly, within the water loops, it also extends to relationships between terminals vs. pipes, pipes vs. pipes, pipes vs. pipe fittings, and pipes vs. pumps. Clashes can be used to infer connectivity between elements as well as containment when a portion of one element is inside another.

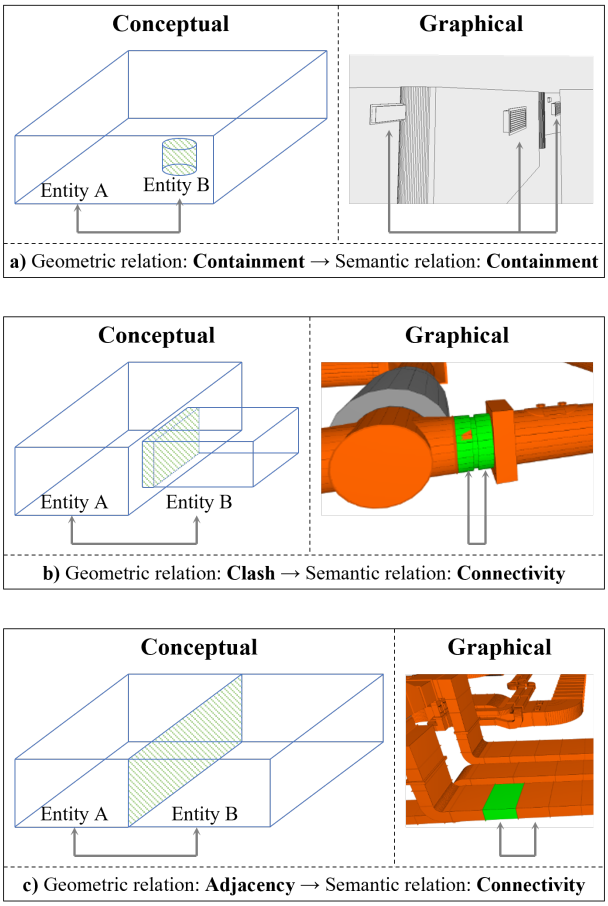


Figure 2: Conceptual and graphical illustration of a containment relationship.

Adjacency

Two solids share an adjacency relationship when the surfaces of one solid's boundary representation lie in the same plane and intersect with the surfaces of the other solid's boundary representation. Figure 2 c) illustrates the adjacency between two elements that frequently appear in pairs of ducts and ducts and systems, especially in rectangle-shaped ducts. Unlike the previous geometric relationships, a threshold value (0.01 m in this case) is used to check adjacency. This value sets the allowable distance between the planes of the intersecting surfaces that define the adjacency. If the distance between these planes exceeds this value, intersection and adjacency are not detected.

GRC tool

To detect the presence of the previous geometric relations among the solid representations of the BIM elements, a geometric relationship-checking tool, named Geometric Relation Checker (GRC), is introduced. The performance of this tool is validated on BIM elements from the architecture and construction domains. GRC is a domain-agnostic application that can be used to identify predefined geometric relationships in any IFC element pair. To achieve the necessary execution speed, the GRC tool is implemented in C++ and uses binary space partitioning to carry out fast Constructive Solid Geometry (CSG) set operations on polyhedra, required for clash detection, using multithreaded executions. GRC can not only identify relationships including containment, clash, and adjacency, but also the surfaces at which the related elements intersect or touch each other. Moreover, a threshold of the gap between each element can be set to define the adjacency and to address potential element mismatching issues.

Oriented bounding box

If surface errors are present in the solid representations of the compared BIM elements, certain Geometric Relation-checking processes like clash and containment detection cannot be performed. Surface errors refer to surface inversions and missing surfaces that form wholes in the boundary representations. To address such scenarios, the study introduced the oriented bounding box (OBB) method to carry out the relevant geometric relationship checks. The GRC tool identifies these cases and automatically routes the problematic elements to the OBB-based clash-checking process. Furthermore, inflation is used during the OBB-based clash-checking process to expand the bounding box in the direction of each face, with the inflation threshold aligned with the gap threshold used by the GRC tool. This method ensures that the OBB-based clash-checking process complements the GRC method, facilitating comprehensive geometric relationship checks for all elements within the IFC file.

Entity linking via semantic checking - Step 2b

Although numerous IfcRelationship entities exist within the BIM context to depict relationships among MEP entities, they often fail to encompass all entities within air loops and water loops. The previously described geometric relationship check processes were introduced to address this gap, enabling the deduction of missing semantics between pairs of MEP elements. Thus, the proposed GRC-OBB-BIM workflow streamlines the extraction of existing and potential containment/connectivity pairs, enhancing the linkage between HVAC elements and helping to finalize the generated HVAC topology.

HVAC topology establishment - Step 3

The HVAC topology establishment process was divided into knowledge graph generation, pathfinding, and error detection, as analyzed in the following subsections. Firstly, knowledge was transferred directly on the basis of the detected geometric relationships to the knowledge graph. At the same time, the path-finding algorithm was used to enrich the semantic linking between the critical HVAC components (e.g., terminal2VAVbox, VAVbox2system, system2pump). The knowledge graph generation was based on the integration of two open-source ontologies, named BrickSchema Brick Consortium (2019) and FSO Kukkonen et al. (2022), as described next.

Knowledge graph generation

BrickSchema, developed by Brick Consortium, standardizes semantic descriptions of the buildings' assets/devices and their inner relationships, which was introduced into this work to represent the HVAC components with a series of Brick entities (`brick:Room`, `brick:Terminal_Unit`, `brick:AHU`, `Brick:DOAS`, `brick:Chiller`, `brick:HX`, `brick:Pump`, and so on). As a supplement, the Flow Systems Ontology (FSO), which focuses on interconnected systems and their energy flows, was incorporated to model both pipework and ductwork segments by using FSO entities (`fso:Segment` and `fso:Fitting`). Furthermore, within the scope of geometric relationships among HVAC elements, `brick:isAssociatedWith` is used to denote connectivity relationships, that is, connectivity pairs resulting from clashes and adjacencies because these detected geometric relationships were downstream/upstream direction agnostic. Regarding the containment pairs, `brick:isLocationOf` was introduced to indicate the containment relationships with specific directions (e.g. space - isLocationOf - system).

Path-finding

In this study, a path-finding algorithm was utilized to discover routes between distinct entities within the knowledge graph by analyzing multi-hop connections. These routes were established among terminal units (such as air terminals), source systems (such as AHUs), and potential nodes (like VAV boxes) that might exist between the source and terminal units within air loops. Similarly, connections were formed between AHUs/FCUs, pumps, and heat exchangers or chillers within the water loops. These routes collectively formed the final HVAC topology. In this topology, the `brick:isFedBy` relation was used to replace the `brick:isAssociatedWith` relation between the critical HVAC components connected via path-finding. The flow direction along these connection links was determined by the ID of the connected elements, wherein each element's ID contains a specific string indicating whether it pertains to an air supply or an air return unit. Figure 3 illustrates how path-finding contributes to the generation of an HVAC topology instance.

In terms of algorithmic methodology, the Depth-First Search (DFS) path-finding algorithm was introduced to convert the generated knowledge graph into the final HVAC topology. DFS prioritizes depth, exploring as deeply as possible along each branch of the graph before backtracking. Additionally, DFS offers the flexibility to omit particular nodes from the search process. For example, when seeking a route from Terminal 1 to AHU 1, other AHUs can and should be excluded. It is essential to prevent path-finding processes from entangling within water loops and air loops. Two simple approaches to accomplish this are splitting the knowledge graph into distinct subgraphs or assigning a specific attribute to knowledge graph links to distinguish connection relationships.

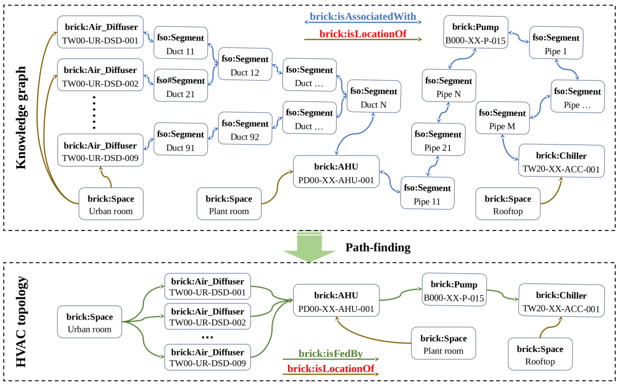


Figure 3: Illustration of the path-finding process to transform the knowledge graph into the HVAC topology.

BIM error handling

Although the preceding workflow's processes are executed automatically, errors within the BIM data can result in either zero or multiple paths detected, from the supply to terminal units. Alternatively, it might be found that one air terminal unit is linked to multiple AHUs or FCUs. These errors manifest in four distinct types, as illustrated in Figure 4, which include:

(a) An abnormal clash is an incorrect intersection among MEP elements that is due to design errors.

(b) A missing element is the absence of an MEP element in a sequence of connected elements.

(c) A geometric mismatch is a misalignment of MEP elements that leads to connection absence.

(d) A space containment issue occurs when a building space volume does not contain or does not intersect, at least, with a solid geometric representation of a terminal unit.

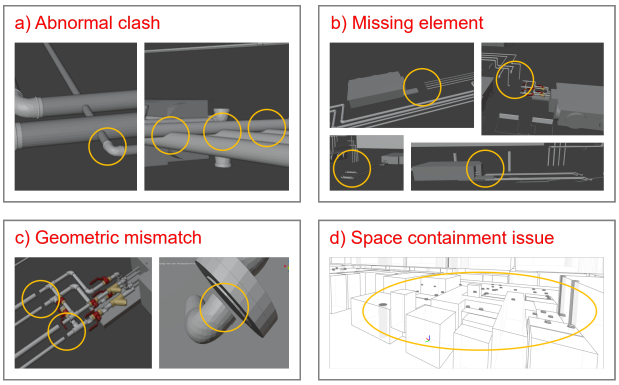


Figure 4: BIM error-type illustration examples

Error types (a), (b), and (c) impact the established routes from the supply to the terminal units within the HVAC domain, while error type (d) influences the containment connection between terminal units and their corresponding building spaces. Visual inspection is applied on the detected paths from supply to terminal units, and on the building space volume geometries, to pinpoint these errors. Subsequently, manual corrections are implemented on the produced knowledge graph to rectify the affected graph nodes and edges.

Case study

The introduced process was applied to BIM data from a new building, named One Pool Street (OPS) located at the UCL East Campus. OPS is a multi-function building that uses a state-of-the-art energy supply and management system. This work concentrated on the automatic generation of the HVAC topology of the podium part of the OPS building. The HVAC topology of this building part is ideal for demonstration purposes, as its energy supply system is complex and includes a set of diverse ventilation units. For cooling purposes, several Air Handling Units (AHUs) are installed with chilled water and hot water units that supply thermostatic air via hundreds of air terminals during the summer and winter periods. Additionally, Fan Coils (FCUs) are also used in some rooms to meet their cooling and heating requirements. Door heaters (DHs) and air curtains (ACs) are also installed for heating purposes only. Besides, Mechanical Ventilation Heat Recovery (MVHR) systems are installed to recycle the waste heat and improve energy performances in several building spaces. Consequently, the manual generation of the HVAC topology for these complex systems is a time-consuming and error-prone process. Hence, the OPS podium was selected as a case study to illustrate the effectiveness of the proposed workflow.

Figure 5 illustrates the 3D representations of the contents found in the ARC-IFC and MEP-IFC files, which encompass both air loops and water loops of the OPS podium. In terms of data quality, the Level of Development (LOD) for the IFC4 file varies between LOD 300-400 (following US standards) and LOD 4-5 (according to UK standards). This file comprises over 22,000 MEP elements, posing a significant challenge due to memory limitations on standard computers when handling MEP-IFC files. To address this challenge, it was imperative to divide the MEP-IFC file of the OPS building into two smaller IFC files, specifically focusing on air loops and water loops.

Results and discussions

In this section, the performance of the proposed workflow and its limitations are also discussed.

Established HVAC topology

The proposed workflow was conducted on a laptop computer, and the geometric relationship checking process consumed the most computing resources of the workflow. The geometric relationship checking between air terminals and spaces took several minutes and involved approxi-

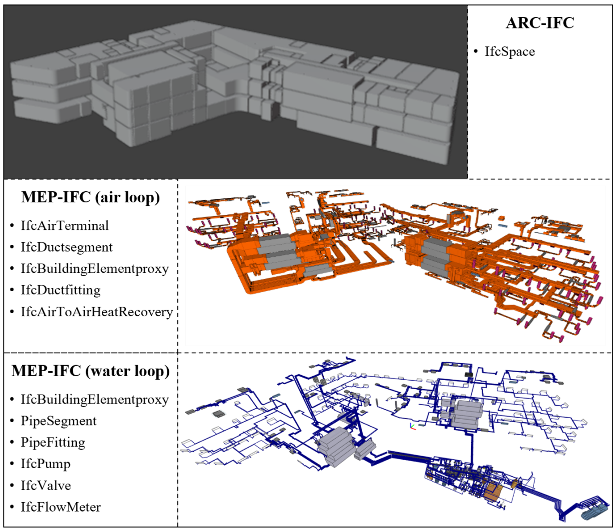
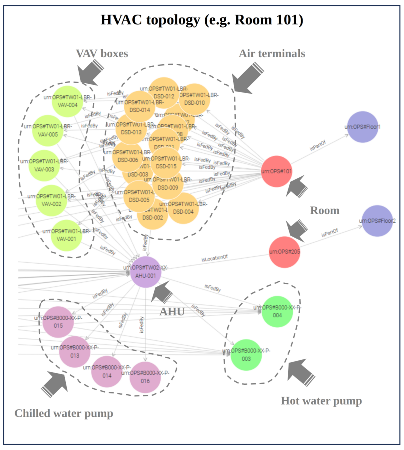


Figure 5: Visual representation of the content of ARC-IFC and MEP-IFC files.

mately 84,000 element pairs. The checking process among the HVAC elements took a few hours and involved approximately 38,000,000 pairs within air-loop elements and 101,000,000 pairs within water-loop elements. In addition, the path-finding process to link terminals, systems, and pumps, took dozens of minutes.

Figure 6 displays a part of the generated HVAC topology referring to Room 101. In this figure, the return air terminals are hidden because they are not as significant as the supply air terminals from the energy simulation perspective. Within the air loops, the air terminals (displacement diffusers) were identified successfully, and the VAV boxes within the duct network were also detected, defining the paths from the system (AHU) to air terminals (DSDs). Since the air terminals had IDs with the special string (DSD: displacement supply diffuser), the directional relations `brick:Feeds/brick:isFeedBy` were determined, linking the air terminals with the system. In the context of water loops, Figure 6 highlights the connection of four chilled water pumps and two hot water pumps with this AHU. However, the uncertain status of the valves, particularly the bypass valves, presents a challenge. This uncertainty makes it impossible to ascertain which pumps are on standby and which are operational, or to determine their positioning relative to the AHU, whether upstream or downstream. Hence, in the HVAC topology, the directional relations `brick:Feeds/brick:isFeedBy` were determined to illustrate the links between pumps and AHUs. Similarly, this approach was also applicable to establishing links from heat exchangers to pumps. The entity inventory of the original IFC files and the HVAC topology are both depicted in the right portion of Figure 6. The podium section of the OPS building comprises approximately 8,800 air-loop MEP entities and over 13,900 water-loop MEP entities, which are integrated into the proposed workflow. Approximately 7,000 of the air-loop entities are modelled using boundary representations and require ad-



Entity inventory

IFC Class	Number
IfcSpace	167
IfcAirTerminal	503
IfcDuctSegment	7804
IfcBuildingElementProxy	425
IfcAirToAirHeatRecovery	13
IfcValve	729
IfcPipeFitting	6455
IfcPipeSegment	6614
IfcFlowMeter	17

IfcClass to Topology

BrickSchema	Number
brick:Room	167
brick:Air_Diffuser	288
brick:VAV	140
brick:FCU	54
brick:AHU	9
brick:DOAS	9
brick:Chiller	1
brick:HX	3
brick:Hot_Water_Pump	8
brick:Chilled_Water_Pump	4

Figure 6: The generated HVAC topology for Room 101 (the exhausted air terminals are hidden) and the inventory of IFC entities and relevant entities.

ditional checks by the GRC tool for potential surface errors. However, this step will become unnecessary in the future if boundary representations are substituted with parametric ones.

MEP element - Space containment issues

Accuracy is one of the critical indicators for validating the performance of this HVAC topology-generation workflow. In this case, the HVAC design drawings of the OPS building were also obtained and acted as a reference point for validation. Figure 7 shows the statistical results of the semantic linking between the air terminals and the podium spaces. More than 70% of the air terminals were linked to their correct spaces. However, about 18% of them were not linked to any spaces because there are design errors in the geometry of the spaces in the IFC file, e.g. some spaces are not extended high enough to reach the ceiling surfaces. Additionally, elements with surface errors caused delays when conducting the GRC executions since pairwise element clash checking requires at least one of the elements in a pair, to be surface error-free. Meanwhile, nearly 8.5% of the air terminals intersected with more than one space, requiring manual work to decide the correct space. The reasons for these misalignments are mainly surface errors (requiring the OBB check) and design errors (human errors), both of which are related to the data quality of the geometric representations of the elements. The suggested workflow can efficiently identify the geometric containment relationships between air terminals and spaces, as long as the input IFC file reaches LOD-4 level or beyond.

MEP element connectivity issues

Another critical aspect of an HVAC topology pertains to the connection between air terminals and systems. Figure 8 illustrates the efficacy of the proposed workflow in identifying paths from air terminals to systems within the air loops. Initially, only half of the air terminals were accurately linked to the correct AHU, MVHR, and FCU, as depicted in Figure 8 (a), before addressing the errors. This discrepancy arose due to the dense arrangement of ducts,

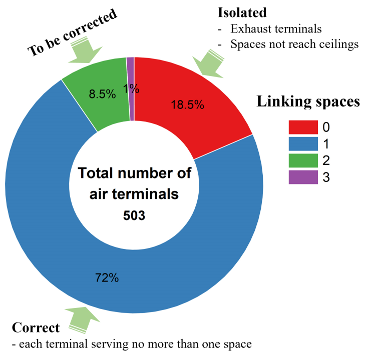


Figure 7: The distribution of air terminals and the number of their linking spaces.

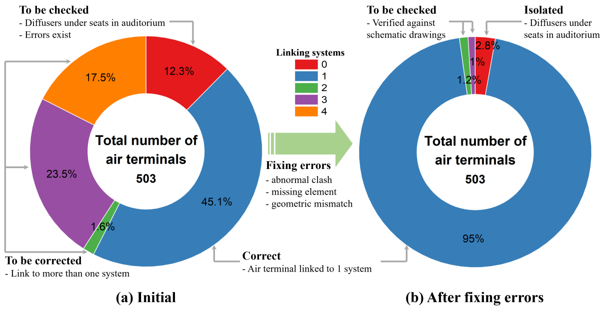


Figure 8: The distribution of air terminals and the number of their linking systems.

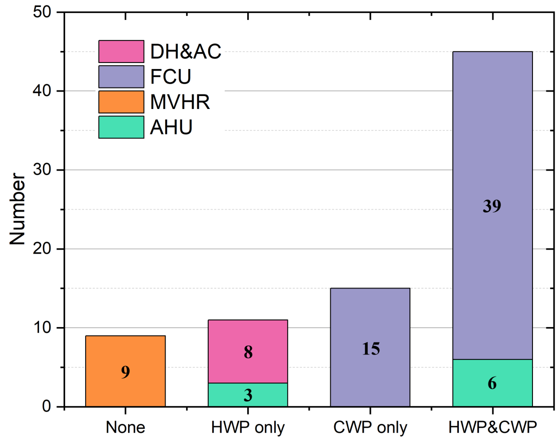


Figure 9: The distribution of units and their HWP/CWP connections.

often closely positioned near the systems. The GRC tool's capability to manage such placements is hindered by surface errors affecting these elements. Consequently, in situations where both elements of a pair exhibit surface errors, rendering clash detection impractical, the OBB method serves as an alternative, albeit less precise than the GRC tool. Furthermore, around 12% of air terminals faced challenges in establishing connection paths with systems. This hurdle arose from certain terminals being designed independently from other air-loop elements, such as diffusers situated beneath seats in auditoriums. Additionally, some instances involved design errors, evident in Figure 4 b) and c), where gaps exist between these elements and adjacent ducts. Adjusting the GRC threshold for adjacency checking proved effective in addressing this issue, although it carries the risk of detecting connections that do not actually exist.

Moreover, the initial results revealed that approximately 40% of air terminals remained connected to more than one system due to BIM design errors as previously defined. Most of these instances stemmed from the alternative OBB method, which erroneously identified nonexistent intersections between bend ducts positioned in parallel and very close proximity to each other. After manual efforts to correct the detected errors in the generated knowledge graph, the path-finding process was reexecuted on the updated knowledge graph. The statistical results of this process can be found in Figure 8 (b). Notably, 95% of the air terminals were accurately linked to the systems (AHU/MVHR/FCU), while only 2% still required verification against schematic drawings. It is advisable to verify the remaining 2.8% (isolated air terminals) against schematic drawings to ensure the completeness of the HVAC topology.

Furthermore, Figure 9 illustrates the results (after error correction) of the water-loop units in the established HVAC topology. In the OPS building, there are a total of 80 water-loop units, including AHU, MVHR, FCU, DH, and AC, within the water loops, supplied by Hot Water Pumps (HWPs) or Chilled Water Pumps (CWPs). Using the path-finding method, it is possible to determine connections between each unit and HWPs, CWPs, or both.

As depicted in Figure 9, among the 9 AHUs, 6 are linked to both HWPs and CWPs, while the remaining 3 are solely connected to HWPs. None of the 9 MVHRs are connected to either HWPs or CWPs. Of the 54 FCUs, 39 are connected to both HWPs and CWPs, with the remaining 15 solely linked to CWPs. All DHs and ACs are connected exclusively to HWPs. Additionally, this path-finding method could be extended to examine the 93 Radiators (RADs) in the OPS building for their connections to HWPs, although this is not detailed here due to length constraints. Finally, all mentioned results have been validated against the schematic drawings.

Limitations and implications

The effectiveness of the framework was assessed through a case study involving a complex HVAC system, revealing various challenges related to the accuracy and completeness of the HVAC topology when dealing with imperfect MEP-BIM data. Initial inaccuracies were attributed to human errors in BIM design and surface errors in the solid geometric representations of certain MEP-BIM elements, which hindered the effectiveness of the GRC tool. Additionally, a complication arose from adjacent duct entities in the BIM model that were not physically connected, a discrepancy not detectable through geometric relationship checking alone. While this study successfully identifies and locates these errors, manual correction remains necessary and time-consuming.

Conclusions

To automate the creation of an HVAC topology, this study introduced a framework utilizing ARC-IFC and MEP-IFC files as BIM input data sources. Initially, semantic checks were conducted using relevant IFC relationship classes to extract connectivity pairs. A hierarchical checking workflow was then employed to infer missing semantic relations. In the workflow's initial stages, the Geometric Relation Checking (GRC) tool facilitated pairwise adjacency, clash, and containment relation checks among MEP and space BIM entities. If surface errors were reported in the geometric representations of BIM elements by the GRC tool, the clash and containment detection were carried out using the OBB process. Ultimately, this process yielded a set of connectivity and containment pairs. Subsequently, a knowledge graph-based HVAC topology, encompassing air loops and water loops, was generated by applying the DFS path-finding algorithm to the obtained connectivity and containment pairs. A case study involving a newly constructed building was conducted to test the framework's performance, demonstrating its feasibility in various and intricate modern HVAC systems.

In the future, significant improvement can be achieved by incorporating a module that converts the IFC geometries to watertight solid boundary representations. This can speed up the process substantially by reducing the need for surface error checking. Meanwhile, beyond the current building case, the plan is to implement the framework in other scenarios to enhance its generalization capabilities. Besides, further investigation will be undertaken to determine how the generated HVAC topology can be leveraged to enhance BEM models that contain information related solely to the building fabric. This effort aims to support the automatic establishment of active components within the HVAC domain, thereby supporting the development of BEM2BEM research.

Acknowledgement

This present study was supported by the Horizon Europe Project (101069658).

References

Brick Consortium (2019). Brick A uniform metadata schema for buildings. `https://brickschema.org/`.

Deng, Z., Chen, Y., Yang, J., and Causone, F. (2023). AutoBPS: A tool for urban building energy modeling to support energy efficiency improvement at city-scale. Energy and Buildings, 282:112794.

Farzaneh, A., Monfet, D., and Forgues, D. (2019). Review of using building information modeling for building energy modeling during the design process. Journal of Building Engineering, 23:127–135.

González, J., Soares, C. A. P., Najjar, M., and Haddad, A. N. (2021). BIM and BEM methodologies integration in energy-efficient buildings using experimental design. Buildings, 11(10):491.

Kamel, E. and Memari, A. M. (2019). Review of BIM's application in energy simulation: Tools, issues, and solutions. Automation in construction, 97:164–180.

Katsigarakis, K., Lilis, G. N., and Rovas, D. (2021). A cloud IFC-based BIM platform for building energy performance simulation. In Proceedings of the 2021 European Conference on Computing in Construction. University College Dublin.

Kukkonen, V., Kücükavci, A., Seidenschnur, M., Rasmussen, M. H., Smith, K. M., and Hviid, C. A. (2022). An ontology to support flow system descriptions from design to operation of buildings. Automation in Construction, 134:104067.

Lilis, G. N., Giannakis, G., and Rovas, D. V. (2017). Automatic generation of second-level space boundary topology from IFC geometry inputs. Automation in Construction, 76:108–124.

Mavrokapnidis, D., Lilis, G. N., Katsigarakis, K., Korolija, I., and Rovas, D. (2023). Semi-automated extraction of HVAC system topology from imperfect Building Information Models. IBPSA.

Ramaji, I. J., Messner, J. I., and Mostavi, E. (2020). IFC-based BIM-to-BEM model transformation. Journal of Computing in Civil Engineering, 34(3):04020005.

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., and Pauwels, P. (2021). BOT: The building topology ontology of the W3C linked building data group. Semantic Web, 12(1):143–161.

Rose, C. M. and Bazjanac, V. (2015). An algorithm to generate space boundaries for building energy simulation. Engineering with Computers, 31:271–280.

Seidenschnur, M., Kücükavci, A., Fjerbæk, E. V., Smith, K. M., Pauwels, P., and Hviid, C. A. (2022). A common data environment for HVAC design and engineering. Automation in Construction, 142:104500.

Wu, Z., Cheng, J. C., and Wang, Z. (2022). An Ontology-Based Framework for Building Energy Simulation in the Operation Phase. In International Conference on Computing in Civil and Building Engineering, pages 351–366. Springer.

Ying, H. and Lee, S. (2021). Generating second-level space boundaries from large-scale IFC-compliant building information models using multiple geometry representations. Automation in Construction, 126:103659.

NAVIGATING BUILDING STANDARDS AND BUILDING WARRANT PROCESSES IN SCOTLANDBUILDING INFORMATION REPOSITORIES

Daria Belkouri[1], Ho Man Yau[1], Theodoros Dounas[2], Jamie Yorkston[1]
[1]Robert Gordon University, Aberdeen, United Kingdom
[2]University of Antwerp, Antwerpen, Belgium

Abstract

Our paper explores the improvements needed in navigating building standards and workflows for building warrant processes in Scotland. We examine this optimization through the lens of user experience, information and data flow, with the directive of improving the service provided to citizens by the Scottish government. We utilized the development of a web application and a digital process map as vehicles for engagement, developing knowledge for improvements and optimization through qualitative methods. The paper concludes with a discussion and proposals for improving the process, including the use of linked data, BIM, ontologies, and blockchain in an upgrade of the process.

Introduction

Governments establish laws, rules, and regulations relating to the built environment in order to safeguard the well-being and safety of their citizens. In the UK, the multitude and complexity of technical requirements and associated legislation for reviewing construction designs and providing building warrants as well as inspections during construction ensure that the building warrant issued for the structure is constructed according to design (Fauth et al. 2023). Acquiring a building warrant holds significant importance in any construction project, serving as a foundational requirement for its lawful execution. On a global scale, the issuance of construction permits remains a complex and error-prone procedure, often characterised by a substantial time commitment and influenced by subjective judgments (Malsane et al., 2015). One contributing factor to the limited implementation of quick complete automation in public administration is the absence of a comprehensive legislative framework, as well as a lack of societal acceptability. Furthermore, building officers responsible for overseeing building permit processes express reservations about the integration of digitization and the use of digital tools such as Building Information Modelling (BIM) models (Fauth, 2023). Additional reasons might be the lack of comprehensive and commonly agreed-upon standards for the processing of information, the lack of interoperability between data pools and databases, and the need to keep the process accessible to non-experts. For this particular reason, the acquisition of a building warrant holds significant importance in any construction project, serving as a fundamental requirement for the lawful execution of the project.

This research project involves the development of a map for navigating the building standards (BS) and building warrant (BW) process of the Scottish government, in the auspices of a larger project that aims to improve the digital delivery of the building standards and regulations. We used two vehicles of engagement to enhance our understanding of the building standards and the building warrant process: a prototype of a web application for the BS and a procedural map for the BW. The paper examines the knowledge and output from these two engagement tools.

Background

Building Standards in Scotland and the United Kingdom – relationship between government and local authorities

The digitisation processes of managing data in the Architecture, Engineering, and Construction (AEC) industry have been recently widely encouraged (Hobeika et al. 2022; European Commission 2021). The Building Standards division in Scotland is driving the digital transformation of building standards and the building warrant process. The traditional way of manually handling and assessing the building warrant applications can be prone to error - an activity that requires a high level of effort from all parties from different disciplines involved (Ullah et al. 2022). There has been a need to improve the building warrant completion process recognized, and this research acknowledges the many disjointed data landscapes – i.e fragmentation between the planning process and design, and a lack of orchestration between data pools/data silos needed to create a well-coordinated digitized system that has all the required functions to smoothly handle the process of building permit acquisition – from online submission, review and management of the applications. In 2017, the e-development platform was introduced (and managed by the Scottish Government in partnership with all Scottish local and planning authorities), providing the ability to digitally submit planning and building warrant applications. The portal is a complimentary digital service established by the Scottish Government in collaboration with all Scottish local authorities. It enables users to fill out and submit online building warrant applications, completion certificates and other relevant forms relating to the entire building standards system process. Yet, the current process of applying for approval documents - building warrants and completion certificates, comes with a myriad of complexities associated with the process that involves many stages and stakeholders along the journey. The e-development portal acts as a digital transition point between the applicants and local authorities but comes with its limitations (for example the size of data that can be uploaded onto the portal or communication of the inspection services) that need to be overcome by using complimentary data transfer options and direct email correspondence between the verifiers and applicants/agents outside the e-development portal.

The building standards as an information repository

This project aimed to create innovative processes for digitising the delivery of building standards information and processes associated with building warrant application, up to obtaining the completion certificate at the end of the permission journey. Additionally, the project focused on understanding the experiences and journeys of various stakeholders as they navigated building regulations and building warrant processes, which involved the provision of inspection services.

To achieve the objective, the research team started with the creation of a persona corresponding to each stakeholder group, which was then prioritised for testing purposes – with a focus on the less technically inclined. A customer journey mapping exercise followed, which determined the measures for success. In the second phase, we developed a minimum viable product (MVP) app and a map of the building warrant processes, using digital frameworks including Figma and WIX.

We focused on creating a web application prototype that will allow a range of stakeholders to navigate the building standards on their mobile phone or laptop/desktop computer in a swift and efficient, manner. The prototype allowed the usability testing of the navigation of the information repository of the Scottish Building Standards while allowing for various types of linkages to be developed following linked data and embedding approaches.

Research Questions

The project seeked to answer the following questions to achieve the objective, by conducting a thorough investigation into the requirements of navigating the building standards and the inspection services throughout the building permit journey.

- How can we improve the navigation of the expanded universe of the Scottish Buildings Standards, for all stakeholders, using an application, enhancing the enabling vectors, and minising restrictions in the day-to-day usage of the building standards?
- How can the building warrant journey and inspection services in Scotland be improved and streamlined?

Methodology

We used a mixed methods approach involving both the development of personas from social sciences but also design science research where we developed a computing prototype and a digital map that we then used as an engagement tool to arrive at recommendations. Our adopted methods for this study were based on a hybrid model of combining the development of digital prototypes - web applications and a digital map of the BW process with the engagement of a focus group of professionals that would use the app and then respond live to questions and a discussion, and then interviews with lay people who would also use the app and respond to a semi-structured set of questions. Essentially, the process of developing computer prototypes was interwoven with qualitative data collection methods derived from social science and user experience research. The framework diagram (Figure 01) is available on Github [1].

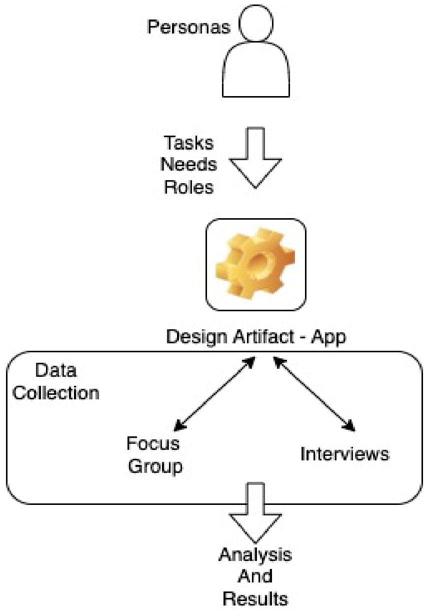


Figure 1: Overall Framework Diagram

Vehicles for research

Personas

Our qualitative methods for this study encompassed various stakeholder engagement activities (focus groups, interviews and, questionnaires) as well as the creation of PERSONAS for each key user group which helped to provide context to [web and mobile] application design for Building Standards navigation and streamline processes involved in the building warrant completion process. The engagement strategy was tailored to pinpoint the specific profile of each user type [for lay people, the general public, and professionals] and their anticipated interaction with the digital tools [web and mobile application] for data management and to develop a series of improved workflows for building warrant completion.

The method of personas creation in user experience research helps to define and understand users' needs, goals, pain points, challenges, and behaviours. The user profiles help make more informed decisions and designs that takes into account diverse user perspectives fostering a user-centred design approach. In this study, through our stakeholder engagement strategy and iterative design processes, we have created different user profiles of professionals: agents [architects, architectural technologists, surveyors, and engineers], verifiers - building officers, consultants [environmental, fire etc] and the general public in order to tailor the proposed digitisation technology and address the various user behaviours and diverse preferences. During the iterative process of engagement (using various qualitative methods – focus groups, interviews, and questionnaires) the stakeholders were asked about their occupation and objectives related to their work [Context]. They were also requested to provide information on the challenges and obstacles related to the building warrant process [Actions, Motivations, and Pains] as well as best practices and motivations [Values] that could drive the improvements in the process (Figure 2). The information gathered from personas helped to inform the design approach of our engagement tools. Essentially our methods are firmly based on social sciences, using however the design of the web application and the map of the digitised process as design artefacts and engagement tools, with which we build knowledge via the engagement with the participants. Further knowledge building took place through a search through the design space for the web application and the gaps highlighted in the map of the digitised process of the building warrant.

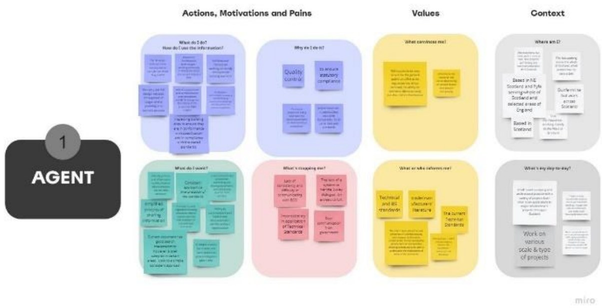


Figure 2: Personas map showing and example of an agent involved in the building warrant process.

Web App

Application Development Stage

Figure 3 showcases the initial web prototype developed in Figma. The proposed mobile application prototype was presented and tested through focus group interviews with professionals in the building construction industry. This prototype was being tested and feedback from different personas was. This prototype was mainly focused on testing the performance of the UI and UX. The structure of the prototype followed the proposed swimlane diagram and the customer journey issued by the Scottish Government: Stage 1 Before you apply for a building warrant; Stage 2 Apply for a building warrant; Stage 3 Local authority assesses your building warrant application; Stage 4 Building warrant is granted by your local authority; Stage 5 Building work (or conversion) starts; Stage 6 Changes to the building warrant design; Stage 7 Building work (or conversion) is complete; Stage 8 Local authority accepts your completion certificate.

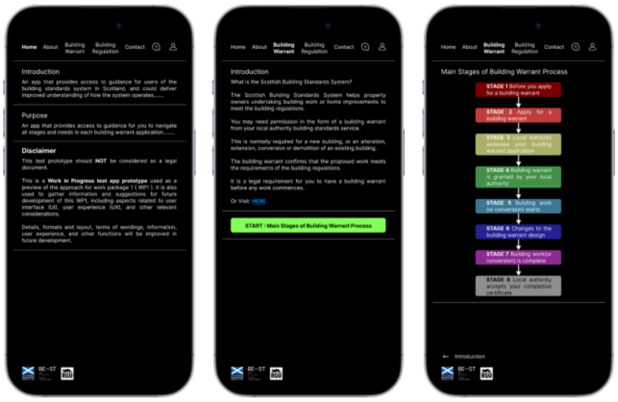


Figure 3: Showcasing Screenshots from Figma Prototype

The feedback from the participants was positive and overall UI and UX were easy to understand and follow. However, there were few participants who pointed out the accessibility of the application. There would be an issue for users to access from the office since they would be mainly accessing from desktops/laptop devices rather than mobile devices. Therefore, based on the feedback, the approach of the development is also considered to improve the accessibility of the prototype.

The Web Application (Web App) was further developed using the wix.com platform, utilized in this project in order to maximise the accessibility for different users from mobile devices, tablets, and desktops that is able to optimize accessibility across various platforms. The web app prototype focused on navigating the building standards for building warrant applications for the general public. The structure of the application is illustrated in Figure 4, while its interface on Figure 5.

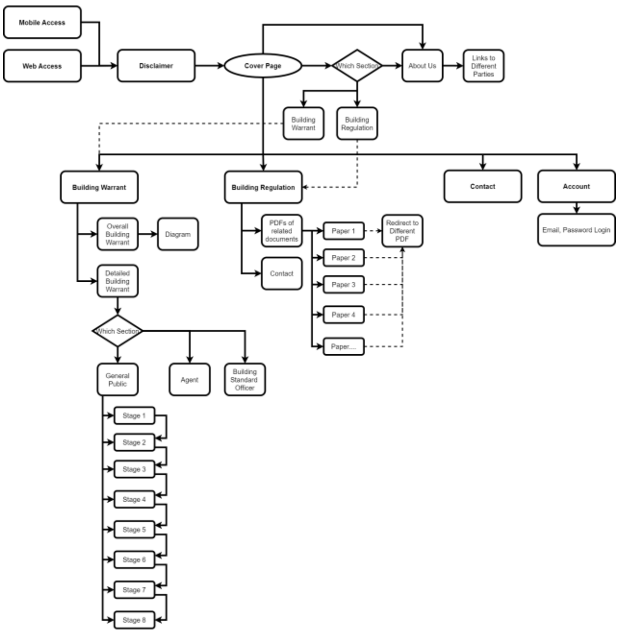


Figure 4: Overall WebApp prototype structure using Wix.com

In order to make the web app more interactive and user-friendly, simplified languages and guidelines of the application, highlighting key actions, and related building standard documents are being added during the web app development (Figure 5). The web app prototype is available on Wix.com [02].

Figure 5: Showcasing Screenshots (Left: mobile, Right: website) from WebAPP (Wix.com) Prototype

The Building Warrant Process

Investigating and recreating the swimlane diagram helps visualise and explain the whole process of all parties involved in building warrant applications. Figure 6 illustrates the full process of the building warrant application, and Figure 7 showcases an overview process of building warrant application. Both application processes followed the stages written in the document issued by the Scottish Government, with meetings and discussions with members from the Building Standards Division of the Scottish Government to prepare and finalise the process of building warrant application.

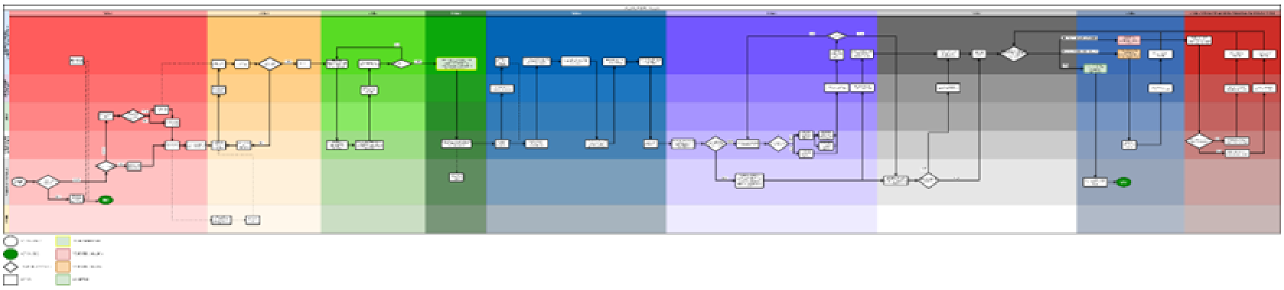


Figure 6: Whole Process of Building Warrant Application

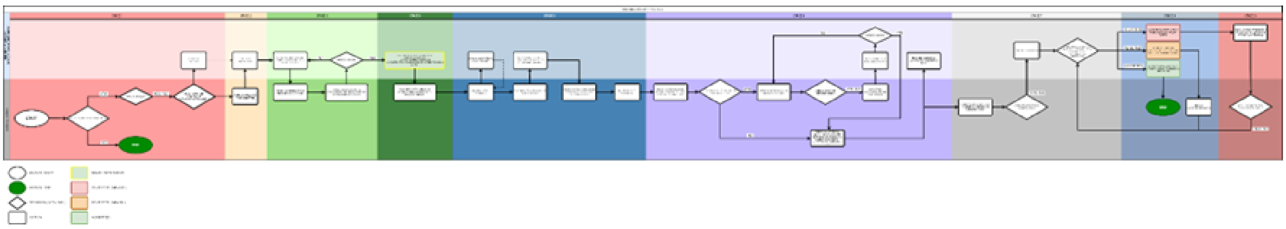


Figure 7: Overview Process of Building Warrant Application

The swimlane diagram uses mainly three shapes to identify the actions for the readers to understand. (1) an Oval shape, representing the starting and ending action, (2) a Diamond shape, representing the decision needed to be made, and (3) a Square, Action happened.

Figures 6, and 7 illustrate the whole process of building warrant application, the process was divided into 8 + 1 stages:

1. Before you apply for a building warrant.
2. Apply for a building warrant.
3. Local authority assesses your building warrant application.
4. A building warrant is granted by your local authority.
5. Building work (or conversion) starts.
6. Changes to the building warrant design.
7. Building work (or conversion) is complete.
8. Local authority accepts your completion certificate.
9. Local authority rejects your completion certificate.

Moreover, the swimlane diagram includes all sections carried by all parties being involved in the process, from a top darker colour to button lighter colour: (1) Building standard officer/ Local authority; (2) eDevelopment Scotland; (3) Agent, a professional who handles the application on behalf of the client; (4) Shared action (Agent/ General public), actions are similar on both parties; (5) Member of the public; and (6) Other, Specialist that involved during the application.

Through several structured dialogues with members from the Digital Strategy and the Building Standards Division, our team analysed the BW process and developed a swimlane diagram of it. While analysing the process of building warrant applications, it was essential to identify all actions (being carried out by different parties), and stakeholders being involved in the application process. Despite the rich documentation published by the Scottish government, none of it was comprehensive and integrated, confusing and misleading its users. Furthermore, the application varies between different councils in Scotland, each council has a slightly different set of explanations and technical requirements. Moreover, eDevelopment Scotland mainly acts as a digital mailbox between applicants and councils, i.e there is no central storage of the data. eDevelopment lacked connections between both platforms and applicants and may lead to a slower process of applications, as the applicants submit to eDevelopment, but receive answers from each particular council which is responsible for their application. Furthermore, building regulations may be interpreted subjectively by different building standard officers, leading to inconsistent application of rules. This creates confusion and unfairness for the applicants.

Data collection and analysis

Focus Group

The initial data collection was conducted through an open invitation online focus group that was distributed through social media channels (LinkedIn and Eventbrite). The focus group participants that signed up for the workshop involved an expert group representing the built environment industry including architects, technologists, quantity surveyors, building physicists, building inspectors, and sustainability consultants working on a variety of project types and scales (interiors, existing buildings, retrospective works, or large-scale education projects) offering full design services throughout all RIBA stage (Figure 8).

Figure 8: Focus group participants

Questionnaire

The questionnaire was issued in the advanced phase of developing the engagement tools timeline which allowed the research team to refine the questions and gather valuable feedback for both work packages. The questionnaire respondents were asked to review the streamline diagram embedded into the survey. The majority of respondents were aware of the overall process of getting a building warrant permission, however only 60% were familiar with the official documents explaining the process. This suggests that an online platform or indeed an interactive diagram with embedded links/images/explanatory documents and videos could be a future useful improvement in making the customer journey through the building warrant process more approachable and easier to understand when asked about the challenges associated with the permission process the participants indicated 'Contacting officers to discuss points raised " and "Technical innovation / alternative approaches", timescales - 'lengthy process', individual interpretation of technical observations and general interaction with the local authority as issues contributing to the complexity of the process. The direct impressions from the diagram indicated that the simplification could be considered as one of the participants commented: 'Far too complex, much simpler in practice/not needed to be explained in that much detail'. Additional comment was raised about the wording of certain stages in the process not adhering to the RIBA Plan of Work Stages. This, however, was predetermined in the official document released by the Scottish Government - 'Making a quality building warrant application - what you need to know' *gov.scot (www.gov.scot).*

The issues of acceptability and potential client-agent relation issues were also raised due to the possibility of progress tracking ability and potential liability of delays in the process being shifted to the agents: 'Any direct involvement from the client in this process would likely only result in further client-agent relation issues. [...] I feel it only opens the chance for a client/BCO to be able to lay blame on the agent unnecessarily. This whole idea could result in overconfident homeowners not appointing an agent.'

However, the overall functionality, clarity, and usefulness of the building warrant timeline were positively rated noting that the relevant information was included within the diagram and that the schematic was relatively easy to follow with legible content. Most of the participants were also in agreement that the best practice to move forward in the process would be an online platform to comprehensively process the application with improved liaison with the building inspector (and local authority prior to submitting the application); reminder of scheduled inspection, automated reporting on the status of the application with the ability to track progress and be aware of the 'next steps' in the application. The mobile [web] application was also mentioned together with notification of the required actions and responsibilities for those involved in the process as useful features that could be embedded in future digitisation of the customer journey diagram (Figure 9).

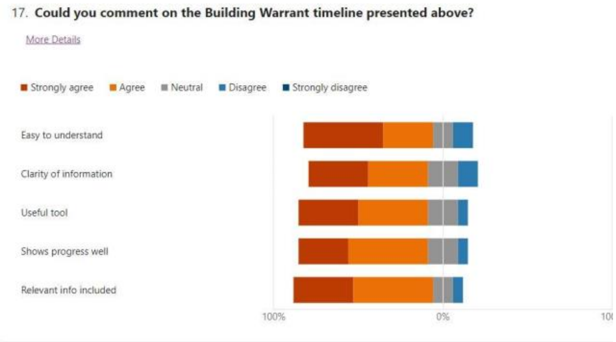


Figure 9: Extract from the survey with participants' comments on the diagram

Most of the participants were also in agreement that the best practice to move forward in the process would be an online platform to comprehensively process the application with improved liaison with the building inspector (and local authority prior to submitting the application); reminder of scheduled inspection, automated reporting on the status of the application with the ability to track progress and be aware of the 'next steps' in the application. The mobile [web] application was also mentioned together with notification of the required actions and responsibilities for those involved in the process as useful features that could be embedded in future digitisation of the customer journey diagram (Figure 10).

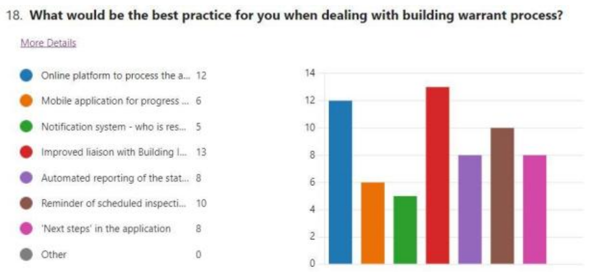


Figure 10: Results from the questionnaire pointing to the best practice

Results and Discussion

In this project, we designed and published a web app prototype focused on providing a platform for the general public to navigate the Scottish building warrant application processes and other related documents of the Scottish building regulation. The prototype streamlined the application process and made the application process more accessible and interactive for members of the public. Through prototype testing and a series of interviews, it shows the potential of web application prototypes to enhance the accessibility and user-friendliness of the navigation process of Building Standards and Building Warrants for the general public. The outcomes of this phase of the prototype show a huge potential for further exploration and development for impact. Through the interviews, we have observed that by simplifying the content of the building standard and building warrant application into illustrations and or plain language we can help a broader audience (members of the public and professionals) to understand easier, without requiring professional high level skills. These enhanced by interactive elements in the app provide clarity to the user experience.

The proposed web app prototype would be beneficial for users to navigate and understand the process of the building warrant applications and other building standards. However, for this project, we recognize that the approach of the research has the following limitations. The navigating framework mainly focused on the members of public access. The needs and usage habits of professionals such as: architects, architectural technologists, building standards authorities, etc. will be different compared to lay people. This prototype used a no-coding platform as the approach of the framework, thus not allowing full customization of all parameters as a full-from the ground -up application would provide.

Beyond the limitations presented here, the proposed web application prototype and interviews show the potential of developing a building standards hub or platform for different stakeholders to navigate the Scottish building regulations integrated with the process of building warrant application. There is different research (Kim, et al, 2020., Zhong, et al, 2018., Zentgraf, et al, 2023) showing that the frameworks are available to fill the gap and to develop a more effective navigation platform for building standards.

We summarise here the directions that this platform could encapsulate: Personalised user profiles where one can save searches, encourage users to revisit topics, easily picking up their trail from a previous visit, along with bookmarking frequently used documents or sections of the documents. Additionally the use of linked data approaches between the text of the regulation and building standard and real examples or guidelines of application would enhance the experience of the user and reduce ambiguity in the regulations. Further, a focus on navigation on building types could be developed. At the moment, when navigating the standards, the user has to choose between domestic or non-domestic regulations and there are cases where categories span both texts. Any platform of course would have to Consider mobile and web applications (user dependent). In order to maximise the flexibility of this application, both mobile and web applications for navigating Scottish building regulations and warrant content enhance the accessibility in the building environment community. Mobile applications such as smartphones, or tablets allow users to access the navigation app in different locations; websites otherwise were designed aimed at office users. Cross-platform allows users to choose their preferred platform and enhance accessibility. Although this prototype has not explored the potential of a Progressive Web App (PWA), such an approach allows offline access in the mobile background that benefits users offshore, and in highland, areas that have no internet connections; Installing the app into the mobile devices, PWA also allow users to download the app and access the content even when there are no internet connections (web. dev, 2023).

Finally, Design compliance tracking tools. Implementing tracking tools in the future development of the web app enables monitoring updates of building regulations and responses from building standard officers. For example, questions on clarifying unclear sections of action during applications or anticipating results of the submission of applications. Furthermore, a direct mailbox as a single point of contact between the user and the authorities, helps individuals to keep track of all actions that have been taken. Additionally, automated alerts are another potential feature that would help guide the applicants to stay on track and generate alerts when deviations are being detected. Such features may help to minimise errors during application.

Conclusions

Within this research study, we analysed the navigation of the building standards of the Scottish government and identified the actions and stakeholders involved during the application process of a building warrant in Scotland. We used a web application and a digital schematic diagram appropriately developed. Our work produced a number of key critical inflection points in the building warrant application and developed a set of guidelines for improving the navigation of the building standardfs.

Further work could include holistic implementation of thorough digitization and standardisation process in the planning process, ie starting from planning permission to the building warrant, a journey that should include a set of newer technologies such as building information modelling, linked data and potentially the use of blockchain for the data that need it.

In this context, different potential digital applications may assist and improve the technological barriers during building warrant application processes. Blockchain and Ontology offer a different approach for future improvement, in improving the quality of the application processes, and user-friendly experience. Blockchain offers a secure and shared system for storing information (Dounas, 2022). Ontology helps to organise and standardise information and data management (Zentgraf, 2023). These approaches provide a huge potential in improving the current system of building warrant applications, improving the application processes, and the security of the data.

Authorship & acknowledgements

HMY and DB wrote most of the report and the present paper, DB developed the engagement methods -personas, focus group, interviews, and the questionnaire, conducted data analysis and provided recommendations; HMY build the engagement tools – web & mobile application and streamline diagram; conducted analysis of the engagement tools and provided recommendations; JY developed the outline of the process for the building warrant in WP2 and contributed to the personas, questionnaire and data analysis, TD conceived, organised and directed the research, determined the structure of outputs and their quality and contributed to the authorship of the report and papers. HMY and DB can both claim first authorship of this paper.

References

Building Standards (2016) The customer journey. [online] Available from: the+customer+journey.pdf (www.gov.scot)

Dounas, T., Lombardi, D. 2022. Blockchain for Construction. Blockchain Technologies. Springer.

European Commission. 2021. The EU in 2021. General report on the activities of the European Union. Published in accordance with Article 249(2) of the Treaty on the Functioning of the European Union, https://op.europa.eu/webpub/com/general-report-2021/en/

Hobeika, N., Liempt, J, V., Noardo, F., Ohori, K, A. 2022. Geobim Information To Check Digital Building Permit Regulations. The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences XLIII-B4-2022:529-535, DOI:10.5194/isprs-archives-XLIII-B4-2022-529-2022

Judith Fauth, Gabriele Pasetti Monizza & Giada Malacarne (2023) Understanding processes on digital building permits – a case study in South Tyrol, Building Research & Information, 51:5, 518-532, DOI: 10.1080/09613218.2023.2178372

Kim, I., Choi, J., Teo, E. A. L., Sun, H. 2020. Development of K-BIM e-Submission prototypical system for the openBIM-based building permit framework. Republic of Korea. Journal of Civil Engineering and Management. Available at: https://journals.vilniustech.lt/index.php/JCEM/article/view/13756

Malsane, S., Matthews, J., Lockley, S., Love, P. E. & Greenwood, D. 2015. Development of an object model for automated compliance checking. Automation in Construction, 49(PA), 51-58.

Ullah, K., Witt, E., Lill, I. 2022. The BIM-Based Building Permit Process: Factors Affecting Adoption. Advancing Digitalisation in Construction: Responding to New and Emerging Drivers and Changes, DOI:10.3390/buildings12010045

Web.dev. 2023. Progressive Web Apps. Available at: https://web.dev/explore/progressive-web-apps

Zentgraf, S. et al. 2023. OntoBPR: Ontology-based workflow and concept for building permit reviews. 30th EG-ICE: International Conference on Intelligent Computing in Engineering. Available at: https://www.researchgate.net/publication/372252159_OntoBPR_Ontology-based_workflow_and_concept_for_building_permit_reviews

Zhong, B., Chen, G., Luo, H., Xing, X. 2018. Ontology-based framework for building environmental monitoring and compliance checking under BIM environment. China. Building and Environment Volume 141, Pages 127-142. Available at: https://www.sciencedirect.com/science/article/abs/pii/S0360132318303123

[01] Github. Overall Framework Diagram. Available at: https://github.com/arlav/ScottishBuildingStandardsWP0102

[02] Wix.com Web App prototype. Scottish Building Standards HUB. Available at: https://scotbwwp0102rgu.wixsite.com/webapp

HARMONIZING INPUT DATA FOR DIGITAL BUILDING TWIN: ENERGY EFFICIENCY USE CASE

Audrey Bouet[1], Nicolas Bus[1], Stoyan Danov[2], Eloi Gabaldon[2], Edgar Martínez-Sarmiento[2], Nicolas Pastorelly[1] and Guillaume Picinbono[1]
[1] C.S.T.B, Sophia Antipolis, France
[2] CIMNE, UPC Campus Terrassa, Terrassa (Barcelona), Spain

Abstract

In this article, we present a "Harmonizer module", based on RML mapping language, SPARQL[1] and RDF, which allows importing heterogeneous data sources into a unique Building Energy Analysis-oriented ontology. This work is being performed in the context of the H2020 BIGG project[2]. The Reference Architecture Framework designed by the project leverages the specific ontology which is used as a harmonized input for the unique BIGG Data Analytics Toolbox that creates various building-energy related insights. The Harmonizer module receives inputs from different kind of building energy analysis data (such as: location, building structure, sensor data, time series), formats them and maps them to the concepts defined in the BIGG ontology. This paper will introduce the objectives of the BIGG project, a description of the BIGG ontology before describing the workflow and the functionalities of the Harmonizer module, it's implementation and current results.

Introduction

Prominent parts of the National Energy Efficiency Action Plans of the EU countries are dedicated to strategies for increasing the energy efficiency in buildings. These plans include overviews of countries' building stock, policies for stimulating the renovation rate, including estimates and projections about the energy savings and carbon dioxide emission reductions from applications of different measures.

The increasing availability of data from smart meters and digital sensors in buildings has enormous potential to support the more efficient use of buildings. However, this requires the capability to collect, harmonize, and jointly process data from a variety of different sources. This problem is tackled by the H2020 BIGG project which has the objective to substantially facilitate the deployment of energy-related big data analytics for buildings and to support a large variety of business cases. BIGG develops a Reference Architecture and an Analytics Toolbox, supported by a common data model (BIGG Ontology) and a Harmonizer module for aligning data from different sources and providing them as an input for the common Analytics Toolbox.

This paper focuses on the presentation of the Harmonizer module designed for the conversion of external heterogeneous data formats into RDF complying with a targeted ontology. It has been created and exemplified in the context of the BIGG ontology for Energy Oriented Building Digital Twin but can be applied to any other ontology. The BIGG ontology is briefly addressed, but a detailed discussion on it is out of the scope of this paper.

Related works

The BIMERR (Boutouni et al., 2021) project is a research project aimed at developing and promoting the use of Building Information Modelling (BIM) in the construction industry. This project provides a tool called BIMERR Interoperability Framework (BIF) to ensure data exchange among applications. BIF utilizes mechanisms that enable semantic and syntactic interoperability. BIF uses the RMLMapper library (Dimou et al.,2014) and benefits from RML mapping to convert and align heterogeneous data into interoperable RDF data. Unfortunately, BIF neglects how to convert data based on an external taxonomy to a common interoperable taxonomy.

Similar attempts to populate Digital Building Twin ontology from heterogeneous data have been initiated through the COGITO platform (Katsigarakis et al., 2022). COGITO mainly uses transformation tools based on RML mapping to transforms data concerning building description such as IFC and store them in a knowledge graph. Data collected by sensors are stored separately in a dedicated time series database. This approach doesn't benefit from knowledge graph to ensure interoperability between static data and timestamped observations.

Several solutions have been proposed to map semi-structured data to RDF. The SweoIG taskforce has built a list of implementations also called RDFizer (Iglesias et al., 2020). Most of them are dedicated to a specific format or model. We focused our analysis on those who reach the project requirements: extensible language with call to user defined functions, model agnostic and fully compliant with W3C recommendations.

RML is one of the approaches that allows mapping between various sources toward RDF. RML is an extension of the R2RML vocabulary to describe logical

[1] SPARQL. https://www.w3.org/TR/sparql11-query/
[2] Building Information aGGregation, harmonization and analytics platform, H2020 funded project. https://www.bigg-project.eu/

sources. It generates RDF from JSON, CSV, or XML sources.

SPARQL-GENERATE (Lefrançois et al., 2017) is based on a query language that queries the combination of an RDF dataset and a set of documents. Various formats can be supported thanks to the extensible set of SPARQL 1.1 binging functions and SPARQL-Generate iterator functions. SPARQL-Generate is appropriate for engineers that are familiar with SPARQL query language.

OpenRefine[3] (formerly Google Refine) helps make sense of tabular data through a set of tools to clean, transform and merge several datasets together. RDF Refine is an extension of Open Refine meant to align the tabular data loaded in Open Refine with existing ontologies, while reconciliating the obtained URIs with third party RDF data sets. The data management part of Open Refine as well as the alignment towards RDF in RDF Refine are fully performed through a Web user interface. Unfortunately, RDF Refine does not allow to map CSV input files to RDF. Open Refine can also handle JSON data but does not implement multi-level iterators that make the mapping complicated when processing deep structures.

The BIGG H2020 Project

General presentation of the BIGG project

The BIGG project aims at demonstrating the application of big data technologies and data analytic techniques for the complete building life cycle of more than 4000 buildings in 6 large-scale pilot test beds. The proposed solutions will be deployed and tested as a pilot with country validation of at least two business scenarios in Spain and Greece. For that, the project has developed: 1) The Open Source Architecture (BIGG Data Reference Architecture 4 Buildings) for collecting, funneling, processing and exchanging data from different sources (smart meters, sensors, BMS, existing data sets); 2) An interoperable buildings data specification (the BIGG Ontology); 3) An extensible, open, cloud-compatible service (BIGG Data Analytics Toolbox of service modules) for batch and real-time analytics that supports effective decision and policy making.

In that context the reference architecture framework (RAF) (Pastorelly et al., 2022) has been designed (see Figure 1) and implemented, and the BIGG consortium has developed open source BIGG components enabling big data analytics for different business cases: Benchmarking and building portfolio management; Energy certification in residential and tertiary buildings; Building life-cycle data interoperability: from planning to renovation; Energy Performance Contract support; Building optimisation for occupants; Flexibility potential of buildings.

Presentation of the overall architecture

The Reference Architecture Framework (RAF) can be seen as a dedicated implementation of the Digital Twin concept. The detailed RAF concept is presented in Figure 1.

The RAF consists in a pipeline of several components which can be orchestrated in different ways. The first step uses BIGG ingestion modules (1) that are software components acquiring specific data (2) from different custom sources using standard protocols like HTTP or MQTT. These ingested data are stored in data lakes for later processing and/or directly pushed through an event message bus like Apache Kafka[4] to the harmonization module (3). The harmonization module is flexible enough to only require a standardized mapping file (4) to enable conversion of the initial specific data to the BIGG Standard Data Model 4 Building (5) required as input by the BIGG Data Analytics Toolbox (6). Harmonized data are instances of the BIGG data model, expressed in RDF format, containing only information needed for a specific use case. They can be stored in "harmonized data lakes" (ex: triple stores) and/or directly pushed through an event bus to the toolbox. The toolbox leverages several data analytics and AI/ML curated technologies to process the building/energy related data and produce insights output and push them back into the BIGG Standard Data Model 4 Building (7). The Toolbox uses machine learning techniques to produce valuable predictions. The knowledge created by the toolbox can be stored in "harmonized lakes shores" and is used by specific dashboard and external system software components (8) that will transform back the harmonized data into the data formats required by external dashboards or systems.

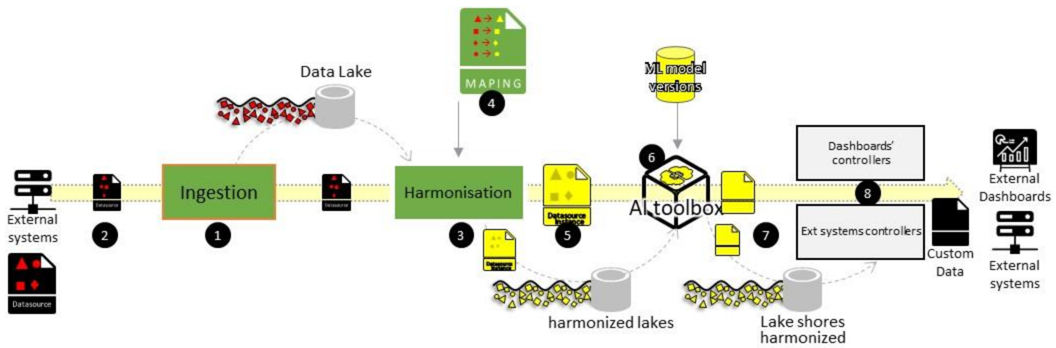


Figure 1: BIGG RAF detailed.

[3] http://openrefine.org/

[4] https://kafka.apache.org/

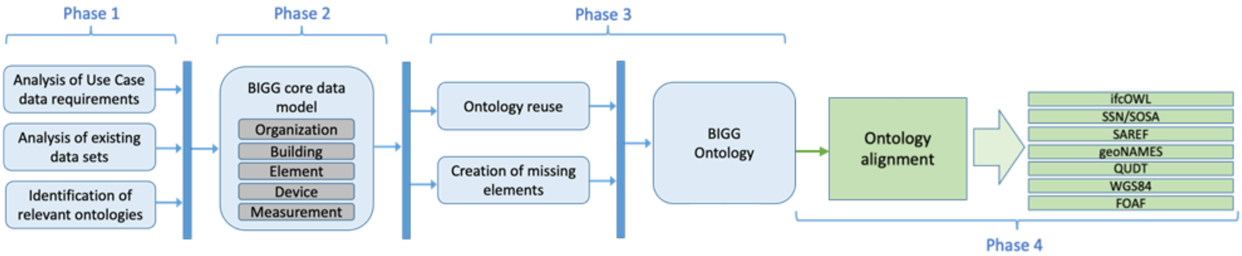


Figure 2: Methodological process for development of the BIGG Ontology.

BIGG Ontology for Energy Oriented Building Digital Twin

The creation of the BIGG ontology is rooted in the necessity to ensure harmonised input for the common Analytics Toolbox that allows developing data-driven services by using heterogeneous datasets generated within different business activities in the built environment. Conceived initially as a common data model developed in UML, it evolved to a graph data structure described in the Web Ontology Language (OWL) to benefit from the standardisation and data processing power offered by the semantic technologies.

The development process of the BIGG ontology followed a methodological approach similar to that suggested in (Radulovic et al., 2015). It is schematically illustrated in Figure 2 and has four phases, including the ontology alignment as a last step (once validated by the implementation of the business cases, BIGG ontology was aligned with already existing ontologies concepts coming from ifcOW, SAREF, SSN/SOSA, BOT, etc, and properly documented and published).

Overall, the BIGG ontology provides a standardised way to describe and organise data for buildings and applications requiring comparison and analysis of data related to performance of buildings, their systems, elements, and their usage.

Table 1: Main classes of the BIGG Ontology

Class	Description
bigg:Building	A building for which data is provided
bigg:BuildingSpace	A space that can represent one or more rooms, floors, or zones of a Building
bigg:Element	Generic element of the building
bigg:Device	Any meter, sensor, or actuator that can capture a signal or assume a state
bigg:Sensor	A collection of Measurements from the same Device
bigg:Measurement	Any time series record registered by a Device.

Currently, the BIGG ontology contains 100 classes, 181 DataProperties and 108 ObjectProperties. Table 1 provides a description of some of the main classes.

The BIGG Harmonizer

The Digital Building Twin involves a multitude of data models and formats coming from different sources such as open data providers, asset management software and sensors. Regarding data formats, relational databases and XML are still present, Open Data portals heavily rely on CSV, and web APIs on JavaScript Object Notation (JSON). The RDF data model is used as a federated model to reach semantic interoperability and querying of data having heterogeneous formats (Michel, 2017).

Presentation of the overall concept and input data

The BIGG harmonizer aims at converting any data in scope of energy performance of building, that fit with the BIGG Ontology, into RDF. Data that are not needed for use cases are filtered. As possible inputs in our approach, we focus on JSON for data coming from sensors (timeseries) and asset management software (building and systems description) as well as CSV coming from Open Data (geolocation) and RDF (data aligned on SAREF, SOSA or IFC) for alignment with existing ontologies (see Figure 3).

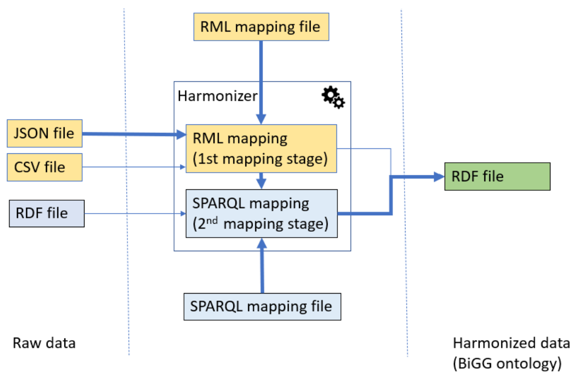


Figure 3: General workflow of the Harmonizer module

The harmonizer must generate data respectful of W3C recommendations and BIGG Ontology by providing the following functionalities: (F1) Converts JSON files into RDF compliant with BIGG Ontology; (F2) Converts CSV files into RDF compliant with BIGG Ontology; (F3) Aligns Data described using standard ontologies covering the same scope (ifcOWL, SAREF, SSN/SOSA, GeoNames, QUDT, WGS84, FOAF); (F4) Allows to map input objects with BIGG classes; (F5) Allows to map input attributes with BIGG data properties; (F6) Interprets implicit links between objects through object properties; (F7) Reconciles input values with open registers, BIGG taxonomies and enumerations; (F8) Materializes data context and generate 5-stars data. The next parts describe how the harmonizer provides these functionalities.

Converting input formats to RDF (F1, F2)

The Digital Building Twin necessitates high-quality data known as 5-star data, as originally introduced in (Berners-Lee et al., 2009). Frequently, the raw data obtained by the Digital Building Twin necessitates cleaning and enrichment to enable full interoperability. The levels established by Tim Berners-Lee are: availability (level 1), structuring (level 2), openness (level 3), universal identification (level 4), and linkage to other data (level 5). In our case, raw data is mainly provided as JSON, but the same methodology could be applied to CSV-formatted data. In this project, the 5-star level is achieved by following the RDF specifications.

First stage of the harmonizer extracts inputs from a JSON file and a RML mapping file in order to produce a first RDF file matching the BIGG scope (see Figure 3). The RML mapping files are composed of iterators to identify data sources and class mapping declarations that refer to iterators. Mapping files can either be produced manually or by using dedicated tools that simplify the mapping process such as RMLEditor[5] or Matey[6]. This procedure is detailed in the Implementation section.

The RML processing is based on the RMLMapper library developed by Ghent University. Firstly, the relevant parts of the JSON are identified according to iterators declared in the mapping file. Then the mapping itself is performed by creating, for each mapping data, an instance of the mapped ontology class, by generating a URI, and by generating triples according to data properties. Each URI is generated from the context (provider or building owner and building) and the local ID defined in the JSON.

Interpreting implicit link (F6)

The underlying structures of JSON are arrays and trees. Relations between objects are implicitly declared by using the native parent-child relation provided by the tree structure. As BIGG Ontology defines several relations, the harmonizer needs to match the parent-child relation between two JSON objects with one of the BIGG Ontology relation. Figure 4 illustrates how two parent-child relation from the same JSON file can be interpreted as distinct relations in the BIGG Ontology.

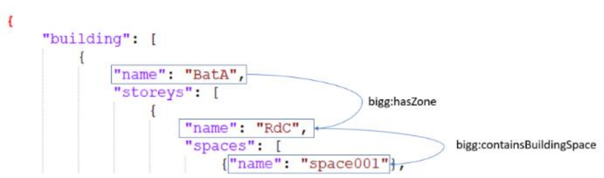


Figure 4: Sample of JSON file describing building structure.

The materialization of implicit links is also ensured by the harmonizer's first mapping stage. Object properties (as declared in the mapping file) map instances with each other's by using their URI. As for instance generation, JSON samples corresponding to the subject and the object of a triple are identified using absolute or relative expressions. As URI are always generated by using the same convention, URIs remain consistent. This way, URIs generated when populating instances match with URIs generated when building relations between instances.

Reconciliating data with enumerations (F7)

The BIGG harmonizer uses GeoNames as a universal register for public buildings. By reconciling GeoNames buildings with BIGG building descriptions, the project allows for the federation of data provided by building owners with data provided by the Open Data community. This is achieved by querying the GeoNames endpoint to retrieve the building's public URI, official name, and address, using the building location as a reference. The harmonizer's first stage generates a graph, which is then used to execute a federated SPARQL query in the second stage to replace the temporary building URI with the GeoNames URI and insert complementary information requested from GeoNames. Alternative public endpoints can be used to provide data on building height, roof inclination, and envelope materials.

The BIGG Ontology is a valuable resource that provides taxonomies for the classification of various objects, including buildings, spaces, and sensors. Within the context of building usage classification, an excerpt of the building usages taxonomy is presented in Figure 5, which outlines the primary services provided by a building.

In situations where the classification of raw data differs from the BIGG classification, the harmonizer's second stage is used to translate between the two. For example, the French building classification term "Université" must be converted to its corresponding term in the BIGG taxonomy, namely "University".

Property	Taxonomy 1st level	Taxonomy 2nd level
buildingSpaceUseType	EducationAndResearch	ExtracurricularEducationCenter
		Laboratory
		MilitaryOrPoliceAcademy
		OtherEducationAndResearch
		Preschool
		PrimarySchool
		ResearchCenter
		SecondarySchool
		University
		(empty)
	Healthcare	DayCareHospital
		HomelessShelter

Figure 5: Sample of one of the BIGG taxonomies classifying building space usages.

A thesaurus is embedded in the BIGG ontology. It employs the SKOS ontology to describe taxonomies (skos:scheme), translations (rdf:label "Université"@fr, "University"@en) and cross-taxonomy semantic links

[5] RMLEditor, https://rml.io/tools/rmleditor/

[6] Matey, https://rml.io/yarrrml/matey/#

(<medical center> skos:match <hospital>). In the harmonizer's second stage, the reconciliation process is accomplished through a federated query that utilizes the thesaurus to translate from the taxonomy used in the raw data to the harmonized taxonomy.

Materializing data context (F8)

When describing energy consumption as JSON time series (Figure 6), the contextual information regarding the source (Who), location (Where), and method (How) of measurement is often overlooked or indirectly specified in the file name. A standardized approach is required to extract and retrieve this contextual information from the pre-existing database or other data sources. The identification of data providers and building owners can be incorporated into the mapping file (in the SPARQL mapping), which can be further utilized by the harmonizer to enhance the graph. In instances where the timestamp is absent, the current date and time can be used to trace the incoming data. As an additional improvement, the harmonizer can also extract pertinent data, such as building and sensor identification, from the file name to further augment the graph.

```
{
    "_id" : "60f7fcd6e594e218826ca9f2",
    "datetime" : "2021-07-20T21:00:00.000Z",
    "cups" : "ES0031406110149001EJ0F",
    "consumptionKWh" : 12.0,
    "obtainMethod" : "Real"
}
```

Figure 6: Example of JSON time series missing building or provider context

Aligning data (F3)

In addition to the raw data collected by the Digital Twin as a JSON file, the harmonizer enables the ingestion of RDF data that implements standardized ontologies.

The BIGG Ontology has been aligned with several existing ontologies related to Digital Building Twins, such as BOT (for building topology description), SAREF (for energy system description), and SSN/SOSA (for sensor and timeseries description).

This alignment consists of triples that declare corresponding classes and corresponding predicates from one ontology to the other. The data alignment is generally performed on a given opportunity via a semantic reasoner embedded in the triple store but, if necessary, the harmonizer can materialize class and properties alignments by substituting the original vocabulary with BIGG vocabulary.

The BOT ontology is a minimal ontology that defines relationships between the sub-components of a building. It has been suggested as an extensible baseline for use with more domain-specific ontologies, following general W3C principles of encouraging reuse and keeping the schema no more complex than necessary. The BIGG Ontology defines some classes that are equivalent to BOT classes. For example, bigg:BuildingElement, which describes building components, is equivalent to bot:Element, and bigg:Zone is equivalent to bot:Zone. With respect to object properties, bigg:hasZone is defined as an equivalent property to bot:containsZone, and bigg:containsBuildingSpace is an equivalent property to bot:hasSpace.

SAREF is an ontology supported by the ETSI SmartM2M standard that enables interoperability among IoT projects and can be extended to any IoT vertical domains, such as smart buildings or energy. The BIGG Ontology reuses a few SAREF concepts, such as saref:Device, saref:Sensor, and saref:UnitOfMeasure, to describe sensor networks and time series. In the BIGG Ontology, a sensor is located (bigg:isContainedInSpace) in a space, observes a building element (bigg:featureOfInterest), and provides measurements (bigg:hasMeasurements).

There are two ways to extend the harmonizer's alignment capabilities. The first way is to extend the ontology with new class and property equivalences. The second way is to add SPARQL mapping queries that will convert classes and properties from one vocabulary to another (2nd mapping stage in Figure 3).

Implementation

The harmonizer is a generic way to convert any building energy related data into BIGG digital twin compliant data. As it would be too much effort to code serializers for each kind of source by using generic programming language, the harmonizer is based on the RML mapping language that simplify mapping specification for data providers.

Choice of the mapping language

The BIGG project focuses on RML as a mapping language for the following technical reasons: 1) RML is based upon RDF that is consistent with the harmonized data; 2) RML mappings are reproducible and maintainable; 3) RML is an extension of a W3C specifications: R2RML (Das et al., 2012); 4) RML is not limited to one type of input format and can easily handle JSON, CSV and XML format. But mostly, in the scope of the BIGG project, it was important that the mapping files can be generated (or at least updated) by non-informatician persons responsible for the different business cases. In this perspective, we have concluded that RML was the best option, thanks to the YARRRML language and Matey tools that offers the most suitable solution of our project. To this, we added a second mapping stage based on SPARQL to align URIs controlled values and taxonomies.

Development of the mapping files

YARRRML[7] is a human-readable mapping language developed by the University of Ghent, it allows to define rules and to convert them into RML or R2ML mapping language. Therefore, the generated RML mapping file contains rules to be used to transform input data into an RDF format. The web-based tool named Matey offers the

[7] https://rml.io/yarrrml/spec/https://rml.io/yarrrml/spec/

```json
{
    "building": [
        {
            "name": "BatA",
            "storeys": [
                {
                    "name": "RdC",
                    "spaces": [
                        {"name": "space001"},
                        {"name": "space002"}
                    ]
                },
                {
                    "name": "R+1",
                    "spaces": [
                        {"name": "space101"},
                        {"name": "space104",
                            "devices": [
                                {"name": "ThSensor104-01"},
                                {"name": "HdSensor104-02"}
                            ]
                        }
                    ]
                }
            ]
```

```yaml
prefixes:
 bigg: "http://bigg-project.eu/"
 i: "http://bigg-project.eu/instances/"
 grel: "http://users.ugent.be/~bjdmeest/function/grel.ttl#"

sources:
    data1:
        access: data.json
        referenceFormulation: jsonpath
        iterator: "$.building[*]"
    data1bis:
        access: data.json
        referenceFormulation: jsonpath
        iterator: "$.building[*]"
    data2:
        access: data.json
        referenceFormulation: jsonpath
        iterator: "$.building[*].storeys[*]"
    data3:
        access: data.json
        referenceFormulation: jsonpath
        iterator: "$.building[*].storeys[*].spaces[*]"
    data4:
        access: data.json
        referenceFormulation: jsonpath
        iterator: "$.building[*].storeys[*].spaces[*].devices[*]"

mappings:
  building:
    sources: data1
    s: i:$(name)
    po:
      - [a, bigg:Building]
      - [bigg:name, $(name)]
      - p: bigg:guid
        o:
            function: grel:string_md5
            parameters:
                - [grel:valueParameter, $(name)]
      - p: bigg:hasSpace
        o: i:space_$(name)~iri

  buildingSpace:
    sources: data1bis
    s: i:space_$(name)
    po:
      - [a, bigg:BuildingSpace]
      - [bigg:name, space_$(name)]
      - p: bigg:hasSubSpace
        o: i:space_$(storeys[*].name)~iri
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix bigg: <http://bigg-project.eu/> .
@prefix i: <http://bigg-project.eu/instances/> .
@prefix grel: <http://users.ugent.be/~bjdmeest/function/grel.ttl#> .

i:BatA rdf:type bigg:Building ;
    bigg:name "BatA" ;
    bigg:guid "4c5ec575c534efc17014c19dc6c2a53f" ;
    bigg:hasSpace i:space_BatA .

i:BatB rdf:type bigg:Building ;
    bigg:name "BatB" ;
    bigg:guid "37e5fc91cc8c10cda8397f490b803c45" ;
    bigg:hasSpace i:space_BatB .

i:space_BatA rdf:type bigg:BuildingSpace ;
    bigg:name "space_BatA" ;
    bigg:hasSubSpace i:space_RdC, i:space_R%2B1, i:space_R%2B2 .

i:space_BatB rdf:type bigg:BuildingSpace ;
    bigg:name "space_BatB" .

i:space_RdC rdf:type bigg:BuildingSpace ;
    bigg:name "space_RdC" ;
    bigg:hasSubSpace i:space001, i:space002, i:space003, i:space004 .
```

Figure 7: Example of a JSON input file (top left) with the corresponding YARRRML mapping (right) and the RDF file created by the mapping (bottom left)

possibility to write YARRML rules in a user-friendly way.

Matey user interface consists of 4 windows. Top left window contains input data (for example a JSON file). Top right window is the YARRRML editor, where user can write the YARRRML mapping script. Pushing the "Generate RML" button will automatically translate YARRRML script to RML and fill in the bottom right windows. Finally, pushing the "Generate LD" button will fill in the bottom left window with the RDF file resulting from the mapping of the input data using RML rules.

In the case of the BIGG harmonizer, Matey is used to develop and test the mapping rules in YARRRML language, generate RML rules and export them into a file. The YARRML mapping file is composed of three main parts (see Figure 7):

<u>The declaration of prefixes:</u> Some prefixes and namespaces are predefined by default in YARRRML, and a set of customs ones can be added by adding the '*prefixes*' collection to the root of the document (for example prefixes for the BIGG Ontology).

<u>The declaration of sources:</u>

Under the tag "sources" are listed all the "entry points" into the input data. Each entry point is defined by the file name, the format and an iterator acting as a regular expression to filter data. For example:

- Iterator: "$.building[*]" will iterate over all "building" entities in the file.
- Iterator: "$..contents[?(@.type=="SITE"&& @.categoryCode=="BUILDING")]" will iterate over all "contents" entities at any level in the JSON tree, and which have specific values of "type" and "categoryCode" properties.

<u>The mapping rules:</u>

Finally, the '*mappings*' collection contains the declaration of the mapping rules explaining how to translate the current entity pointed by the iterator into triples (subjects, predicates, objects) of the BIGG Ontology. For the example of Figure 7, the mapping named "building" takes as an input source the iterator "data1", and instantiate an object named "$(name)" as a bigg::Building, which as a bigg::name egal to "$(name)" and has a relation to a space declared as "space_$(name)". Instantiation of the bigg::BuildingSpace will be done in the next mapping rule.

Providing the harmonizer as a service

SDM-RDFizer is an interpreter of mapping rules that allows the transformation of (un)structured data into RDF knowledge graphs. The current version of the SDM-RDFizer assumes mapping rules are defined in the RDF Mapping Language (RML). The SDM-RDFizer

```
python harmonizer.py --input inputFile [--mapping RMLFile] [--sparql SparqlFiles] [--output outputFilename]
```

Figure 8: Command line calling the Harmonizer module.

implements optimized data structures and relational algebra operators that enable an efficient execution of RML triple maps even in the presence of big data. SDM-RDFizer is able to process data from heterogeneous data sources (CSV, JSON, RDB, XML) processing each set of RML rules (TriplesMap) in a multi-thread safe procedure.
RocketRML (Simsek et al., 2019) is also an implementation of the RML mapper specification based on NodeJS. Whereas RocketRML appears to be faster than RMLMapper library, it only supports a subset of the RML specification that is needed for our use cases.
RMLMapper developed in java (*rmlmapper.jar*) is the reference implementation of the RML mapper specification as it covers all the features. Furthermore, it is more widely used and benefits from a more active community regarding version control insights (Heyvaert et al., 2018). For those reasons, RMLmapper is the one used in BiGG to process RML mapping rules on a selected dataset to generate an RDF document. The integration of this library into a Python module allows to experiment the harmonization of input data from multiple sources automatically. The Python module created to harmonize input data with the BIGG Ontology is composed of two stages, the first one is related to the RML mapping file, and the second one is related to the alignment of standards ontologies with the BIGG Ontology.
The conversion step is the first stage of the Python module, it corresponds to the use of the java library to convert an input JSON file into an RDF file thanks to the mapping rules defined in the RML file. The module can output two serialization formats of RDF, the TTL (Turtle) or the JSON-LD format.
The second stage of the Python module, corresponds to the use of SPARQL queries in order to add, translate or complete the RML stage. The second stage can also be used to align data based on standardized Ontology into the BIGG compliant RDF. For instance, it allows to align data based on standards ontologies like ifcOWL, SOSA or SAREF, with the BIGG Ontology.
The execution and test of the Python module can be done in a Jupiter Notebook, with the command line of Figure 8.

Experimentation with data provided by pilot sites

The harmonizer process can be illustrated with a simple example of a building containing storeys, spaces, and some devices. Figure 7 shows the corresponding JSON input file on the top left, and the YARRRML mapping file on the right. Once the RML mapping file is generated, we can use it to convert the input file into an RDF file aligning with the BIGG Ontology (bottom left).
The Python module allows the use of the RML mapping rules to align the input data with the BIGG Ontology into an RDF document. Automatically, it generates instances of BIGG objects like bigg:Building, bigg:BuildingSpace and bigg:Device with the related relations (Figure 9).
The same approach has been successfully applied with input data from other sources: group, systems and sensors structure exported from exploitation control software, or time series coming directly from sensors. Each specific data is mapped to instantiate a specific part of the BIGG ontology.
In a second phase, the harmonization module is also used to create rules to include import data in standard format (e.g. ifcOWL) into an instance of BIGG model. Indeed, the harmonizer second stage allows to execute SPARQL queries to specify correspondences between ontologies (Figure 10).

IFCOWL	BIGG
:IfcBuilding	bigg:Building
:IfcSpace	bigg:BuildingSpace
:IfcZone	bigg:Zone
:IfcDistributionElement	bigg:Device

Figure 9: Example of Alignment between ifcOWL Ontology and BIGG Ontology

Discussion & perspectives

The BIGG harmonizer implements and orchestrates transformations on heterogeneous data. This article demonstrates how the harmonizer can convert any data covering building and sensors description as well as energy related time series. It is fully compliant with W3C recommendations by composing an RML transformation with SPARQL transformations. This solution is extensible as any data providers can add his own mapping files to process his own data stream to the unified BIGG data toolbox. The main limitation of the current implementation is the use of files as data buffers between providers and the digital twin.

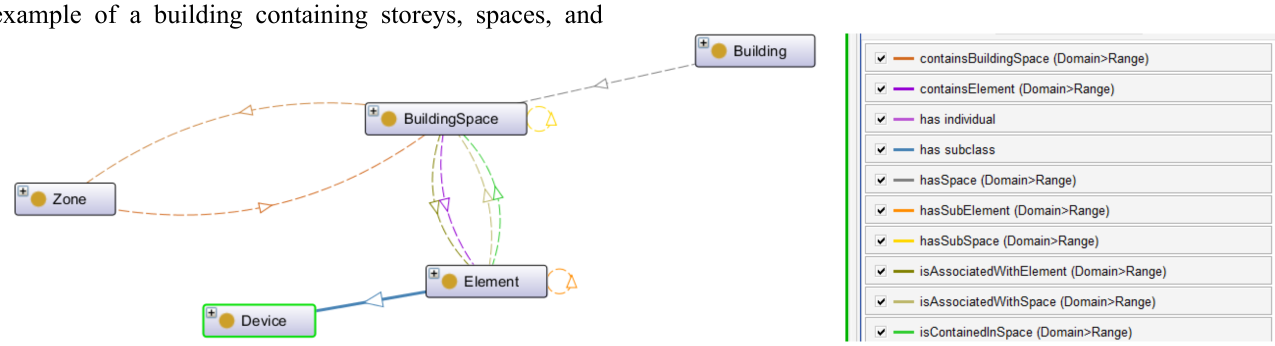


Figure 10: Visualization of objects from the BIGG Ontology to be instantiated with the Python module.

At this stage of the project, a dozen of RML mapping files have been developed for real input data (several initial tests were done with manually developed input files), covering 5 different Use Cases. Two of the involved partners were able to generate these mapping file quite easily using Matey tool, and three other partners have been trained to learn YARRRML language and are now able generate their own mapping files.

Matey tool and YARRRML language allowed us to develop most of the mapping we need. We encountered mapping difficulties on only one input file, which have a more complex structure: the depth of element decomposition depends on the type of element. This requires implementing some “Conditions” in YARRRML and managing “Iterators” in a different way. These aspects have been investigated during the last months of the project and seems very promising.

From a data integration point of view, the main objective of the BIGG project was to design, develop, and test a specific architecture allowing to address as many use cases as possible through a common data analysis pipeline. By designing a dedicated BIGG ontology, and developing a generic data harmonization tool, based on RML mapping language, to feed it from heterogeneous data sets, we achieved our objective, while using standard, open, and extensible technical solutions.

The RAF describes BIGG components and state-of-the-art techniques to coordinate BIGG components, may the actual architecture deployment on a local server or on a cloud infrastructure. The next step is the integration of the harmonizer into the whole stream-oriented architecture. Each BIGG component implements its business logic using specific technologies and specific supporting open sources libraries or frameworks (e.g., Kafka (Noac'H et al., 2017) (Hiraman et al., 2018)). Then business logic code is containerized using Docker (Wan et al., 2018). Components get their features accessible via APIs (Application Programming Interface), with event message compatible interface. For the latter point, the Kafka event streaming message system and RMLStreamer (Min et al., 2022) has been chosen to allow the platform to process continuous data streams.

References

N. Boutouni, F. Lampathaki, S. Kousouris, A. Tsitsanis, G. Vafeiadis (2021) The BIMERR Interoperability Framework: Towards BIM Enabled Interoperability in the Construction Sector.

A. Dimou, M. Vander Sande, P. Colpaert, R. Verborgh, E. Mannens, and R. Van de Walle. RML: A Generic Language for Integrated RDF Mappings of Heterogeneous Data. In Proceedings of the Workshop on Linked Data on the Web, Seoul, Korea, 2014.

K. Katsigarakis, G. N. Lilis, D. Rovas, S. González-Gerpe, S. Bernardos, A. Cimmino, M. Poveda-Villalón, R. García-Castro (2022) Digital Twin Platform Generating Knowledge Graphs for Construction Projects.

N. Pastorelly. BIGG, D2.2 - Initial technical specifications and preliminary design of BIGG Architecture building blocks. H2020-LC-SC3-EE-2020-1/LC-SC3-B4E-6-2020.

Radulovic, F., Poveda-Villalón, M., Vila-Suero, D., Rodríguez-Doncel, V., García- Castro, R. and Gómez-Pérez, A., 2015. Guidelines for Linked Data generation and publication: An example in building energy consumption. Automation in Construction, 57, pp.178-187.

F. Michel. Integrating heterogeneous data sources in the Web of data. Other [cs.OH]. Université Côte d’Azur, 2017. English. ffNNT : 2017AZUR4002ff. fftel-01508602v3f

Berners-Lee, T. 2009. Linked data: Design issues. Accessed March 02, 2023. https://5stardata.info/en/

E. Iglesias, S. Jozashoori, D. Chaves-Fraga, D. Collarana, and M.E. Vidal, 2020. SDM-RDFizer: An RML Interpreter for the Efficient Creation of RDF Knowledge Graphs

M. Lefrançois, A. Zimmermann, N. Bakerally, 2017. SPARQL-Generate. A SPARQL extension for generating RDF from heterogeneous formats.

S. Das, S. Sundara, R. Cyganiak, R2RML: RDB to RDF Mapping Language, 2012. https://www.w3.org/TR/r2rml/

U. Simsek, E. Karle, and D. Fensel. RocketRML - A NodeJS implementation of a use-case specific RML mapper. In Proceeding of the 1st International Workshop on Knowledge Graph Building, 2019.

Heyvaert, P., De Meester, B., Dimou, A., Verborgh, R. (2018). Declarative Rules for Linked Data Generation at Your Fingertips!. In: , et al. The Semantic Web: ESWC 2018 Satellite Events. ESWC 2018. Lecture Notes in Computer Science(), vol 11155. Springer, Cham. https://doi.org/10.1007/978-3-319-98192-5_40. https://rml.io/yarrrml/matey/

L. Noac'H, A. Costan, and L. Bougé, "A performance evaluation of Apache Kafka in support of big data streaming applications", in Big Data (Big Data), 2017 IEEE International Conference on, 2017, pp. 4803-4806.

Hiraman B., Chapte M. and Abhijeet C. (2018). A Study of Apache Kafka in Big Data Stream Processing. 1-3. 10.1109/ICICET.2018.8533771.

Wan X., Guan X., Wang T., Bai G. and Choi B. 2018. Application deployment using Microservice and Docker containers: Framework and optimization. Journal of Network and Computer Applications. 119. 10.1016/j.jnca.2018.07.003.

S. Min Oo, G. Haesendonck, B. De Meester, A. Dimou. RMLStreamer - an RDF stream generator from streaming heterogeneous data. The Semantic Web – ISWC 2022. Springer International Publishing, (2022)

TOWARD A WORKING THEORY FOR HUMAN-DATA INTERACTION IN THE BUILT ENVIRONMENT

Jennifer Li [1], Gulben Calis[2], Diego Calvetti[3], Peter Demian[4], Mohamad Kassem[5], Mahir Msawil[6], May Shayboun[7] and Ranjith Soman[8]
[1]Northumbria University, Amsterdam, The Netherlands; [2]Ege University, Bornova, Turkey; [3]CONSTRUCT/Gequaltec, Porto University, Porto, Portugal; [4]Loughborough University, Loughborough, United Kingdom; [5]Newcastle University, Newcastle, United Kingdom; [6]Northumbria University, Newcastle, United Kingdom; [7]Halmstad University, Halmstad, Sweden; [8]TU Delft, Delft, The Netherlands.

Abstract

Human-Data Interaction is central to the dynamic nature of the construction industry and its growing reliance on substantial data sets. In this paper, we evaluate literature on the topic of Human-Data Interaction across several technologies and concepts within the built environment with the aim of identifying underlying aspects that can lead to an underpinning theory of Human-Data Interaction in the built environment to support advancement of the research in the field. Those aspects were identified as trust, game theory, empowerment of humans, human control, safety, accessibility, enhancing understanding, and the three pillars of Human-Data Interaction of agency, legibility, and negotiability.

Introduction

The evolution from specialised to ubiquitous and interconnected computing has instigated a profound transformation within the construction sector impacting every facet of the design, construction and operation of buildings and infrastructure. During the design phase, the adoption of tools such as Building Information Modelling (BIM) has revolutionised approaches by enabling collaborative planning and precise 3D modelling in real time. This technological advancement continues into the construction phase with the use of devices like smartphones and tablets along with devices and sensors that monitor progress and site conditions. These innovations contribute to decision making and efficient resource management. In the operations phase, integrating building technologies and data analytics plays a crucial role in optimising maintenance, energy usage and overall facility management. This promotes sustainability and cost effectiveness.

The widespread use of various devices and networks, along with their technological interactions, results in the creation of extensive data trails with considerable consequences. Data are under constant revision and extension or transformation and do not only concern the individual who provided the data or about whom the data is, but also other stakeholders that might have different interpretations of the data (Hornung *et al.*, 2015).

Considering the dynamic nature of the construction sector and its growing reliance on substantial data sets, it becomes crucial to delve into the study of Human-Data Interaction (HDI), particularly in the construction informatics domain. The aim of this paper is to make progress towards establishing a theory of HDI to underpin subsequent research into construction informatics.

HDI described as interdisciplinary encompassing "*data visualization, user interface design, and interaction as well as psychology, behavioral science, and human cognition*" (Widjojo *et al.*, 2017, p. 3). In this paper, we uncover some of these aspects through investigations into HDI across the built environment (BE). In the next section, HDI is positioned in the context of its evolution from HCI and offer an understanding of the role of data in the construction sector. The methodology describes how the study was conducted and is followed by the findings of a literature review on HDI across several technologies and concepts in the BE. Analysis of the identifies potential elements that could contribute to development of a theory of HDI for the BE leading to a roadmap to support development of such a theory. The paper end with a discussion and conclusions.

From Human-Computer Interaction (HCI) to Human-Data Interaction (HDI)

Human-Computer Interaction (HCI) has been a cornerstone in our interaction with technology, focusing on the design and use of computer technology alongside the interactions between humans and computers (Dix, 2017). HCI's evolution has shifted from viewing computers merely as devices to a complex understanding of these interactions (Grudin, 1990), encompassing psychology (Card, 2018), hardware, software, interface, and even deeper organisational aspects. This expanded view transcends the simple operator-hardware relationship, leading to a more in-depth exploration of user-computer system interrelationships, particularly as systems become more integrated within organisations (Bowers and Rodden, 1993).

An emerging dynamic ecosystem, marked by both collaborative and competitive interactions, revolves around data creation and usage by individuals (Brown, 2013). This dynamic ecosystem along with advanced machine learning (ML) and artificial intelligence (AI) has initiated a transition towards HDI (Crabtree and Mortier, 2015). Unlike HCI, HDI focuses on the interaction between humans and data, moving beyond traditional interfaces. HDI's realm extends to the subtle and often passive engagement of individuals with complex infrastructures that are typically misunderstood or ignored (Haddadi *et al.*, 2013). Although the concept of *data* is nebulous (as will be discussed below), it is arguably less important than the elements of *human interaction*. HDI views data as a boundary object, open to diverse interpretations and relevant to a broad spectrum of stakeholders, including those it concerns, collectors, legal custodians, and users (Mortier *et al.*, 2015). The challenge of HDI lies not just in usable ("user-friendly") interfaces but in making complex data comprehensible, actionable, and ethically managed. Its core themes include *Legibility, Agency*, and *Negotiability*. *Legibility* aims to make data and algorithms transparent and understandable, balancing intellectual property protection. *Agency* empowers individuals to manage their data, including opting in or out, without necessitating constant engagement but offering the choice for those interested or concerned. *Negotiability* addresses the evolving perceptions and relationships around data, societal responses, social norms, legal frameworks, and the ambiguity in data's subject and meaning (Mortier *et al.*, 2015).

As the era of big data in the construction sector progresses, HDI's themes become more pertinent, signifying a shift from mere interaction with technologies (e.g., BIM, IoT, Digital Twins, etc.) to a deeper engagement with data and its implications.

What is data and where is it produced in the construction sector?

According to the Mirriam-Webster (2023) dictionary, data is "*factual information (such as measurements or statistics) used as a basis for reasoning, discussion, or calculation*", "*information in digital form that can be transmitted or processed",* and *"information output by a sensing device or organ that includes both useful and irrelevant or redundant information and must be processed to be meaningful*".

The definition in data science might be extended beyond the use of data, as a data application acquires its value from data and produces more data as a result (Loukides, 2011). The meaning is that data do not only function as input in data driven applications or services but also outputs data that need to be processed or transmitted. On the other hand, big data can be defined as an overall term for a "*complex and substantial collection of data, which needs advanced engineering strategies and analytics systems to process, store and manage*" (Li *et al.*, 2023). Big data includes a data transformation flow and is produced in the construction sector from sites and procedures within the entire construction process, consisting of modelling, designing, planning, scheduling, and management. Big data in construction, therefore, has different sources such as BIM as a main source of data supporting planning, design, building and operation (Li *et al.*, 2023), in addition to sensors and embedded devices, operational and maintenance data including text, audio, video and more (Bilal *et al.*, 2016; Li *et al.*, 2023). In this context, there is a role of project participants in data input, acquisition, modification, and integration of information leading to information exchange as well as integration (Li *et al.*, 2023). Furthermore, construction digital twin is distinguished from BIM because of its connection to the physical twin and updates when the physical one changes throughout the process (Sacks *et al.*, 2020). The digital twin would allow for 'data centric' construction management with a flow of information between the physical and virtual models in a cycle of Plan-Do-Check-Act where 'check' marks the significant difference with current construction control if the data can be capably interpreted and automatically produce accurate and comprehensive information (Sacks *et al.*, 2020). So, data function in a process of input, processing, output in a flow of various types of information in constant interaction with the human at any point of any process.

Methodology

This inductive study conducted a literature review of HDI across several themes related to the BE and the focus of the HDI Committee of the European Council on Computing in Construction (EC3). The framework under which the literature review was conducted stemmed from the mission of the HDI Committee (EC3, n.d.) across the three research perspectives in Figure 1. This is devised to avoid the risk of developing an unordered and incomplete list of research directions omitting important areas.

HDI Research Perspective	Description
i) Understanding and evaluation (end user-oriented)	• Users and their behaviours. • Individual and societal implications of emerging interactions. • Novel/emerging uses of ICT and AI-related applications.
ii) Development (technology-oriented)	• Platforms, architectures or component technologies. • Interactions between human actors and the data systems within projects and the built environment. • Testing methods for existing technologies and design methodologies.
iii) Foundation (theory building-oriented, supported by or supporting i) and ii)	• Theories, methodologies and methods involved in emerging HDI. • Paradigms and frameworks for analysis, design and application of HDI. • Ethics and Regulations.

Figure 1: Factors of the HDI Committee Mission

Literature specific to the concept of HDI and the six themes of blockchain; machine learning (ML) and deep learning (DL); robotics for industrialised construction; sensored construction sites; smart buildings; and virtual reality (VR) and augmented reality (AR) were reviewed. These themes follow the topics of the HDI Committee's seminal white paper (Kassem and Kifokeris, forthcoming) as key considerations for the BE. We acknowledge these themes are not exhaustive to the concept of HDI in the BE, however, they provide a springboard from which to commence discussions on development a theory of HDI *for* the BE. The purpose of the review was to identify similarities across these themes that could begin to form a working theory for HDI in the BE and specifically computer informatics. Inclusion criteria included papers considering the factors above with a particular focus on users, their behaviour, their interaction with data, and the testing and/or development of technology that integrates both user data and humans. The findings of the study were consolidated to propose key elements for consideration when developing a theory of HDI in the BE and are supported by a roadmap to support achieving that theory.

Literature review

This section presents the literature reviewed across the six themes and is followed by Figure 2 showing the HDI research perspectives across each theme. The themes are ordered alphabetically with no one theme having more importance over another. Literature on each of these themes with a specific focus on HDI is limited within and outwith the BE. That which is present has been carefully selected and reviewed to highlight the components that could be relevant to overarching HDI theory in the BE.

References	i) Understanding and evaluation (end user-oriented)	ii) Development (technology-oriented)	iii) Foundation (theory building-oriented, supported by or supporting 'i' and 'ii')
Blockchain			
Becherer et al. (2020)	Human-centred design to generate trust between the system and humans. Focuses on operational risk management.	Public-key infrastructure (PKI) that gives control to an individual over their data is centrally managed with blockchain.	Humans put trust into the blockchain-based system.
Lockwood (2021a; 2021b)	Increasing agency of human actors over their personal data. Focuses on privacy and its importance in society, with a particular focus on surveillance.	SSI and Web3 gives agency and control to the user over personal data. Blockchain can facilitate persistence, portability, consent and interoperability of data.	The three pillars of HDI: agency, legibility, negotiability.
Calvetti et al. (2021)	Facilitation of contractual arrangements for sensored construction sites.	Considers relationships between workers, sensors, systems and methodologies, performance, human factors, and on-site conditions where individual data are collected.	Increased transparency for users of the system.
Hunhevicz et al. (2021); Wang et al. (2022)	Changes in ideas of ownership where it is shifted from the human (or organisation) to the asset.	no1s1 – a prototype for self-owning assets facilitated by a decentralised autonomous organisation (DAO).	Removal of profit-seeking intermediaries returning some agency to the individual.
Li et al. (2019)	A framework emphasising the need for societal input into development and deployment of blockchain-based applications in construction.	Blockchain and smart contracts as integrating technologies for construction sector applications alongside BIM, IOT, GIS etc.	A sociotechnical implementation framework, considering the dimensions of technology, process, society, and policy.
Machine Learning (ML) & Deep Learning (DL)			
Urquhart et al. (2019)	Framing regulatory, interactional and technical challenges of Adaptive Architecture.	Adaptive buildings and integrated IoT devices need to be ensured by legal rights informed by how occupants and their data interact with the buildings.	Physical and information security; establishing responsibility; understanding flows, collection, use and control of personal data; sensitive personal data and monitoring routine activities.
Schia et al. (2019)	AI trust is conditioned by knowing the AI is based on the right data and how and why the output was made.	A framework including technology, process and culture to determine the success of implementation.	The control of humans on the input creates a sense of trust. The more automated the process, the less humans have control.
Mosqueira-Rey et al. (2023)	Humans act as teachers of an ML system depending on the level of involvement of the human in the learning process.	Humans can also be at the end of the process trying to understand how and why ML decisions were made.	HITL, explainability, usable and useful AI.
Hanafy (2023)	User iterates multiple queries for optimal output whether by choosing best version or changing terms.	General purpose generative text-to-image AI used in architecture interior and exterior design.	Augmenting creativity and visualisation in architecture. Iterative trial and error and produce various choices.
Robotics for Industrialised Construction			
Sun et al. (2023)	Investigating and evaluating the impact of robots on psychological and physical safety of human agents.	Robotics technology requires a complicated integration of multidisciplinary technologies and their safety needs to be considered in regulations.	Physical risks imposed on human agents. Visual and cognitive distraction of involved human agents. Cognitive load and negative emotional state.
Rodrigues et al. (2023)	Classifying human traits, identifying human roles, understanding the main factors that impact HRI. Proposing safety mechanisms for HRI that involve the human side.	Classifying robot traits, identifying robot types, and categorising level of robot autonomy (LoRA). Proposing safety mechanisms for HRI that involve the robotic side.	Human-related factors in respect to safety: (i) mental workload, (ii) situation awareness, (iii) trust in automation, and (iv) (physical) ergonomics.
Liu et al. (2021)	Proposing a human-centred collaborative framework that enables robots to interpret human agents' cognitive loads to improve HRC.	Enabling robotic agents to capture and predict human psychological behaviour through developing machine learning models.	Acknowledging the need for reducing the cognitive and mental load imposed on human agents involved in working with robotic agents.
Fu et al. (2024)	Understanding the impacts of HITL collaboration in off-site construction that employ robots.	Alluding to the need for training various robot types in actual work environments rather than in experimental environments.	Recognising the need for respecting the value of human skills and strengths to empower human agents over robotic agents.
Zhang et al. (2023)	Understanding risks imposed on human agents in HRC in on-site construction (e.g., physical and psychology-related risks).	Understanding risks imposed on robotic agents in HRC in on-site construction (e.g., program failure and maintenance-related risks).	focus is needed on safety and management issues of HRC in on-site construction with an implicit emphasis on empowering the human side.
Sensored Construction Sites (SCS)			
Tang et al. (2020)	Evaluates worker PPE compliance using computer vision.	Develops a new vision-based system for PPE compliance.	N/A
Park et al. (2016)	Evaluates detection of unsafe conditions and collects and analyses the trajectories of workers with respect to potential safety hazards.	Develops a Bluetooth and BIM system for safety monitoring.	N/A
Kim et al. (2021)	Evaluates compliance levels and informs site managers about non-compliance.	Develops an autonomous detection method based on sensor technology for hard hat non-use.	N/A
Jebelli et al. (2018)	Evaluates early detection of workers' stress feasibility.	Develops a wearable EEG-based field stress recognition procedure.	N/A
Aryal et al. (2017)	Evaluates workers' real-time fatigue.	Develops a multi-sensor approach using heart rate monitor, infrared temperature sensors and an EEG sensor.	N/A
Shanti et al. (2022)	Evaluates workers' safety compliance at-height activities.	Develops a novel system integrating DL and drones to monitor workers in real-time Personal Fall Arrest System components compliance such as safety harness, lifeline, and helmet.	N/A
Smart Buildings			
Konstantakopoulos et al. (2019)	Incorporates HITL modelling by creating an interface to allow building managers to interact with occupants and potentially incentivize energy efficient behaviour.	Develops a gamification application to motivate humans toward energy efficient behaviour.	Introduces a smart building social game concept based on a friendly competition between occupants. Game theory concepts are used to learn models of players' decision-making in residential buildings.
Kim et al. (2022; 2023)	Incorporates user's baseline thermostat usage and allows users to set up thermostat schedules.	Develops a cloud-based eco-feedback and gaming platform (MySmartE app) that aims to promote energy conserving thermostat adjustment behaviours in multi-unit residential buildings.	Introduces a mathematical framework through a new methodology for personalized eco-feedback design integrated with a collaborative social game to assist residents in enhancing thermostat use while promoting community-level energy-savings.
Jenkins et al. (2019)	Enables occupants to take an active role in monitoring and managing plug loads via a mobile application.	Develops web and mobile applications for building administrators and users for plug load management.	Incorporates occupant engagement, gamification, automation, and machine learning for enabling occupants to take an active role in monitoring and managing their plug loads.
Virtual Reality (VR) & Augmented Reality (AR)			
Widjojo et al. (2017)	Human perception, cognition and decision-making are deeply subjective. Systems should aim to enhance humans' understanding of big datasets through enhanced visualisation.	Systems must be scalable in terms of quantity of data and dimensionality/complexity. The focus of a technological system must be in simple accessibility of data to the human.	Systems for Visual Analytics need to exploit the respective strengths of humans (perception and reasoning) and computers (computation and visualisation) and consider meaningful ways for humans to understand data through visualisation.
Alhakamy et al. (2021)	Humans can potentially interact with data using their whole bodies.	Tested two styles of interacting with data: using gestures and body movements in a large (semi-immersive) environments versus a joypad.	The mode of interacting with data affects the easing and conclusions drawn from the data.
Schiavi et al. (2022)	N/A	N/A	It is important to develop data flows between the systems used in construction (BIM, CDE, Digital Twins) and VR/AR systems.
Grübel et al. (2022)	N/A	Situated analytics allows the anchoring of data in its spatial context, which is particularly important in built environment applications of AR/VR.	N/A

Figure 2: Literature considered against the framework of HDI research perspectives

Blockchain

Blockchain was established as a technology to challenge capitalist power structures that have existed for many years (Ekblaw *et al.*, 2016). Research has demonstrated its potential to challenge many economic structures, business models, operations and much more. For the BE, and particularly the construction sector, the focus of research generally is heavily rooted in data and how they can be leveraged to make efficiencies and increase productivity whilst still maintaining an advantage over competitors (Li and Kassem, 2021). Such capitalist systems have resulted in challenges including predatory business models (e.g., underbidding for projects, using projects funds for cashflow), the justification of competitive advantage to limit information sharing, and a general lack of trust between contracting parties (Li and Kassem, 2021). Combined with new insights gained from [big] data, a new level of exploitation of individual actors from such actions drove enactment of the General Data Protection Regulation (GDPR) and highlighted the need to rethink companies' collection and use of data. Blockchain is presented as a way to support this rethinking, particularly from a Web3 perspective where focus is on giving *agency* back to the individual.

Lockwood's (2021a; 2021b) research focuses on self-sovereign identify (SSI) as an inevitable aspect of Web3. Privacy and its importance in society is central to the discussion and understanding that centralisation removes agency of personal data. The solution focuses on *agency*, *legibility* and *negotiability*, and the principles for an SSI technology: existence of the user; control and access by the user; transparency within the system; persistence of identities; portability of information and services about the data; interoperability to ensure wide usability; consent from users to use their identity; minimalization of disclosure of claims; and protection of users' rights. From a trust perspective, Becherer *et al.* (2020) present an exploration of human data engineering (HDE) for operational risk management with blockchain. Human-centred design of a blockchain-based system generates trust between it and humans. Public key infrastructure (PKI) is used to manage data centrally. While this goes against the decentralised nature of blockchain, its key HDI considerations put trust and human-centricity at the centre of the big data-designed system.

Specific to construction, Calvetti *et al.* (2021) address HDI for sensored construction sites (SCS) proposing the use of blockchain to facilitate data in contractual arrangements by investigating the perceptions of HDI and SCS. A use case enhances the relationships between workers, sensing technologies, performance, systems and methodologies, human factors, and on-site conditions where data are collected about the individual. Standards of transparency between workers and companies of data collection is raised as a result of implementing their proposed HDI information process. Li *et al.* (2019) proposed a socio-technical framework encompassing dimensions of technology, process, policy and society that gives consideration to the users of the system and the interactions they have with the technology to achieve the desired outputs and benefits. Finally, Hunhevicz *et al.* (2021) and Wang *et al.* (2022) present an alternative to the current model of ownership of assets shifting it away from the human (or organisation) to the asset. *no1s1* (no one's one) is a prototype for self-owning assets facilitated by a decentralised autonomous organisation (DAO) that increase cooperation and coordination between transacting parties whilst removing profit-seeking intermediaries returning some agency to the individual.

Machine Learning (ML) and Deep Learning (DL)

Much of AI subfields function by studying data and making decisions such as ML, computer vision, optimisation, and knowledge-based systems (Abioye *et al.*, 2021). AI has been gradually advancing in the construction sector with different aspects including health and safety, BIM, supply chain management and other (Abioye *et al.*, 2021). The subfields of AI can vary but they share one characteristic of being dependent on input data and result in output data whether this output is a prediction, a recommendation, or an automatic procedure (Allen, 2020). Human-in-the-loop ML (HITL-ML) is a concept situating the human in the process of ML in terms of the human's role in developing an AI system which is determined by the level of control humans and machines have in the learning process, in explaining the outcome to humans, and in the usability and usefulness of AI systems (Mosqueira-Rey *et al.* 2023). An HDI theory should explicit the relationship of its principles between the human and the data at any point where interaction takes place.

Urquhart *et al.* (2019) focused on how inhabitants interact with their buildings that adapt with the environment and its inhabitants through a series of IoT devices integrated with ML or DL technologies. The analysis discusses the challenges of adaptive architecture in dimensions of physical and information security, establishing responsibility, rights over personal data, sensitivity of visible emotions and bodies, and continuous monitoring. From a user perspective, Schia *et al.* (2019) presented an AI scheduling system at an early stage of implementation and conducted interviews about the system and compared it to other well-implemented digital tools. Interviewees with low AI knowledge working in construction projects indicated that trust in AI is conditioned by knowing the AI is based on the right dataset of BIM, and visualisation of why and how the output was made will help understanding of causality and trust. Moreover, AI technology and other digitalised applications are successfully used when they fulfil maturity in terms of technology, process, and culture. Hanafy (2023) investigated the use of general-purpose technology text-to-image generative AI images was used for architecture. The interaction of the user with the AI system showed that there is a behaviour of refinement as the user did not opt to create the perfect query from the start but built the best one by choosing the best version of adding extra terms. The study found the most suitable aspect of such a tool to be in the early design phase. These tools can be seen as augmenting creativity and visualisation within architecture. In interior design, the tool has potential to allow the user flexibility in design to explore different choices which indicates that the users interact with the output in a form of judgement and learning.

Robotics for Industrialised Construction

In the field of robotics research, Bock (2015) and Sun *et al.* (2023) predicted that the future of construction projects will rely on both robotic and human agents working together to execute construction-related activities. Since the functional operation of robots pivots on embedded sets of data within them (i.e., data are their lifeblood), human-robot interaction (HRI) can be considered a specific instance of HDI. Despite the plethora of research efforts on robotics in construction, only a limited number of studies could be analysed with the goal of distilling a set of specific themes that underlie HRI focusing on the human (Fu *et al.*, 2024).

The psychological and physical safety of human agents in HRI has been the focus of three studies. In their conceptual study, Sun *et al.* (2023) reveal HRI could result in negative psychological impacts on involved human agents including (i) anxiety and acute stress, (ii) cognitive load, (iii) negative emotional state, and (iv) visual and cognitive distraction. These impacts tend to cause increased likelihood of physical accidents, thereby impacting the physical safety of human agents. In the same vein, Rodrigues *et al.* (2023) underline the importance of understanding HRI to protect health and safety of human agents as well as to create an enhanced task environment. Resonating with Sun *et al.* (2023), the authors identify human-related factors that can impact the effectiveness of HRI: (i) mental workload, (ii) situation awareness, (iii) trust in automation, and (iv) (physical) ergonomics. The materialisation of any factor highlighted by Sun *et al.* (2023) and Rodrigues *et al.* (2023) may undermine the perceived benefits of employing robots on construction sites. To this effect, Liu *et al.* (2021) make an attempt at reducing the cognitive and mental load imposed on human agents involved in HRI. They proposed a solution that enables robots to recognize psychological signals of human brains with the goal of adjusting their robotic performance. Such an adjustment is argued to enhance the physical and psychological safety of the involved human agents, thereby improving the efficiency of human-robot collaboration (HRC).

Scholarly suggestions to empower human agents over robotic agents are evident in two studies. Within the context of HRC for modular construction manufacturing (MCM), Fu *et al.*, (2024) place emphasis on the impact of 'human-in-the-loop collaboration' concluding that involved human agents can take on a range of collaborative roles and interact with robotic agents at varying levels. These roles and levels depend on a combination of factors: requirement of the MCM tasks, capacity of robotic agents, and preferences of the human agents. The latter alludes to the value of humans in HRC. Zhang *et al.*, (2023) illuminate this in a recent review on HRC for on-site construction where human agents are tasked with working with robots. They suggest designing human-led systems for HRI to empower humans to control and manipulate robotic agents.

These studies draw our attention to the importance of empowering the human side of the HRI as a proactive measure to enhance interaction efficiency. It can be inferred that the optimization of HRI as a specific instance of HDI relies on human-centric system designs that need to take account of four induced factors: psychological health, physical safety, explicit human value, and empowerment of human agents. These factors can be translated into the three HDI principles as: *legibility*: empowerment of human agents; *agency*: human value and empowerment of human agents; *negotiability*: psychological health and physical safety.

Sensored Construction Sites (SCS)

Information communication technologies (ICT) such as wireless sensors networks, wearables, CCTV, robotics and mobile BIM technologies are increasingly adopted on construction sites (Rossi *et al.*, 2019). Construction sites are challenging environments due to their inherently dynamic, complex, and dangerous nature involving complex interactions between various parties, each with human actors, equipment, and products.

HDI-relevant studies exist across several areas such as safety (e.g., monitoring, compliance, inspection), worker ergonomics (e.g., stress, fatigue), productivity measurement/monitoring, and activity monitoring/ recognition. Common themes include the development and evaluation of technologies like computer vision, sensor technology, and machine learning for safety and compliance monitoring on construction sites. There is a notable gap in addressing human-centred aspects. Many studies do not document participant feedback, assumptions about data ownership, transparency, or algorithm bias. Where participant involvement is mentioned, it often lacks depth in their perspectives or considerations about data matters. Another notable gap is the lack of theories produced or used to support the studies within this domain. This highlights a significant gap in the current research, emphasising the need for more human-centred approaches than human interactions (e.g., consequences/implications/feelings from or when using the developed systems, user acceptance), and data matters (e.g., ethical use, bias, consent, confidentiality, compliance with regulations, etc.).

While the list of papers reviewed is not exhaustive, the trend of insufficient attention to human-centred aspects in safety-related research is indicative of a broader issue in the field. Indeed, safety is a domain where human-centricity is expected to be prominent, and the lack of depth in addressing HDI aspects in these studies suggests a likely widespread gap in the field at large. This trend likely reflects a general oversight in the digitalisation of construction sites, underscoring the need for more comprehensive research that considers HDI, particularly in areas as critical as safety.

Developments in this field will often collect data actively or passively, including workforce location, movements, gestures, physiological levels, and physical efforts (Joshua and Varghese, 2014; Aryal *et al.*, 2017; Park *et al.*, 2017; Jebelli *et al.*, 2018; Tang *et al.*, 2020; Kim *et al.*, 2021; Shanti *et al.*, 2022). However, distinguishing between task analysis and the analysis of the individual performing the task is a delicate matter. It is challenging to determine which data pertain to the project tasks and which data truly reflect the physical and motor abilities of the workers themselves (Calvetti *et al.*, 2021). This complexity underscores the fine line between performance metrics and personal data.

Smart Buildings

As technology continues to advance, a growing interest is observed for integrating cutting-edge technologies and interconnected systems into buildings to enhance their efficiency, functionality, and sustainability. These *smart buildings* leverage various sensors, devices, and automation to create intelligent and responsive spaces that optimize resources, improve occupant comfort, and reduce environmental impact. Subsequently, smart connected devices and advancements in sensing, actuation, and communication bring new modes of interaction within the BE across different contexts and scales. The most common robust examples that illustrate this interaction in smart buildings are interactive systems, which require IoT devices to enable occupants to interact with the building environment actively. Among these systems, automatic climate controllers autonomously regulate indoor air temperature, relative humidity, and air quality in a designated BE (Favero *et al.*, 2022; Lee *et al.*, 2019). Lighting control systems fine-tune the luminous environment and ensure the desired illumination level (Rossi *et al.*, 2015; Cho *et al.* 2020) whereas automatic window systems respond dynamically by automatically opening and closing windows based on fluctuations in indoor air temperature, relative humidity, CO_2 levels, VOC concentrations, as well as PM2.5 and PM10 particles (Cheng *et al.*, 2016; Kim *et al.*, 2019). Most of these systems aim at improving occupant satisfaction, perceived service quality and indoor air quality in buildings.

Recent advancements in this domain evolved around user experience and user-centric features, such as intuitive or guided wayfinding systems, interactive kiosks, and personalised environmental controls. For example, Target retail stores in the United States installed LED lighting systems integrated with Visible Light Communication capabilities. Through the coupling of these LEDs with Visible Light Positioning and the incorporation of image sensors on smartphones, in-store shoppers can navigate in the stores via maps and find the locations of specific products via the directions sent by the retailer's app (Halper, 2019). The EDGE building in Amsterdam exemplifies advanced utilisation of human-building interaction (HBI) data to enhance the user experience via a seamless integration of individual preferences and real-time environmental adaptability. For example, occupants at EDGE find themselves effortlessly connected through an integrated system, with the application aligning with their daily agendas and informing the building of their arrival. The building's advanced system promptly detects incoming vehicles, guiding occupants to designated parking areas. Users have access to a diverse range of work environments and the system, finely tuned to individual preferences, adjusts lighting and temperature within these environments to ensure maximum comfort and efficiency (Randall, 2015).

The current state of smart buildings shows that incorporating human knowledge, behaviour, needs, and preferences into the operational phase is crucial for achieving optimal performance in buildings. Even in automated buildings where users lack direct control over systems, integrating the human dimension through feedback mechanisms can enhance perceived control, leading to higher satisfaction levels. Therefore, HBI emerges as a viable alternative, introducing consensus-based decision-making for building operations to potentially minimize conflicts. It is noteworthy, however, that despite endeavours to create 'smart buildings', this objective remains unattainable without 'smart users'. Users need to be active participants in the building ecosystem rather than passive recipients, engaging with the data available to them in the most effective manner.

Virtual Reality (VR) and Augmented Reality (AR)

With its focus on 3D space on a large scale, the BE is a natural application domain for VR and AR, arguably comparable to entertainment and tourism. Literature on VR/AR within and outwith the BE disciplines presents applications to visualise the 3D BE or abstract data which is somehow related to the design/construction/ operation of buildings but with no natural 3D mapping. Data visualisation is often broadly classified into *information visualisation* and *scientific visualisation*; the former focuses on abstract data, whereas the latter deals with datasets mapped to space and time.

In their position paper, Widjojo *et al.* (2017) define HDI in slightly narrower terms than elsewhere in the literature, drawing an analogy to HCI, but focusing specifically on data. They coin the term *Visual Analytics (VA)* as the integration of human efforts and computation in the analysis of data. They argue that VR technology is particularly suited to VA as it overcomes the challenge of limited display real estate. Coupled with innovative interaction design, VR can facilitate the visual analysis of complex data, such as high-dimensional and multivariate data. The authors highlight the importance of understanding human cognition when exploiting it for VA. They go on to cite various interaction patterns that can inform the design of VR systems for HDI, such as selection, manipulation and viewpoint control. In the case of 3D spatial data, clearly of interest in the BE, Widjojo *et al.* highlight three components of the immersion enabled by VR that are particularly beneficial: head tracking, field of regard, and stereoscopic rendering. These help enable a sense of presence and spatial awareness. The paper culminates with a research agenda framed in an abstract way: research on the required understanding of a specific data set, and on the VR visualisation/interaction functionality that would enable this understanding.

Outside the BE, Alhakamy *et al.*, (2021) study how users perceive correlation and causation in numerical data. They compare two styles of interacting with the data: using gestures and body movements in a large (semi-immersive) environment versus a joypad, similar to the controller used in games. It is reported that participants tended to agree more with statements that portray correlation and causation in data after using the semi-immersive system. The conclusions are presented in the light of theories of embodied learning/cognition.

Two applications of VR/AR in the BE seem implicitly to address the interaction of humans with data, but without explicitly using the term HDI: data for construction safety (Schiavi *et al.*, 2022) and data fused from digital and physical twins to provide spatial context in a smart city (Grübel *et al.*, 2022). The general themes emerging from these are that both VR and AR offer benefits for BE applications in particular because of the 3D orientation of most digital content; VR is used preferentially for office uses whereas AR is more often used with tablets and mobile devices 'in the field'; data flow between the systems used in design, construction and operation of the BE and VR/AR systems remains an issue.

Potential elements for a theory of HDI in the built environment (BE)

Following a review of the literature, the papers were mapped to the framework in Figure 1, as shown in Figure

2. *Understanding and evaluation* focused on what the studies are trying to achieve from the technologies and their potential impact on the end-user in the context of HDI. *Development* focused on any development of the technologies to achieve improvements in HDI for the end-user. *Foundation* considered any proposed frameworks or theories specific to HDI that could lead to an underpinning theory of HDI for the BE that may be applicable across all technologies and concepts.

Trust appears to be a central facet of a system when considering interaction of humans and data. This is applicable in terms of human-centred systems as well as humans knowing the correct data are being used when data-driven decisions are made *about* or *for* them without human input, especially when related to wellbeing. ***Game theory*** is considered as a way to encourage engagement with data. This relates to ***empowerment of humans*** when collaborating with technologies (e.g., robots, surveillance systems) and moving toward automation that (partially) removes ***human control***. Humans are often wary of new technologies and can result in an unwillingness to collaborate. This links to potential issues of ***safety*** of the human from both physical and psychosocial perspectives when humans work with hardware (e.g., robots, wearables) and/or software where concerns may be less detectable. ***Accessibility*** considers how humans interact with data and affects how they make sense of correlation and causation of those data. ***Enhancing understanding*** of data via visualisation can increase accessibility and willingness to engage. The three pillars of ***agency***, ***legibility***, and ***negotiability*** appear across much of the literature, though agency appears more than the other two. Acknowledgement and inclusion of these pillars could be as a result of research offered by Mortier *et al.* (2015) representing seminal work in the field and offering empirical contributions to the term.

Discussion and conclusions

With the pervasive nature of technology and the increasing volume of data produced and then used to drive decision-making in the BE, consideration is required in terms of how humans interact with data, at which levels this takes place, and for what purposes those data are used. Analysis of literature on the theme of HDI sees the emergence of key issues requiring further attention before a theory of HDI for the BE can be conceived. Of the technological themes and concepts evaluated (blockchain; ML and DL; robotics for industrialised construction; sensored construction sites; smart buildings; and VR and AR), there are some interesting elements that can begin to form the basis of such underpinning theory including trust, human empowerment and control, safety, accessibility and understanding of data, and of course agency, legibility and negotiability. However, we are at the beginning of the journey to propose a robust theory for HDI as highlighted by the literature and the many gaps that exist when analysed against a framework of HDI research perspectives spanning i) understanding and evaluation (end user-oriented), ii) development (technology-oriented), and iii) foundation (theory building-oriented). It seems that research has a grasp of understanding and evaluation in terms of the HDI benefits a system or concept is attempting to do for humans, but developments in technology that specifically consider and integrate HDI concepts and foundations to support theory building are somewhat lacking. Also missing across the literature is acknowledgement of education, policy and standardisation that would have implications for HDI, and consideration of the Digital Services Act.

Encapsulating the induced human-driven considerations in a theory will likely extend our understanding of HDI within the context of the analysed technologies and concepts. As an initial step toward building such a theory, future work is proposed as follows: 1) broaden the analysis of literature that considers elements of HDI, whether or not so-called, to further understand the challenges of HDI and the extent to which current technological systems and concepts address issues of HDI within and outwith the BE; 2) establish the array of purposes for which data are collected and/or used by technologies in the BE, for example, to make decisions on behalf of humans or to influence human behaviour, that will enhance our understanding of the true implications of humans' interaction with data; 3) survey individuals and organisations about production, collection, processing and storage of data to achieve such purposes as identified in 2) to identify the benefits, challenges and areas requiring attention in future studies; 4) consider the education, policy and standardisation aspects surrounding HDI; and 5) evaluate any emerging theory against technologies and concepts such as those in this paper and beyond (e.g., BIM, IoT, GIS) to establish its robustness and suitability for computer informatics in the BE.

References

Abioye, S.O. *et al*. (2021) 'Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges', *Journal of Building Engineering*, 44, 103299.

Alhakamy, A.A. *et al.* (2021) 'Show Me How You Interact, I Will Tell You What You Think: Exploring the Effect of the Interaction Style on Users' Sensemaking about Correlation and Causation in Data', in *Designing Interactive Systems Conference 2021*, pp. 564-575.

Allen, G. (2020) *Understanding AI technology*. Joint Artificial Intelligence Center (JAIC), The Pentagon United States.

Aryal, A. *et al.* (2017) 'Monitoring fatigue in construction workers using physiological measurements', *Automation in Construction*, 82, pp. 154–165.

Becherer, M. *et al*. (2020) 'Engineering a Trustworthy Private Blockchain for Operational Risk Management: A Rapid Human Data Engineering Approach Based on Human Systems Engineering', In Handley, H. and Tolk, A. (eds.) A Framework of Human Systems Engineering: Applications and Case Studies, Wiley Online Library, pp. 205-223.

Bilal, M. *et al*. (2016) 'Big Data in the construction industry: A review of present status, opportunities, and future trends', *Advanced engineering informatics*, 30(3), pp. 500-521.

Bock, T. (2015) 'The future of construction automation: Technological disruption and the upcoming ubiquity of robotics', *Automation in Construction*, 59, pp. 113–121.

Bowers, J., and Rodden, T. (1993) 'Exploding the interface: Experiences of a CSCW network', *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, pp. 255–262.

Brown, I. (2013) 'The Economics of Privacy, Data Protection and Surveillance', *SSRN*, Scholarly Paper 2358392. https://papers.ssrn.com/abstract=2358392.

Calvetti, D. *et al*. (2021) 'Human Data Interaction in Sensored Sites, Challenges of the Craft Workforce Dimension', *EC3 Conference 2021 (Online)*, 26-28 July, pp. 173-180.

Card, S. K. (2018) *The Psychology of Human-Computer Interaction*. CRC Press.

Cheng, Z. *et al.* (2016) 'Satisfaction based Q-learning for integrated lighting and blind control', *Energy and Buildings*, 127, pp. 43–55.

Cho, Y. *et al.* (2020) 'Platform design for lifelog-based smart lighting control', *Building and Environment*, 185, p. 107267.

Chowdhury, S.N. and Dhawan, S. (2016) 'HDI based data ownership model for smart cities', in *2016 International Conference on Recent Trends in Information Technology (ICRTIT)*, 8-9 April, Chennai, India: IEEE, pp. 1–5.

Crabtree, A., and Mortier, R. (2015) 'Human Data Interaction: Historical Lessons from Social Studies and CSCW', in N. Boulus-Rødje, G. *et al*. (Eds.), *ECSCW 2015: Proceedings of the 14th European Conference on Computer Supported Cooperative Work*, 19-23 September, Oslo, Norway, pp. 3–21.

Dix, A. (2017) 'Human–computer interaction, foundations and new paradigms', *Journal of Visual Languages & Computing*, 42, pp. 122–134.

EC3 (n.d.) *Human-Data Interaction Committee*. Available at: https://ec-3.org/governance/technical-committees/human-data-interaction-committee/ (accessed: 9 January 2024).

Ekblaw, A. et al. (2016) 'Bitcoin and the myth of decentralization: Socio-technical proposals for restoring network integrity', *2016 IEEE 1st International Workshops on Foundations and Applications of Self* Systems (FAS*W)*, Augsburg, Germany, pp. 18–23.

Favero, M. *et al.* (2022) 'Human-in-the-loop methods for occupant-centric building design and operation', *Applied Energy*, 325, p. 119803.

Fernandez, L. (2021) 'Teaching students how to frame human-computer interactions using instrumentalism, technological determinism, and a quadrant learning activity', *Frontiers in Computer Science*, 3, p. 771731.

Fu, Y. *et al.* (2024) 'Human-robot collaboration for modular construction manufacturing: Review of academic research', *Automation in Construction*, 158, p. 105196.

Grudin, J. (1990) 'The computer reaches out: The historical continuity of interface design', *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 261–268.

Grübel, J. *et al*. (2022) 'The hitchhiker's guide to fused twins: A review of access to digital twins in situ in smart cities', *Remote Sensing,* 14(13), 3095.

Haddadi, H. *et al*. (2013) *Human-data interaction*. (UCAM-CL-TR-837). University of Cambridge.

Halper, M. (2019) 'Target finally confirms that Acuity provides the IoT lighting', *LEDs Magazine*. Available at: https://www.ledsmagazine.com/leds-ssl-design/networks-controls/article/16699076/target-finally-confirms-that-acuity-provides-the-iot-lighting (accessed: 8 January 2024).

Hanafy, N.O. (2023) 'Artificial intelligence's effects on design process creativity: A study on used AI Text-to-Image in architecture', *Journal of Building Engineering*, 80, p. 107999.

Hunhevicz, J.J. *et al*. (2021) 'no1s1 – a blockchain-based DAO prototype for autonomous space', *Proceedings of the 2021 European Conference on Computing in Construction [Online]*, 26-28 July, pp. 27–33.

Hornung, H. *et al*. (2015) 'Challenges for Human-Data Interaction – A Semiotic Perspective', in: Kurosu, M. (eds) *Human-Computer Interaction: Design and Evaluation*. HCI 2015. Lecture Notes in Computer Science, 9169. Springer, Cham.

Jebelli, H. *et al.* (2018) 'EEG-based workers' stress recognition at construction sites', *Automation in Construction*, 93, pp. 315–324.

Jenkins, C. *et al.* (2019) 'Effective management of plug loads in commercial buildings with occupant engagement and centralized controls', *Energy and Buildings*, 201, pp. 194-201.

Joshua, L. and Varghese, K. (2014) 'Automated recognition of construction labour activity using accelerometers in field situations', *International Journal of Productivity and Performance Management*, 63(7), pp. 841–862.

The remaining references in this paper are available at this link: https://drive.google.com/file/d/1U30mDhy6BluXQPXuyopZ1VoPsSErsxGa/view?usp=drive_link.

AUTOMATED COMPLIANCE CHECKING SITUATION IN HEALTH AND SAFETY MANAGEMENT IN UK'S INFRASTRUCTURE SECTOR

Zijing Zhang[1,3,4], Algan Tezel[2], Maxwell Fordjour Antwi-Afari[1], Tala Kasim[1], and Doug Potter[3]
[1]Aston University, Birmingham, UK
[2]University of Nottingham, Nottingham, UK
[3]Arcadis Consulting, UK
[4]University College London, London, UK

Abstract

Despite intensive research efforts in developing automated compliance checking (ACC) systems for improved health and safety (H&S) outcomes, existing research has mainly focused on the technical aspects and the building sector, whilst research focusing on the social and organisational aspects of ACC in the infrastructure sector is lacking. To address this gap, this study used a case study to explore the issues and readiness to implement ACC in H&S in a design organisation in the UK's infrastructure sector. The findings show there is low readiness with issues in four main aspects, where suggestions were proposed. This research contributes to the body of knowledge by identifying the gaps and proposing ways to achieve better readiness for implementing ACC in design organisations in the UK's infrastructure sector.

Introduction

The architectural, construction and engineering (AEC) industry is widely exposed to numerous health and safety (H&S) concerns (Anwer et al., 2021). Fatalities in the AEC industry are significantly higher than the all-industry rate in the UK. H&S concerns are especially prominent in infrastructure projects that are complex, dynamic and involve many workers and plant on site (Alhammadi et al., 2022).

In the design stage, safety by design has proven to play a key role in achieving better H&S outcomes in infrastructure projects by identifying H&S hazards during the design stage (Hardison and Hallowell, 2019). With the technological advancement in the AEC industry, there have been many digital tools developed to improve the industry's H&S performance, including knowledge-based systems, hazard visualisation, safety training, and automated compliance checking (ACC) systems.

Among these developments, ACC systems dealt with compulsory regulatory requirements and have shown promising results in improved H&S compliance (Zhang et al., 2013a). Despite some existing ACC research for H&S management that mainly focused on the technical aspect, little is known regarding the social and organisational aspects of implementing such ACC systems, which focuses on the readiness of ACC adoption, and inter- and intra-organisational factors affecting such readiness. To address this research gap, this study aims to explore the ACC situation in H&S aspects in the UK's infrastructure sector, using a design consultancy organisation as an example. Specifically, the objectives are to identify the gaps and issues, and evaluate the readiness of ACC for H&S at a design consultancy organisation in the UK's infrastructure sector and propose some potential improvements. As part of an ongoing project, this paper will outline the initial findings about the current situation, readiness and gaps in implementing ACC in H&S aspects in the UK's infrastructure sector.

The remainder of the paper is structured as follows. The next section presents the literature review related to the research objectives. The third section details the methodology used in this paper. Next, the results and findings are presented. The fifth section discusses the results. The last section concludes this paper with remarks for future work.

Literature review

Digitalisation in the construction industry

Digitalisation is the gradual process of widespread adoption of digital technologies to generate new revenue streams or to improve the value-generation capacity of existing business workflows (Bajpai and Misra, 2022). With the advent of the "digital revolution", digital transformation has been one of the key trends in businesses over the last few decades. In the AEC industry, digitalisation is also seen as one of the primary means to improve the industry's overall value delivery performance, and eventually "modernise" construction by today's standards.

With external factors such as changing market requirements, technological advancements, decreasing hardware and software costs, increasing complexity of projects, a new generation of digitally adept workforce, growth of start-ups, supportive and demanding trends (e.g., digital mandates) induced by governments, policymakers and clients, digitalisation in construction is progressing at a pace. This is also due to internal drivers such as the strategic rhetoric for digital transformation to be the main solution for inefficiencies, increased profitability expectations with efficiencies gained through digitalisation and building a "modern" company image.

To this end, several factors and adoption frameworks for digitalisation have been outlined in the literature. Many studies of technology adoption in construction use theories drawn from the established body of knowledge in information systems. The digitalisation adoption process has been studied from a socio-technical, organisational and psychological (individual) perspective (Sepasgozar et al., 2016). For construction organisations, among the many factors affecting digitalisation, external factors (e.g., legislation and regulation, mandates, and pressure from clients and competitors, market demand, and standards and specifications) (Aghimien et al., 2022; Liu et al., 2023) are important. Also, organisational factors (e.g., organisational culture, structure, leadership, and internal processes) (Zulu et al., 2023) are pivotal. Finally, availability of resources (human resources, hardware and software, organisational knowledge and experience, integration of different systems and platforms across the supply chain, standardisation of data, systems, and workflows) (Li et al., 2018) come to fore.

Automated compliance checking

ACC has been an area attracting global research interests and commercial development efforts for more than 60 years. In the ACC literature, most research efforts focused on technical aspects, including two main themes: target design model and rule representation.

Research on the target design model topic mainly focused on developing various lightweight data schema or data retrieval or query methods to extract useful data from the design model due to the difficulty of data exchange and filtration for ACC tasks. Examples include using graph databases to represent design information (Ismail et al., 2017a). Nevertheless, efficient retrieving and querying BIM model data remains a challenge. Methods that enable spatial query of Industry Foundation Classes (IFC) data are especially scarce.

Rule representation methods aim to represent building rules in a machine-readable form by interpreting and capturing rules. Early research developed decision tables, object-oriented and logic-based representations (Yabuki and Law, 1993; Fenves et al., 1995; Han et al., 1998).

More recently, researchers have proposed domain-specific languages to represent building requirements, including text-based languages (Lee et al., 2015; Sydora and Stroulia, 2020) and visual programming languages (Preidel and Borrmann, 2016; Kim et al., 2019). They both aimed to develop easy-to-use methods for domain experts to write codes that represent building requirements. Other researchers focused on semantic web technology-based methods, which used query or rule languages such as SPARQL (Jiang et al., 2022) or SWRL (Beach et al., 2015) to represent building rules. Nonetheless, most existing rule representations are not able to capture the full meaning of the building rules and the rule interpretation process still relies on some manual efforts (Zhang et al., 2023b).

Several studies have focused on rule classification and organisation to support rule representation. Solihin and Eastman (2015) classified building requirements into four classes based on their computational complexity. Zhang et al. (2022) proposed a more comprehensive four-criteria (i.e., semantic constructs, intensity, self-contained or linked explanatory, prescriptive or performance-based) classification to categorise building rules.

The recent advancements in natural language processing (NLP) and machine learning (ML) have made full automation of compliance checking possible. Zhang and El-Gohary (2017) used NLP techniques to automatically generate a logical representation of the building codes. Nisbet et al. (2023a) proposed a rule-based approach using RASE to automatically generate SPARQL queries for rule representation. Nevertheless, ACC based on NLP methods has not yet achieved 100% accuracy. Extra efforts are still required to review and check the results. Also, many methods can only deal with quantitative rules.

There have also been several studies assessing ACC implementation readiness and efficiency. For example, Beach et al. (2020) used questionnaire surveys to understand the obstacles of adopting ACC in the UK. Results showed that the top three challenges are: shared open standards for regulation clauses are lacking, no tool can be used for complete pre-submission checks, and difficulties of making brief and regulatory requirements contractually enforceable. They also provided a roadmap based on interviews with experts, highlighting that 1) there have been some interest in ACC adoption from the UK government and government commitment would be crucial; 2)ACC checking results should be used with expert review; and 3)ACC may be more suitable to be used for conventional projects instead of multi-use, complex ones. Zou et al. (2022) conducted a case study to assess New Zealand's offsite manufacturing industry's readiness for ACC implementation. They suggested that improving readiness of ACC requires assessing ACC systems in different scenarios, further improving technical maturity and promoting education and training. Their later research captured lessons learned from the implementation of ACC globally (Zou et al., 2023), which highlighted the relatively low readiness for ACC adoption and the importance of governments' role in promoting ACC adoption.

Automated compliance checking for health & safety management

There have not been many research efforts on using ACC for H&S regulations. One of the earliest ones were Zhang et al. (2013a) and Zhang et al. (2013b), where rule sets for ACC against Occupational Safety and Health Administration (OSHA) regulations were developed (Occupational Safety and Health Administration, 2023) based on existing safety in design best practices. These two studies selected geometry-related rule sets, such as workways and egress rule sets. Geometrical attributes were used for compliance checking, including the dimensions of holes in slabs and openings in walls. Similarly, Qi et al. (2014) developed rule sets for fall protection using both Solibri Model Checker (Solibri, 2024) and BIM Server as model-checking platforms. More recently, a study by Getuli et al. (2017) used parametric tables to represent Italian Construction H&S normative texts.

The literature shows that the studies regarding ACC for H&S regulations are limited, and they have mainly focused on the technical aspects. To the best of the authors' knowledge, there has been no practical implementation of the developed ACC systems on H&S aspects in the AEC industry. Especially, there has been no substantial research on ACC with respect to the Construction (Design and Management) Regulations 2015 (CDM), which are important H&S regulations in the UK's infrastructure sector and particularly no research on this from a social and organisational perspective.

ACC affecting parameters and success factors

Drawing on insights from prior scholars, the effectiveness of ACC for H&S regulations is subject to key factors. Kamara et al. (2002) highlighted "effective knowledge representation" as a crucial element, the need for standardised approaches, minimising inconsistencies, and fostering interoperability. Emphasising the importance of "oversight for performance-based criteria", Amor and Dimyadi (2021) stressed clear guidelines and human expert involvement to reduce subjectivity in H&S compliance assessments. This is echoed by Zhang et al. (2023a), where the importance of interpreting ambiguous clauses correctly was highlighted. Fuchs and Amor (2021) highlighted the challenge of "accuracy in information classification", suggesting the importance of mapping criteria against information models in ACC. Advocating for "adaptable information modelling approaches", Ismail et al. (2017b) and Nawari (2019) suggested fostering interoperability and collaboration in diverse construction projects. In addition, ensuring the "accurate generation of BIM data" is emphasised by Ismail et al. (2017b), who also suggest robust quality control should be involved to reinforce reliability in H&S compliance assessments. Beach et al. (2020) underscored the "quality and transparency of compliance information", urging advanced techniques in NLP and ML for ACC. Streamlining "quality assurance and control processes", as identified by multiple scholars, is another success factor for ACC. The process must involve strategic approaches to balance process validation and practical implementation, thereby enhancing the overall H&S compliance assessment (Beach et al., 2020; Amor and Dimyadi, 2021). Addressing these factors collectively can elevate the reliability, efficiency, and overall effectiveness of ACC systems tailored for H&S regulations.

Methodology

In this paper, the authors adopted a qualitative approach. An exploratory case study method was used to explore and gain a deeper understanding of the current issues and requirements of automated H&S regulatory compliance in the infrastructure sector. The exploratory case study method is suitable for this study as it allows more in-depth understanding of a scarcely researched topic in its own context (Yin, 2009). It also allows collecting and analysing both primary (interviews, questionnaire surveys) and secondary data (documents) (Eriksson and Nilsson, 2008; Zuo et al., 2013), which is the case for this study.

To achieve the research objectives, the authors targeted one organisation to 1) have a snapshot of the current issues, gaps and requirements in digital H&S management in the infrastructure sector, and 2) gain a more in-depth understanding of their digital H&S management processes and assess their readiness for ACC in H&S aspects. This organisation is a large multi-national design consultancy in the UK's infrastructure sector, which usually takes principal designer and designer duties as specified in the CDM regulations.

Primary and secondary data was collected, including data collected from a 5-month observation while working within the organisation and 50 questionnaire surveys answered by project managers (PMs) (or other senior employees working on 21 small- medium- or large-sized projects). Five individual interviews with CDM designer managers were also conducted to complement the data gathered from the questionnaires. In addition, one design risk management schedule (DRMS) and one H&S compliance audit document were collected for review.

All collected data was then analysed using thematic analysis to elicit the emerging themes. The thematic analysis process generally includes 6 steps, namely data familiarisation, coding, generating themes, reviewing themes, defining and naming themes and writing up (Braun and Clarke, 2006). The next section presents the results and findings.

Results

Based on the data analysis, the authors found that overall, some gaps exist in implementing ACC in the H&S aspects of the UK's infrastructure sector and some existing issues need to be addressed before ACC is ready to be implemented. This section first presents the general H&S management process of the organisation, followed by four themes that emerged from the collected data, namely regulation, technology, human factors and culture, and external environment aspects, as presented in Table 1 and detailed below.

General Design Risk Management Process

As a design organisation, the most important H&S management processes are those relating to the design risk management (DRM) process. In the organisation, the DRM processes generally include obtaining relevant pre-construction information, identification of constraints, production of a DRM schedule (DRMS) to identify hazard and risk levels, developing mitigations with an emphasis on hazard elimination or reduction, reviewing the DRMS, and handing over the residual risks register to the principal contractor. To facilitate the implementation of the DRM processes, there are some organisational-level guidance documents available, which are applicable to all projects regardless of sub-sector (e.g., highways, water) or project size. Some documents need updating, resulting in partial implementation and some fragmentation of processes on some projects. Based on the observation, how well the process is followed in practice and whether the actual processes comply with the CDM regulations mainly depends on the quality of DRM (the risks identified, proposed mitigations, whether the risk levels are addressed) and the process of reviewing DRMS. The DRMS is typically reviewed at an Integrated Design Review (IDR) with the CDM principal designer manager, Design Manager and design discipline leads. As such, the DRM processes heavily rely on DRM experts' competency (e.g., skills, experience, and knowledge). Although several lessons learned workshops are held periodically, more workshops would capture more tacit knowledge of the experts to assist the development of a unified and comprehensive digital CDM compliance checking tool, which will complement the current DRM system.

Table 1: Main themes and subthemes emerged

Themes	Sub-themes
Regulation	Understanding and interpretation of regulations
	Metrics for assessing compliance against regulations
Technology	Awareness of ACC and the required technical capabilities
	Related digital systems and technical capabilities
	Data availability for ACC
Human factors and culture	Competency and upskilling of employees (CDM duty holders)
	Standardisation for risk mitigation methods and severity ratings at an organisational level
	Culture and behaviour changes in safety by design and CDM compliance
	Awareness of digital technology and tools
External environment	External initiatives from government bodies and client organisations

Regulation Aspects

A good understanding of the CDM regulations is imperative in achieving CDM compliance. From the interview and questionnaire survey responses, staff within the design team have various levels of knowledge, experience and understanding of the CDM regulations. Employees who specialise in the CDM and DRM (e.g., CDM principal designer managers) have excellent knowledge, while some other designers only have a basic understanding of CDM regulations. It was also observed that some ambiguous expressions in the CDM clauses lead to differences in understanding and interpretation of the regulations.

A similar issue is that as the CDM regulations are performance-based and not very descriptive in nature, it is difficult to produce specific metrics that directly assess CDM compliance. Ideally, detailed metrics would help the design team understand what specific measures or aspects of hazard mitigation to pay attention to. This resonates with the suggestion by Amor and Dimyadi (2021) on providing clear guidelines. The current CDM Compliance Audit form refers to the general clauses provided in the CDM regulations, such as "ensure the client makes suitable arrangements for managing a project", which renders the interpretation subjective based on the auditor's experience. However, this subjectivity is alleviated as the CDM Compliance Audit makes reference to the check sheets in DRMS as evidence.

As such, first, for organisations adopting a similar approach, improving the granularity of their compliance audit forms will better support their ACC efforts by reducing the potential subjectivity in assessment. Second, improving the metrics and providing detailed duty definitions for the requirements to ensure standardisation in practice will be useful.

Technology Aspects

Generally, the technology for the key tasks of ACC (i.e., data retrieval from and semantic enrichment of BIM models, knowledge representation and natural language processing) is relatively mature, although no system has achieved ACC fully automatically. Within the organisation, although there are existing expert systems for H&S management, no ACC system or similar is currently in use. In addition, employees across the design delivery workstreams are not familiar with ACC nor the underlying technologies of ACC. Apart from several experts, they also lack awareness regarding what can be achieved using ACC and what level of automation can be achieved based on the organisation's current technological capabilities.

In addition, the questionnaire responses show that three main methods for H&S management on projects exist, with different levels of digital implementation and technical capabilities. These methods include 1) A spreadsheet based DRMS; 2) A spreadsheet based DRMS linked to hazard triangles on drawings with reference number; and 3) A GIS-driven digital CDM system, where the spreadsheet-based DRMS can be imported or exported. All three methods can be used on various sizes of projects, while the third method is mainly used on large projects. Technical skills and experience in developing digital DRM systems may be helpful for developing an

ACC system, as there are similar elements (such as functionalities for manipulating BIM information). Some of the data required for ACC may also be available in the DRM systems. Currently, new processes are in development across the organisation to cater for different project sizes.

Nevertheless, data availability seems to be an issue in automating CDM compliance checking. Although all projects must be CDM compliant, data required for ACC is stored across several standard forms used on the projects which are not directly linked. This makes it less efficient to demonstrate compliance, especially when considering the lack of definition and ambiguous nature of some of the clauses. Despite being complex to acquire, high quality data is crucial for successful implementation of ACC, as suggested by Ismail et al. (2017b).

Human Factors and Culture Aspects

DRM is an essential part of achieving CDM compliance. It requires competent CDM duty holders to make endeavors to finish the tasks to a high quality. From the interviews with designers and CDM managers and the authors' observations, the current situation and gaps are summarised, as follows.

First, some more junior members of staff may face challenges in fulfilling their tasks due to potential gaps in skills, experience, and/or knowledge. For instance, designers might encounter difficulty in identifying all risks or suggesting appropriate mitigation methods. This suggests a need for more support and efforts for training and upskilling some of the duty holders involved. There can also be inefficiencies in determining the competency of employees when appointing competent duty holders. It was noted that a new digital competency system had been developed to improve recording of the skills, experience and knowledge of employees which will improve the efficiency of the project personnel selection process.

Second, at a national and organisational level, the standardisation for risk mitigation methods and severity ratings for different risks should be improved. Various designers have employed diverse severity and mitigation approaches. This presents a challenge in compliance verification, particularly when considering automated checks. It was noted that work to address this matter is being championed by the organisation at a company and national level.

Third, it was observed that some designers tend to see the CDM compliance as mostly a safety related matter and do not fully understand the health implications engendered by construction work. There needs to be a cultural change at a national level to truly embracing the benefits of improving health and safety by design.

Fourth, there have been varied awareness and expectations of digital tools in H&S. Some have high expectations of such tools, thinking their functionality will be so powerful that most requests can be achieved. Others, however, are more skeptical about digital tools, arguing that the results produced by digital systems may not always be reliable and must be reviewed by experts. Regarding their expectations for ACC, they expect ACC systems to not only check compliance but also highlight where the non-compliance lies and suggest how to achieve compliance. In addition, Some CDM experts expressed their concerns regarding the use of digital technologies in CDM compliance, as such automation may result in the future generation of designers lacking basic knowledge regarding H&S risk identification and mitigation.

External Environment Aspects

There has been increasing support from the UK government in digitalisation generally and the H&S aspects in the construction industry specifically. The UK government published a new digital strategy in 2022, which set out the visions and continuous support the government will provide. The Health and Safety Executive (HSE) has established the Discovering Safety programme, where there have been a handful of projects focusing on improving H&S performance in the construction industry (Health and Safety Executive, 2023a). Fundings were provided for H&S technology companies to develop novel solutions via the Industrial Safetytech Regulatory Sandbox. Recently, there has been further funding secured for the HSE's Regulatory Technology sandbox to demonstrate use cases of ACC and help innovators bring new ACC products to market (Health and Safety Executive, 2023b). The HSE also recently commissioned ACC experts in the UK to deliver an ACC workshop for regulators, showing their great interest in ACC and their aspirations in achieving ACC.

There are also various initiatives driven by large client organisations in the UK's infrastructure sector, such as National Highways' Digital Roads 2025 initiative. They work with various industrial and academic partners to improve outcomes in the whole life cycle of their projects. In H&S aspects, this initiative aimed to achieve enhanced onsite safety using data and digital tools, reducing fatalities and proactively managing risk by the end of 2023. This highlights National Highways' commitment to future digitalisation in construction H&S.

Overall, the favourable external environment provides great opportunities for developing tools or systems for automated CDM regulatory compliance and client or internal process compliance.

Discussion

The results and findings of this research highlighted the current lack of readiness for developing and implementing ACC systems in H&S aspects in design organisations of the UK's infrastructure sector. There has not been much literature focusing on the social and organisational aspects of ACC. In comparison with the limited existing literature, findings of this research is consistent with previous findings in the UK construction sector in general, which mirrors similar earlier findings in New Zealand's offsite manufacturing industry (Beach et al., 2020; Zou et al., 2022). This relatively low level of readiness is attributed to various causes, including regulation, technology, human factors and culture and external environment aspects. In this section, the authors discuss the existing issues and suggest some potential solutions that could improve such readiness.

First, the difficulty of having unified interpretations of H&S regulations in general and CDM regulations in specific has hindered the development of ACC tools for such regulations. The possible confusion and multiple interpretations were highlighted in the recent HSE digital regulatory compliance workshop report (Nisbet et al., 2023b). To alleviate this issue, design organisations could initiate more detailed internal review meetings and workshops on the CDM regulations with the wider CDM team and with the assistance of the legal team. Such internal review meetings aim to 1) share knowledge related to the interpretation and practical CDM compliance experience; 2) agree on an organisational-level standard of unified interpretations for ambiguous expressions in the CDM regulations; 3) standardise some elements related to CDM compliance such as risk severity definitions (risk matrix); 4) better define the detailed matrix for CDM compliance checking; and 5) produce a "traffic light" system that highlights which CDM regulations can (or are easier to) be checked automatically.

More fundamentally, a CDM regulations review from the ACC perspective by regulators at the HSE could be helpful. This review could 1) analyse the previous incidents that are directly or indirectly occurring because of the lack of understanding of ambiguous expressions in the CDM regulations (e.g., so far as is reasonably practicable); 2) check if any clause needs to be updated; 3) revise clauses that are not very clear and/or provide official guidance and examples for those clauses; 4) suggest what changes can be made to make the CDM regulations both human- and machine-readable, which is aligned with the HSE's digital regulatory compliance agenda.

Regarding the technological aspect, design organisations could analyse technological capabilities to automate CDM regulation compliance, which will highlight the gaps and opportunities. This can be reviewed in line with digital DRM systems to identify information gaps, and how the DRM and ACC systems can help each other to achieve enhanced CDM compliance outcomes. Care should be taken to consider the different sizes of projects, as the suitability of the CDM compliance processes and methods may be different. In addition, existence of a multitude of digital systems, platforms, and initiatives used across different projects make it more difficult for standardising and integrating (both from a technology and process perspective) an ACC approach in practice. There seems to be a "noise" in digitalisation in the AEC industry, where many ongoing efforts and ideas, and external factors compete for attention and resources.

Nevertheless, the advancement in technology cannot guarantee the successful adoption of digital CDM compliance solutions alone. It faces similar challenges as adopting other digital technologies in the infrastructure sector, where changing the culture and people's behaviour is one of the most prominent issues. Design organisations should take the initiative to promote and highlight the importance and the value of the health and safety by design approach, with evidence from other pioneering organisations to demonstrate feasibility and incentives for designers to facilitate behavioural changes. Design organisations could also organise more specific training sessions for their employees to showcase the available digital innovations in this area. This would highlight what functions are achievable and/or only speculative, to calibrate the expectations from digital tools in general and ACC in specific. Senior managers should understand and communicate clearly what ACC means, how it is performed and what could be practically expected from such systems. This would also guide their subordinates' conduct.

In addition, despite the encouraging external environment for digitalisation in the infrastructure sector, developing a new digital solution in the commercial world requires business justification. Hence, such an ACC system needs to demonstrate its tangible benefits such as efficiency, productivity gain, and cost-savings for the design organisation. This needs to be supported by organisational investment and resources (e.g., testbed projects for pilot and feasibility studies). If the existing ACC systems around the world are proven to be efficient, government policy incentives or mandates could also help achieve digital compliance. Some government initiatives to review the regulatory framework and regulation clauses could also facilitate its digitalisation process. The current regulatory framework could be improved to provide more guidance on achieving improved H&S outcomes. Firstly, the UK BIM framework has little guidance and focus on H&S aspects (CIRIA, 2023). Currently, there is only PAS 1192-6 available (British Standards Institution, 2024), with ISO 19650-6 out for consultation. More comprehensive guidance needs to be provided. Secondly, the CDM 15 regulations are performance-based and objective-oriented documents where guidance on paths to compliance is lacking. Some design examples could be provided to further explain and clarify how the requirements could be met.

Conclusions

This research presents an exploratory single case study of a design consultancy in the UK's infrastructure sector to understand the issues, gaps and readiness for implementing ACC in its H&S aspects. The results show that the readiness for implementation could be improved.

The issues and gaps mainly lie in aspects that include regulation, technology, human factors and culture, and the external environment. Suggestions are proposed for the design organisations to raise employees' awareness of technological advancements and enhance organisational knowledge sharing, training, standardisation and facilitation of culture changes, especially related to the H&S management and CDM compliance domain. The authors also suggested that the regulatory framework and regulations could be reviewed by regulators to facilitate digitalisation.

This study is one of the first studies looking at the readiness of implementing ACC in the H&S aspect of the UK's infrastructure sector. Unlike most existing ACC studies that focus on technical issues of ACC systems, it studied ACC from a socio-technical perspective to understand the readiness and how to facilitate better implementation from social and organisational perspectives. The results would be helpful for design organisations in the UK's infrastructure sector towards better development and implementation of digital regulatory compliance tools for CDM regulations.

This paper has limitations. An exploratory single case study is conducted in this paper, which may not reveal the wider H&S management situation in the UK's infrastructure sector. Our future study will conduct more case studies in similar contexts and look at the topic from different analysis units (e.g., sectoral, supply chain, organisational, departmental, team and individual) to obtain a better assessment and understanding of the situation and propose more comprehensive suggestions to improve this situation.

Acknowledgments

Appreciation is expressed to the company and all participants for their support in this study. The authors would also like to thank UKRI, as this investigation was developed through a Knowledge Transfer Partnership (KTP) project supported by Innovate UK.

References

Aghimien, D., Aigbavboa, C., Oke, A. & Aghimien, L. 2022. Latent institutional environment factors influencing construction digitalization in South Africa. *International Journal of Construction Education and Research,* 18**,** 142-158.

Alhammadi, S. A., Tayeh, B. A., Alaloul, W. S. & Jouda, A. F. 2022. Occupational health and safety practice in infrastructure projects. *International journal of occupational safety and ergonomics,* 28**,** 2631-2644.

Amor, R. & Dimyadi, J. 2021. The promise of automated compliance checking. *Developments in the built environment,* 5**,** 100039.

Anwer, S., Li, H., Antwi-Afari, M. F. & Wong, A. Y. L. 2021. Associations between physical or psychosocial risk factors and work-related musculoskeletal disorders in construction workers based on literature in the last 20 years: A systematic review. *International Journal of Industrial Ergonomics,* 83**,** 103113.

Bajpai, A. & Misra, S. C. 2022. Evaluation of success factors to implement digitalization in the construction industry. *Construction Innovation*.

Beach, T. H., Hippolyte, J.-L. & Rezgui, Y. 2020. Towards the adoption of automated regulatory compliance checking in the built environment. *Automation in construction,* 118**,** 103285.

Beach, T. H., Rezgui, Y., Li, H. & Kasim, T. 2015. A rule-based semantic approach for automated regulatory compliance in the construction sector. *Expert Systems with Applications,* 42**,** 5219-5231.

Braun, V. & Clarke, V. 2006. Using thematic analysis in psychology. *Qualitative research in psychology,* 3**,** 77-101.

British Standards Institution. 2024. *PAS 1192-6:2018* [Online]. Available: https://knowledge.bsigroup.com/products/specification-for-collaborative-sharing-and-use-of-structured-health-and-safety-information-using-bim?version=standard [Accessed 19/01/2024].

Ciria. 2023. *UK BIM Framework* [Online]. Available: https://www.ukbimframework.org/ [Accessed 19/01/2024].

Eriksson, P. E. & Nilsson, T. 2008. Partnering the construction of a Swedish pharmaceutical plant: Case study. *Journal of Management in Engineering,* 24**,** 227-233.

Fenves, S. J., Garrett, J. H., Kiliccote, H., Law, K. H. & Reed, K. A. 1995. Computer representations of design standards and building codes: US perspective. *The International Journal of Construction Information Technology,* 3**,** 13-34.

Fuchs, S. & Amor, R. Natural language processing for building code interpretation: A systematic literature review. 2021 2021. 11-15.

Han, C. S., Kunz, J. C. & Law, K. H. 1998. Client/server framework for on-line building code checking. *Journal of computing in civil engineering,* 12**,** 181-194.

Hardison, D. & Hallowell, M. 2019. Construction hazard prevention through design: Review of perspectives, evidence, and future objective research agenda. *Safety Science,* 120**,** 517-526.

Health and Safety Executive. 2023a. *Discovering Safety* [Online]. Available: https://www.discoveringsafety.com/ [Accessed 18/01/2024].

Health and Safety Executive. 2023b. *Discovering Safety wins further funding following the success of the world's first industrial safetytech regulatory sandbox* [Online]. Available: https://www.discoveringsafety.com/index.php/news/discovering-safety-wins-further-funding-following-success-worlds-first-industrial-safetytech [Accessed 18/01/2024].

Ismail, A., Nahar, A. & Scherer, R. 2017a. Application of graph databases and graph theory concepts for

advanced analysing of BIM models based on IFC standard. *Proceedings of EGICE*, 161-173.
Ismail, A. S., Ali, K. N. & Iahad, N. A. A review on BIM-based automated code compliance checking system. 2017 2017b. IEEE, 1-6.
Jiang, L., Shi, J. & Wang, C. 2022. Multi-ontology fusion and rule development to facilitate automated code compliance checking using BIM and rule-based reasoning. *Advanced Engineering Informatics,* 51**,** 101449.
Kamara, J. M., Augenbroe, G., Anumba, C. J. & Carrillo, P. M. 2002. Knowledge management in the architecture, engineering and construction industry. *Construction innovation,* 2**,** 53-67.
Kim, H., Lee, J.-K., Shin, J. & Choi, J. 2019. Visual language approach to representing KBimCode-based Korea building code sentences for automated rule checking. *Journal of Computational Design and Engineering,* 6**,** 143-148.
Lee, Y. C., Eastman, C. M. & Lee, J. K. 2015. Automated rule-based checking for the validation of accessibility and visibility of a building information model. *Computing in Civil Engineering 2015*.
Li, L., Su, F., Zhang, W. & Mao, J. Y. 2018. Digital transformation by SME entrepreneurs: A capability perspective. *Information Systems Journal,* 28**,** 1129-1157.
Liu, H., Yu, H., Zhou, H. & Zhang, X. 2023. Research on the Influencing Factors of Construction Enterprises' Digital Transformation Based on DEMATEL-TAISM. *Sustainability,* 15**,** 9251.
Nawari, N. O. 2019. A generalized adaptive framework (GAF) for automating code compliance checking. *Buildings,* 9**,** 86.
Nisbet, N., Zhang, Z. & Ma, L. Automated generation of SPARQL queries from semantic mark-up. 2023 2023a. European Council on Computing in Construction.
Nisbet, N., Zhang, Z., Williams, T. & Dobos, J. 2023b. Regulation Mapping-making human-readable requirements machine-operable. *In:* EXECUTIVE, H. A. S. (ed.) *Discovering Safety.* Health and Safety Executive.
Occupational Safety and Health Administration. 2023. *Regulations (Standards - 29 CFR)* [Online]. Available: https://www.osha.gov/laws-regs [Accessed 08/12/2023].
Preidel, C. & Borrmann, A. 2016. Towards code compliance checking on the basis of a visual programming language. *Journal of Information Technology in Construction (ITcon),* 21**,** 402-421.
Sepasgozar, S. M. E., Loosemore, M. & Davis, S. R. 2016. Conceptualising information and equipment technology adoption in construction: A critical review of existing research. *Engineering, Construction and Architectural Management,* 23**,** 158-176.
Solibri. 2024. *Solibri Model Checker* [Online]. Available: https://www.solibri.com/ [Accessed 11/01/2024].
Solihin, W. & Eastman, C. 2015. Classification of rules for automated BIM rule checking development. *Automation in construction,* 53**,** 69-82.
Sydora, C. & Stroulia, E. 2020. Rule-based compliance checking and generative design for building interiors using BIM. *Automation in Construction,* 120**,** 103368.
Yabuki, N. & Law, K. H. 1993. An object-logic model for the representation and processing of design standards. *Engineering with Computers,* 9**,** 133-159.
Yin, R. K. 2009. *Case study research: Design and methods*, sage.
Zhang, J. & El-Gohary, N. M. 2017. Semantic-based logic representation and reasoning for automated regulatory compliance checking. *Journal of Computing in Civil Engineering,* 31**,** 04016037.
Zhang, S., Teizer, J., Lee, J.-K., Eastman, C. M. & Venugopal, M. 2013a. Building information modeling (BIM) and safety: Automatic safety checking of construction models and schedules. *Automation in construction,* 29**,** 183-195.
Zhang, S., Teizer, J., Perez, E. & Mcdonald, M. Automated safety-in-design rule-checking for capital facility projects. 2013 2013b.
Zhang, Z., Ma, L. & Nisbet, N. 2023a. Unpacking ambiguity in building requirements to support automated compliance checking. *Journal of Management in Engineering,* 39**,** 04023033.
Zhang, Z., Nisbet, N., Ma, L. & Broyd, T. 2022. A multi-representation method of building rules for automatic code compliance checking. European Conference on Product and Process Modeling 2022, 2022 Trondheim, Norway.
Zhang, Z., Nisbet, N., Ma, L. & Broyd, T. 2023b. Capabilities of rule representations for automated compliance checking in healthcare buildings. *Automation in Construction,* 146**,** 104688.
Zou, Y., Guo, B. H. W., Papadonikolaki, E., Dimyadi, J. & Hou, L. 2023. Lessons Learned on Adopting Automated Compliance Checking in the AEC Industry: A Global Study. *Journal of Management in Engineering,* 39**,** 04023019.
Zou, Y., Wu, Y., Guo, B. H. W., Papadonikolaki, E., Dimyadi, J. & Hung, S. N. 2022. Investigating the New Zealand off-site manufacturing industry's readiness for automated compliance checking. *Journal of Construction Engineering and Management,* 148**,** 05022013.
Zulu, S. L., Saad, A., Ajayi, S., Unuigbe, M. & Dulaimi, M. 2023. A thematic analysis of the organisational influences on digitalisation in construction firms. *Journal of Engineering, Design and Technology*.
Zuo, J., Read, B., Pullen, S. & Shi, Q. 2013. Carbon-neutral commercial building development. *Journal of Management in Engineering,* 29**,** 95-102.

A METHODOLOGICAL APPROACH TO ASSET INFORMATION MANAGEMENT VIA KNOWLEDGE GRAPHS AND LARGE LANGUAGE MODELS

Marta Boscariol[1,*], Silvia Meschini[2] and Lavinia Chiara Tagliabue[2]
[1] Department of Management, Università degli Studi di Torino, 10134 Turin, IT
[2] Department of Computer Science, Università degli Studi di Torino, 10149 Turin, IT
[*] marta.boscariol@unito.it

Abstract

Tackling the need of large organizations for a proactive Asset Information Management (AIM) System, a methodological approach to knowledge management applied to built assets portfolios is proposed. It aims at synergically leveraging Knowledge Graphs (KGs) and Artificial Intelligence (AI) technologies to enable analytics on input data. In the theorized pipeline Large Language Models (LLMs) are meant to be used both in the graph creation phase, extracting data from unstructured sources and organizing them according to domain ontologies, as tested on a use-case sample, and in the knowledge extraction phase via queries.

Introduction

Asset Management (AM) is defined by international standardization as "coordinated activity of an organization to realize value from assets", where an asset is identified as an "item, thing or entity that has potential or actual value to an organization" (ISO 55000, 2014). It can be applied to the construction sector in terms of balancing of financial aspects, extraordinary interventions, operations and maintenance, in order to provide the most cost-effective, sustainable, well-engineered and efficient solution. The operational phase of buildings and infrastructures (ISO 19650-1, 2019), represents that time span when a continuous stream of minor interventions take place and the asset itself both delivers value to the organization and represents a source of costs. Asset owners with large portfolios typically manage numerous facilities, having the majority of them in the operational phase (UK BIM Framework, 2021).

An AMS, "whose function is to establish the asset management policy and asset management objectives" (ISO 55000, 2014), bases its performance on data exploitation, since assets can be considered data-rich environments, or recollections of a multitude of data and documents depicting information about each building's lifecycle (LC) step. From a data management perspective, future availability and accessibility of information is required, for at least as long as the life of the asset, jointly with pre-agreed Master Data Management (MDM) (ISO 8000-100, 2016) procedures internal to the organization, to support the effective management of data quality and deliver benefits across various LC phases.

Provided the output of this investigation is suggesting an Asset Information Management (AIM) approach to support the managerial processes, information ought to be regarded as a precious asset itself, whose structure and consistency provide measurable performance indicators and insights for decision-making purposes (KPMG, 2021). The main challenges associated with managing built asset data are related to their heterogeneous and siloed nature, being their structure and storage format quite often not regulated and then handled by multiple users. These dynamics determine a loss of value related to the information managed per each step of the process (Eastman, 2011), proving to be obstacles to the extraction of valuable and complete information in a timely manner. Managers and other stakeholders involved should be enabled to store, access and query information regarding their portfolio, ranging from documents to spatial information, 3D models or activity-related materials, no longer dealing with inaccurate, incomplete or even undiscoverable pieces of documentation. Inappropriate creation and maintenance of data sources, as well as interoperability issues foster higher management costs and a lack of trust in the final user, who's trying to gain a deeper insight. Recently, with the advent of Big Data (Ajah and Nweke, 2019), innovative technology solutions enable organizations to improve operational efficiency (Gomolluch, 2023), strengthen stakeholder relations, reduce costs and provide demonstrable legal compliance. Big Data Analytics, as a topic, has found some application also in the built environment to foster decision making (Farghaly et al., 2017, Demirdöğen et al., 2023). Limiting their implementation with BIM tools might not be the most efficient way to manage such intense streams of data, due to inadequate storage capabilities of those technologies. The goal of the current research is to suggest a holistic (Gomolluch, 2023) practical approach to knowledge management, applied to extended built assets portfolio, such as buildings and private or public infrastructures, through the adoption of Knowledge Graphs (KGs), as well as Large Language Models (LLMs) automating otherwise lengthy manual workloads. This proposition stems from a real use-case scenario, which is the extended built assets portfolio managed by the University of Turin. In this public institution an AMS (Di Giuda et al., 2023) is currently under development, dealing with the above mentioned limitations related to siloed documental resources and data quality. The proposed methodology is willing to enable the extraction of knowledge from heterogeneous data sources and their proactive interrogability, therefore improving the understanding of how real estate assets are used, along with decision-making at every stage of the asset's LC.

Literature background

Scientific literature in the field of AM in the built environment lacks a unitary vision and covers the topic of Information Management through sectorial applications. AM is often investigated employing BIM (Farghaly et al., 2017, Moretti et al., 2022) or through Digital Building Logbook as documental repositories of a building's lifecycle data, with a focus on sustainability and energy efficiency (Signorini et al., 2023, Malinovec Puček et al., 2023). When it comes to applying KGs and Linked Data to the built environment (Pauwels et al., 2022), it is often related to the definition of ontologies fit for specific use cases and their application to BIM models: sensors and mechanical systems (Quinn et al., 2020, Eleftheriou et al., 2022); building components (Sobhkhiz and El-Diraby, 2023); operation and maintenance (Xie et al., 2022); thermal comfort and energy analysis (Esnaola-Gonzalez et al., 2022); construction phase (Farghaly et al., 2021), even considering the integration of open source BIM models (IFC based) (Elshani et al., 2022). Others investigated link validation in ICDDs (Information Containers for Linked Document Delivery), meant for linking information in documents with Resource Description Framework (RDF) data (Hagedorn, Liu, et al., 2023, Hagedorn, Pauwels, et al., 2023), or exploited Shapes Constraint Language (SHACL) applied to distributed Linked Building Data for building model checking and constraint validation (Nisbet et al., 2023). Differently, KGs as a representation of Digital Twins has been studied in other sectors, mainly in industry 4.0 (Yahya et al., 2021, Tamašauskaitė and Groth, 2023), while text-based knowledge discovery techniques (i.e., Text Mining, Text processing, and Natural Language Processing) are hardly applied to the building sector. Natural Language Processing (NLP), which is branch of Artificial Intelligence (AI) has found application in text analysis and compliance checking related to contracts and legislation (Locatelli et al., 2023). In the field of AI and LLMs, algorithms are being developed to deal with the extraction of content from multiple kind of sources, both structured databases (Sequeda et al., 2023) and text-like documents, proving promising in assisted and semi-automated KG construction pipelines (Zhu et al., 2023).

Research context and aim

In compliance with studies stating that organizations can benefit from implementing AMS conforming to the ISO 55000 normative series (Almeida et al., 2022, Maletič et al., 2023), the proposition of a practical AIM approach should consider some recurrent characteristics in big organizations managerial praxis:

- Widespread organizations usually store large volumes of data and documentation, heterogeneous in types and formats, in centralized repositories. So called data lakes ingest all types of data from any source, ranging from structured (database tables, .xlsx sheets) to semi-structured (XML, webpages) to unstructured (images, audio files, PDFs), in their original raw form. Due to data lakes scalable architecture, a variety of applications can be compatible (Sawadogo and Darmont, 2021), from big data analytics to machine learning (ML) or predictive analytics, therefore greater value can be extracted from stored data.
- Inside an organization, analogous or related tasks are not always systematically carried out using harmonized terminology or producing consistent documentation, resulting in misalignments in both structure and content of files. Semantic enrichment of raw data would ensure clear interpretation and understanding, enabling the adoption of a graph structure (Pan et al., 2024) to interconnect individual records in a network and overcome inconsistencies.
- Structuring a uniform framework of machine-readable data generates benefits (Gomolluch, 2023) both in the short term, providing a Knowledge Base (KB) agile to interrogate, and in the long term, posing the basis for further ML implementations, which might include prediction algorithms and automatizations, optimizing managerial processes.

The theorized approach aims at multiplying the advantages provided by AI-powered technologies in Information Management, especially in decision-making processes at a managerial level. In a dual sided technological pipeline, from one side there's an AI system automatically extracting information from document sources and structuring a KB as a graph, while on the other side there's a human user interrogating another AI agent, capable of performing neural-symbolic reasoning (Zhang et al., 2021). This system, enriched by structured knowledge and logic-based inference from ontologies, based on the relevant facts retrieved from the KB, should enable informed decision making.

Methodology

As previously stated, the core idea of this research project deals with the development of a methodology to gather all sources of knowledge regarding multiple built assets in a unique interrogable environment, structured as a KG. The methodology (*Figure 1*) is composed of two parts: an *automated text-to-KG pipeline* and a *chatbot for human-machine interaction*. The former is responsible for the extraction of every bit of information existing in the organization's documents corpora, for their labeling according to pre-established sets of terms and for their interrelation to generate a KB with a graph structure. The latter assists users in finding the information they require from the KG. It should act as user-friendly interface and extract information from the KG through an NLP agent, which transforms natural language queries into query language (e.g. SPARQL), drawing inference and returning search results in human readable form. Its purpose is to perform those analytics typically carried out manually and visualizing them through graphical forms and set Key Performance Indicators (KPIs) into the AMS digital boardroom, although its proposition will be subject to future developments.

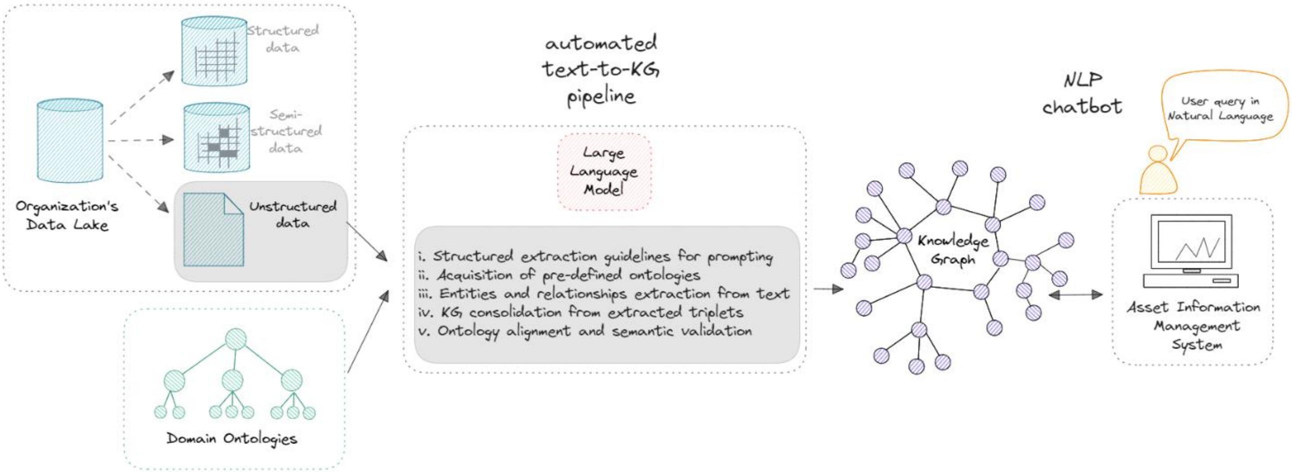


Figure 1. Theorization of an Asset Information Management approach

Prior to the tools development stage, some considerations ought to be made, regarding the intended scope of an AIM digital tool, according to the organization's administrative activities related to its use. For this purpose, a team of experts should identify one or more Domain Ontologies relevant to AIM scopes, creating those that match the organization's requirements, formulated in terms of competency questions (Keet, 2020). In addition to specifically created ontologies, existing and validated ones could be reused.

The definition of the methods to be applied in order to extract a KG from the enterprise Data Lake, based upon heterogeneous data sources, both structured and non-structured, strongly depends on the use case scenario and it has been chosen to deal with text documents. The suggested methodology aims at integrating what seems to be two distinct technologies, LLM and KGs, which though aim both at capturing a detailed representation of a portion of the world. Their different knowledge representation formats complement each other, overcoming individual limitations. A so called *text-to-KG* pipeline will be implemented, exploiting literature-tested solutions, and it involves:

- a Large Language Model (LLM), which acquires world knowledge by pre-training on massive text corpora and is capable of understanding and transforming human language, although with severe likelihood of hallucinations or logical leaps, showing limitations in domain-specific performance, trustworthiness, controllable generation or content quality assessment.
- KGs, which directly store factual knowledge as interconnected networks via relationships of real-world entities and provide structured, symbolic knowledge and deterministic logic, but struggle at handling nuanced language. They ground LLMs in factual knowledge, ensuring truthful text generation, while, on the other hand, LLMs capabilities in text processing and fact extraction mitigate their rigidity.
- Ontologies, which represent structured formal knowledge, including commonsense knowledge lacking in LLMs, and are used to define concepts, properties, relationships, constraints, and axioms in a machine-readable format, using standards like RDF and Web Ontology Language (OWL). They play the role of "organizing principle", as they provide agreed sets of hierarchies and terminologies, harmonizing the content handled and permitting entity disambiguation, as well as describing with the required level of detail the domain of interest. Ontologies act as a bridge between the unstructured textual knowledge of LLMs and the structured world knowledge represented in KGs, improving symbol grounding and enabling formal logic reasoning (Keet, 2020), such as inference through querying protocols.

The exploitability of the LLM-driven *text-to-KG* pipeline is strictly dependent on the domain ontology, as, in the first place, they will help the AI algorithm extract the correct entities and relationships. Although their outlining is an activity that requires time and the involvement of many stakeholders, it can be seen as a more formal task, upon which the technological pipeline can be built, as in the two steps described onwards.

Domain Ontologies development: the organizing principle

AM information ought to be systematized with the help of domain-specific ontologies, through classes, their properties and relations. Since "an ontology is a formal, explicit specification of a shared conceptualization" (Studer et al., 1998), the task of ontology engineering has to be demanded to a group of domain experts, jointly with professionals competent in RDF and OWL ontologies matter. Considering that concepts to be harmonized are already in use among the organization's practitioners, a bottom-up approach is suggested (Keet, 2020). Core concepts and relationships, as they tend to appear in datasets, are to be stated and formalized in the main ontology first, along with formal definitions via ontological axioms. Local ontologies and their alignment will be implemented only after the whole framework and its applicability have been validated, resulting in an incremental approach.

It is to be remarked that the core aim of the methodology is to create KGs conforming to a specific ontology, in order to guide accordingly to the schema all queries posed to the resulting graph, as well as to process a large number of documents though multiple prompts, due to computational size limitations.

LLMs for Knowledge Graph generation from text: the technological pipeline

Given recent advancements of LLMs and foundation models, they could be leveraged to alleviate time-consuming tasks, dealing with a great amount of unstructured files stored in the organization's repositories. To better understand how the *text-to-KG* pipeline could prove successful, a disaggregation of its procedure (*Figure 1*) is therefore outlined:

i. Creation of Structured Extraction Guidelines, whose scope it's the definition of the types of entities, properties and relationships the LLM needs to recognize and extract, referring to the formal content of the use-case ontology. Guidelines will be supported with input samples from the target domain and examples of the expected output structure in ontology language, to teach the LLM how to correctly encode human logic and to avoid inconsistencies in information extraction.
ii. Use of a LLM to transform unstructured text to KG in compliance with a pre-defined ontology. This task could be accomplished using either a LLM pre-trained with standard ontologies, a LLM fine-tuned with a custom ontology, a LLM prompted with a custom ontology, or a hybrid of a fine-tuned (with a custom ontology) and pre-trained (with a standard ontology) LLM. A comparative analysis investigating the most fit-for-purpose model should be conducted, evaluating benefits and limitations of each of the four plausible approaches outlined, taking into account token costs, response times, training and tuning datasets required. A hybrid approach is considered more suitable for high-quality KG construction (Wang et al., 2023), relatively to its accurate extraction of domain ontology components.
iii. Extraction of entities and relationships from large volumes of text sources via LLM, automating the process of KG extraction through text chunking, extraction of triplets, their normalization and enrichment. Subsequently text mining techniques and NLP algorithms (Van Assche et al., 2023, Zhao et al., 2023) operate on entities and their relationships.
iv. KG consolidation, merging nodes and edges extracted from all sources into a unified set and enhancing them with metadata and provenance information, to allow topological analysis on communities and centralities. The resulting KG is ready to be uploaded to a storage database, becoming a unitary source of data, over which complex analytics and queries could be carried out. Further integrations of departmental KGs, deriving from other domain ontologies, are allowed through an incremental approach.
v. Ontology alignment or ontology matching (OM) takes ontologies as input and determines as output a set of correspondences between semantically related entities, dealing with semantic heterogeneity problems (Li et al., 2019) and enabling navigation over KGs.

A few-shot prompting framework has been tested, to guide LLMs in performing NLP tasks, finalized at KG creation, exploiting GPT-3.5 (gpt-3.5-turbo), despite it being a closed-source model. The authors tried to define an ad-hoc chaining prompts to guide the LLM through a series of smaller, more manageable tasks to ultimately achieve a complex goal, which is the conversion of an unstructured paragraph into a structured graph, where each node corresponds to an entity, and the interconnecting edges depict the relationships linking them, matching those defined in a sample ontology. An exploratory trial-and-error experiment has been conducted, feeding a use-case textual source, in which university spaces are listed and assigned to organizational units and their occupants, and a reduced version of the RealEstateCore (REC) Ontology. The prompts engineered go through the following steps:

1. Task definition: the LLM is given a role as expert and is told to perform information extraction from textual documents, converting it into triplets made as: Source_Node, Relationship, Target_Node. Additional instructions are here provided, concerning how to maintain consistency in entities and handle numerical data, property formats, naming conventions.
2. Ontology ingestion: the LLM is fed with an ontology in RDF compatible .ttl format, being instructed on which entities and relationships to look for in the following tasks. Due to tokens limitations, a small ontology, composed only of use-case related entities and relationships, has been exploited.
3. Unstructured paragraph input: the LLM is given as input the text containing data aimed at being converted to KG. It is preferably accompanied by an example of how information should be correctly extracted, stating a small paragraph and the desirably outputted list of triples that could be identified in that paragraph.
4. RDF serialization: the LLM is asked to serialize the list of extracted entities and relationships according to RDF format, in this case in .ttl, enabling compatibility with graph databases.

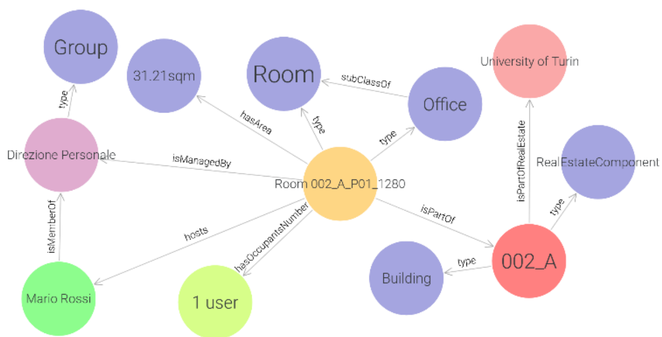


Figure 2. Visualization of nodes connected to a room entity in Ontotext GraphDB.

The resulting file was imported into an open source graph store, namely Ontotext GraphDB, where triplets consistency with the original NL source has been verified. As an example, nodes and relationships connected the one of the room entities extracted from text has been visually displayed as a graph (*Figure 2*).

Discussion

Theoretical contributions

Literature demonstrates that a single technology platform can be exploited to create accurate and transparent records of organizations' reference data (including buildings, assets, people, and contracts) and it can be integrated in internal business processes (e.g. contract and project management, strategic planning, environmental reporting, preventative and reactive maintenance) (Hanley and Brake, 2016, Geisler et al., 2021). However, a strategic approach to implementation, which must align to business process change and strong data governance, is required.

From an Information Management perspective, the process as a whole should be applied proportionally to the scale and complexity of the organization and of its asset portfolio, reflecting concepts and principles from ISO 19650-1 (UK BIM Framework, 2021). Accordingly, agreed methods and procedures should be internally standardized, to ensure the production of compliant set of data and information deliverables. Agreed schemata would ensure compatibility when federating data produced by different internal departments, which can then be used and maintained during the asset's operational life for analysis, reporting and decision-making.

Practical implications and limitations

The proposed methodological approach to knowledge management, applied to built assets' portfolios, aims to leverage a KG structure and AI technologies to enable analytics on input data for a proactive Asset Information Management System in large organizations. From an enterprise management perspective, a KG approach could solve data analytics challenges related to data quality or complex properties, exploiting intrinsic characteristics of Data Lakes storing the sources of knowledge related to accessibility, enabling scalability through a domain oriented structure and integration across all domains of the organization's network.

The main advantages deriving from the exploitation of semantic web technologies, in particular of KGs and ontologies, can be recognized in the automatic interrelation of information content, without having a human operator search for a specific document in a multitude of folders, and in its being univocally classified, reducing data creators' margins of personalization. With an incremental approach, from a single KB might stem several use-case applications, as multiple domain-specific ontologies can coexist in the same KG, putting in place an iterative construction of the KG introducing "just enough" semantics as each use case demands.

Integrating ontologies enable formal logical inference, complementing the inductive reasoning of LLMs with description logics and enhancing their generalization capabilities, as symbol grounding explicitly links language to formal conceptual representations. Moreover, a coexistence of standard and custom ontologies is feasible, as well as both independent and cross equivalence queries across domains. Additional benefits could be given by the adoption of FAIR (Findable, Accessible, Interoperable, Reusable) Data Principles, as they provide a data management framework, enabling findability of data and metadata through persistent digital object identifier (DOI); accessibility to trusted repositories; interoperability of knowledge representation language and dictionaries; reusability of data, thanks to accurate provenance information and clear usage license (Martorana et al., 2022).

Other potential advantages of using an automated methodology for KG construction from unstructured data at scale resides on saving a great amount of hours otherwise required for manual reading, labeling and linking of documents, as for inconsistencies and redundancies checking in their content. LLM integration further differentiates this methodology from traditional KG creation, accomplishing the main aim of the final KG, namely inference and query answering according to explicitly defined schema and properties, along with ensured high volume scalability, easier model retraining, debugging and dataset or prompt fixings.

According to the authors observations during the testing phase, the accuracy of the generated knowledge graph clearly relies on the underlying model leveraged, as the extracted information might be affected by biases and hallucinations, given by the LLM's ability to correctly interpret its tasks. Prompt development involved an interactive process of writing, testing, and refinement, based on feedback and experimentation. Tailoring the prompt to use-case specific application required giving the LLMs sufficient context to generate the desired output. This includes writing clear step-by-step instructions, as detailed as possible about the outcome format. Instead than zero-shot prompting (when the LLM relies solely on its pre-existing knowledge and on the instructions in the prompt to perform the task), adopting few-shot learning enabled the LLM to recognize patterns from little examples provided and to generalize to unseen task variations, also avoiding time-consuming data labeling. It emerged how the model tends to assign all relationships extracted from text to the ontology proposed, even those not standardized by it, and it struggles to infer transitive and inverse relationships from the ontology, if not explicitly declared in the example as well. Other limitations encountered testing this prompting framework rely on computational limits, which, in the case of GPT-3.5, are set to 4'096 tokens per prompt. Therefore textual sources to be analyzed must be pre-processed and split into smaller size compliant chunks, taking care of adequately preventing the loss of context and the separation of related entities (Carta et al., 2023). In order to ensure scalability for larger datasets, an algorithm devoted to chunking text and automatically and iteratively prompting the LLM should be implemented.

The proposed methodology presents intrinsic limitations, related to the time and expertise required to develop a proper domain ontology, as well as to the implementation and testing phase of the technological pipeline. It will be highly dependent on the application use-case, mainly caused by the uncertainty generated by the sample size and to the unreliability of its data sources. Due to the lack of prior research studies on this precise topic, it will be difficult to compare results and asses their validity, requiring the establishment of validation methods.

Future directions

The potential offered by this approach in the field of Knowledge Representation and Reasoning (KRR) is manifest, since KGs have emerged as key technologies for representing knowledge in a variety of domains and supporting intelligent applications, such as chatbots, question answering systems, and recommendation systems (Hu et al., 2023). An NLP agent devoted to answering a natural language question, by reasoning over facts in a KB, could potentially deal with text-to-SPARQL tasks and multi-hop reasoning, appropriately querying the organization's KG on user’s behalf. These technologies should provide also faster access to continuously-updating data like sensor readings, getting closer to real-time analytics.

Other future developments might consider multiple domains integration, differential privacy policies, implementing user-based restrictions on visualization, as well as the integration of ML algorithms for predictions. Reinforcement learning from human feedback and automatic updates to the KG as new document are ingested, will make the methodology more versatile.

As previously stated, the research project will leverage a practical use-case methodological approach, not only suggesting a plausible use of digital technologies to tackle AM processes inside the University of Turin, but also aiming at providing a scalable, generalizable and reusable set of tools, which might solve similar knowledge-related problems in other domains. Relatively to the applicative use-case, textual documents containing information regarding ownership and occupancy of the building are employed. The implementation of a *text-to-KG* will extract and interconnect factual knowledge, applying a customized ontology as “organizing principle”. The resulting KG’s consistency will be evaluated through recall and precision metrics based on confusion matrix entries, obtained from manual evaluation campaigns.

Aligning with past projects at the University of Turin, where Business intelligence (BI) tools have been implemented as means of data-driven decision making (Di Giuda et al., 2023), an integration of AI and BI would be desirable, to drive faster, automated and predictive insights. The use of AI powered analysis tool could overcome traditional BI tools limitations regarding analysis at a more granular level, benefitting the end user in terms of automating data cleaning to ensure accuracy in the procedure, integrating multidimensional analysis as well, to detect clusters or similarities across hierarchies, attributes and metadata stored in the graph database.

Conclusions

The paper tackled the proposition of a methodological approach to AIM, leveraging technological pipelines whose task is the automation of truthful knowledge extraction from a large organization’s documental corpora, to enable informed decisional processes. The pivotal element is the human user, as it determines specific information needs guiding the ontology creation, as well as practical decision-making goals and problem-solving context defined by the organization it represents. KGs provide structured knowledge by encoding concepts and relationships in machine-readable ontologies, therefore they can deliver accurate, safe and responsible AI applications, as they drive reasoning and semantics. However, the creation of an enterprise KGs manually might involve gigantic and cost-inefficient workload, hence the automation of ontology extraction from unstructured corpora through LLMs could be a viable solution. As it has been outlined, the incorporation of ontologies in the pipeline can significantly strengthen LLMs reasoning capabilities and structure accurate KGs, overcoming present challenges in reasoning and symbol grounding, granting a significant increase in the value of an organization’s knowledge graph. The authors tested a semi-automated approach to KG construction for the AIM domain, leveraging few-shot LLM prompting with GPT-3.5 model. The chaining prompting technique dealt with triplets extraction from unstructured sources, semantically aligning them with a pre-processed ontology acting as T-Box. The resulting KG showed GPT-3.5 proficiency in NLP tasks applied to closed-domain settings, taking advantage from clearly stated tasks explanation, although hallucinating ontology classes. Since being tested only on a small dataset, the authors intend to conduct more in-depth experimentation, to properly evaluate reliability and fitness to the intended purpose of the approach, qualifying replicability and generalizability, also evaluating alternative LLMs. Future developments will likely integrate other AI systems on the inference side of the pipeline, capable of dealing with continuously changing data and forecasting plausible event scenarios.

Acknowledgements

It is to be remarked that sensitive data contained in sources provided by the University of Turin have been anonymized prior to AI tools exploitation. The authors also want to thank Dr. Mirko Locatelli for sharing his knowledge on NLP and for engaging in enriching conversations on the topic.

References

Ajah, I. A. & Nweke, H. F. (2019) Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications. Big Data and Cognitive Computing. Multidisciplinary Digital Publishing Institute, 3(2), p. 32.

Almeida, N., Trindade, M., Komljenovic, D. & Finger, M. (2022) A conceptual construct on value for infrastructure asset management. Utilities Policy, 75, April, p. 101354.

Buchholz, M. & Lützkendorf, T. (2023) European building passports: developments, challenges and future roles. Ubiquity Press, 4(1), pp. 902–919.

Carta, S., Giuliani, A., Piano, L., Podda, A. S., Pompianu, L. & Tiddia, S. G. (2023) Iterative Zero-Shot LLM Prompting for Knowledge Graph Construction. arXiv.

Demirdöğen, G., Işık, Z. & Arayici, Y. (2023) BIM-based big data analytic system for healthcare facility management. Journal of Building Engineering, 64, April, p. 105713.

Di Giuda, G. M., Meschini, S., Accardo, D., Gasbarri, P., Tagliabue, L. C. & Scomparin, L. (2023) BIM-GIS and BI integration for facility and occupancy management of university assets: the UniTo pilot case. In. Florence, Italy.

Eastman, C. M. (ed.) (2011) BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors. 2nd ed, Hoboken, NJ: Wiley.

Eleftheriou, O., Dimara, A., Kotis, K. & Anagnostopoulos, C.-N. (2022) Saving Energy with Comfort: A Semantic Digital Twin Approach for Smart Buildings. In Proceedings of the Third International Workshop on Semantic Digital Twins. Hersonissos, Greece: CEUR (CEUR Workshop Proceedings).

Elshani, D., Wortmann, T. & Staab, S. (2022) Towards Better Co-Design with Disciplinary Ontologies: Review and Evaluation of Data Interoperability in the AEC Industry. In Proceedings of the 10th Linked Data in Architecture and Construction Workshop. Hersonissos, Greece: CEUR (CEUR Workshop Proceedings), pp. 43–52.

Esnaola-Gonzalez, I., Bermúdez, J. & Aceta, C. (2022) An Ontology-Driven Approach to Support Data Analysts with Thermal Comfort Problems in the Built Environment. In Proceedings of the 10th Linked Data in Architecture and Construction Workshop. Hersonissos, Greece: CEUR (CEUR Workshop Proceedings), pp. 8–19.

Farghaly, K., Abanda, H., Vidalakis, C. & Wood, G. (2017) BIM Big Data System Architecture for Asset Management: A Conceptual Framework. In Lean and Computing in Construction Congress - Volume 1: Proceedings of the Joint Conference on Computing in Construction. Heraklion, Crete, Greece: Heriot-Watt University, pp. 289–296.

Farghaly, K., Soman, R. K. & Whyte, J. (2021) Visualizing real-time information through a construction production control room. In. European Council on Computing in Construction (Computing in Construction), pp. 415–422.

Geisler, S., Vidal, M.-E., Cappiello, C., Lóscio, B. F., Gal, A., Jarke, M., Lenzerini, M., Missier, P., Otto, B., Paja, E., Pernici, B. & Rehof, J. (2021) Knowledge-Driven Data Ecosystems Toward Data Transparency. Journal of Data and Information Quality, 14(1), p. 3:1-3:12.

Gomolluch, A. B., M. Block, D. Gogolin, S. (2023) Potential of holistic asset information management. In Life-Cycle of Structures and Infrastructure Systems. CRC Press.

Hagedorn, P., Liu, L., König, M., Hajdin, R., Blumenfeld, T., Stöckner, M., Billmaier, M., Grossauer, K. & Gavin, K. (2023) BIM-Enabled Infrastructure Asset Management Using Information Containers and Semantic Web. Journal of Computing in Civil Engineering. American Society of Civil Engineers, 37(1), p. 04022041.

Hagedorn, P., Pauwels, P. & König, M. (2023) Semantic rule checking of cross-domain building data in information containers for linked document delivery using the shapes constraint language. Automation in Construction, 156, December, p. 105106.

Hanley, B. P. & Brake, D. J. (2016) Putting asset data at the heart of organisational decision-making using an Integrated Workplace Management System. In Asset Management Conference (AM 2016), pp. 1–7.

Hu, N., Wu, Y., Qi, G., Min, D., Chen, J., Pan, J. Z. & Ali, Z. (2023) An empirical study of pre-trained language models in simple knowledge graph question answering. World Wide Web, 26(5), pp. 2855–2886.

ISO 8000-100 (2016). Switzerland: ISO. ISO/TC 184, Automation systems and integration, Subcommittee SC 4, Industrial data.

ISO 19650-1 (2019). London: British Standards Institution.

ISO 55000 (2014). Switzerland: ISO. ISO/PC 251 Asset management.

Keet, M. (2020) An Introduction to Ontology Engineering.

KPMG (2021) The value of Information Management in the construction and infrastructure sector. A report commissioned by the University of Cambridge's Centre for Digital Built Britain (CDBB), p. 104.

Li, H., Dragisic, Z., Faria, D., Ivanova, V., Jiménez-Ruiz, E., Lambrix, P. & Pesquita, C. (2019) User validation in ontology alignment: functional assessment and impact. The Knowledge Engineering Review. Cambridge University Press, 34, January, p. e15.

Locatelli, M., Pattini, G., Pellegrini, L., Meschini, S. & Accardo, D. (2023) Fostering the consensus: a BERT-based Multi-label Text Classifier to support agreement in public design call for tenders. TEMA: Technologies Engineering Materials Architecture, 9(1), pp. 62–73.

Maletič, D., Grabowska, M. & Maletič, M. (2023) Drivers and Barriers of Digital Transformation in Asset Management. Management and Production Engineering Review, 14(1), pp. 118–126.

Malinovec Puček, M., Khoja, A., Bazzan, E. & Gyuris, P. (2023) A Data Structure for Digital Building Logbooks: Achieving Energy Efficiency, Sustainability, and Smartness in Buildings across the EU. Buildings. Multidisciplinary Digital Publishing Institute, 13(4), p. 1082.

Martorana, M., Kuhn, T., Siebes, R. & Ossenbruggen, J. van (2022) Aligning restricted access data with FAIR: a systematic review. PeerJ Computer Science. PeerJ Inc., 8, July, p. e1038.

Moretti, N., Xie, X., Garcia, J. M., Chang, J. & Parlikad, A. K. (2022) Built environment data modelling: a review of current approaches and standards supporting Asset Management. IFAC-PapersOnLine. (5th IFAC Workshop on Advanced Maintenance Engineering, Services and Technologies AMEST 2022), 55(19), pp. 229–234.

Nisbet, N., Zhang, Z. & Ma, L. (2023) Automated generation of SPARQL queries from semantic mark-up. In. European Council on Computing in Construction (Computing in Construction), pp. 0–0.

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J. & Wu, X. (2024) Unifying Large Language Models and Knowledge Graphs: A Roadmap. IEEE Transactions on Knowledge and Data Engineering pp. 1–20.

Pauwels, P., Costin, A. & Rasmussen, M. H. (2022) Knowledge Graphs and Linked Data for the Built Environment. In Industry 4.0 for the Built Environment: Methodologies, Technologies and Skills. Cham: Springer International Publishing (Structural Integrity), pp. 157–183.

Quinn, C., Shabestari, A. Z., Misic, T., Gilani, S., Litoiu, M. & McArthur, J. J. (2020) Building automation system - BIM integration using a linked data structure. Automation in Construction, 118, October, p. 103257.

Sawadogo, P. & Darmont, J. (2021) On data lake architectures and metadata management. Journal of Intelligent Information Systems, 56(1), pp. 97–120.

Sequeda, J., Allemang, D. & Jacob, B. (2023) A Benchmark to Understand the Role of Knowledge Graphs on Large Language Model's Accuracy for Question Answering on Enterprise SQL Databases. arXiv.

Signorini, M., Moretti, N., Merino, J., Daniotti, B. & Parlikad, A. (2023) Digital-twin based data modelling for digital building Logbook implementation. IET Digital Library, January, pp. 130–137.

Sobhkhiz, S. & El-Diraby, T. (2023) Dynamic integration of unstructured data with BIM using a no-model approach based on machine learning and concept networks. Automation in Construction, 150, June, p. 104859.

Studer, R., Benjamins, V. R. & Fensel, D. (1998) Knowledge engineering: Principles and methods. Data & Knowledge Engineering, 25(1), pp. 161–197.

Tamašauskaitė, G. & Groth, P. (2023) Defining a Knowledge Graph Development Process Through a Systematic Review. ACM Transactions on Software Engineering and Methodology, 32(1), p. 27:1-27:40.

UK BIM Framework (2021) Information management according to BS EN ISO 19650 - Guidance Part 3: Operational phase of the asset life-cycle.

Van Assche, D., Delva, T., Haesendonck, G., Heyvaert, P., De Meester, B. & Dimou, A. (2023) Declarative RDF graph generation from heterogeneous (semi-)structured data: A systematic literature review. Journal of Web Semantics, 75, January, p. 100753.

Wang, S., Sun, X., Li, X., Ouyang, R., Wu, F., Zhang, T., Li, J. & Wang, G. (2023) GPT-NER: Named Entity Recognition via Large Language Models. arXiv.

Xie, X., Moretti, N., Merino, J., Chang, J. Y., Pauwels, P. & Parlikad, A. K. (2022) Enabling building digital twin: Ontology-based information management framework for multi-source data integration. IOP Conference Series: Earth and Environmental Science. IOP Publishing, 1101(9), p. 092010.

Yahya, M., Breslin, J. G. & Ali, M. I. (2021) Semantic Web and Knowledge Graphs for Industry 4.0. Applied Sciences. Multidisciplinary Digital Publishing Institute, 11(11), p. 5110.

Zhang, J., Chen, B., Zhang, L., Ke, X. & Ding, H. (2021) Neural, symbolic and neural-symbolic reasoning on knowledge graphs. AI Open, 2, January, pp. 14–35.

Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., Liu, P., Nie, J.-Y. & Wen, J.-R. (2023) A Survey of Large Language Models. arXiv.

Zhu, Y., Wang, X., Chen, J., Qiao, S., Ou, Y., Yao, Y., Deng, S., Chen, H. & Zhang, N. (2023) LLMs for Knowledge Graph Construction and Reasoning: Recent Capabilities and Future Opportunities, May.

ENHANCING BUILDING MAINTENANCE EFFICIENCY THROUGH BIM-LIDAR INTEGRATION

Hamidreza Alavi[1], Weiwei Chen[1,2], Mudan Wang[1], Soheila Kookalani[1] and Ioannis Brilakis[1]
[1]University of Cambridge, United Kingdom
[2]University College London, United Kingdom

Abstract

Facility Maintenance Management constitutes a significant portion of building costs, emphasizing the need for innovative approaches to enhance efficiency. This paper introduces a novel methodology for transforming point cloud data into Industry Foundation Classes (IFC) models. This method involves a meticulous point cloud classification and a subsequent conversion into a visually appealing mesh representation. This mesh is then linked to IFC entities, enabling the creation of comprehensive and semantically rich BIM models. A case study on the Civil Engineering Building at the University of Cambridge demonstrates the methodology's effectiveness, showcasing improved visualization and supervision of maintenance tasks. The contribution lies in the refined point cloud classification, mesh conversion, and seamless integration with IFC, providing a novel and efficient solution to existing challenges. The results underscore the potential for revolutionizing maintenance practices in existing buildings.

Introduction

The aging of buildings accentuates the imperative for effective maintenance, constituting over 65% of total building costs (Matos *et al.*, 2020; Alavi, Bortolini and Forcada, 2022). Despite this, facility managers frequently depend on conventional methods which can be inefficient and time-consuming (Juricic, Krstić and Čulo, 2021; Alavi and Forcada, 2022).

The integration of BIM with various technologies allows for the creation of 3D information models for facility maintenance and administration work, ultimately enhancing productivity. However, one major challenge in utilizing BIM for maintenance activities is the lack of a BIM model for existing buildings (Alavi and Forcada, 2019). This paper addresses this gap by introducing a novel "Scan-to-BIM" methodology. The implementation of LiDAR technology confidently converts point cloud data into a mesh format, providing a solid foundation for creating extremely accurate as-built BIM models (Antah *et al.*, 2021). The proposed methodology aims to enhance maintenance activities in existing buildings where there is no BIM model available. The paper aims to outline the utilization of LiDAR technology to obtain point cloud data of existing buildings and convert it into a mesh format, which can serve as a foundation for assigning relevant Industry Foundation Classes (IFC) information to different building components.

This method enables the development of an as-built BIM model that precisely depicts the current state of the building. Once created, this model can be integrated with a Computerized Maintenance Management System (CMMS) to simplify maintenance tasks. Each maintenance request from the CMMS can be associated with the applicable components in the as-built BIM model, making it easy for maintenance personnel to find and evaluate maintenance needs.

Literature Review

To enable effective maintenance activities during the O&M stage, a diverse range of data is required in BIM (H. Alavi *et al.*, 2021; Fan *et al.*, 2023). Furthermore, Nicolle and Cruz (Nicolle and Cruz, 2011), emphasize the need for precise data on installed components in buildings to achieve effective BIM implementation in the O&M phase. These kinds of data might not be available in many existing buildings due to imperfect and deficient, obsolete, or disintegrated building information (Gursel *et al.*, 2009). The significant difference between enabling BIM for FM in new buildings and existing buildings is the lack of as-built, CAD files as well as insufficient and outdated information of the buildings (Becerik-Gerber *et al.*, 2012; Hamidreza Alavi *et al.*, 2021). Missing this kind of information might lead to ineffective building management, uncertain process results and time loss or cost increases in facility management processes (Alavi and Forcada Matheu, 2022).

In recent years, the integration of BIM and LiDAR has emerged as a promising approach to create as-built BIM models for existing buildings and maximize building maintenance efficiency (Shin and Cha, 2023). This integration offers a robust framework for capturing detailed building information and leveraging it for effective maintenance processes. One of the key advantages of this integration is the ability to create 3D information models for facility maintenance (Wang, Guo and Kim, 2019; Liu, 2022). By integrating LiDAR data in BIM, it becomes possible to enhance the productivity of maintenance and administration work in facilities. This approach allows for a more rapid detection of faults or potential faults and a significant reduction in maintenance costs.

Multiple studies have explored the benefits and limitations of using LiDAR technology in combination with BIM for facilities management. These studies have highlighted the importance of accurately capturing the existing building conditions through LiDAR scans and converting them into BIM models (Brilakis *et al.*, 2010; Juan, Hwang and Kim, 2020; Liu, 2022). The study conducted by Hong Kong researchers focused on utilizing Scan-to-BIM, which combines terrestrial laser scan data with BIM, to perform quantity take-offs for estimating building maintenance costs (Sing *et al.*, 2022). This demonstrates the potential of LiDAR technology and BIM integration in enhancing maintenance activities by providing precise information for cost estimation and planning. Additionally, another study emphasizes the automation of the Scan-to-BIM process through semantic labeling. This study proposes an integrated system called Auto-Scan-To-BIM that automates the generation of a complete BIM model from scanned point cloud data. The Auto-Scan-To-BIM system utilizes region segmentation methodology and enhanced plane boundary line detection methods to accurately identify and analyze building elements from the point cloud data (Suprun *et al.*, 2022). Furthermore, research has also focused on the integration of BIM and LiDAR technology for quality management during the maintenance phase (Begić and Galić, 2021). Moreover, scan-to-BIM technology supports the creation of accurate as-built models that can be used by facility managers and stakeholders in maintenance operations. The integration of scan-to-BIM technology extends the value of BIM to the lifecycle of buildings, contributing to improved facility management and maintenance practices. The practicality of implementing the Scan-to-BIM approach in maintenance projects for existing buildings has not been widely explored, despite its potential benefits for creating accurate as-built models and improving facility management practices (Sing *et al.*, 2022). To address these limitations, this article proposes a methodology for enhancing maintenance activities in existing buildings using LiDAR technology and the scan-to-BIM process.

Methodology

The methodology involves utilizing LiDAR technology to capture point cloud data of the existing building. It is important to note that when processing point clouds, several challenges arise, including occlusions, misclassification, and identifying hidden elements. Occlusions occur when objects obstruct parts of the building, leading to incomplete or distorted point cloud data. Misclassification refers to the incorrect labeling of point cloud points, which can result in inaccuracies in subsequent modeling processes. Similarly, identifying hidden elements, such as structural components obscured from view, poses challenges in capturing comprehensive building information. To address these challenges, innovative approaches, such as deep learning and machine learning algorithms, were used to improve the identification of hidden elements and enhance the overall accuracy of the resulting BIM model.

The existing point cloud data is then processed and converted into a mesh representation of the building. Then, the methodology oversees creating IFC instance models to store building components created out of a laser scan. It is divided into four parts, as illustrated in Figure 1: (i) point cloud classification; (ii) point cloud conversion; (iii) the IFC model creation; and (vi) CMMS integration and visualization.

The first step in the methodology is point cloud classification. During this step, the point cloud data is classified and labeled to identify different building elements such as walls, floors, ceilings, doors, windows, and some MEP elements. Models created using the data contained in point clouds can be used for constructing an accurate representation of existing facilities by measuring their geometry and appearance.

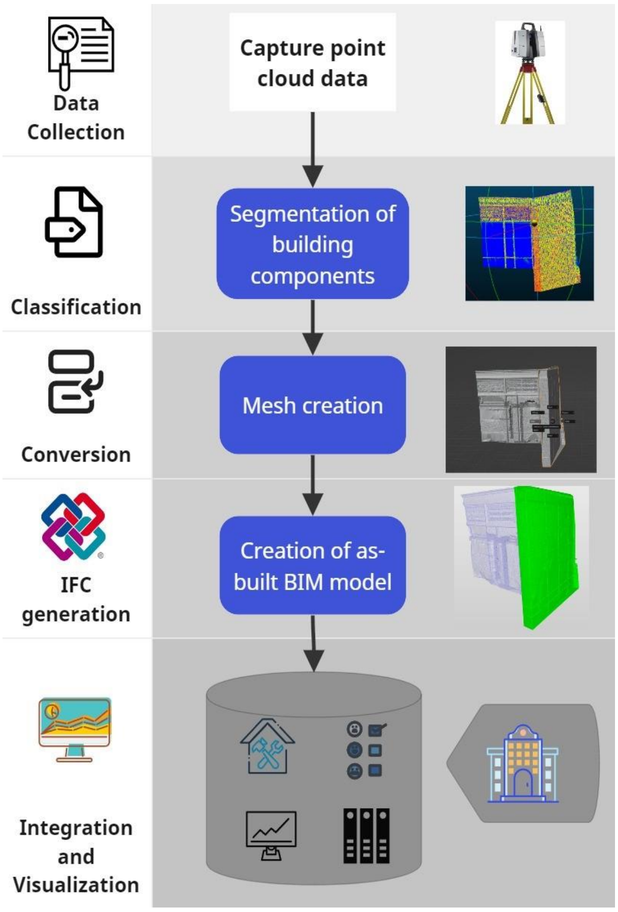


Figure 1. Methodology

The second step is point cloud-to-mesh conversion. During this step, the classified point cloud data is converted into a mesh representation of the building. This mesh representation allows for a more visually appealing and interactive view of the building, making it easier for maintenance workers to identify different elements and their locations. The third step in the methodology is the generation of an IFC model. During this step, the mesh representation of the building is assigned relevant Industry Foundation Classes entities. This involves mapping the building elements identified in the point cloud data to their corresponding IFC entities, such as walls, floors, doors, and windows. Once the relevant IFC entities are assigned, a complete IFC model is generated for the entire building. This IFC model contains all the necessary information about the building, including its geometry, spatial relationships, and attributes. The final step in the methodology involves using Dynamo in Revit to link maintenance requests from the building maintenance management system to the IFC model. By integrating the building maintenance with the generated IFC model, each maintenance request can be linked to the corresponding element in the building model. A filter can be applied to visualize the priority of each maintenance request on the model, dynamically representing their level of urgency or importance with different colors or indicators.

Case Study Implementation

To demonstrate the effectiveness of this methodology, a case study was conducted on the Civil Engineering Building in University of Cambridge, United Kingdom. The Cambridge University Engineering Department's (CUED) Civil Engineering Building, completed in 2019, stands as a pioneering example of modular and flexible laboratory and research infrastructure on the eastern fringe of the University's West Cambridge Campus. Encompassing approximately 5000m², the facility seamlessly integrates laboratory, research, and office spaces. LiDAR technology was utilized to capture a detailed and accurate point cloud of the building, shown in Figure 2.

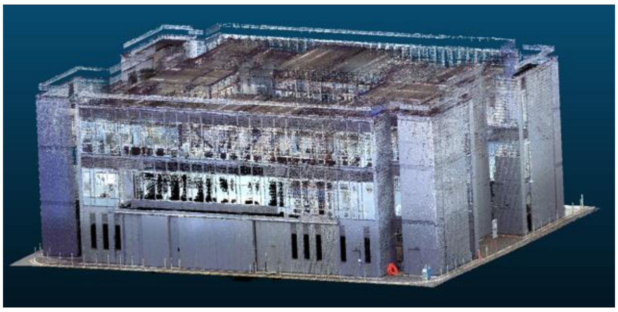


Figure 2. Point Cloud of the Civil Engineering Building

Point Cloud Classification

The point cloud data was then processed and classified to identify various building elements. The classification of the point cloud data enabled the identification of different building elements such as walls, floors, ceilings, doors, windows, and some MEP elements. Once the point cloud was segmented and the noise omitted, the next step involved the conversion of the classified point cloud data into a mesh representation of the building.

Point Cloud Conversion

The point cloud was converted into a mesh representation of the building. Using the innovative Python Ball Pivoting Algorithm, we were able to convert the raw point cloud data into a detailed mesh structure. This algorithm effectively handles challenges such as occlusions and hidden points, guaranteeing that the resulting mesh-based geometry accurately portrays the intricate details of the building's structure. The mesh representation provided a more intuitive and visually appealing view of the building, making it easier for maintenance workers to understand and interact with the model. Figure 3 shows the created mesh based on the point cloud data.

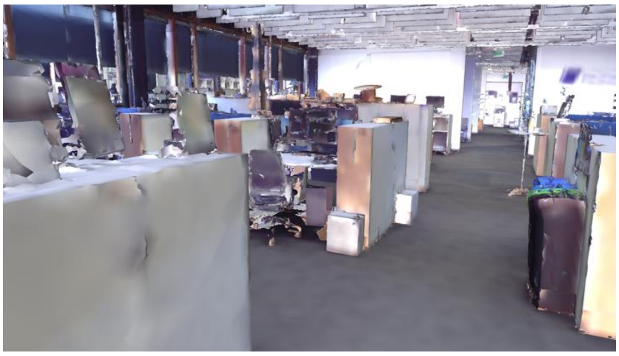


Figure 3. Mesh of the building

IFC Creation

Next, the classified point cloud data was mapped to relevant IFC entities, such as walls, floors, and doors. The IFC model was created using the ifcopenshell package in Python. This package allows for the manipulation and extraction of data from IFC files, enabling the assignment of relevant IFC entities to the classified point cloud data. The ifcopenshell library provides a set of tools and functions for working with IFC files, including the ability to create, read, and modify IFC models programmatically. Using the ifcopenshell package, the classified point cloud data was mapped to their corresponding IFC entities, such as walls, floors, doors, and windows. This process involved creating IFC elements for each building component, ensuring that the IFC model accurately represented the geometry and attributes of the existing building. The ifcopenshell package facilitated the generation of a complete IFC model that contained all necessary information about the building, allowing for spatial relationships and properties to be incorporated into the model.

By leveraging the capabilities of ifcopenshell in Python, the methodology successfully assigned relevant IFC entities to the classified point cloud data, resulting in the creation of a comprehensive and semantically rich IFC model for the existing building. Figure 4 illustrates ifcopenshell package for assigning relevant IFC to segmented mesh.

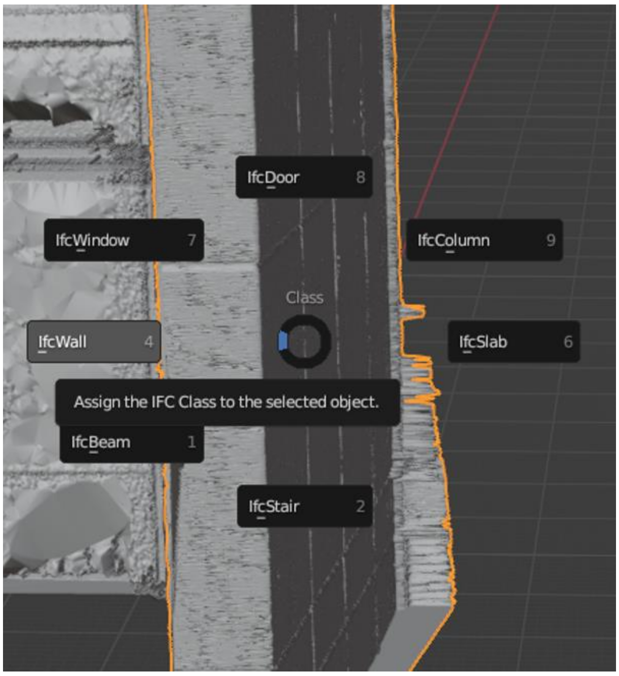


Figure 4. Process of assigning relevant IFC to mesh

CMMS Integration and Visualization

The process of linking maintenance requests from the CMMS to the generated IFC model involved several steps. First, the maintenance data obtained from the CMMS in an Excel file format was imported into Dynamo, a visual programming extension for Autodesk Revit as shown in Figure 5. Each row in the Excel file represented a different maintenance work order, with columns containing specific information such as issue, priority, and equipment codes. Table 1 shows examples of maintenance requests received by facility management department of west campus.

In this study, the shared parameter was utilized to allow BIM models to contain EquipmentID and Priority parameters. The shared parameter is a Revit term that can be added to the Revit family for custom data fields. It can also be accessible for any project due to holding parameters in a separate file. This enables the integration of maintenance data from the building maintenance management system with the BIM model, allowing for better coordination and planning of maintenance activities.

Table 1. Examples of maintenance requests for West Campus

Issue	**Date reported to Facilities**	**Priority**
Air Conditioning is still playing up	07 March 2023	Medium
remove alarm on back door	08 August 2023	low
repaint bike shed & compounds where w/shop exit is and where skips are	09 August 2023	low
The gas boiler in the plant room wont switch on	18 October 2023	medium
AC not working in the workshop	20 October 2023	medium
light inside study room not working	23 October 2023	medium
A leaking pipe in office	30 October 2023	high

To map maintenance requests to the BIM model for visualization, the following steps can be taken. First, the Excel file obtained from CMMS is imported in Dynamo, a visual programming extension for Autodesk Revit. In this file, each row corresponds to a different maintenance work order, and each column contains specific information such as issue, priority, and equipment codes. The parameters from the file are then grouped by type and mapped to each equipment (depending on the code) using the SetParameterbyName node in Dynamo. This step

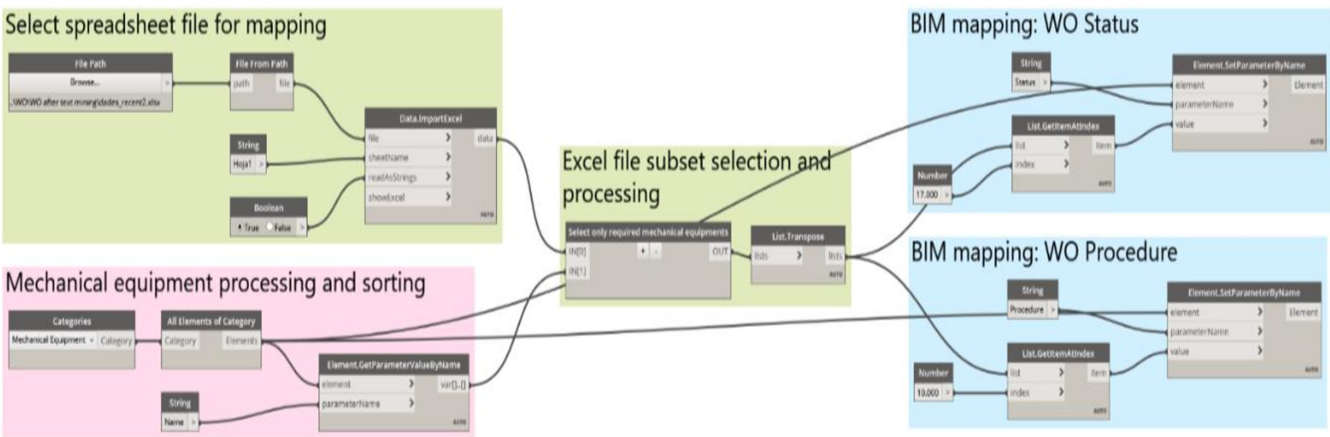


Figure 5. Dynamo code to import the maintenance requests to the mechanical equipment of the IFC model

involves linking the maintenance information to the corresponding elements in the BIM model. Using a Dynamo script, the information from the selected columns of the file can be saved in each parameter. This script contains two key Python code blocks. The code block "Reorganize room objects into Revit ordering" creates a list of all rooms and their room numbers, which are then organized to match the element numbering set by Revit. Developing custom Dynamo scripts specifically for this purpose made integrating maintenance activities into the BIM model easier. These scripts, using unique Python code blocks, showing our dedication to creating solutions that meet the specific needs of maintenance tasks in facility management. By using these custom scripts, we ensured that maintenance information was accurately mapped to corresponding rooms in the BIM model, improving the efficiency of maintenance planning and execution.

This integration allowed for the linking of each maintenance request to the corresponding element in the building model, with the ability to visually represent the priority of each maintenance request on the model. By dynamically indicating the urgency or importance of maintenance requests with different colors or indicators, the maintenance team gained an efficient and effective tool for managing and prioritizing maintenance activities.

Results and Discussion

The use of the proposed methodology for enhancing maintenance activities in existing buildings without a BIM model can greatly improve efficiency and accuracy. By utilizing LiDAR technology to capture point cloud data, maintenance workers can have access to detailed and accurate information about the building's geometry. This data is then transformed into a mesh representation using the Python Ball Pivoting Algorithm to ensure precise modeling that includes intricate details, such as holes in walls, thus enhancing the fidelity of the resulting BIM model. This BIM model can then be linked to the CMMS, enabling maintenance requests to be directly connected to the corresponding elements in the model.

The methodology involves using the ifcopenshell package in Python to convert the mesh representation into an IFC file. Unlike conventional approaches, ifcopenshell offers a versatile and general solution, providing greater flexibility in manipulating and extracting data from IFC files. This contributes to advancing the field by offering a more accessible and adaptable framework for generating IFC models from point cloud data.

In addition, we have developed custom Dynamo scripts designed to integrate maintenance tasks into Revit, enhancing the usefulness of BIM models in facility management. The unique Python code blocks within these scripts, especially the "Reorganize room objects into Revit ordering" code, demonstrate our dedication to creating solutions tailored for maintenance activities. This customization enables accurate mapping of maintenance information to respective rooms in the BIM model, leading to significant improvements in maintenance planning and execution efficiency.

The proposed visualization tool aligns with the case study's findings, illustrating how the seamless integration of LiDAR technology with BIM models effectively facilitates more efficient and accurate maintenance practices. By implementing the visualization tool, Facility Management and Maintenance teams can better organize and prioritize maintenance tasks. The clear visualization of maintenance requests directly on the 3D model of the building enables FMM personnel to assess the urgency and allocate resources effectively. Figure 6 illustrates examples of two maintenance requests on the West Campus of Cambridge University. As shown in the figure, each maintenance request is depicted by a colored marker on the BIM model. For instance, a red marker indicates a high-priority maintenance request, while a yellow marker represents medium-priority requests. A leaking pipe in the office is an example of a high-priority maintenance request and is represented as a red marker in the BIM model. On the other hand, a light in the study room constitutes a medium-priority request and is symbolized by a yellow marker. This color-coded representation not only improves the efficiency of maintenance workers in addressing urgent issues promptly but also allows for better planning and allocation of resources to address less urgent maintenance requests. In addition, this integration provides maintenance workers with a clear understanding of the location and status of each maintenance request, facilitating their decision-making process and streamlining the overall maintenance workflow.

The process of implementing the visualization tool involves the utilization of LiDAR technology to capture comprehensive point cloud data, enabling the creation of a detailed and visually rich BIM model. This model, integrated with the CMMS, enhances the clarity and accessibility of maintenance requests, streamlining the overall maintenance workflow while ensuring accurate documentation and spatial representation of the building's elements.

The comparison of the current maintenance management approach, where maintenance work orders are received in Excel files, with the proposed visualization tool clearly demonstrates the potential for improvement. The visualization tool provides a more intuitive way to manage maintenance tasks, moving away from traditional tabular data towards a spatially enriched representation of the building's elements. This transition can significantly enhance the FMM's ability to plan, coordinate, and execute maintenance activities effectively.

The implementation of the visualization tool, as evidenced by the successful integration of LiDAR technology, point cloud data, and BIM models in the case study, presents a compelling opportunity to revolutionize the management of maintenance tasks in existing buildings.

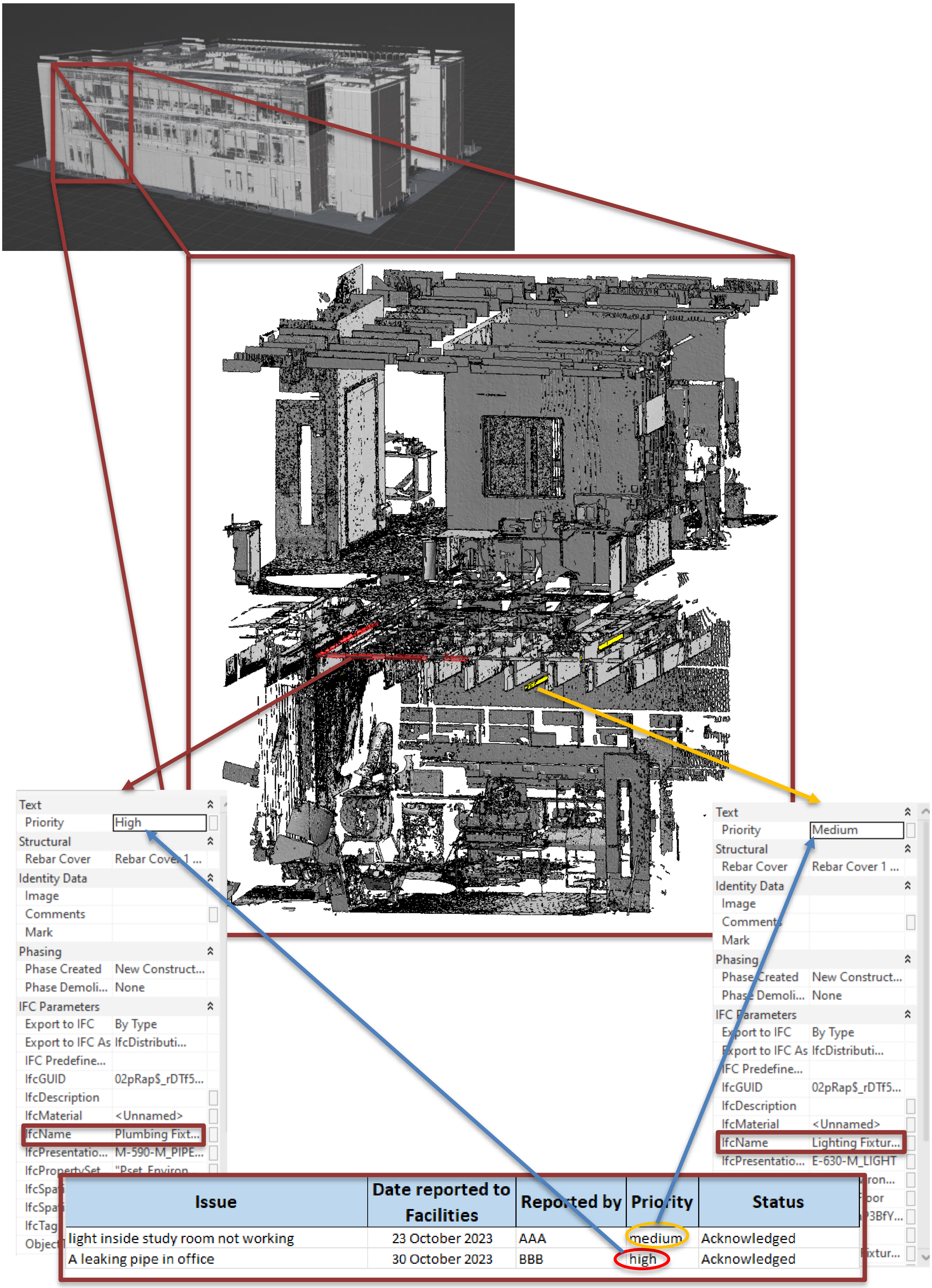


Figure 6. Example of the equipment visualization based on the priority of maintenance requests

The visualization tool not only fosters a more organized and informed approach to maintenance but also elevates the overall understanding of the building's conditions and facilitates well-informed decision-making for maintenance and rehabilitation efforts. This methodology not only improves the maintenance process but also enhances the overall management of the building. With the integrated system of LiDAR technology, point cloud data, and BIM models, facility managers can easily access and analyze important information about the building's condition, identify potential issues, and make informed decisions regarding maintenance and repairs. This approach also enables better communication and collaboration among stakeholders involved in the maintenance process.

Conclusion

The use of LiDAR technology and point cloud data to enhance maintenance activities in existing buildings without a BIM model offers significant advantages.

It allows for the creation of detailed and accurate as-built BIM models, which can be linked to maintenance management systems to streamline the maintenance workflow. By assigning relevant IFC entities and color-coding maintenance requests based on priority level, maintenance workers have a clear understanding of the location and status of each maintenance request, improving efficiency and effectiveness. Moreover, the integration of point cloud data and BIM models provides facility managers with valuable information for better decision-making and overall building management. For instance, various kinds of data can be linked to IFC models to improve maintenance activities. This includes information about equipment specifications, maintenance history, warranty details, and any other relevant data that can assist in planning and executing maintenance tasks. By leveraging the power of LiDAR technology and point cloud data, facility management teams can transform the way they approach maintenance in existing buildings. This innovative approach not only improves the accuracy and efficiency of maintenance tasks but also enhances the overall management of building conditions and facilitates informed decision-making processes. The combination of LiDAR technology, point cloud data, 3D models created with BIM in the context of maintaining activities for existing buildings presents a significant opportunity to revolutionize traditional approaches towards conducting property upkeep while improving their overall performance as well as longevity.

The development of custom Dynamo scripts highlights the potential of this methodology in the future. These scripts form the basis for potential projects, such as creating APIs for seamless integration with emerging digital twin platforms. Future work could focus on advancing the use of BIM and point cloud data in maintenance activities for existing buildings. One potential future direction is the refinement of automated algorithms and machine learning techniques to streamline the process of converting point cloud data into parametric BIM models. This advancement can significantly reduce the manual effort required, making the generation of BIM models from point cloud data more time-efficient and scalable. Additionally, future work in this area could explore the integration of Internet of Things data with BIM models to enable real-time monitoring of building components and systems. This integration would allow for proactive maintenance, where sensor data from IoT devices triggers maintenance alerts directly within the BIM model, enabling facility managers to address issues before they escalate. Moreover, the role of BIM in predictive maintenance can be explored in existing buildings. By leveraging historical maintenance data, sensor information, and building performance data within the BIM environment, predictive maintenance strategies can be employed to anticipate and plan for maintenance activities, maximizing the operational efficiency of existing buildings.

Acknowledgments

This work was supported by an EU Commission-funded project, the AEGIR project [grant number 101079961], and the OMICRON project, EU H2020 [grant number 955269].

References

Alavi, Hamidreza *et al.* (2021) 'BIM-based Augmented Reality for Facility Maintenance Management', in *Proceedings of the 2021 European Conference on Computing in Construction*, pp. 431–438. Available at: https://doi.org/10.35490/EC3.2021.180.

Alavi, H. *et al.* (2021) 'Enhancing occupants' comfort through BIM-based probabilistic approach', *Automation in Construction*, 123. Available at: https://doi.org/10.1016/j.autcon.2020.103528.

Alavi, H., Bortolini, R. and Forcada, N. (2022) 'BIM-based decision support for building condition assessment', *Automation in Construction*, 135. Available at: https://doi.org/10.1016/j.autcon.2021.104117.

Alavi, H. and Forcada Matheu, N. (2022) *Building information modeling for facility managers*, *TDX (Tesis Doctorals en Xarxa)*. Universitat Politècnica de Catalunya. Available at: https://doi.org/10.5821/dissertation-2117-375223.

Alavi, H. and Forcada, N. (2022) 'User-Centric BIM-Based Framework for HVAC Root-Cause Detection', *Energies*, 15(10). Available at: https://doi.org/10.3390/en15103674.

Alavi, S.H. and Forcada, N. (2019) 'BIM LOD for facility management tasks', in *Proceedings of the 2019 European Conference on Computing in Construction*, pp. 154–163. Available at: https://doi.org/10.35490/EC3.2019.187.

Antah, F.H. *et al.* (2021) 'Perceived Usefulness of Airborne LiDAR Technology in Road Design and Management: A Review'. Available at: https://doi.org/10.3390/su132111773.

Becerik-Gerber, B. *et al.* (2012) 'Application areas and data requirements for BIM-enabled facilities management', *Journal of Construction Engineering and Management*, 138(3), pp. 431–442. Available at: https://doi.org/10.1061/(ASCE)CO.1943-7862.0000433.

Begić, H. and Galić, M. (2021) 'A Systematic Review of Construction 4.0 in the Context of the BIM 4.0 Premise'. Available at: https://scite.ai/reports/10.3390/buildings11080337.

Brilakis, I. *et al.* (2010) 'Toward automated generation of parametric BIMs based on hybrid video and laser scanning data', *Advanced Engineering Informatics*, 24(4), pp. 456–465. Available at: https://doi.org/10.1016/j.aei.2010.06.006.

Fan, S.-L. *et al.* (2023) 'Augmented reality-based facility maintenance management system', *Facilities*, 41(13/14), pp. 769–800. Available at: https://doi.org/10.1108/F-04-2022-0059.

Gursel, I. *et al.* (2009) 'Modeling and visualization of lifecycle building performance assessment', *Advanced Engineering Informatics*, 23(4), pp. 396–417. Available at: https://doi.org/10.1016/j.aei.2009.06.010.

Juan, J., Hwang, K.-E. and Kim, I. (2020) 'A Study on the Constructivism Learning Method for BIM/IPD Collaboration Education'. Available at: https://doi.org/10.3390/app10155169.

Juricic, B.B., Krstić, H. and Čulo, K. (2021) 'Theoretical Analysis and Comparison of the Thermal Performance, Construction Costs, and Maintenance Complexity between a Conventional and an Intensive Green Roof'. Available at: https://doi.org/10.1155/2021/5559467.

Liu, X. (2022) 'Research on Construction of Bill of Quantities of Prefabricated Buildings Based on BIM'. Available at: https://doi.org/10.26689/jard.v6i5.4259.

Matos, R. *et al.* (2020) 'Strategies to Support Facility Management Resourcing Building Information Modelling'. Available at: https://doi.org/10.23967/dbmc.2020.131.

Nicolle, C. and Cruz, C. (2011) 'Semantic Building Information Model and multimedia for facility management', *Lecture Notes in Business Information Processing*, 75 LNBIP, pp. 14–29. Available at: https://doi.org/10.1007/978-3-642-22810-0_2.

Shin, H.J. and Cha, H.-S. (2023) 'Proposing a Quality Inspection Process Model Using Advanced Technologies for the Transition to Smart Building Construction'. Available at: https://doi.org/10.3390/su15010815.

Sing, M.C.P. *et al.* (2022) 'Scan-to-BIM technique in building maintenance projects: practicing quantity take-off'. Available at: https://doi.org/10.1108/ijbpa-06-2022-0097.

Suprun, E. *et al.* (2022) 'Digitisation of Existing Water Facilities: A Framework for Realising the Value of Scan-to-BIM'. Available at: https://scite.ai/reports/10.3390/su14106142.

Wang, Q., Guo, J. and Kim, M. (2019) 'An Application Oriented Scan-to-BIM Framework'. Available at: https://doi.org/10.3390/rs11030365.

AN ONTOLOGY CONCEPT FOR THE TOPOLOGICAL ABSTRACTION OF INFRASTRUCTURE NETWORKS

Jens Bartnitzek[1], and Al-Hakam Hamdan[1]
[1]A+S Consult GmbH, Dresden, Germany

Abstract

Infrastructure planning requires the use of various geometric representations of traffic nets thus exchanging different data models during a project that all refer to the same domain object. In order to provide a common data basis for a BIM-based workflow, this article proposes a topological model for representing traffic nets based on the parametric planing objects. Therefore, a concept for connecting the topology with alignment geometry data is described. Furthermore, the proposed concept allows the abstraction of multiple macro topologies for various use cases that all refer to the same lower level micro topology. Additionally, the approach is implemented by utilizing Semantic Web Technologies and therefore developing an OWL ontology concept for representing traffic topologies. Subsequently the concept is exemplified through an ontology that defines the topology representation of a simple rail switch.

Introduction

Due to the inherent curvature of the earth of the engineering alignment and the large-scale expansions of rail networks, different types of georeferenced representations of the rail network are required. The alignment planning is based on projected geodetic coordinate reference systems such as Universal Mercator Projection (UTM) where the site planning is separated from the height planning. Consequently, a common base model is required to which the various geometric representations could refer. This is especially required when applying methods of Building Information Modeling (BIM) in infrastructure planning, whereby various stakeholders and domain-specific project participants are involved.

One possible approach would be the use of a model that represents only the topological relations of the infrastructure network. This way, different geometric representations could reference or be linked to the same consistent topology, thus preserving data integrity in digital planning processes.

An infrastructure topology allows disparate geometric datasets to integrate with a single logical network representation, avoiding inconsistencies. The topological connectivity also enables sophisticated network analysis for tasks like routing, vulnerability assessment, and optimization. Since the topology is decoupled from physical geometry, it supports flexibility in aligning to different coordinate projections or levels of detail.

A new concept for modeling a topology that represents traffic nets of any kind is proposed in this article. Theoretically, this concept could be implemented in a variety of data structures, e.g. as an SQL database or using BIM standards such as the Industry Foundation Classes (IFC). Since topologies can be represented well as graphs and in order to integrate the underlying logic-based concepts in the model, the developed topology concept is implemented as a graph-based ontology using the Web Ontology Language (OWL). The proposed concept is demonstrated on an exemplary ontology that defines the topology representation of a simple rail switch.

Related Work

Various ontological approaches exist that aim for a topological representation of constructions. According to Gruber (2009), an ontology in the context of computer and information sciences is defined as a set of representational elements that can be used to model a domain of knowledge or discourse. These representational elements include classes, properties, and relationships or relations. Furthermore, the representational elements of an ontology contain information about their meaning as well as constraints for their logically consistent application. Ontologies can be created in various formats, of which the most prominent one is OWL that is also W3C-standardized[1]. OWL is based on the Resource Description Framework (RDF)[2] and Resource Description Framework Schema (RDFS)[3], which are also W3C standardized and allow the formulation of logical statements about any things (referred to as resources in RDF) in the form of triples consisting of subject, predicate and object. In this regard, OWL adds further concepts that enable the representation of knowledge on the basis of description logic (Baader et al., 2017). Since based on description logic, OWL-based ontologies are divided into two separate systems, one of which describes the Terminological Component (TBox) and the other the Assertional Component (ABox). In the TBox the knowledge concepts used in the ontology are defined including classes and properties. It establishes the hierarchy of classes and the specifications on properties, such as domain and range specifications, class equivalences, and property characteristics. The ABox contains assertions about instances of the classes defined in the TBox, essentially the factual knowledge. By reasoning the ABox against a TBox by an OWL-compatible reasoning engine, implicit information can be inferred based on the asserted facts in the ABox.

For buildings the Building Topology Ontology (BOT) (Rasmussen et al., 2021) can be used to define topological relations between digital representations of building zones and elements. Consequently, no geometry is involved in the topology, however references to geometry data can be

[1]https://www.w3.org/OWL/ (accessed: April 03, 2024)
[2]https://www.w3.org/RDF/ (accessed: April 03, 2024)
[3]https://www.w3.org/TR/rdf-schema/ (accessed: April 03, 2024)

made through links on BOT resources. A similiar approach can be realized for bridges with the Bridge Topology Ontology (BROT) (Hamdan and Scherer, 2020). Both approaches have in common that the structure of a construction is realized through classes for zones and components. Comparatively, a traffic net is described on a more abstract level and wider scale consisting of stations and connecting tracks.

An ontology for representing this kind of structure has been developed by Lorenz et al. (2005) as Ontology of Transportation Networks (OTN). Thereby, not only stations and their respective connections are modeled but also corresponding land areas or built structures. Furthermore, the connections could be typed via classes, e.g. a connection can be classified as waterway, road or railway. Another ontology for representing the topology of rails has been developed as Rail Topology Ontology (RTO) by Bischof and Schenner (2021). RTO is an OWL representation of the RailTopoModel (RTM), which has been released as International Railway Standard 30100 and is defined as UML model. Compared to OTN, RTM and RTO allow the definition of different aggregation levels such as nano, micro or macro for modeling the topology in various detail levels, e.g. for visualizing the network in a broader scale. However, parallel definition of higher level topologies that refer to the same lower level topology are usually not supported by RTM or lead to inconsistencies. Furthermore, OTN as well as RTM and RTO define no alignment strategies in their respective topology concepts for connecting topological data objects exactly to related geometry background by using the geometric alignment stations.

Besides ontological approaches, also general concepts for realising topologies in infrastructure have been developed. For instance, according to Bendfeldt (2005) a railway station can be divided:

- by operational technology in the areas of train transportation and train formation,
- operationally in the areas of route junctions and track groups or
- by type of transport in the areas of rail passenger transport and rail freight transport

Furthermore, Bendfeldt (2005) investigates how stations can be broken down into elementary nodes and elementary stations in order to derive standardization's and design principles for optimized operational and infrastructural aspects.

Methodology

In order to develop the topology and subsequently the prototypical ontology, core requirements must first be identified that have to be fulfilled by the resulting concepts. In this regard, a fundamental principle of traffic planning that needs to be considered for the topology is the need for specific traffic routes that result from the distribution of potentials. A location for a potential is defined as the position of an individual settlement or business location as well as an agglomeration of these locations. Topologically these (combined) locations are abstracted as nodes. Traffic potentials exist between the nodes, which as directed edges describe the potential for a traffic relationship in the form of travel or goods transportation between the nodes. Lill's law of travel (see eq. 1) provides a simplified way of evaluating this potential (Lill, 1891).

$$V_{AB} := k \cdot \frac{P_A \cdot P_B}{d(A,B)^2} \quad (1)$$

The size of the possible traffic flow V_{AB} is proportional to the product of the size of the agglomerations P_A, P_B to be connected and inversely proportional to the square of their distances $d(A,B)$. Furthermore, k is an additional constant correction value (Lill, 1891).

The logical network of potentials is realized with a physical network of traffic infrastructure across several modes of transport. Track-guided traffic routes form an essential part of the physical network, but require a reference to the associated geometric representation. At the lowest topological level, the node-edge model represents the exact physical network, while other topological models generalize it through a topological abstraction. A topology can have a geometric representation, but it does not have to. Furthermore, the representation may differ from the physical realization. For instance, the geometric representation of a station can, for example, be realized geometrically via its station building, the arithmetic center of its points or an arbitrarily set station position.

Consequently, the following requirements for a infrastructure topology can be determined:

1. The topology should consist of multiple abstraction layers.
2. The lowest abstraction layer should represent the exact physical infrastructure network.
3. Higher abstraction layers could generalize the topological abstraction of the infrastructure network.
4. Since the lowest abstraction layer represents the exact physical infrastructure network, it should be possible that an optional geometrical representation could be mapped to the layer components.

The infrastructure topology can therefore have a geometric representation, but does not have to. This representation may differ from the physical realization. The geometric representation of a station can, for example, be realized geometrically via its station building, the arithmetic center of its switches or an arbitrarily set station position. This creates a strict separation of the levels:

- The route alignment is the level of physical realization of the network
- The topology is a logical level that can reference the route alignment in a specialization, but does not correspond to it.

After this initial step of identifying the topology requirements, a topology concept will be developed that aims to fulfil them (see section "Concept of the Infrastructure

Topology"). The concept should be applicable on various data structures. For an initial prototypical implementation of the resulting topology concept the research team has decided to develop an OWL ontology. This allows for representing the topologies as graphs which are flexible to modify, especially since new edges can be added to nodes without violating previously defined data schemes. Furthermore, OWL supports logic-based modeling allowing the integration of underlying knowledge concepts and their automatic reasoning. Additionally, by using OWL as machine-interpretable standard, the resulting topology concepts can be shared and connected with other knowledge domains. For developing the prototypical ontology the Linked open terms methodology (Poveda-Villalón et al., 2022) is used, which is structured in the four steps (1) ontology requirements specification, (2) ontology implementation, (3) ontology publication and (4) ontology maintenance. For proposing the initial concept in this article, only the first two steps are processed, followed by a publication of the prototype ontology as draft. The requirements for the ontology should be the same as for the overall topology concept (see requirements 1. - 4. above). During the ontology implementation it has been considered to reuse components of OTN and RTO (see section "Related Work") directly in the newly developed ontology. However, since the structure of both ontologies differs greatly from the abstracted topology concept proposed in this paper, only the development of corresponding alignment ontologies is considered for future research.

Concept of the Infrastructure Topology

The proposed topology is an abstraction of a infrastructure network realized through logical nodes and edges that are separated from its geometric references. Therefore, the topology is defined as directed graph, which consists of the following elements:

- a set $V = \{v_0, ..., v_n\}$ of nodes
- a set $E = \{e_0, ..., e_m\}$ of directed edges between the nodes

Fig. 1 demonstrates an example of this abstract model consisting of 5 nodes and 8 edges that result in a directed graph.

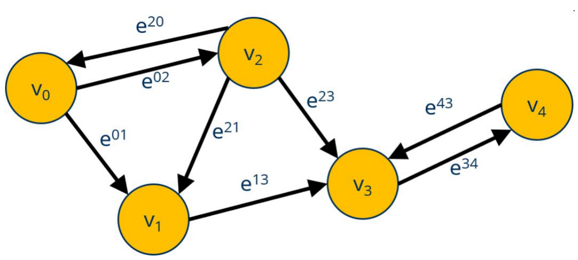


Figure 1: Directed graph consisting of 5 nodes and 8 edges

Topologically the edge $e_i := e^{jk} := (v_j, v_k)$ is unique between the nodes v_j and v_k. If there are more direct paths between v_j and v_k these are summarized topologically to e^{jk}. Through the graph the edges have access to their associated nodes and the nodes have access to their linked edges.

The proposed topology utilizes microtopologies and macrotopologies as utilized levels for abstracting a route network.

Microtopologies

The microscopic topology (microtopology) of a transportation network describes the first topological abstraction of the network from the geometry to a node-edge model. It is also the only geometric abstraction. Fig. 2 presents an exemplary microtopology of a simple switch, in which the two possible paths shown between the nodes are physically defined via the main or branch track. However, an exact geometric location and demarcation between the switch and the rest of the track can only be realized by a topological representation for the start and end of the switch. The double crossing switch results in four possible paths, only two of which are described by the alignment of the two tracks. Thus the geometric descriptions of the two additional paths are therefore required as alignments and are later assigned to the switch object.

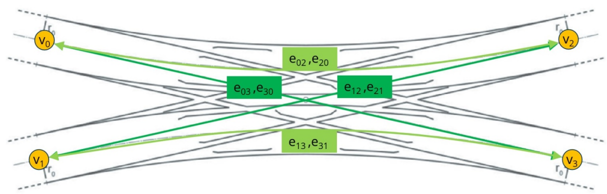


Figure 2: Exemplary microscopic topology of a single switch and a double crossing switch. The directions are indicated by double arrows. Background Figures from (Fendrich and Fengler, 2014)

If the topological decomposition of the switch is omitted, the geometric links to the physical network at these points are incomplete or only implicitly contained via the switch parameters. In addition, the switch is only positioned at a virtual intersection point, which is irrelevant for topological references back to the geometry.

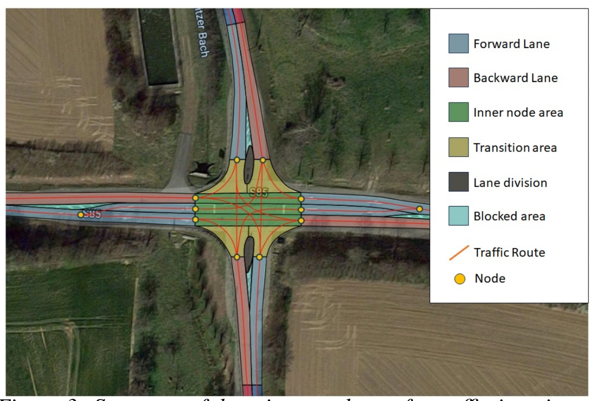


Figure 3: Structure of the microtopology of a traffic junction of two roads (Fig. from Wollschläger (2021) overlaid with the microtopological nodes)

The abstraction corresponds to the definition of junctions in road engineering. As shown in Fig. 3, there is also a strict separation between the geometric construction and the topological abstraction. The references to each other are established via nodes on the geometric lanes and edges as sections on the lanes.

Macrotopologies

The macroscopic topology (macrotopology) of a transportation network describes a topological abstraction of an existing topological network. It is therefore based on either a microscopic or macroscopic topology. It has only topological and no geometric references. Any number of macrotopologies can be constructed for a microtopology. An exemplary macrotopology is shown in Fig. 4.

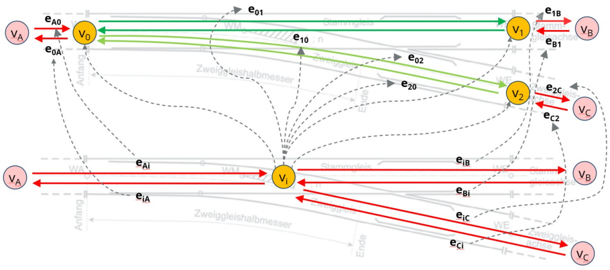


Figure 4: Abstraction of a switch as node v that aggregates nodes v_0, v_1, v_2

In the example a simple switch is defined through 3 inner nodes v_0, v_1 and v_2. According to the described data model, these nodes have back references to their linked edges, e.g. node v_0 to the edges e_{0A}, e_{01} and e_{02} with exact geometric back references.

When the microtopology is abstracted to a macrotopology, a macro node v is generated as a switch representation from the micro nodes v_0, v_1 and v_2. So that the macro node can form back references to the geometry without defining them itself, it must link all the associated micro nodes v_0, v_1 and v_2. For completeness, all inner edges e_{01}, e_{10}, e_{02} and e_{20} are also linked. Thus v itself is a directed graph.

The macro edges refer to their corresponding micro edges, e.g. e_{Ai} refers to e_{A0}, and if available their corresponding enclosed micro nodes.

All levels must be linkable with each other for a uniform, consistent topology approach. By linking them, arbitrarily abstract topology levels implicitly gain access to the underlying microtopology and thus to the physical geometry.

Structure of an OWL Topology

Based on the proposed concept of an infrastructure topology the Infrastructure Topology Ontology (INTO)[4] has been formalized using OWL. INTO provides the necessary TBox to support the modeling of microtopologies and all macrotopologies at higher abstraction levels as well as additional linking components for connecting the topology with geometry. In this regard, micro- and macrotopologies are specializations of the abstract definition of a graph consisting of nodes and edges according to Fig. 1.

Table 1 contains an overview of the prefixes used for the realization of the ontology. The prefix *into* refers to a namespace that is used by all classes and properties specifically created for INTO. In this regard, the INTO ontology uses terminology from RDFS (prefix: *rdfs*) and OWL (prefix: *owl*).

Table 1: Namespaces and prefixes used in this article

Prefix	Namespace
rdfs	https://www.w3.org/2000/01/rdf-schema#
owl	https://www.w3.org/2002/07/owl#
into	https://app.korfin.de/ontology/into#

An important implementation aspect is that edges are not formalized as OWL object properties but instead as named individuals. The primary reason for this design decision is that not only nodes are referenced by lower and higher topology abstraction levels but also edges, thus requiring a form of linking mechanism for statements. One possible solution would be the use of the RDF extension RDF-star[5], which allows the direct linking of complete RDF statements to other resources. However, RDF-star is not interpretable by most OWL reasoning engines, thus reducing the overall applicability of the ontology in an OWL context. The same applies to alternative graph models, e.g. the graph database Neo4J[6]. For this reason edges have been modelled as individuals according to the ontology design pattern "N-ary Relations"[7]. Thus, each instance of *into:edge* is connected to two instances of *into:node* via subproperties of *into:connectsToNode* but in addition can also be referenced by other instances of *into:TopologicalElement* of higher or lower topology layers. Direct connections in the form of auxiliary object properties are inferred by reasoning corresponding property chain axioms. The core concepts of INTO are shown in Fig. 5. There, each topology, be it a micro- or macrotopology, is represented as instance of *into:Topology*. Corresponding nodes (*into:Node*) and edges (*into:Edge*) are assigned to the topology trough respective object properties (*into:containsNode* or *into:containsEdge*). In this regard, the *into:Topology* instance is used as container object for storing the topological components. The actual topology is defined by assigning *into:Node* instances as start- and end-node to instances of *into:Edge* via the object properties *into:startsAtNode* and *into:endsAtNode*, thus implicitly defining the edge direction. The properties *into:startsAtNode* and *into:endsAtNode* are sub properties of *into:connectsToNode*, which is an inverse object property of *into:connectsToEdge*. Thus, by using a property chain axiom that infers the object property *into:isConnected* if an instance of *into:Node* is connected via *into:connectsToEdge* to an instance of *into:Edge* and that edge representation connects to a different instance of

[4] `https://app.korfin.de/ontology/into`

[5] `https://w3c.github.io/rdf-star/cg-spec/editors_draft.html` (accessed April 03, 2024)

[6] `https://neo4j.com/` (accessed: April 03, 2024)

[7] `https://www.w3.org/TR/swbp-n-aryRelations/` (accessed: April 03, 2024)

into:Node via *into:connectsToNode* it is possible to model direct connections between nodes through reasoning.

The symmetric property *into:inverseToEdge* allows the definition of an edge that has an inverse direction compared to another edge and thus opposing start and end points.

Both classes *into:Node* as well as *into:Edge* are subclasses of the generalizing class *into:TopologicalElement*. Furthermore, *into:Node* and *into:Edge* have additional subclasses that are used for topological elements in macro-topologies (*into:MacroNode* and *into:MacroEdge*) and microtopologies that are linked with a geometry-based alignment (*into:AlignmentNode* and *into:AlignmentEdge*).

An essential part of the proposed topological concept is the aggregation of topological elements, e.g. the aggregation of inner nodes in a macro node or macro edge. Therefore, each instance of *into:TopologicalElement* can aggregate a number of other instances of the same class via the object property *into:aggregatesTopologicalElement*. On the other hand the inverse relation via *into:isAggregatedBy-TopologicalElement* is also possible. Thus, inner nodes and edges as well as outer nodes and edges can be realized for each topological element.

OWL Realization of Microtopology

Modeling a microtopology is realized through alignment nodes (*into:AlignmentNode*) as specialized subtype of a node and alignment edges (*into:AlignmentEdge*) as specialized subtype of a common edge. These are the only types of nodes or edges, which are related to the alignment via alignment sections and thus are related to the rail geometry.

For each alignment section two alignment edges can be modeled, since in the directed, labelled graph each direction is represented through its own edge. Therefore, for each instance of *into:AlignmentEdge* a reference to the respective instance of *into:AlignmentSection* is necessary as well as a definition of the direction via the data property into:direction.

OWL Realization of Macrotopology

Macrotopologies are defined via macro nodes (*into:MacroNode*) as specialized subtype of nodes as and macro edges (*into:MacroEdge*) as specialized subtype of edges. Macro nodes and macro edges connect respectively the corresponding nodes and edges of lower topology levels. For simplified calculations, the inner nodes are separated from the outer nodes and the inner edges from the outer edges in a macro node. A macro edge also has this separation, although it cannot have any outer nodes. Macro nodes and macro edges describe an independent, directed, labelled graph.

Thus infrastructure elements can be separated and abstracted from each other. The alignment is abstracted topologically via the microtopology including mapped switches. A topological track plan could be obtained by merging the switches into macro nodes and defining the connecting macro edges. For example, if the macro nodes of all points of a station and possible multi-tracks between the stations are summarized in macro edges, it would be possible to get Line network plans. Furthermore, it is possible to determine comprehensive network maps by combining stations into agglomerations in further macro nodes as well as combining routes including "intermediate stations" into macro edges.

Connecting Topology with Geometry

To define the topological node-edge model independently from the underlying geometry, additional linking elements are used (see Fig. 7). The core elements for this are alignment sections (*into:AlignmentSection*) and alignment stations (*into:AlignmentStation*), which define the alignment positions. Instances of both of these classes can be assigned to an instance of *into:Alignment* that represents the alignment via the object properties *into:partAlignment* for linking alignment sections and *into:containsStation* for linking alignment stations. Additionally, an inverse object property of *into:partOfAlignment* named *into:dividedBy-Section* could be utilized in the modeling process.

An alignment section is defined through a start station and end station both positioned on the same alignment. Therefore, instances of *into:AlignmentStation* are assigned to an *into:AlignmentSection* instance via the properties *into:startsAtStation* and *into:endsAtStation*. Since the stations are defined as OWL individuals they can be reused in the modeling process, e.g. the end of an section is usually the start of the next section on the same alignment.

The configuration of an alignment section consisting of multiple sub-sections from different alignments, which could happen for example through different planning sections or sectional remeasurement, could be realized through a specialized alignment that is linking multiple alignment sections.

Thus the geometry reference is completely described and independent of the geometric construction, i.e. the topology can be defined at any point on the geometry.

The proposed concept can be generalized to point and line objects. Consequently, an instance of *into:AlignmentSection* represents a special line object and an instance of *into:AlginmentStation* a special point object. The further topological view is limited to the geometric reference, which is also defined by the generalization. Via property chain axioms alignment sections can also be assigned automatically to alignments if they define the start or end of a corresponding alignment section.

The connection between the alignment section and the topology is realized through linking instances of *into:AlignmentSection* and *into:Alignment* with micro edges realized through *into:Alignment*. Therefore, the object properties *into:assignedToSection* and *into:assignedToAlignment* are used. Furthermore, the relation *into:assignedToAlignment* between *into:Alignment-Edge* and *into:Alignment* can be inferred if the respective instance of *into:AlignmentEdge* is connected to an in-

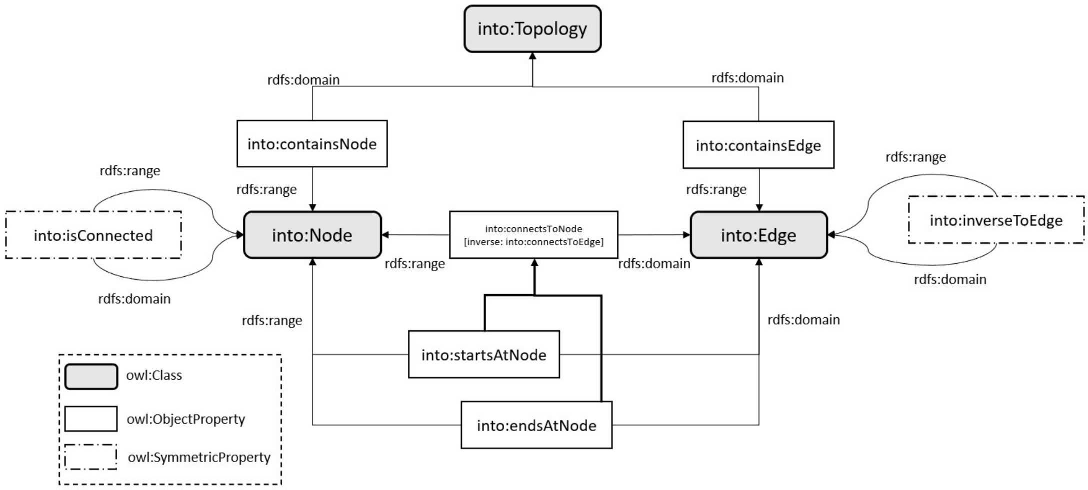


Figure 5: Core concepts of the topology part of the proposed OWL ontology

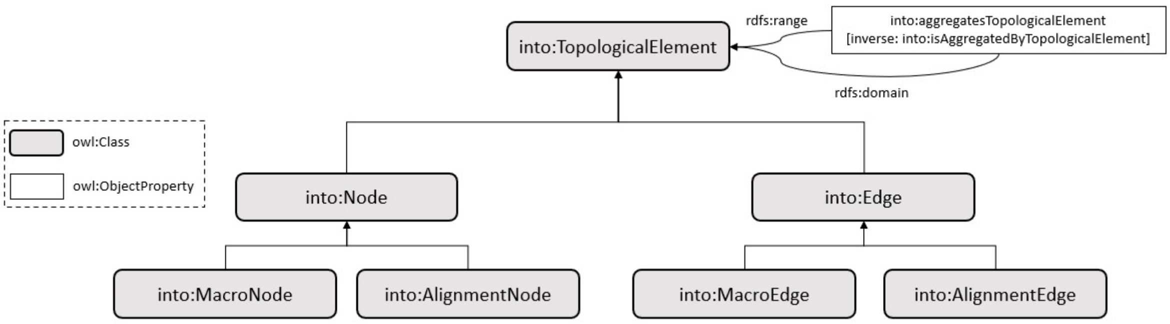


Figure 6: Inheritance concept for topological components

stance of *into:AlignmentSection* via *into:assignedToSection* and that representation of alignment section is linked to into:Alignment via *into:partOfAlignment*. The geometric data can either be linked directly in the OWL graph by using auxiliary ontologies such as the File Ontology for Geometry (FOG) (Bonduel et al., 2019) or by defining a separate link model, e.g. a linkset in an Information Container for Linked Document Delivery (ICDD) according to ISO 21597. Additionally, further management of the alignment representation could be utilized via the Ontology for Managing Geometry (OMG) (Wagner et al., 2019).

Application of the OWL Topology

To demonstrate the developed ontology an exemplary abstraction between two topological levels to represent a simple rail switch have been modeled as ABox using the TBox described in the previous section. The structure of the switch is the same as shown in Fig. 4. The topology example has been modeled as OWL ontology and uploaded in a locally hosted GraphDB[8] triplestore. Fig. 8 shows a graph visualization whereby the focus node is v_i.

By reasoning the modeled assertional graph (Abox) that stores all factual information against the terminology (TBox) of the INTO Ontology, aggregations of topological elements of the macro topology towards a microtopology can be inferred and vice versa. For instance, for the nodes v_0, v_1 and v_2 the outer node v_i has been automatically inferred if v_i assigned these nodes as inner nodes. Furthermore, all inner edges e_{01}, e_{10}, e_{02} and e_{20} are also linked. Consequently, v itself is a directed, labelled graph. By processing the graph in a reasoning engine direct connections (*into:isConnected*) can be inferred for all nodes that belong to the same edge. For example, v_0 is connected to v_1 and v_2, while v_1 and v_2 share no connection with each other.

For a high-performance evaluation of the node at runtime, it is advised to set up a path matrix across all outer nodes of the macro node, i.e. a statement whether v_1 can be reached via v_2. However, this is not part of the ontology model.

The macro edges refer to their corresponding micro edges, e.g. e_{Ai} refers to e_{A0}, and if available their corresponding enclosed micro nodes.

[8] https://graphdb.ontotext.com/ (accessed: April 03, 2024)

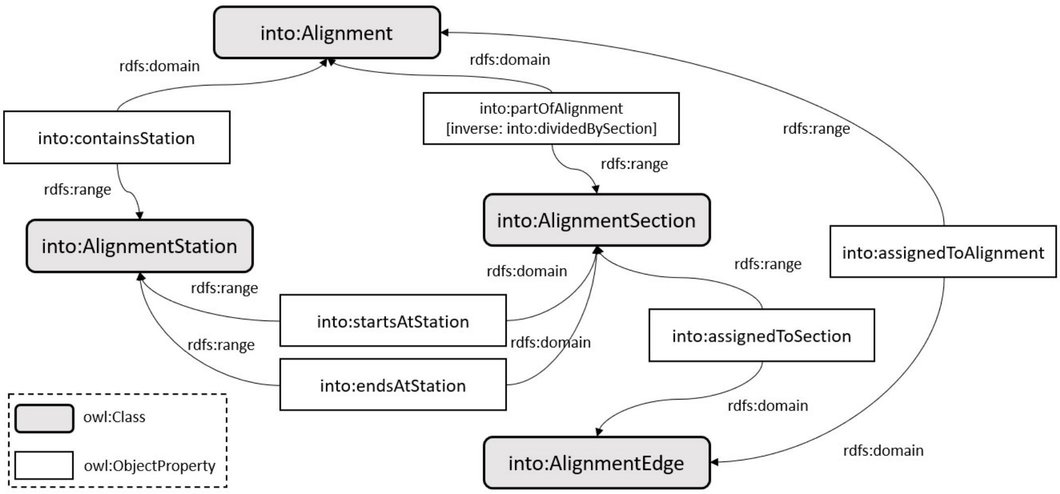


Figure 7: Concept for ontological alignment representation

The presented concept can also be used recursively for all higher-level macrotopologies.
In general the abstraction of the ontology involves the following steps:

1. Definition of the nodes and edges to be abstracted, so that the necessary nodes and edges can be clearly assigned to a higher abstraction level in a node-edge model.
2. Creation of the new macro nodes and internal structure by referencing the associated nodes and edges.
3. Creation of the new macro edges and of the inner structure by referencing the associated nodes and edges.
4. Linking the new macro nodes and macro edges via the back references of the outer nodes and edges.
5. Temporary creation of path matrices for macro nodes and macro edges via the inner structures.

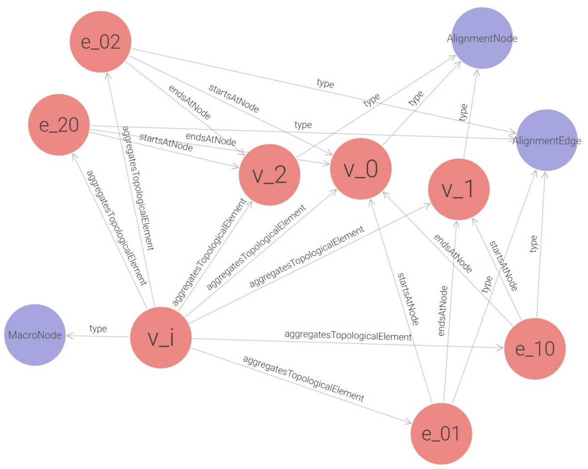


Figure 8: OWL ontology example in GraphDB

Based on the requirements defined in the "Methodology" section various SPARQL queries could be applied that make use of the topology structure and inference mechanisms to provide information about the infrastructure network and validate the ontologies functionality. For instance Query 1 identifies elements of higher and lower abstraction layers for a set of topological elements.

Listing 1: SPARQL Query for Topological Aggregation

```
SELECT ?element ?lowerElement ?higherElement
WHERE {
  {
    ?element into:aggregatesTopologicalElement ?
         lowerElement .
    OPTIONAL {
      ?element into:isAggregatedByTopologicalElement ?
           higherElement .
    }
  }
  UNION
  {
    ?element into:isAggregatedByTopologicalElement ?
         lowerElement .
    OPTIONAL {
      ?element into:aggregatesTopologicalElement ?
           higherElement .
    }
  }
}
```

Based on the requirement that the lowest abstraction layer should represent the exact physical infrastructure network, queries for selecting information about instances of *into:AlignmentEdge*, e.g. in Query 2 could be applied. Thereby, Query 2 is used for determining the starting and ending station of each alignment edge. A similar query could be applied for instances of *into:AlignmentSection*.

Listing 2: SPARQL Query for Alignment Edges

```
SELECT ?edge ?start ?end
WHERE {
  ?edge a into:AlignmentEdge ;
          into:startsAtStation ?start ;
          into:endsAtStation ?end .
}
```

Conclusions

In this paper a concept for a topology that represents traffic nets has been proposed. In addition a corresponding prototypical ontology has been formalized in OWL and illustrated in an example of representing a simple switch.

The topology can be structured in multiple abstraction levels. In this regard two level types are used: (1) Microtopologies that describe the traffic net at the lowest abstraction level and reference to a geometric representation and (2) Macrotopologies that describe a higher topological abstraction of the traffic net and therefore are based on lower microtopologies or macrotopologies. An important aspect is that multiple macrotopologies can refer to the same microtopology in parallel for various use cases, allowing the use of domain specific topologies for e.g. train transportation or route junctions.

One key aspect is that edges are not classified as OWL object properties when modeling the topological concept in an ontology. The main reason for this is that individuals that represent edges need to reference other edges in order to abstract the topology in multiple levels of different hierarchies. A combination of RDF-star and OWL has been considered but was not applied due to inference issues when processing the ontology in an OWL reasoning engine.

The topology can be used in various user scenarios. For example, a model for data exchange can be made available for traffic flow analyses and simulations. The topology can also act as the foundation of further BIM-supported infrastructure planning by providing a basic geometry-independent model. It is subject of future research how the developed ontology could be aligned with other ontologies used in infrastructure and construction planning, e.g. the Bridge Topology Ontology (BROT) (Hamdan and Scherer, 2020) or the Rail Topology Ontology (Bischof and Schenner, 2021). Furthermore, geodata ontologies such as the GeoSPARQL Ontology (OGC GeoSPARQL Standards Working Group, 2024) could be utilized for georeferencing the components of the topology. In order to ensure the practicability of the topology concept and the ontology, both developments must be applied in comprehensive test cases as part of future research.

References

Baader, F., Horrocks, I., Lutz, C., and Sattler, U. (2017). An Introduction to Description Logic. Cambridge University Press.

Bendfeldt, J.-P. (2005). Möglichkeiten der Standardisierung in der Infrastrukturplanung von Eisenbahnknoten. Eurailpress Tetzlaff-Hestra.

Bischof, S. and Schenner, G. (2021). Rail topology ontology: A rail infrastructure base ontology. In International Semantic Web Conference, pages 597–612. Springer.

Bonduel, M., Wagner, A., Pauwels, P., Vergauwen, M., and Klein, R. (2019). Including widespread geometry formats in semantic graphs using rdf literals. In Proceedings of the 2019 European Conference on Computing in Construction, volume 1 of Computing in Construction, pages 341–350, Chania, Greece. European Council on Computing in Construction. 10.35490/EC3.2019.166.

Fendrich, L. and Fengler, W. (2014). Handbuch Eisenbahninfrastruktur. Springer-Verlag.

Gruber, T. (2009). Ontology. In Liu, L. and Özsu, M. T., editors, Encyclopedia of Database Systems, pages 1963–1965, Boston, MA. Springer US. 10.1007/978-0-387-39940-$9_1$318.

Hamdan, A.-H. and Scherer, R. J. (2020). Integration of bim-related bridge information in an ontological knowledgebase. In Proceedings of the 8th Linked Data in Architecture and Construction Workshop-(LDAC), pages 77–90.

Lill, E. (1891). Das Reisegesetz und seine Anwendung auf den Eisenbahnverkehr mit verschiedenen auf die Betriebsergebnisse des Jahres 1889 bezugnehmenden statistischen Beilagen in Tabellen und bildlicher Form. Spielhagen u. Schurich.

Lorenz, B., Ohlbach, H. J., and Yang, L. (2005). Ontology of transportation networks. University of Munich, REWERSE Project, Tech. Rep. A1-D4.

OGC GeoSPARQL Standards Working Group (2024). Geosparql ontology. https://opengeospatial.github.io/ogc-geosparql/geosparql11/index.html , accessed: 2024-01-30.

Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., and García-Castro, R. (2022). Lot: An industrial oriented ontology engineering framework. Engineering Applications of Artificial Intelligence, 111:104755. https://doi.org/10.1016/j.engappai.2022.104755.

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., and Pauwels, P. (2021). Bot: The building topology ontology of the w3c linked building data group. Semantic Web, 12(1):143–161.

Wagner, A., Bonduel, M., Pauwels, P., and Uwe, R. (2019). Relating geometry descriptions to its derivatives on the web. In Proceedings of the 2019 European Conference on Computing in Construction, volume 1 of Computing in Construction, pages 304–313, Chania, Greece. European Council on Computing in Construction. 10.35490/EC3.2019.146.

Wollschläger, J. (2021). Implementierung von parametrischer geometrie, topologie und semantik für verkehrsknotenpunkte in georeferenzierten bim-projekten.

AIPDORCS: ARTIFICIALLY INTELLIGENT PRELIMINARY DESIGN OF REINFORCED CONCRETE STRUCTURES

Alon David Argaman and Rafael Sacks
Technion – Israel Institute of Technology, Haifa, Israel

Abstract

Leveraging 106 engineers' expert assessments of preliminary structural design in 48 reinforced concrete building models, we compiled two experimental Graph Neural Network (GNN) tools to demonstrate feasibility for automated classification of structural schematic layouts, a key step toward building generative artificial intelligence (AI) tools for design. Contributions include a robust project database, a model-to-graph conversion tool, and a structural design scoring application. Acknowledging limitations related to modelling assumptions and a relatively small dataset, this research clarifies the opportunity and the obstacles to AI-driven advancements in preliminary structural design.

Introduction

The construction industry has seen remarkable advances in computational tools, from 3D modelling to finite element analysis. One critical aspect has remained relatively untouched: the preliminary design phase, a creative step where engineers propose conceptual structural layouts using heuristic rules (Caprani, 2009). Despite the availability of sophisticated tools for detailed design, the quality of results is often constrained by suboptimal preliminary designs. While optimization algorithms like genetic algorithms are available for engineering design (Katoch et al., 2021), they face challenges in execution simplicity, adapting to changing standards, and precise programming. This research explores whether deep learning can be used at the preliminary design stage to simulate the experience of senior engineers, offering faster, cheaper, and more accurate solutions. Focusing on reinforced concrete (RC) structures, we address the scarcity of AI applications in preliminary design, particularly within the construction industry (Afzal et al., 2020). The research aims to leverage recent developments in AI and machine learning to better automate the preliminary design process, with a specific focus on structural plans classification.

Background

Engineering Design and Automation

Engineering design, a creative and iterative process, has seen a surge in automation using AI techniques, aiming to enhance efficiency, reduce costs, and improve design quality. Expert systems, an early AI development in the 1970s and 1980s, simulated human decision-making by using a knowledge base, and were applied in various engineering fields to make decisions and solve problems (Rychener, 1988). However, the emergence of more advanced AI algorithms, such as machine learning, neural networks, and deep learning, has surpassed the limited use of expert systems (Medsker, 1995).

The AEC (Architecture, Engineering, and Construction) industry has witnessed significant advancements in AI applications, notably in architectural layout design and generative design. The use of Generative Adversarial Networks (GANs) for architectural conceptual design (As et al., 2018), floor plan generation through rectangular dual finding algorithm (Wang et al., 2018), and automated apartment plan generation using genetic algorithms (Laignel et al., 2021) exemplify the industry's adoption of AI. AI's influence extends beyond architecture to construction, with applications like predicting labour costs (Huang and Hsieh, 2020), estimating assembly costs of flooring systems (Elhegazy et al., 2022), and employing Deep Convolutional Neural Networks (CNNs) for automated assembly of lunar construction components (Zhou et al., 2020). Start-up companies like “FIRMUS” (Abecasis and Amar, 2019), “Daisy” (Selvaraj et al., 2019), and “Structure Plus” (Jozi et al., 2019) further demonstrate the feasibility and effectiveness of incorporating AI in the AEC industry, enhancing design efficiency and optimizing structural elements.

Artificial Intelligence and Structural Design

Over the years, AI has advanced significantly, and the integration of Building Information Modelling (BIM) has become commonplace. The rationale for employing AI in preliminary structural design is rooted in the recognition that this phase demands substantial creative input, combining engineering expertise, wisdom, and judgment (Samuel and Weir, 1999). While traditionally deemed difficult to simulate with rule-based tools, recent developments showcase the ability of advanced tools, such as Artificial Neural Networks (ANNs), to emulate these qualities by drawing inspiration from human brain neurons and incorporating engineering experience through prior data collection. Noteworthy applications include the use of ANNs for tall building design (Anwar et al., 2015), design response grammars and evolutionary algorithms for conceptual layouts creation (Boonstra et al., 2020), and reinforcement learning for optimizing plane frames (Hayashi and Ohsaki, 2020). Recent examples span diverse areas, such as accelerating the exploration of shell structure topological design (Zheng et al., 2020), designing RC columns with ANNs surpassing traditional design charts (Charalampakis and Papanikolaou, 2021), automating shear wall design for residential buildings using a method called StructGAN (Liao et al., 2021), and recommending early-stage structural design with minimal errors (Ampanavos et al., 2022). These advances collectively highlight the potential of AI to revolutionize the structural design phase, offering expedited and precise design solutions while minimizing costs and computational resources.

BIM and Graph-based Representation of Structures

Building Information Modelling (BIM), introduced in the seminal article "The Use of Computers Instead of Drawings in Building Design" (Eastman, 1975), has evolved into a crucial platform for detailed geometry exploration, interference resolution during early design stages and more, as evidenced by a study on modular housing design (He et al., 2021). Parametric design, integral to BIM, involves defining model families or element classes with rules controlling parameters. This approach facilitates the integration of AI in structural design by enabling computers to comprehend and control structural models. Unlike humans, computers rely on parametric objects, aligning with the principles of object-oriented programming. While the IFC schema (ISO 16739-1, 2018) is the standard for BIM data exchange, challenges persist, leading to inaccuracies and information loss (Sibenik and Kovacic, 2020). Recent solutions, such as invariant signatures (Wu et al., 2021a), and semantic enrichment via machine learning (Bloch and Sacks, 2018), aim to address these issues. Additionally, a shift towards graph-based representation in structural design, exemplified by studies like structural optimization using genetic algorithms and ANNs (Chang and Cheng, 2020), indicates a growing trend in data processing methods, offering potential for further exploration and improvement.

Graph Neural Networks in Engineering

The advantages of graphical representation in BIM models lead to proposal of GNNs for reliable and efficient structural design using AI. GNNs, introduced in 2009 (Scarselli et al., 2009), have demonstrated significant power in various domains, such as financial networks and molecular structures (Xu et al., 2022). Practical applications of GNNs in fields like internet traffic optimization (Bernárdez et al., 2023), stock classification (Xu and Zhang, 2023), and predicting microbe-disease associations (Jiang et al., 2023) are emerging. In the context of BIM models, which inherently possess complex relationships and dependencies, GNNs offer a natural solution for their non-Euclidean structure (Wang et al., 2021). Notably, the Graph Convolution Network (GCN) appears promising for classifying graphs representing BIM models, leveraging convolution to learn from node connections, akin to classical CNNs but tailored for graph structures (Wu et al., 2021b).

Gaps in Knowledge and Research Aims

Despite advances in graphical representation-based AI applications, those learning directly from human input remain absent. Moreover, the scarcity of solutions for AI-driven preliminary building design, emphasizes the need for research in this crucial domain (Huang and Fu, 2019). Employing the Design Science Research (DSR) methodology (vom Brocke et al., 2020), this study aims to show the feasibility of automated preliminary classification of structural plans using GNNs and engineers' experience. In other words, can AI effectively learn from the experience and expertise of human engineers through a combination of BIM models and engineer surveys? Answering this research question involves developing a tool that transforms basic BIM structural models into actionable information, enabling accurate classification based on a knowledge base of pre-labelled engineers' structural plans.

AIPDORCS Concept

Interviews were conducted with five senior structural engineers, each with at least a decade of experience, to establish existing preliminary design practices. The resulting process was meticulously mapped in a BPMN diagram, summarizing the preliminary design process from the architect's requirements until the completion of all preliminary design tasks by the structural engineer, as shown in Figure 1. Interestingly, the interplay of architectural requirements and structural solutions matches Maher's model of co-evolutionary design (Maher and Poon, 1996). This visualization underwent validation by eight senior structural engineers. The engineers

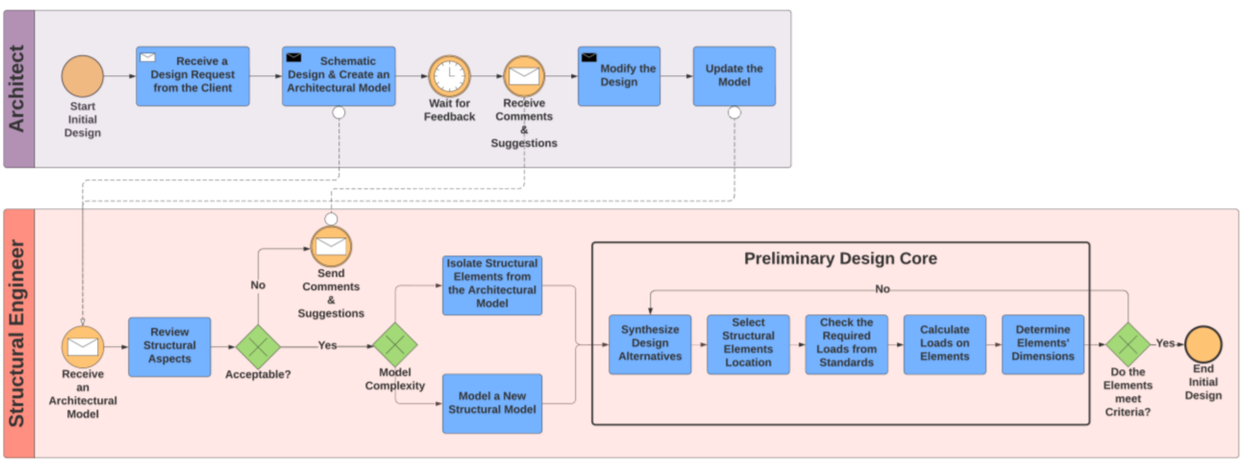


Figure 1: The process of Preliminary Structural Design, BPMN diagram.

emphasized the iterative nature of the process, likening it to a Taylor series, where each step contributes to convergence. They revealed diverse preferences in material selection, structural elements, and evaluation methods, showcasing the complexity of design choices. Notably, the interviews underscored the absence of a singular truth in structural design, with engineers prioritizing factors differently. A new concept emerged from this nuanced understanding, aiming to enhance the structural design process through AI by incorporating diverse insights and automating key aspects. We call this new concept "*AIPDORCS*", an acronym for *"Artificially Intelligent Preliminary Design of Reinforced Concrete Structures."* This concept comprises three stages, as shown in Figure 2. The first stage, *Data Collection & Preparation*, involves collecting and preparing data, converting structural plans into appropriate formats, addressing challenges in BIM model collection by utilizing 2D plans, and creating structured data tables for graph construction, all of which are essential for training a GNN model in the subsequent stage. The second stage, *Structural Plans Classification*, focuses on training the GNN model through engineer surveys, employing data tables and structural plans obtained from BIM models. Finally, the third stage, *Structural Design Generation*, delves into employing optimization algorithms like genetic algorithm or deep learning to evolve and refine solutions based on the previously trained classifier, iteratively until achieving optimal engineering solutions. The first two stages are the focus of this study.

AIPDORCS Development (Stages 1 and 2)

Structural Plans Collection

Initial attempts to gather the vast amount of data needed for training the supervised learning model, i.e. acquiring BIM models from large engineering firms, faced challenges due to legal and copyright issues. Requests for schematic plans also met with limited success. An alternative approach involved leveraging academic resources – final project plans from undergraduate students. These plans, encompassing diverse structures, were obtained with permissions from graduates, yielding 201 approved plans. The variation in quality within student designs proved advantageous, offering a spectrum for training the AI model. Subsequently, a database was constructed, filtering out non-relevant plans and ensuring anonymity by removing personal information. This database comprised around a hundred structural plans, predominantly in PDF format.

Plans to Models

Structural plans in PDF or DWG format were deemed unsuitable for training the GNN model based on engineers' experience, due to their lack of essential information. The chosen solution involved manual conversion of 2D plans, of typical residential and commercial buildings, into intelligent 3D BIM models using Autodesk Revit software (Autodesk, 2022). Recognizing the absence of flawless automatic conversion applications, a meticulous workflow was developed for manual modelling, addressing challenges of human error and maintaining a clean, uniform model database.

The modelling process started with loading a premade template into each project, resolving scale issues, and utilizing Autodesk Revit's "Snap" tool to identify and correct discrepancies. The modelling workflow included distinct steps for modelling walls, columns, beams, and slabs. Each element type was carefully modelled, considering specific rules such as ignoring windows and small openings, simplifying door modelling, addressing various complexities in columns, and modelling a separate slab for each space. Completing the models included crucial steps like "Join Geometry" automation for correct connectivity between structural elements, element numbering for referencing, manual checks by the researcher to ensure database quality, and sheet creation for later survey reference. Each model was assigned the following global parameters: building function, number of floors and gross height. Plans deemed unsuitable for the study, such as bridges or small houses, were excluded. In total, 48 projects were modelled.

Models to Data

In the transition from BIM models to usable data, Dynamo (Autodesk, 2020), a visual programming language for Autodesk Revit, was employed to automate the conversion process into structured CSV files. The feature engineering aspect of this stage focused on creating an equal number of features for each structural element. Geometric properties such as dimensions in three directions and volume were considered, avoiding redundancy. Noteworthy is the decision to exclude element location coordinates, deeming them unnecessary for preliminary design as the focus was on the geometry of elements and their connections.

Feature engineering aligned with accepted design practice, was primarily based on the calculation of the "limitation of slenderness." A uniform assumption of 30 MPa concrete with limestone aggregate was made for all elements. The Dynamo script categorized objects into groups of nodes, which facilitated the extraction of global parameters, adjusted gross height, and extracted geometric properties and connections for beams, walls, columns, and slabs. Unique challenges arose in feature engineering for columns, where complex cross-section shapes were converted to equivalent rectangular cross-sections. Similarly, for slabs, irregular shapes were converted to equivalent rectangular slabs based on the largest inscribed circle's diameter, and the distance of the slab's centroid from nearest vertical support elements.

In total, each CSV file extracted for each structural element type included three geometric dimensions, volume, automatically numbered element ID, and a list of connections to different elements, forming a foundation for subsequent analysis and classification stages.

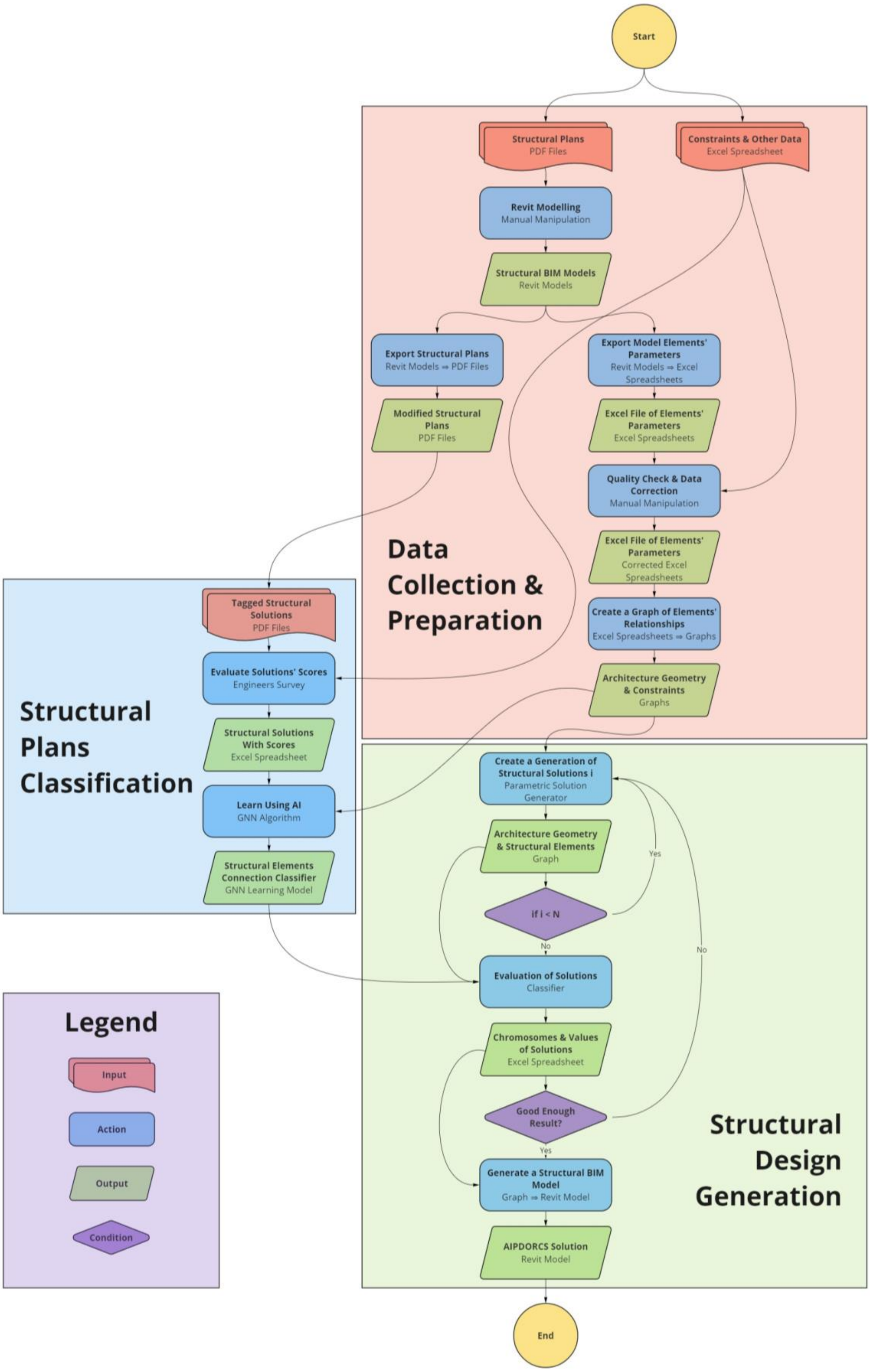


Figure 2: AIPDORCS Concept Diagram

Engineers' Challenge

The *"Engineers' Challenge"* online portal was compiled to facilitate labelling of structural plans for training the GNN model. Structural engineering experts were asked to use the portal to review plans presented to them at random from the set of 48 projects, and to give scores to the overall design and to annotate the individual design elements with their design critique (by selecting design faults from a list). The result was a comprehensive dataset for automated structural design with AI.

Compilation of the portal tool required identifying the critical aspects of preliminary structural design, designing an effective and efficient user experience, and considering engineering discretion while crafting questions. Given the unique nature of the survey – requiring feedback on 48 distinct structural plans – back-end project database management and user interface considerations were paramount. No available survey compilation tool could satisfy the requirements, and thus the portal was programmed from scratch utilizing JavaScript through the Wix API (Velo, 2006).

The portal's dialogs cover project data, element comments, and general comments with an overall score. The survey is available online at the AIPDORCS's website (Argaman, 2020). Note the addition of comment properties to each element, improving results by incorporating engineering aspects such as cross-sectional area, cost, simplicity of execution, and element slenderness. Users could provide feedback on specific elements, offering a comprehensive view of the structural plans. Moreover, the general comments allow users to review the global plan as a whole, considering comments such as 'inadequate vertical support', 'uneconomical design alternative' and 'implementation difficulty'.

The survey generated valuable responses, with 112 feedbacks obtained, including scores for all 48 models. The results and analysis of the survey contributed significantly to the subsequent stages of the AIPDORCS process.

Data to Graphs

At this stage, we utilized Python routines to construct a graph database for training the GNN model using CSV files from Revit models and *Engineers' Challenge* responses as input. To form graphs from the CSV files, preliminary data consolidation and processing were required. Each structural element from each BIM model is represented as a graph node containing features derived from exported CSV files. The natural and error-free conversion involved creating a file listing all elements as graph nodes with their respective features, incorporating both geometric and survey properties.

Creating graph edges, representing connections between nodes, was a more intricate process. The edges capture the relationships between structural elements. Given that the graph is homogeneous, i.e. all elements shared the same number of properties and dimensions, we could represent all structural elements with a single node type, named "Structural Element", as shown in Figure 3. This figure also illustrates the graph creation process from BIM models, focusing on a segment of *Project 001* from the AIPDORCS database.

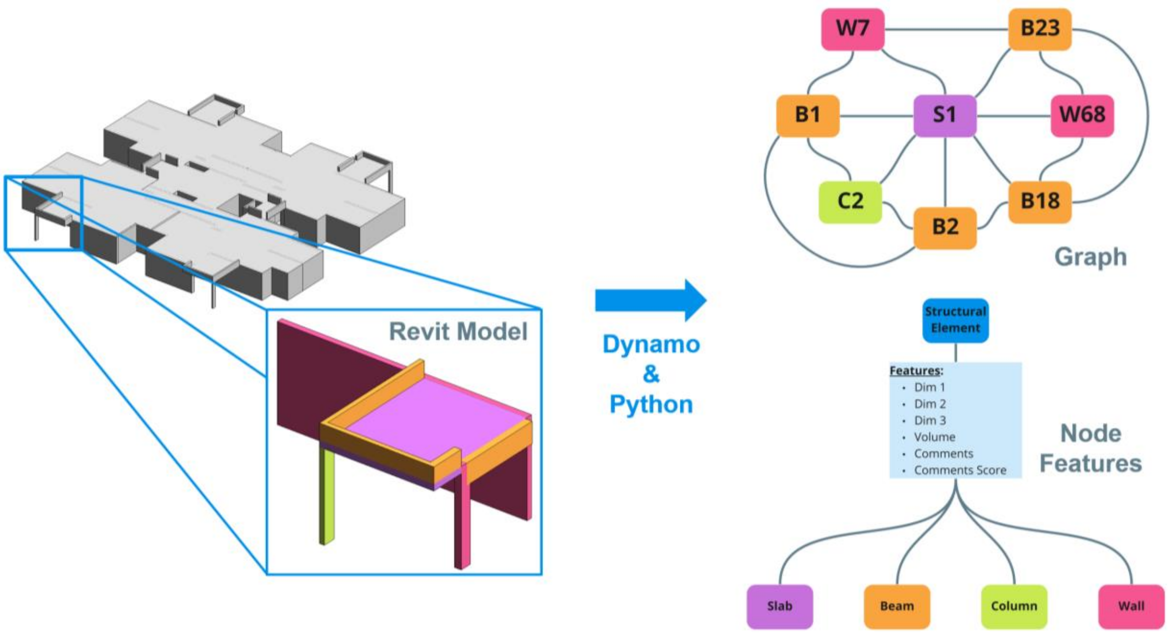


Figure 3: An illustration of the graph creation process, based on Project 001's BIM model

The Python code for this stage is organized into three files:

1. *data2graph*: The main file, orchestrating the process of building a graph database from CSV files.
2. *classes*: Defines classes for structural elements, graph nodes, and graph edges, adhering to object-oriented programming standards.
3. *websiteData*: Manages functions related to extracting and processing information from the *Engineers' Challenge* CSV file.

The code follows object-oriented programming standards, ensuring readability and maintainability for future research, and appears on AIPDORCS GitHub repository (Argaman, 2022). It uses standard Python libraries for data science, with a focus on DGL (Deep Graph Library, 2018) for creating graphs and training the GNN model.

Upon running the code, graphs are generated, exemplified by Figure 4, which shows the graph extracted for *Project 001* from the BIM models database. This visual representation exhibits element numbering and their geometric connections, providing a means for validation against the original structural plans.

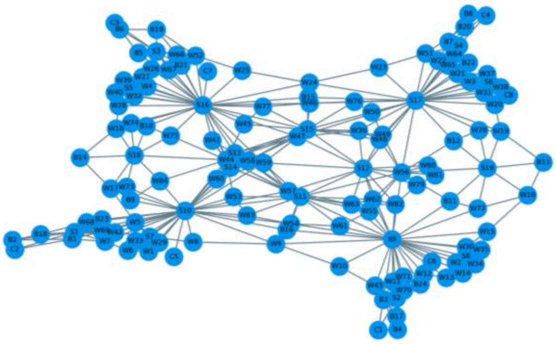


Figure 4: Graph of Project 001's BIM model

Graph Neural Network

Two experiments were conducted utilizing GNNs, or more specifically GCNs, for classifying structural design plans with a pass/fail grade according to a certain threshold. Experiment 1 was based on geometric model graphs, where 48 graphs correspond to 48 projects, with graph nodes containing only geometric features (four features: Dim 1, Dim 2, Dim 3, and volume). Figure 5 shows the ANN structure of experiment 1.

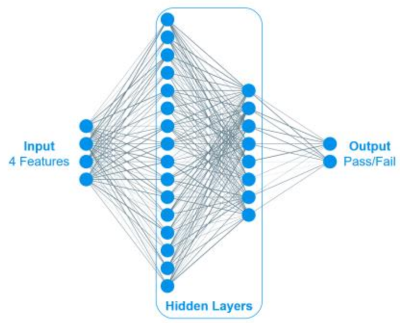


Figure 5: Experiment 1 artificial neural network structure

Experiment 2 is based on engineering feedback graphs, where 106 graphs correspond to 106 valid portal responses, with graph nodes containing all 14 features (geometric properties, survey comment properties, and points subtracted for comments). Figure 6 illustrates the ANN structure of Experiment 2.

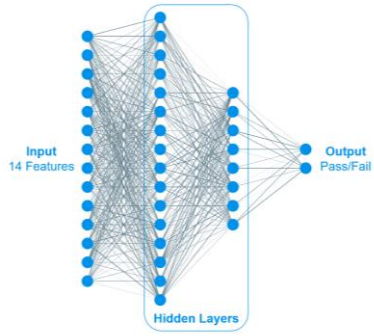


Figure 6: Experiment 2 artificial neural network structure

The architecture of the networks for both experiments is illustrated in Figures 5 and 6 respectively. The networks are structurally identical, with variations in the number of features in the input layer based on the experiment. Careful experimentation and evaluation are essential in determining the most suitable hyperparameter values for achieving the best results, and avoiding overfitting. In both experiments, the output labels were the *Engineers' Challenge* overall scores, converted to a binary value by the average score threshold. The data was split to an 80% training set and a 20% test set, and the following hyperparameters were selected by testing and tuning: number of hidden layers: 2, hidden dimensions: 16, activation function: ReLU, learning rate: 0.005, and number of epochs: 200.

Results

Engineers' Challenge

A total of 112 responses were gathered through the portal, six of which were deemed invalid and filtered out. The analysis of the survey results revealed insights into the engineers' perspectives and highlighted their attention to specific elements, as well as the distribution of feedback across different categories.

When comments were given on specific elements, the majority of engineers provided an average of three comments on each, demonstrating a conscientious approach to the task. Additionally, 58 written comments, categorized as "Additional Comments", required manual conversion to relevant element features, showcasing the depth of insights shared by engineers. The distribution of element comments illustrates that issues related to the size of the element's cross-section were the most frequently noted, emphasizing the significance of this aspect in the eyes of experienced engineers.

Furthermore, analysis of overall scores revealed a predominantly positive trend, with most models receiving scores above 70. Outliers with lower scores were investigated, leading to the identification and correction of errors in feedback or potential typing mistakes. The survey results also prompted an examination of the relationship between the extent of engineers' years of experience and the scores they assigned. The findings suggested a deviation in scores based on experience, with less experienced engineers tending to give higher scores. However, the scores did not exhibit a consistent trend, emphasizing the need for further data collection to draw more conclusive insights. Overall, the *Engineers' Challenge* portal provided a valuable dataset for understanding expert opinions on structural plans, laying the groundwork for subsequent experiments with GNNs.

Experiments

The outcomes of the GNN experiments exhibited variability across runs, influenced by both the inherent randomness in the code and the model's instability arising from limited data in the database. Focusing on the most promising runs, the analysis employed standard data science performance metrics, derived from a binary classification's confusion matrix: Accuracy, precision, recall, and F1-score.

Table 1: Experiments performance comparison

	Accuracy	Precision	Recall	F1 Score
Experiment 1	60.0%	57.1%	80.0%	66.7%
Experiment 2	68.2%	75.0%	88.2%	81.1%

The results, detailed in Table 1, highlight Experiment 2's superior performance compared to Experiment 1 across all metrics. The larger dataset and richer feature set in Experiment 2 contributed to its enhanced precision, recall, F1-score, and overall accuracy. These findings underscore the potential of GNNs in structural design classification, suggesting that leveraging extensive and diverse datasets can significantly enhance the performance of GNN-based classifiers, offering promising prospects for their application in preliminary structural design assessments.

Discussion

The AIPDORCS concept introduces an innovative approach, setting it apart from existing works in the realm of structural engineering. Integrating AI with human expertise, this concept redefines the preliminary structural design process, offering a unique solution. In the modelling phase, which produced 48 BIM models, opportunities for improvement were identified. They included refining geometric connections and expanding the features of structural elements, which would yield a heterogeneous graph. Unequal scoring patterns revealed in survey responses necessitate further considerations, like re-scoring based on comments and normalizing scores by engineers' experience clusters. The structural plan classification experiments yielded promising results, particularly with the utilization of engineering feedback graphs, showcasing the potential of leveraging survey data for precise structural design classification. However, employing 5-fold cross-validation emerges as a key recommendation to ensure dataset representativeness, detect overfitting, and bolster model generalization.

The technical and research limitations present crucial aspects to address. Technical limitations, such as support for specific building patterns and materials, indicate areas for improvement and expansion. Research limitations, including a small dataset and a limited number of feedbacks, underscore challenges related to generalizability and robustness. Issues like varying scales of overall scores and limited time and resources further emphasize the complexities and constraints of the study. While the two initial experiments offer valuable insights, the need for additional experiments on a larger and more diverse dataset is recognized. Notably, the use of comments as features in Experiment 2 prompted consideration of the applicability of this approach in a real-world AI tool. These limitations shape the context of the study and offer directions for future research, emphasizing the need for a comprehensive and conclusive evaluation of the proposed approach for scoring structural plans using GNN.

Conclusion

Through the execution of two experiments utilizing GNNs, this study established the basic feasibility of automating structural plans classification, underscoring the potential to enhance accuracy and efficiency in the design process. Nevertheless, further research work is required to reliably answer the research question. Acknowledging concerns among engineers about AI's role in the creative phase of design, this study emphasizes a collaborative approach, asserting that AI can augment human capabilities without replacing them. The concluding sentiment draws inspiration from Charles Darwin's principle of adaptability, encouraging engineers to embrace AI's transformative potential in structural design, foreseeing a future where the integration of AI not only optimizes processes but also fosters innovation.

This study contributes to the field of civil engineering by proposing the AIPDORCS concept, innovatively transforming BIM models into GNN-ready graphs. Additional contributions encompass the development of a robust database, an automated model-to-graph solution, and a scoring program for structural plans. These innovations offer valuable resources, time-saving tools, and educational applications, positioning the research at the forefront of AI-assisted structural design. To conclude, this research marks a valuable step toward a future where AI and human expertise harmoniously shape the landscape of structural engineering.

References

Abecasis, S., and R. Amar. (2019). "FIRMUS | Mitigate risks early. Bid with confidence." Accessed April 18, 2023. https://www.firmus.ai.

Afzal, M., Y. Liu, J. C. P. Cheng, and V. J. L. Gan. (2020). "Reinforced concrete structural design optimization: A critical review." *J. Clean. Prod.*, 260: 120623. https://doi.org/10.1016/j.jclepro.2020.120623.

Ampanavos, S., M. Nourbakhsh, and C.-Y. Cheng. (2022). "Structural Design Recommendations in the Early Design Phase Using Machine Learning." Comput.-Aided Archit. Des. Des. Imperatives Future Now, Communications in Computer and Information Science, D. Gerber, E. Pantazis, B. Bogosian, A. Nahmad, and C. Miltiadis, eds., 190–202. Singapore: Springer Singapore.

Anwar, N., L. Khatiwada, T. Aung, and J. Sy. (2015). "Preliminary Design of Tall Buildings Using Artificial Neural Networks."

Argaman, A. D. (2020). "AIPDORCS | Research | Israel." AIPDORCS. Accessed May 21, 2023. https://alonf4.wixsite.com/aipdorcs.

Argaman, A. D. (2022). "Alonf4/AIPDORCS." Accessed May 21, 2023. https://github.com/Alonf4/AIPDORCS.

As, I., S. Pal, and P. Basu. (2018). "Artificial intelligence in architecture: Generating conceptual design via deep learning." Int. J. Archit. Comput., 16 (4): 306–327. SAGE Publications. https://doi.org/10.1177/1478077118800982.

Autodesk. (2020). "Dynamo." Accessed May 21, 2023. https://dynamobim.org/.

Autodesk. (2022). "Autodesk Revit Software." Accessed May 21, 2023. https://www.autodesk.com/products/revit/overview.

Bernárdez, G., J. Suárez-Varela, A. López, X. Shi, S. Xiao, X. Cheng, P. Barlet-Ros, and A. Cabellos-Aparicio. (2023). "MAGNNETO: A Graph Neural Network-Based Multi-Agent System for Traffic Engineering." IEEE Trans. Cogn. Commun. Netw., 9 (2): 494–506. https://doi.org/10.1109/TCCN.2023.3235719.

Bloch, T., and R. Sacks. (2018). "Comparing machine learning and rule-based inferencing for semantic enrichment of BIM models." Autom. Constr., 91: 256–272. Elsevier B.V. https://doi.org/10.1016/j.autcon.2018.03.018.

Boonstra, S., K. van der Blom, H. Hofmeyer, and M. T. M. Emmerich. (2020). "Conceptual structural system layouts via design response grammars and evolutionary algorithms." Autom. Constr., 116: 103009. https://doi.org/10.1016/j.autcon.2019.103009.

vom Brocke, J., A. Hevner, and A. Maedche (Eds.). 2020. Design Science Research. Cases. Progress in IS. Cham: Springer International Publishing.

Caprani, D. C. (2009). "Preliminary Design of Building Structures - 3rd Year Structural Engineering." 221.

Chang, K.-H., and C.-Y. Cheng. (2020). "Learning to simulate and design for structural engineering." arXiv.org.

Charalampakis, A. E., and V. K. Papanikolaou. (2021). "Machine learning design of R/C columns." Eng. Struct., 226: 111412. https://doi.org/10.1016/j.engstruct.2020.111412.

Deep Graph Library. (2018). "DGL." Accessed May 21, 2023. https://www.dgl.ai/.

Eastman, C. (1975). "The Use of Computers Instead of Drawings in Building Design." AIA J., 63.

Elhegazy, H., D. Chakraborty, H. Elzarka, A. M. Ebid, I. M. Mahdi, S. Y. Aboul Haggag, and I. Abdel Rashid. (2022). "Artificial Intelligence for Developing Accurate Preliminary Cost Estimates for Composite Flooring Systems of Multi-Storey Buildings." J. Asian Archit. Build. Eng., 21 (1): 120–132. Taylor & Francis. https://doi.org/10.1080/13467581.2020.1838288.

Hayashi, K., and M. Ohsaki. (2020). "Reinforcement learning for optimum design of a plane frame under static loads." Eng. Comput. https://doi.org/10.1007/s00366-019-00926-7.

He, R., M. Li, V. J. L. Gan, and J. Ma. (2021). "BIM-enabled computerized design and digital fabrication of industrialized buildings: A case study." J. Clean. Prod., 278: 123505. https://doi.org/10.1016/j.jclepro.2020.123505.

Huang, C.-H., and S.-H. Hsieh. (2020). "Predicting BIM labor cost with random forest and simple linear regression." Autom. Constr., 118: 103280. https://doi.org/10.1016/j.autcon.2020.103280.

Huang, Y., and J. Fu. (2019). "Review on Application of Artificial Intelligence in Civil Engineering." Comput. Model. Eng. Sci., 121 (3): 845–875. https://doi.org/10.32604/cmes.2019.07653.

ISO 16739-1. (2018). "Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries — Part 1: Data schema." ISO. https://www.iso.org/standard/70303.html.

Jiang, C., M. Tang, S. Jin, W. Huang, and X. Liu. (2023). "KGNMDA: A Knowledge Graph Neural Network Method for Predicting Microbe-Disease Associations." IEEE/ACM Trans. Comput. Biol. Bioinform., 20 (2): 1147–1155. https://doi.org/10.1109/TCBB.2022.3184362.

Jozi, A., M. Moradi, and A. B. Gerami. (2019). "Structure Plus." Accessed April 18, 2023. https://www.structure.plus/.

Katoch, S., S. S. Chauhan, and V. Kumar. (2021). "A review on genetic algorithm: past, present, and future." Multimed. Tools Appl., 80 (5): 8091–8126. https://doi.org/10.1007/s11042-020-10139-6.

Laignel, G., N. Pozin, X. Geffrier, L. Delevaux, F. Brun, and B. Dolla. (2021). "Floor plan generation through a mixed constraint programming-genetic optimization approach." Autom. Constr., 123: 103491. https://doi.org/10.1016/j.autcon.2020.103491.

Liao, W., X. Lu, Y. Huang, Z. Zheng, and Y. Lin. (2021). "Automated structural design of shear wall residential buildings using generative adversarial networks." Autom. Constr., 132: 103931. https://doi.org/10.1016/j.autcon.2021.103931.

Maher, M. L., and J. Poon. (1996). "Modeling Design Exploration as Co-Evolution." Computer aided Civil Eng, 11 (3): 195–209. https://doi.org/10.1111/j.1467-8667.1996.tb00323.x.

Medsker, L. R. (1995). Hybrid Intelligent Systems. Boston, MA: Springer US.

Rychener, M. D. (1988). Expert Systems for Engineering Design. (M. D. Rychener, ed.). Academic Press.

Samuel, A., and J. Weir. (1999). Introduction to Engineering Design. Oxford: Butterworth-Heinemann.

Scarselli, F., M. Gori, A. C. Tsoi, M. Hagenbuchner, and G. Monfardini. (2009). "The Graph Neural Network Model." IEEE Trans. Neural Netw., 20 (1): 61–80. https://doi.org/10.1109/TNN.2008.2005605.

Selvaraj, M., K. Borowska, and B. Neuman. (2019). "Daisy- Design AI Systems."

Sibenik, G., and I. Kovacic. (2020). "Assessment of model-based data exchange between architectural design and structural analysis." J. Build. Eng., 32: 101589. https://doi.org/10.1016/j.jobe.2020.101589.

Velo. (2006). "Velo by Wix: Full-Stack Web Dev | Build Professional Web Apps." Accessed May 21, 2023. https://www.wix.com/velo.

Wang, X.-Y., Y. Yang, and K. Zhang. (2018). "Customization and generation of floor plans based on graph transformations." Autom. Constr., 94: 405–416. https://doi.org/10.1016/j.autcon.2018.07.017.

Wang, Z., R. Sacks, and T. Yeung. (2021). "Exploring graph neural networks for semantic enrichment: Room type classification." Autom. Constr., 104039. https://doi.org/10.1016/j.autcon.2021.104039.

Wu, J., H. L. Sadraddin, R. Ren, J. Zhang, and X. Shao. (2021a). "Invariant Signatures of Architecture, Engineering, and Construction Objects to Support BIM Interoperability between Architectural Design and Structural Analysis." J. Constr. Eng. Manag., 147 (1): 04020148. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001943.

Wu, Z., S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu. (2021b). "A Comprehensive Survey on Graph Neural Networks." IEEE Trans. Neural Netw. Learn. Syst., 32 (1): 4–24. https://doi.org/10.1109/TNNLS.2020.2978386.

Xu, K., W. Hu, J. Leskovec, and S. Jegelka. (2022). "How Powerful are Graph Neural Networks?"

Xu, Y., and Y. Zhang. (2023). "Enhancement Economic System Based-Graph Neural Network in Stock Classification." IEEE Access, 11: 17956–17967. https://doi.org/10.1109/ACCESS.2023.3246525.

Zheng, H., V. Moosavi, and M. Akbarzadeh. (2020). "Machine learning assisted evaluations in structural design and construction." Autom. Constr., 119: 103346. https://doi.org/10.1016/j.autcon.2020.103346.

Zhou, C., B. Tang, L. Ding, P. Sekula, Y. Zhou, and Z. Zhang. (2020). "Design and automated assembly of Planetary LEGO Brick for lunar in-situ construction." Autom. Constr., 118: 103282. https://doi.org/10.1016/j.autcon.2020.103282.

MULTI-STAKEHOLDER INFORMATION REQUIREMENTS TO SUPPORT LIFE CYCLE ASSET MANAGEMENT

Conor Shaw[1], Flávia de Andrade Pereira[1,2] James O'Donnell[1]
[1]School of Mechanical and Materials Engineering and Energy Institute, University College Dublin, Ireland
[2]University College London, United Kingdom

Abstract

The current IT paradigm in the built environment is misaligned with sustainability policy. Previously proposed ontologies for Asset Management, such as life cycle analysis, lack complete concepts to cater to a wide stakeholder group. This paper describes a more comprehensive asset management software landscape. It details the initial development steps using the Linked Open Terms methodology including requirements gathering and ontology conceptualisation. A modular ontology landscape is proposed including top-level, domain-wide concepts and modular, application-specific concepts; a scenario suited to the particularly broad domain needs.The work fosters consensus in the domain and we propose alignment/extension with the existing RealEstateCore standard.

Background

The future of asset information management

In line with sectoral goals to reduce emissions in the built environment, various policy mechanisms exist. These include emissions target-setting, as well as a trajectory towards mandatory reporting on how efficiently assets and organisations use resources throughout their lifetime. Examples include the recent Corporate Sustainability Reporting Directive (European Commission, 2023) aimed at organisations generally, as well as the EU's ambitious Level(s) framework (European Commission, 2022) focused specifically on the built environment. Additionally, within financial markets, the value of a building is increased in investors' eyes by obtaining environmental certifications (such as LEED, BREEAM etc.) which evaluate lifetime performance around various metrics. This phenomenon reflects a shift in values where high-consumption, environmentally unfriendly assets become a business risk in a decarbonised future (Dumrose and Höck, 2023). In this context, Asset Managers (AMs) play an important role as the discipline tasked with operating and maintaining built assets, particularly throughout their in-use phase where the majority resources are consumed (Geekiyanage and Ramachandra, 2018). As a result, AMs have a significant influence on the sustainability and resulting value of buildings and infrastructure.

AMs rely on various indicators to make value-based decisions and assess the operational performance of facilities, evaluating decisions by balancing cost, risk and performance (Fang et al., 2022). This approach aligns conceptually with the Life Cycle Sustainability Assessment (LCSA) methodology which evaluates an asset through a holistic financial-social-environmental lens (Figure 1). LCSA provides a comprehensive view of an asset's resource consumption over its entire life, enabling evidence-based decision making supported by long-term value rather than solely focusing on short-term, or initial investment costs (Kehily and Underwood, 2017). This approach contrasts with the typical practice which tends to prioritise immediate investment expenses, often neglecting the broader implications of decisions over time (Grzyl et al., 2017), a tendency which, according to Collier (2018), is part of a much wider phenomenon of short-term-ism in the financialised built environment and current workings of capitalism. To conduct LCSA analysis using IT systems, AMs require structured information about facilities; however, this information is frequently unavailable due to the widely-recognised inadequacy of current information management practices (Gao and Pishdad-Bozorgi, 2019). Additionally, the software required to perform these insightful analytics is often unsuitable. Existing LCSA applications tend to be either too costly, lack the flexibility to meet specific stakeholder needs, or simply be unavailable altogether (Shaw et al., 2024). As a result, AMs currently rely on laborious, error-prone, ad-hoc analyses to support their decisions. Consequently, the potential benefits of LCSA are not being widely realised.

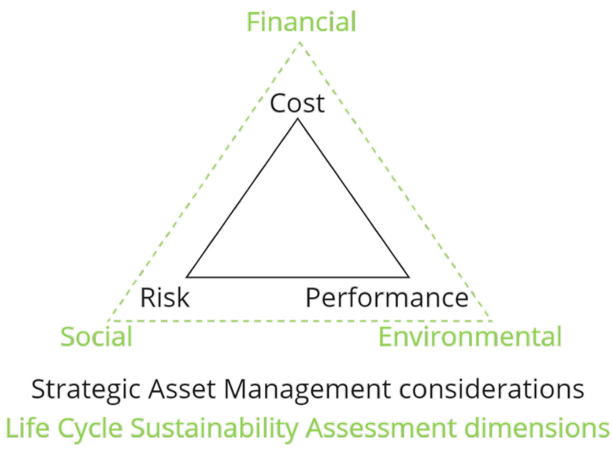


Figure 1: Conceptual overlap between AM decision-making and the holistic LCSA methodology

Ontologies can be used to organise domain knowledge in a formalised manner within IT system, facilitating information management and automation through inheritance and logical reasoning. Furthermore, ontologies can impart consensus semantic meaning to data shared over the web, referred to as Semantic Web Technology (SWT), a direction suggested by experts in the field as a necessary next step in LCSA research given the data-intensive nature of the practice (Salvado et al., 2021). A further advantage of information management using SWT is the potential for extensibility for specialised stakeholder needs. Given the expansive scope of AM activities, this is a logical requirement of a future-oriented IT landscape.

Previous efforts in AM-related ontology development

ISO 15978 (European Committee for Standardization (CEN), 2011) establishes a widely-agreed taxonomy of asset lifecycle phases. It serves as a foundational framework for much of the work in information management to support LCSA activities, involving the classification of costs and resources per phase. In accordance with the standard numerous applications and ontologies have been developed catering to various LCSA-related use cases (Lu et al., 2021), of which relevant works are now discussed.

In terms of envisioning software landscape, recent work by Sobhkhiz et al. (2021) emphasises the necessity of leveraging SWTs, enabled by ontologies, to address the substantial data handling challenges inherent in LCSA. The authors demonstrate the efficiency gains achieved over traditional relational database methods. Other studies, such as Wilde et al. (2022) and Ghose et al. (2022), focus on establishing a foundational, or top-level ontology, to support LCSA activities. These studies enable stock-level analysis, and consequently, the outcomes of these efforts do not support AMs in operational-level decision-making. In the pursuit of multi-scale analysis and aggregation, a recent work of significance is the SLiCE data model by Röck et al. (2024) which supports analysis from individual materials and parts, through building-level and up to the stock-level. Given the broad scope of AM functions, it is reasonable that achieving consensus on a universally shared ontology for the domain remains a challenge; however, the above initiatives are clearly progressing in this direction.

Another promising development are consolidation and alignment activities between standards communities. Xie et al. (2022) propose alignment between the BOT and BRICK ontologies. Their FDM ontology establishes a top-level asset information management concept for data integration in an effort towards a Digital Twin paradigm in future. Hammar et al. (2019) pioneered the development of a now widely adopted ontology tailored for asset owners, with a focus on concepts relevant to smart building applications and tenancy/leasing. Their collaborative effort, backed by a consortium of major asset owners in Sweden, aimed to establish shared domain use cases and describe these in the RealEstateCore ontology. The standard boasts a large user base and actively engages in alignment activities with other leading domain ontologies like Brick (for building automation systems) and BOT (for topological building description),

A number of efforts are of note which address specific analytical AM use cases. On the maintenance side, Katsumi et al. (2022) utilise the Ontology Requirements Specification Document (OSRD) development methodology to identify user requirements for a common AM ontology, drawing from a specific water treatment plant case study. Though focused on maintenance work orders, the study demonstrates a related ambition towards consensus that could potentially align with our own efforts at a later stage. Of particular relevance to our specific objectives of conducting multi-criteria LSCA analysis is the research by Gao et al. (2020) which proposes the LCCA Ontology, tailored for generating machine-learning-based financial life cycle cost predictions at the building level. Though publicly available, this ontology may be overly specific to its application context and not readily adaptable to the granular level of detail required for AM decision making.

Synthesis and research method

This body of research showcases a promising, future-oriented evolution toward effective asset information sharing on the web. However, none of the existing studies have comprehensively addressed multi-stakeholder needs for implementing LCSA-informed AM applications, with a particular gap in operational phase capabilities. A harmonisation of these efforts will be crucial for supporting decision-making and reporting processes in future given the broad expanse of AM activities, and it is promising to see a recognition of this in the active alignment efforts in the community. To this end, the Linked Open Terms (LOT) ontology development methodology (Poveda-Villalón et al., 2022) provides a structured approach for gathering domain requirements for a future ontology landscape. This paper describes our initial steps in this direction, following the research activities as illustrated in Figure 2.

Information requirements to support life cycle asset information management

The following information requirements derive from a combination of research activities. These include extensive reflection on the domain literature, as described in the previous section, supplementary interviews and discussions with AM practitioners and ontology developers, and from our previous practitioner-based research in AM decision-support system development (Shaw et al., 2024). In that study we developed a software prototype and reference architecture to support AMs with a number of fundamental use cases relating to financial life cycle cost analysis. The remainder of this paper builds upon this domain insight, and illustrates our perspective on future AM system requirements to support life cycle asset management more broadly, including a concept for the arrangement of top-level and modular application ontologies.

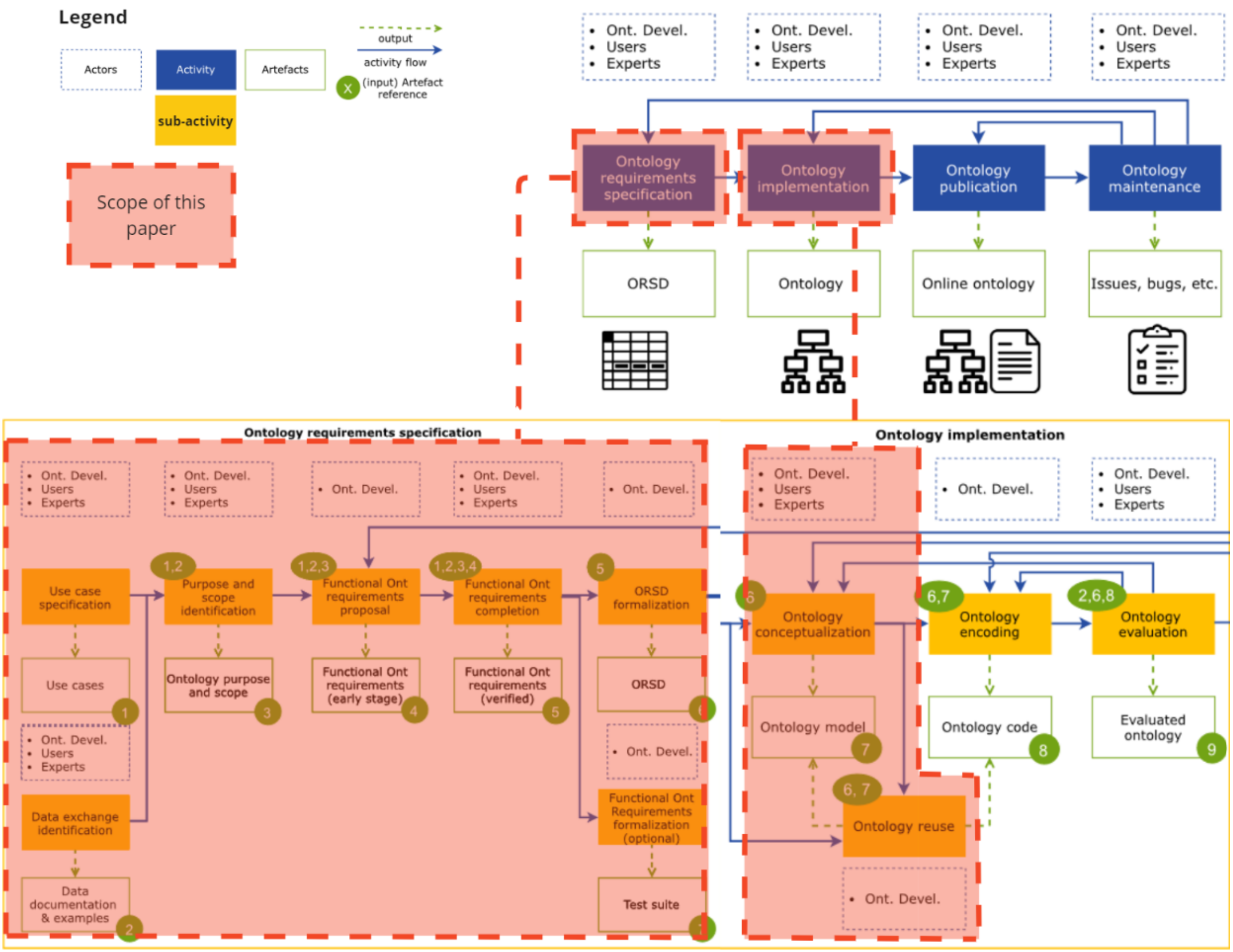


Figure 2: Paper scope - adapted from the generic LOT ontology development methodology, described by Poveda-Villalón et al. (2022)

Concept and use case specification

We envision a democratised IT landscape, and a departure from the current inflexible, proprietary software paradigm. This web-based system should be highly customisable (or *extensible*) to specific user needs, but with a foundational, shared top-level ontology which formalises domain knowledge and logic in line with international standards. Due to the data-intensive nature of the LCSA activities, a main purpose of the system is be to integrate inputs over the web, via either data warehousing or mediation, as set out by Xie et al. (2022). In this way, multiple stakeholders from diverse disciplines could make use of the structured information to suit their specialised analytics needs using additional modular, application-specific ontologies. The modularity and Microservices concepts are described by Pritoni et al. (2021) and Werbrouck et al. (2023), respectively, and our system architecture concept is presented in Figure 3. For the purposes of this study, and in line with the LOT methodology, the following use cases are suitably representative of the multi-stakeholder needs for the asset life cycle information system.

- ***Use Case 1:*** Life cycle analysis at various levels of granularity and across various indicators (ie. finance, energy, condition)
- ***Use Case 2:*** Decision-alternative analysis (or *optioneering*)
- ***Use Case 3:*** Performance gap analysis (against a baseline - supporting performance-based contracting)
- ***Use Case 4:*** Reporting in line with policy frameworks (the EU's Level(s) methodology, for example)

Non-functional requirements

- ***Purpose:*** The system architecture concept aims to define fundamental information categories for life asset management and analysis applications, offering standardised terminology and relationships to integrate information from diverse IT systems and sources.
- ***Scope:*** Both the top-level and modular application-specific ontologies are widely generalisable, ensured by their being based on international standards and remaining relatively abstract and extensible.

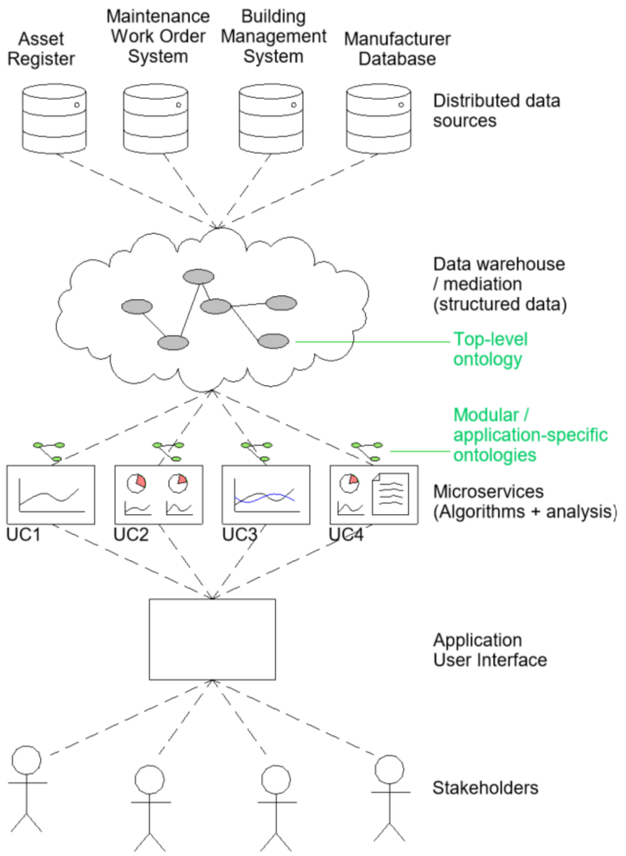


Figure 3: Concept sketch of the proposed system architecture with top-level and modular ontologies identified.

- ***Intended End-Users:*** Asset managers and owners seeking insight on decisions relating to an individual asset or a broader portfolio perspective, and to support performance contract administration. Designers comparing alternative decisions related to building, refurbishment, or maintenance projects. Regulators seeking to audit asset performance in accordance with reporting frameworks.

With the objective of supporting a variety of stakeholders, the proposed system should be ***accessible over the web*** and have an intuitive interface and functionality. Information to support LCSA is typically distributed across multiple, often unintegrated, IT systems, and is a well-established barrier to wider use. Future AM systems should therefore ***integrate information*** using standardised terminology and formalise domain knowledge through logic. Because of the wide variety of contexts to conduct LCC analysis, future systems require the flexibility to ***aggregate information*** consistently, so that assets can be looked at individually or on an aggregated basis across a portfolio, the SLiCE data model (Röck et al., 2024) providing a solid foundation for this. Again, given the wide variety of use cases, a major frustration for practitioners in the current IT landscape is being over-constrained by software. Therefore future AM systems must maintain the ***extensibility to add cost/resource items*** while maintaining a consistent top-down approach based on agreed taxonomies and classification systems. Needless to say, the systems must enable a user to ***conduct life cycle analysis*** based on standardised methodologies, namely ISO 15686-5 (International Organization for Standardization, 2017). A variety a ***'views' of the analysis*** should be possible to visualise including the temporal nature (discount to present value, future value, payback period), and visual nature (graphs, tabular). Due to the data-intensive nature of LCSA activities, an important consideration is validation and trust around the input data. Therefore, future systems need to ***demonstrate data quality*** in terms of provenance and completeness, the Shapes Constraint Language (SHACL) (W3C, 2017) providing a promising technical direction in this regard.

Functional requirements

This sub-section contains excerpts from the Ontology Requirements Specification Document (ORSD). The LOT methodology allows for the use of various requirements gathering approaches. Both Tabular and Competency Question (CQ)-based methods are use in this study. Tabular information includes a specification of Concepts, Relations and Attributes to be encoded in the semantics and logic of the ontology, as well as typical data types, units of measurement and cardinality. An excerpt is shown in Figure 4, denoting between top-level and application-specific aspects.

Since the concept has yet to be expressed in a formal ontology language, we provide informal CQs the proposed system would be expected to answer. CQs support validation during the encoding and testing activities. This list progresses from returning simple attribute values using the top-level ontology to retrieve distributed data, to increasingly complex queries requiring application-specific analytics and modular ontologies.

Top-level:

- CQ1 - What is the [Lifetime] of Asset with [AssetID]?
- CQ2 - Which Assets [List] have a [ResidualValue] > '0'?
- CQ3 - How many Assets [enumerate] have [Condition] in range[1-3]?
- CQ4 - What is the average [AnnualEnergyCost] of Assets with [AssetType]?

Application specific:

- CQ5 - What is the StudyPeriod for the AnalysisEvent [Scenario1]? ***Use Case 1***
- CQ6 - What is the [LifeCycleCost] for SUM[Assets IN AssetResister]? ***Use Case 1***
- CQ7 - What is the percentage breakdown between LifeCyclePhases for Asset [AssetID]? ***Use Case 1***
- CQ8 - Which is the most expensive year given [AnalysisEvent] and which are the Assets being renovated or replaced in that year? ***Use Case 1***
- CQ9 - Which Asset replacement option [list] has the lowest LifeCycleCost over StudyPeriod? ***Use Case 2***

Concepts

Concept	Description
Asset	A very general description of any object (example AHU), or collection of objects (building). Core concept in the ontology.
Cost	Top-level concept, subclasses include various costs (InitialCost, EnergyCost, ResidualCost)
StudyPeriod	Defines the scope of analysis in time
EconomicFactor	Top-level concept, subclasses include Discount and Inflation rates
AnalysisEvent	Specifies a particuar analysis activity

Relations

Concept	Relation	Target	Max cardinality
Asset	hasCost	Cost	1:n
Asset	hasLifetime	Lifetime	1
Asset	hasInstallationDate	InstallationDate	1
Asset	hasCondition	Condition	1
Asset	hasEnergyUse	EnergyUse	1
AnalysisEvent	hasStudyPeriod	StudyPeriod	1
AnalysisEvent	hasEconomicFactor	DiscountRate	1
AnalysisEvent	includesAsset	Asset	1:n

Application specific concepts

Attributes

Concept	Attribute	Description	Value type expected	Max cardinality	Unit
Asset	AssetID	Unique Identifier	String	1	description
Asset	AssetType	Type according to classification system	String	1	description
Asset	InstallationDate	Date of installation	Datetime	1	Datetime
Asset	InitialCost	Capital cost	Float	1	Monetary
Asset	Lifetime	Duration of usable life before requiring decommission/replacement	Int.	1	Years
Asset	ReplacementCost	Future cost of replacement (like for like)	Float	1	Monetary
Asset	RenovationCost	Cost of renovation to prolong lifetime	Float	1	Monetary
Asset	RenovationFrequencey	Number of years before renovation needed to prolong usable life	Int.	1	Years
Asset	AnnualEnergyCost	Cost of energy required to operate the asset	Float	1	Monetary
Asset	AnnualMaintenanceCost	Cost of regular maintenance operate the asset	Float	1	Monetary
Asset	ResidualValue	Positive value of asset at end of usable life (scrappage, upcycling, recycle parts)	Float	1	Monetary
Asset	Condition	Licart rating of an assets operating / quality condition (organisation spepcific, surveyed)	Int.	1	Assessed rating
Asset	AnnualEnergyUse	Sum of energy used over one year	Float	dynamic	kWh
EconomicFactor	InflationRate	Rate of annual inflation	Float	1	%
EconomicFactor	DiscountRate	Rate of discounting future value to present value	Float	1	%

Figure 4: Tabular requirements for modelling financial life cycle cost analysis, gathered per the LOT methodology. Denotes between top-level concepts and those more appropriately stored within modular application ontologies.

- CQ10 - What is the StudyPeriod which makes Option1 outweigh Option2 as the cheaper LifeCycleCost (payback period)? ***Use Case 2***
- CQ11 - What is the percentage reduction in [LifeCycleCost] over [StudyPeriod] for Asset [List] between Date1 and Date2? ***Use Case 3***
- CQ12 - What is Level(s) Indicator [6.1] for [Asset] over [StudyPeriod]? ***Use Case 4*** *- financial life cycle cost according to Level(s)*
- CQ13 - Which Level(s) Indicators [List] have value 'Null'? ***Use Case 4***

Ontology conceptualisation

A conceptualisation activity is carried out based on the above requirements. Concepts are arranged using a diagramming tool. Due to the visual nature, this activity is suitable for collaboration with experts who may not otherwise be familiar with ontology languages. In this case, input was sought from a number of AM and ontology experts via supplementary interviews, who gave input on the domain logic and hierarchical arrangement of concepts. The result of this step is an initial ontology concept diagram (Figure 5).

Potential for reuse and alignment

The LOT authors recommend carrying out an analysis of potential ontology reuse and alignment only after the conceptualisation stage. Table 1 details the considerations in this study for reuse and alignment with existing related ontologies. Based on this assessment it is determined that the RealEstateCore (REC) ontology is most promising for alignment or extension due to the many overlapping concepts, significant existing user-base and the community's active participation in alignment activities. As an exploratory step in this direction, Figure 6 identifies overlapping concepts in the REC ontology. The conclusion here is that our lifecycle information management ontology concept could potentially be achieved with a relatively light extension of the REC ontology. Further investigation will be required, however, to determine the practicality of this proposal.

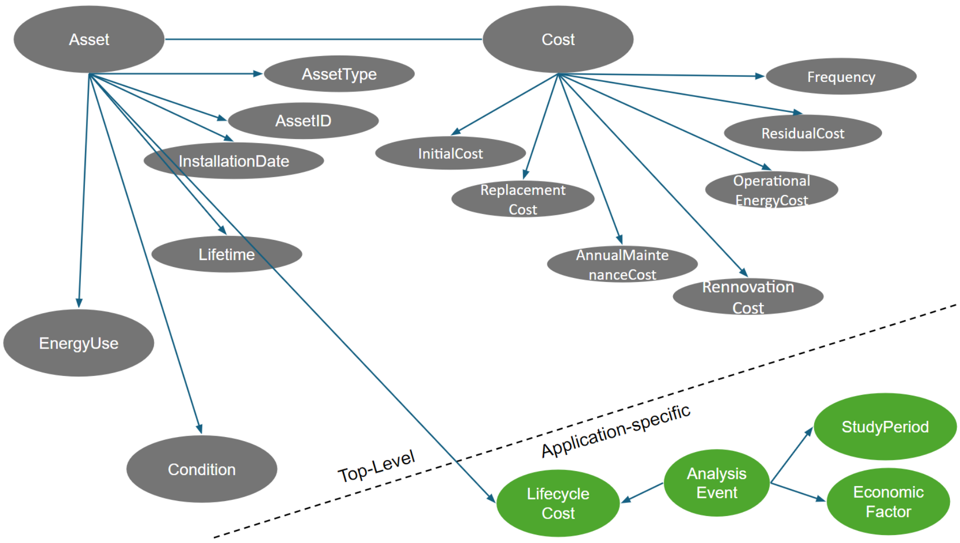


Figure 5: Initial AM Ontology concept denoting top-level and application-specific classes.

Table 1: Existing ontology reuse/extension considerations

Related ontology	Author	Suitable Concepts	Existing user base	Alignment activities	Generality
REC	(Hammar et al., 2019)	Many	Significant	Active	General
SLiCE	(Röck et al., 2024)	Some	Minimal	Active	General
FDM	(Xie et al., 2022)	Some	Unknown	Active	General
BONSAI	(Ghose et al., 2022)	Some	Unknown	Unknown	Macro-specific
LCCA	(Gao et al., 2020)	Some	Unknown	Unknown	Appl.-specific
AMO	(Katsumi et al., 2022)	Few	Under development	Unknown	case-specific

Discussion

The portrayal of the changing regulatory landscape in the built environment at the outset of this paper underscores the growing emphasis on sustainability and resource efficiency. As policy transitions towards mandatory sustainability reporting, the need for robust IT systems which are fit for purpose, becomes paramount. Life cycle analysis has emerged as a means to counteract the phenomenon of short-term-ism in decision-making by viewing impacts of decision over longer time periods. Our previous research portrays the current IT landscape as not fit for purpose to meet future information requirements, and advocates reconfiguration of future systems.

Illustrating our vision for this new software landscape, we outline a scenario whereby various stakeholders access a shared knowledge base, structured semantically around commonly agreed concepts over the web. Through specialised, modular applications, stakeholders leverage this shared knowledge base to suit their specific analytical needs. Following the LOT ontology development methodology, we describe the activities in developing an initial concept, focusing on fundamental AM use cases which serve a variety of key AM stakeholders. This involves defining the scope and use cases, gathering requirements via background research, conceptualising and verifying the ontology logic with experts, and exploring potential alignment with existing efforts. We present excerpts from the Ontology Requirements Specification Document (ORSD), including natural language Competency Questions (CQs) and tabular information, which lay the groundwork for encoding domain knowledge in a formal ontology language. The outcome is an initial conceptualisation of the AM ontology, with a recommendation for alignment with the RealEstateCore ontology, a data model which already describes multiple concepts useful for AM stakeholders and our use cases. Alignment with the RealEstate-

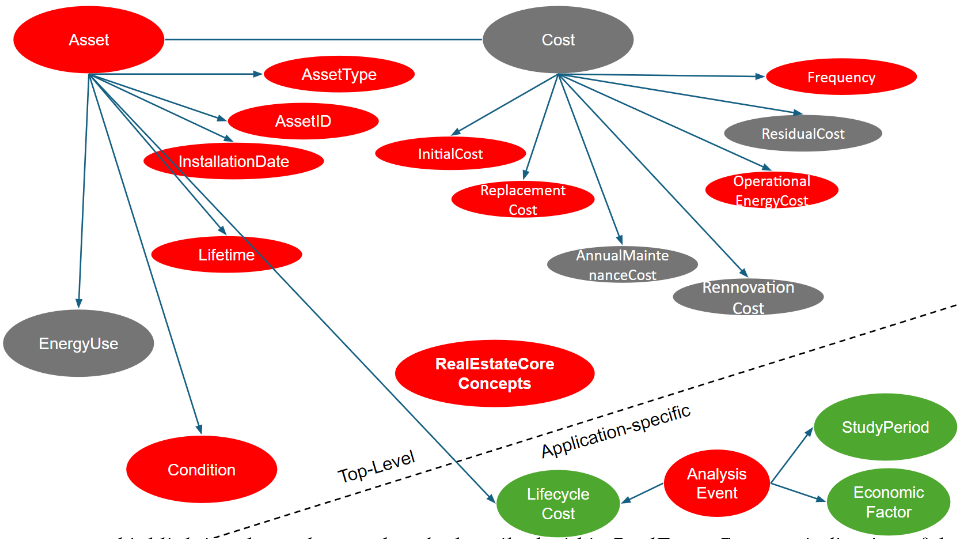


Figure 6: Ontology concept highlighting those classes already described within RealEstateCore, an indication of the extent to which REC would require extension to accommodate our research objectives.

Core ontology presents a strategic opportunity for uptake, given its existing user base and openness to alignment efforts with other ontology communities. Looking ahead, our vision for the future IT landscape aligns with the efforts of researchers in the built environment informatics domain, in particularly the activities of the Linked Building Data community (W3C, 2022).

Though this study is based upon extensive prior practitioner-based research in the AM and LCSA fields, we stop short at validating the ontology through encoding and testing in real-world scenarios, activities which remain as future work. Furthermore, though the selected use cases serve a broad selection of key stakeholders, the outcomes are nonetheless limited to describing those few applications. There are, of course, a vast range of potential uses for such a knowledge base, which is entirely the objective of the extensible and modular approach described; but with the rapidly approaching requirements of the CSRD and other sustainability reporting requirements, if we are to ensure equitable participation, it is of utmost importance to support the domain in managing their asset information particularly small-medium enterprises. Our next endeavors will focus on expanding the ontology concept to cover additional use cases and we will progress through the subsequent stages of the LOT development methodology activities to encode and validate the concept, ensuring its applicability and effectiveness in real-world scenarios.

Conclusions

This paper advocates a paradigm shift in IT systems within the built environment to align with environmental sustainability policy ambition. Proposing a web-enabled technology stack and leveraging the Linked Open Terms (LOT) methodology, we illustrate a asset management ontology landscape fit for purpose, allowing for modularity of specialised stakeholder applications ingesting shared information from a common knowledge base. With a focus on supporting fundamental Life Cycle Sustainability Assessment use cases, our research lays the groundwork for technical ontology development by gathering domain insight, outlining functional requirements and conceptualising the modular ontology landcsape. By reflecting on related research efforts, we highlight the potential for aligning with or extending existing standards such as RealEstateCore and SLiCE data models. This work contributes to fostering consensus within the domain, offering a roadmap for future research and development in enhancing asset lifecycle information management and decision-making processes.

Acknowledgments

The authors thank Julia Kaltenegger (TU/e) and Esra Bektas (TNO) for their input. This work was supported by the CBIM-ETN funded by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860555.

References

Collier, P. (2018). The Future of Capitalism: Facing the New Anxieties. Harper, New York.

Dumrose, M. and Höck, A. (2023). Corporate Carbon-Risk and Credit-Risk: The Impact of Carbon-Risk Exposure and Management on Credit Spreads in Different Regulatory Environments. Finance Research Letters.

European Commission (2022). Level(s): A Guide to Europe's new Reporting Framework for Sustainable Buildings. Technical report.

European Commission (2023). The Commission adopts the European Sustainability Reporting Standards.

European Committee for Standardization (CEN) (2011). EN 15978 Sustainability of construction works. Assessment of environmental performance of buildings. Calculation method.

Fang, Z., Liu, Y., Lu, Q., Pitt, M., Hanna, S., and Tian, Z. (2022). BIM-integrated portfolio-based strategic asset data quality management. Automation in Construction, 134:104070.

Gao, X. and Pishdad-Bozorgi, P. (2019). BIM-enabled facilities operation and maintenance: A review. Advanced Engineering Informatics, 39:227–247.

Gao, X., Pishdad-Bozorgi, P., Tang, S., and Shelden, D. (2020). Machine Learning-Based Building Life-Cycle Cost Prediction: A Framework and Ontology.

Geekiyanage, D. and Ramachandra, T. (2018). Significant Factors Influencing Operational and Maintenance (O&M) Costs of Commercial Buildings. 7th World Construction Symposium.

Ghose, A., Lissandrini, M., Hansen, E. R., and Weidema, B. P. (2022). A core ontology for modeling life cycle sustainability assessment on the Semantic Web. Journal of Industrial Ecology, 26(3):731–747.

Grzyl, B., Miszewska, E., and Apollo, M. (2017). The life cycle cost of a building from the point of view of environmental criteria of selecting the most beneficial offer in the area of competitive tendering. E3S Web of Conferences, 17:00028.

Hammar, K., Wallin, E. O., Karlberg, P., and Hälleberg, D. (2019). The RealEstateCore Ontology. In The Semantic Web – ISWC 2019, volume 11779, pages 130–145, Cham. Springer International Publishing. Series Title: Lecture Notes in Computer Science.

International Organization for Standardization (2017). ISO 15686-5 Buildings and constructed assets - Service life planning - Part 5: Life-cycle costing.

Katsumi, M., Huang, T., and Fox, M. S. (2022). Toward Requirements for an Ontology of Asset Management. In 12th International Workshop on Formal Ontologies meet Industry.

Kehily, D. and Underwood, J. (2017). Embedding life cycle costing in 5D BIM. Journal of Information Technology in Construction (ITcon), 22(8):145–167.

Lu, K., Jiang, X., Yu, J., Tam, V. W. Y., and Skitmore, M. (2021). Integration of life cycle assessment and life cycle cost using building information modeling: A critical review. Journal of Cleaner Production, 285:125438.

Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., and García-Castro, R. (2022). LOT: An industrial oriented ontology engineering framework. Engineering Applications of Artificial Intelligence, 111:104755.

Pritoni, M., Paine, D., Fierro, G., Mosiman, C., Poplawski, M., Saha, A., Bender, J., and Granderson, J. (2021). Metadata Schemas and Ontologies for Building Energy Applications: A Critical Review and Use Case Analysis. Energies, 14(7):2024. Number: 7 Publisher: Multidisciplinary Digital Publishing Institute.

Röck, M., Passer, A., and Allacker, K. (2024). SLiCE: An open building data model for scalable high-definition life cycle engineering, dynamic impact assessment, and systematic hotspot analysis. Sustainable Production and Consumption, 45:450–463.

Salvado, F., de Almeida, N. M., and Azevedo, V. e. (2021). Building Investment Index: A Decision-Making Tool to Optimize Long-Term Investment Decisions. In Rodrigues, H., Gaspar, F., Fernandes, P., and Mateus, A., editors, Sustainability and Automation in Smart Constructions, Advances in Science, Technology & Innovation. Springer International Publishing.

Shaw, C., de Andrade Pereira, F., Hoare, C., Farghaly, K., Hartmann, T., and O'Donnell, J. (2024). Life cycle cost analysis at scale: a reference architecture-based approach. Construction Industry as a Net-Zero Enabler: Driving Circular Economy and Sustainability through Innovation and Change Management, special issue of Built Environment Project and Asset Management. Emerald (Accepted).

Sobhkhiz, S., Taghaddos, H., Rezvani, M., and Ramezanianpour, A. M. (2021). Utilization of semantic web technologies to improve BIM-LCA applications. Automation in Construction, 130:103842.

W3C (2017). Shapes Constraint Language (SHACL).

W3C (2022). Linked Building Data Community Group.

Werbrouck, J., Schulz, O., Oraskari, J., Mannens, E., Pauwels, P., and Beetz, J. (2023). A generic framework for federated CDEs applied to Issue Management. Advanced Engineering Informatics, 58:102136.

Wilde, A. S., Wanielik, F., Rolinck, M., Mennenga, M., Abraham, T., Cerdas, F., and Herrmann, C. (2022). Ontology-based approach to support life cycle engineering: Development of a data and knowledge structure. Procedia CIRP, 105:398–403. Publisher: Elsevier.

Xie, X., Moretti, N., Merino, J., Chang, J., Pauwels, P., and Parlikad, A. K. (2022). Enabling building digital twin: Ontology-based information management framework for multi-source data integration, volume 1101. Journal Abbreviation: IOP Conference Series: Earth and Environmental Science Publication Title: IOP Conference Series: Earth and Environmental Science.

Data Sensing and Acquisition

IDENTIFICATION OF COMPOSITE SENSOR FAULTS IN STRUCTURAL HEALTH MONITORING SYSTEMS USING LONG SHORT-TERM MEMORY NETWORKS

Thamer Al-Zuriqat, Heba Al-Nasser, Kosmas Dragos, Chillón Geck, and Kay Smarsly
Hamburg University of Technology, Hamburg, Germany

Abstract

Fault identification (FI) is an integral part of sensor fault diagnosis in structural health monitoring (SHM) systems. However, current FI approaches often overlook composite sensor faults, i.e. different sensor fault types occurring simultaneously within an individual sensor. As a result, actual fault occurrences in real-world SHM systems may be underestimated. This paper introduces an FI approach utilizing long short-term memory networks, addressing composite faults. The FI approach is validated using sensor data recorded by a real-world SHM system. The results demonstrate the capability of the FI approach to identify composite sensor faults, thus enhancing the reliability and accuracy of fault diagnosis.

Introduction

Structural health monitoring (SHM) is a non-destructive evaluation technique employing data recorded by sensors ("sensor data") to assess structural conditions (Law et al., 2014). SHM aims to improve safety and cost efficiency in structural maintenance through filling gaps of periodic visual inspections (Cawley, 2018). Hardware or software errors, exposure to harsh environmental conditions, degradation, and signal interferences may lead to malfunctions of sensors ("faulty sensors") in SHM systems (Zhang et al., 2018). Consequently, faulty sensors may compromise the outcomes of SHM systems (Steiner et al., 2019).

In SHM systems, sensor fault types include bias, drift, gain, precision degradation, complete failure (constant or with noise), and outliers (Kullaa et al., 2013). Fault diagnosis approaches for SHM systems have been proposed based on either physical or analytical redundancy (Frank, 1990). Physical redundancy entails installing additional, i.e. "redundant", sensors and detecting faults based on sensor data comparisons. However, the high cost, power consumption, and maintenance associated with physical redundancy have been the primary motivation for developing analytical redundancy approaches (Smarsly & Petryna, 2014). In general, analytical redundancy employs mathematical models to characterize a system, leveraging the inherent redundancy present in the sensor data (Al-Zuriqat et al., 2023). Fault diagnosis using analytical redundancy comprises four steps (Patton, 1990):

- Fault detection
- Fault isolation
- Fault identification
- Fault accommodation

In the fault detection step, residuals between sensor data and corresponding "virtual outputs", derived from mathematical models, are evaluated using threshold logic or hypothesis testing (Isermann & Balle, 1997). In case faults are detected, fault isolation involves determining the locations of the faulty sensors. To gain insights into the underlying causes of sensor faults and define strategies for compensating for errors induced by sensor faults, the type or nature of the faults are determined in the fault identification (FI) step. Finally, in the fault accommodation step, data recorded by the faulty sensor is reconstructed using virtual outputs of the mathematical models.

Existing fault diagnosis approaches have mainly focused on detecting, isolating, and accommodating sensor faults. Rao et al. (2007) has presented a concept, originally proposed by Kramer (1992), introducing a null-subspace-based approach for sensor fault detection and isolation, combined with autoassociative neural networks for fault accommodation. Smarsly & Law (2014) have proposed a decentralized fault detection and isolation approach in wireless SHM systems employing artificial neural networks. The approach has been extended from the time domain to the frequency domain and has also accounted for the presence of structural damage (Dragos & Smarsly, 2016). Al-Zuriqat et al. (2023) have introduced an adaptive sensor fault detection, isolation, and accommodation approach for SHM systems using feedforward artificial neural networks with backpropagation, considering single sensor faults occurring simultaneously in individual sensors. Liu et al. (2022) have used stacked gated recurrent unit neural networks to detect, isolate and accommodate sensor faults, and global-local logic to differentiate sensor faults from structural damage.

Among fault diagnosis approaches that include FI, Fritz et al. (2022) have proposed an analytical-redundancy approach, which integrates feedforward artificial neural networks and convolutional neural networks to detect, isolate, identify, and accommodate sensor faults. Additionally, other approaches, based on generalized likelihood ratio (Li et al., 2019), on generalized quasi-natural analogy test principle (Yan et al., 2020), and on set theory and support vector machine (Yu et al., 2014), have also addressed the FI step. However, in the aforementioned studies, only single sensor faults have been considered, i.e. one sensor fault type within each individual sensor. Moreover, not all sensor fault types have been addressed. For example, outliers (Fritz et al. 2022; Li et al., 2019), drift, precision degradation, and complete failure (Yan et al., 2020) have not been included.

Sensors in SHM systems may experience combinations of faults, i.e. faults of different types occurring simultaneously within an individual sensor (hereinafter termed "composite sensor faults"). The plausibility of composite sensor faults occurrence in SHM systems has been corroborated in Li et al. (2023), in which composite sensor faults have been observed during experimentation, specifically gain and drift in one sensor. Composite sensor faults have been investigated in other disciplines for more than two decades, for example, in boiler processes using structured residuals with maximized sensitivity (Qin & Li, 1999), in wastewater treatment plants using Fisher discriminant analysis (Luca et al., 2023), and in lithium-ion batteries using a two-layer identification algorithm (Shen et al., 2024).

Despite the importance of FI in gaining insights into underlying causes of sensor faults within SHM systems, current FI approaches take into consideration single sensor faults, often failing to account for composite sensor faults, which may manifest in real-world SHM systems (Li et al., 2023). Moreover, current FI approaches rely on recognizing unique patterns in sensor data to identify different sensor fault types (Peralta et al., 2022). Nonetheless, composite sensor faults may generate distinct patterns that may be hard to recognize for current FI approaches. In this paper, an approach towards identification of composite sensor faults (ICSF) is proposed. Sensor data with artificially injected composite sensor faults is used to train long short-term memory (LSTM) classification networks, addressing composite sensor faults.

The remainder of the paper is organized as follows: First, composite sensor faults in SHM systems are elucidated. Then, a description of the ICSF approach is introduced. Next, the implementation and validation of the ICSF approach are presented, and the results are discussed. Finally, the paper ends with a summary and conclusions, and an outlook on future research is provided.

Composite sensor faults

Sensor fault types manifest into errors, which have distinct footprints on sensor data. For example, a bias is represented by a constant deviation of the sensor data from the values that should be measured ("actual values", e.g. structural responses or environmental parameters). A drift is characterized by the gradual deviation of sensor data from the actual values over time. A gain occurs when sensor data is systematically scaled by a constant value. Precision degradation is caused by the contamination of sensor data with noise. A complete failure is observed as sensor data consisting of a constant value ("complete failure constant") or noise ("complete failure with noise"), regardless of changes occurring in the actual values. Finally, an outlier manifests as a discontinuous observation in the sensor data that deviates from the actual values at isolated time instances.

The aforementioned sensor fault types are usually addressed as single sensor faults, occurring in individual sensors. However, as mentioned previously, sensors in real-world SHM systems may experience composite sensor faults, an example of which, consisting of drift and an outlier, is illustrated in Figure 1. As a result, full fault identification needs to address both *single sensor faults* and *composite sensor faults*. The description of the ICSF approach, including two phases, each comprising several steps, are presented in the following section.

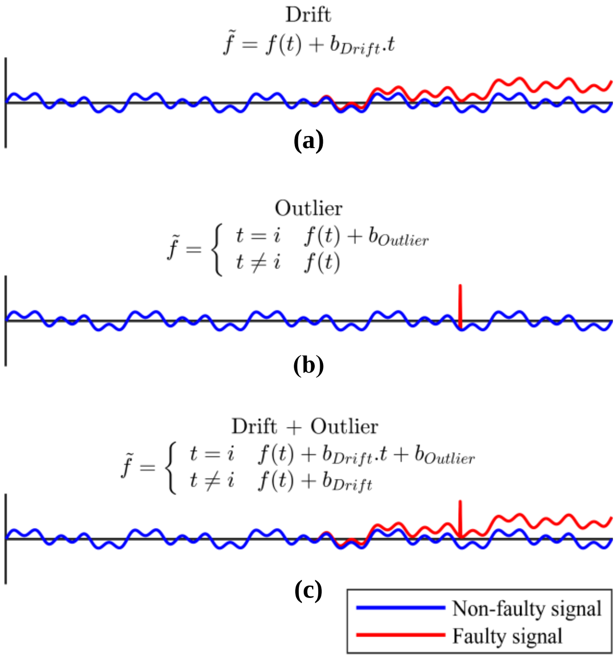


Figure 1: Illustration of a composite sensor fault: (a) a single sensor fault with drift, (b) a single sensor fault with an outlier, and (c) a composite sensor fault consisting of drift and an outlier

Description of the ICSF approach

In this section, the ICSF approach is presented, comprising two phases, (i) preparing the training dataset, and (ii) developing the classification models. A flowchart describing the workflow of the ICSF approach is shown in Figure 2.

Phase 1: Preparing the training dataset

1. Sensor data is recorded by SHM systems within the so-called "data collection period". The total number of data points p recorded by each sensor in the data collection period is representative of the normal operation of the structure.
2. A correlation analysis is conducted on the sensor data to unveil the number of correlated sensors k within a SHM system. Since the classification models are based on mapping sequences of sensor data, correlation between sensors is a prerequisite. Each sensor in the set of correlated sensors is denoted with i (i = 1, …, k). Data recorded by all correlated sensors $f_{1 \to k}$ is stored in matrix **A**. As result, matrix **A** has a length equal to the number of data points p, and a width equal to the number of correlated sensors k, as shown in Equation 1.

$$\mathbf{A}_{p \times k} = \begin{bmatrix} f_{1,1} & f_{1,2} & \cdots & f_{1,k} \\ f_{2,1} & f_{2,2} & \cdots & f_{2,k} \\ \vdots & \vdots & \ddots & \vdots \\ f_{p,1} & f_{p,2} & \cdots & f_{p,k} \end{bmatrix} \quad (1)$$

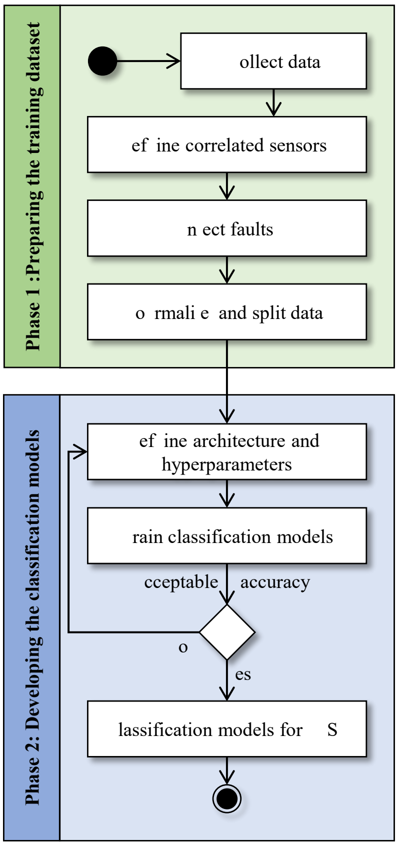


Figure 2: Flowchart of the ICSF approach

3. A total of k input and output datasets are created to train classification models in a later step. The sensor data in matrix **A** are copied into input dataset i (i = 1, …, k), and composite sensor faults are artificially injected into the vector $\mathbf{f}_i$, i.e. in sensor data recorded by sensor i. Each fault is injected via a mathematical model, representing the footprint of the fault on the sensor data (e.g. adding a constant value to the sensor data to simulate bias). Injected sensor fault types are stored in the classification output dataset i. As a result, the input dataset i contains clean data from sensors (1, 2, …, i–1, i+1, …k), and sensor data from sensor i with composite sensor faults.
4. Data in the input dataset i is normalized using a minimum-maximum normalization method, shown in Equation 2, to avoid overfitting issues while training the classification models. Then, the normalized input dataset i, as well as the output dataset i are split into training dataset (70%), validation dataset (15%), and testing dataset (15%). In Equation 2, x represents an arbitrary measurement in the sensor data, while x_{min} and x_{max} denote the minimum and maximum measurements in the sensor data, respectively. The variable $x_{normalized}$ corresponds to the normalized value. Both normalization parameters x_{min} and x_{max} used for the input dataset i are saved to be applied to new sensor data that is fed to the classification model M_i after training.

$$x_{normalized} = \frac{x - x_{min}}{x_{max} - x_{min}} \qquad (2)$$

Phase 2: Developing the classification models

1. The initial architecture of the LSTM classification network is defined. The LSTM classification network has an input layer, hidden layers, and an output layer. In Figure 3, the general architecture of the LSTM classification network is illustrated. The input layer is a sequence input layer with a length equal to the number of correlated sensors k. The hidden layers include successions of LSTM layers and "dropout" layers. In each LSTM layer, memory cells and a set of gates regulate the flow of data. This architecture enables LSTM layers to capture and retain long-term dependencies in sequential data, making them particularly effective for tasks such as sensor data analysis. In each dropout layer, a probability of dropout is defined representing the likelihood (set between 0 and 1) of randomly neglecting measurements passed through the LSTM layer as means of preventing overfitting. The number of LSTM layers as well as the number of dropout layers may vary based on the model accuracy, which is determined in a later step. The output layers contain a fully-connected layer, a Softmax layer, and a classification layer. The fully-connected layer connects every neuron of the last hidden layer to the neurons of the output layer, representing the number of classes in the classification problem. The Softmax layer is used

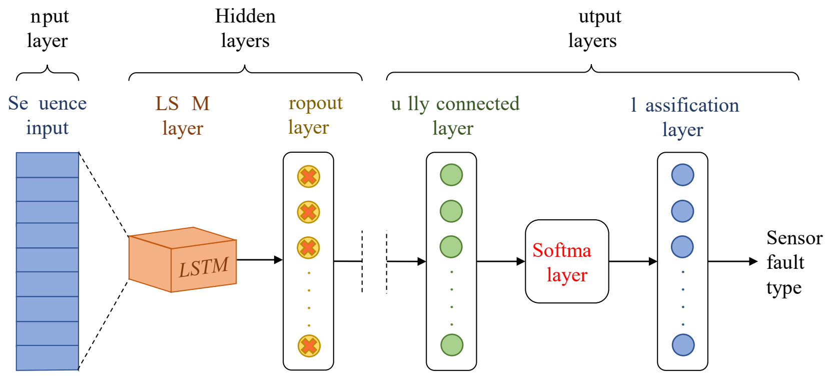


Figure 3: General architecture of the LSTM classification network

between the fully connected layer and the classification layer to convert raw logits into probability distributions, facilitating interpretability, backpropagation, and optimal training. Finally, the classification layer contains all classes of the classification problem, more specifically, the number of sensor fault types c to be included in the FI step. The number of sensor faults c, is calculated using Equation 3, and depends on: (i) the number of single sensor faults m included in each composite sensor fault, and (ii) the total number of single sensor faults n included in all composite sensor faults.

$$c\ (n, m) = \frac{n!}{m!(n-m)} \quad (3)$$

2. One classification model M_i for each correlated sensor i (i = 1, ..., k) is created by training an LSTM using the training dataset i. During training, the training dataset i is fed sequentially (in "batches") to the LSTM classification network, i.e. each batch of the normalized input dataset is fed to the input layer and the corresponding classes of the output dataset are fed to the output layer. The training consists of updating weighted connections between neurons, until the probabilities of the Softmax layer reach a predefined level of accuracy. Thereupon, the classification model M_i is created, capable of recognizing patterns and features in the input dataset i and classifying the sensor data into the corresponding composite sensor fault type. The validation dataset is used to fine-tune the hyperparameters, and prevent overfitting during training.
3. Upon completing training and validating, the accuracy of model M_i is obtained using the testing dataset. For model M_i to be accepted, the classification accuracy must lie above a FI threshold γ. The FI threshold γ is derived following the same rationale as the fault detection threshold γ in previous related research (Al-Zuriqat et al., 2023). Equation 4 introduces the formula to calculate the accuracy (Acc) of the classification model M_i.

$$Acc = \frac{Number\ of\ correct\ preditions}{Total\ number\ of\ preditions} \quad (4)$$

4. Finally, if the accuracy of model M_i is satisfactory, the model is saved to identify faults of sensor i (i = 1, …, k). However, if the accuracy of model M_i is not satisfactory, a different architecture of the LSTM classification network is defined and a new model M_i is trained (back to step 1 of phase 2).

The implementation and validation of the ICSF approach using sensor data from a real-world SHM system installed on a pedestrian bridge is presented in the next section.

Implementation and validation of the ICSF approach

This section describes the implementation and validation of the ICSF approach. To implement the ICSF approach, the programming language MATLAB is utilized to analyze the sensor data, inject faults, and train the classification models. The validation test is carried out using sensor data (acceleration measurements) from a real-world SHM system installed on a pedestrian bridge. First, the pedestrian bridge and the SHM system are presented. Then, the ICSF approach is applied using the acceleration measurements recorded by the real-world SHM system.

Description of the pedestrian bridge and the SHM system

The pedestrian bridge, located in Evosmos, Thessaloniki, Greece, has been a subject of research since the construction in 2016. (Dragos et al., 2020; Smarsly et al., 2022a; and Smarsly et al., 2022b). The pedestrian bridge, is a composite concrete-steel structure, in which a steel structure supports a reinforced-concrete deck. The dimensions of the reinforced-concrete deck are 35.00 m (length) and 4.60 m (width).

Figure 4 illustrates the top view of the pedestrian bridge and the locations of the accelerometers used in the validation test. The SHM system comprises four accelerometers (i.e. S_1, S_2, S_3, and S_4), which have been previously tested and found to be free of faults. The accelerometers are equally distributed along the central axis spanning the length of the bridge, spaced seven meters apart from each other.

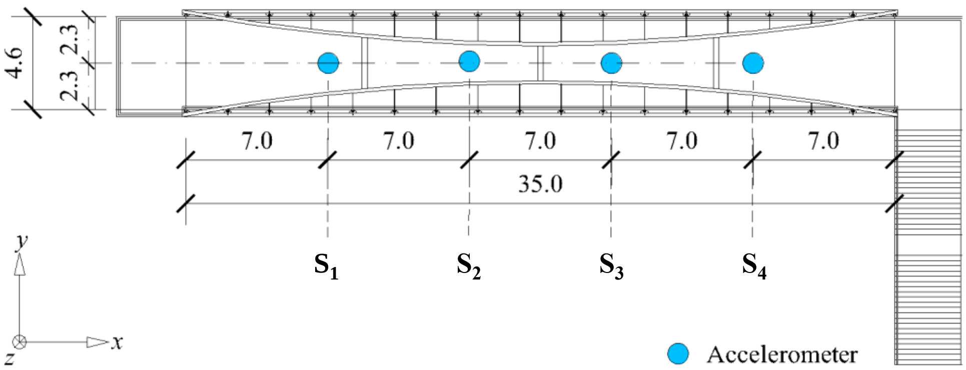


Figure 4: Top view of the pedestrian bridge with the locations of the accelerometers (dimensions are given in meters)

Description of the validation test

For preparing the training dataset (phase 1), acceleration measurements are recorded by the four accelerometers (S_1, S_2, S_3, and S_4) over data collection period of 90 minutes at a sampling rate of 128 Hz. The total number of data points $p = 692{,}628$ is recorded by each accelerometer.

Next, correlations within the acceleration measurements are investigated via a correlation analysis. A strong correlation has been exposed by the Pearson correlation coefficient among all four accelerometers, i.e. $k = 4$. Notably, the lowest correlation coefficient 0.937 has been observed between sensors S_1 and S_4. Then, the acceleration measurements, recorded by all four correlated accelerometers $f_{1 \rightarrow k = 4}$, are stored in matrix $\mathbf{A}_{692628 \times 4}$.

To determine the sensor fault types to be included in the validation test, as well as the size of the classification layer in a later step, a total of five single faults ($n = 5$) are considered, i.e. bias, drift, gain, precision degradation and outliers. Complete failure, either constant or with noise, entails sensors essentially ceasing to operate and can, therefore, hardly be combined with other single sensor faults into composite sensor faults. Combinations of two single faults in the same sensor ($m = 2$) are considered, e.g. drift and outlier at the same time within an individual sensor. As a result, using Equation 3, the total number of composite sensor faults is $c\,(n, m) = 10$. Table 1 provides an overview of the composite sensor faults considered in the validation test.

Table 1: Sensor fault types considered in the validation test

No.	Composite sensor fault
1	Bias + Drift
2	Bias + Gain
3	Bias + Precision degradation
4	Bias + Outliers
5	Drift + Gain
6	Drift + Precision degradation
7	Drift + Outliers
8	Gain + Precision degradation
9	Gain + Outliers
10	Precision degradation + Outliers

Acceleration measurements stored in matrix $\mathbf{A}_{692628 \times 4}$ are copied into four input datasets. Each input dataset is divided equally into 10 subsets, for which each sensor fault type shown in Table 1 is artificially injected. The classes of the composite sensor faults injected into the input dataset i ($i = 1, \ldots, 4$) is stored in the output dataset i ($i = 1, \ldots, 4$).

Next, acceleration measurements, stored in the input dataset i, are normalized using the minimum-maximum normalization method shown in Equation 2. The normalized input dataset i is split into 70 % training dataset (484,838 data points), 15 % validation dataset (103,894 data points), and 15 % testing data (103,894 data points). Both normalization parameters x_{min} and x_{max} used for the input dataset i during training are saved, to be applied to new sensor data that is fed to the classification model M_i after training.

The initial architecture for the LSTM classification network is defined, with the input layer consisting in a sequence input layer with length $k = 4$. The output layer consists of a fully-connected layer, a Softmax layer, and a classification layer with 10 classes, where each class present a composite sensor fault from Table 1. Three hidden layers are defined, each comprising a LSTM layer followed by a dropout layer. To avoid overfitting during training, the probability of all dropout layers has been set to 20 %.

A total of four classification models, equal to the number of the correlated sensors ($k = 4$), are created by training an equal number of LSTM classification networks. Each model M_i classifies sensor fault types of sensor i, using acceleration measurements from the four correlated sensors of the SHM system as input data. A total training time of 69 minutes for each LSTM classification model is recorded.

Based on Equation 4, the accuracy of the models is evaluated using the testing dataset (103,894 data points). Using a FI threshold $\gamma = 0.85$, a model with an accuracy above 85 % is considered acceptable (Al-Zuriqat et al., 2023). The lowest model accuracy has been observed in model M_2, with an accuracy of 90.3 %.

Finally, the fourth step involves saving all trained classification models M_1, M_2, M_3, and M_4 to identify sensor fault types for sensor S_1, S_2, S_3, and S_4, respectively. The trained classification models are tested using acceleration measurements newly recorded by the SHM system, as presented and discussed in the following section.

Results and discussion

This section presents and discusses the results of applying the classification models (M_1, M_2, M_3, and M_4), obtained by the ICSF approach, to newly recorded acceleration measurements. The acceleration measurements used in the validation test have been recorded by the real-world SHM system and correspond to a data collection period of 30 minutes with a sampling rate of 128 Hz. The total number of data points recorded by each of the accelerometers is $p = 230{,}876$ data points. Subsequently, all composite sensor faults introduced in Table 1 have been randomly injected into the newly recorded acceleration measurements. The purpose of the injected faults here is to provide a labeled dataset, which is used to evaluate the accuracy of the classification models. Table 2 presents the results of testing the classification models. To determine the accuracy of the classification models, the number of sensor faults identified is compared to the number of sensor faults injected in the testing dataset.

Table 2: Fault identification results of artificially injected sensor fault types including composite faults

Fault type	Sensor	Injected faults	Correctly identified faults	*Acc*
Bias + Drift	S_4	13,607	13,600	99.9 %
Bias + Gain	S_1	13,522	13,411	99.1 %
Bias + Precision degradation	S_2	13,720	13,632	99.3 %
Bias + Outliers	S_3	5,378	4692	87.2 %
Drift + Gain	S_4	13,627	12859	94.3 %
Drift + Precision degradation	S_1	13,390	12,944	96.6 %
Drift + Outliers	S_2	5,461	261	4.7 %
Gain + Precision degradation	S_3	13,638	13,373	98.0 %
Gain + Outliers	S_4	5,556	4,871	87.6 %
Precision degradation + Outliers	S_1	5,398	4,126	76.4 %
Total	**-**	**103,297**	**93,769**	**90.7 %**

As presented in Table 2, the FI capabilities of models M_1, M_2, M_3, and M_4 have been demonstrated using the newly recorded acceleration measurements with artificially injected faults. Out of 103,279 faults injected into the newly recorded acceleration measurements, 93,769 sensor fault types have been correctly identified, representing a total accuracy rate of 90.7 %.

Upon delving deeper into the results of the identification of composite sensor faults, it is noticeable that the classification model achieved the highest accuracy in identifying the composite sensor fault "Bias + Drift" with an accuracy rate of 99.9%. Conversely, a considerably low accuracy rate was observed in the identification of the composite sensor fault "Drift + Outliers" at 4.8%. The low accuracy of identifying the composite fault type "Drift + Outliers" may be attributed to the subtle patterns of outliers, as they correspond to point singularities in a continuous signal, which may be hard to be identified as features by the LSTM classification network.

Summary and conclusions

Fault diagnosis using analytical redundancy, building on mathematical models, has been gaining increasing interest in SHM, owing to the low cost when avoiding redundant sensors, as compared to traditional physical redundancy approaches. Fault diagnosis encompasses fault detection, isolation, identification, and accommodation. Despite the importance of FI in gaining insights into underlying causes of sensor faults within SHM systems, current FI approaches often fail to account for composite sensor faults, which may manifest in real-world SHM systems.

This paper has presented an approach towards identification of composite sensor faults (ICSF). Sensor data with artificially injected composite sensor faults has been used to train LSTM classification networks, addressing composite sensor faults. The ICSF approach has been implemented using the programing language MATLAB, and validated using acceleration measurements recorded by a real-world SHM system. From the results, it is concluded that the ISCF approach has been proven capable of identifying composite sensor faults, thus enhancing the reliability and accuracy of fault diagnosis in SHM systems.

Future work may involve coupling the ICSF approach with other existing fault detection, isolation, and accommodation approaches to establish a comprehensive fault diagnosis concept. Furthermore, the ICSF approach may be embedded into wireless sensor nodes to identify sensor faults on board the sensor nodes within real-world decentralized SHM systems.

Acknowledgments

This paper is the result of a collaborative effort of authors funded by different sources. The authors would like to gratefully acknowledge the support offered by the German Research Foundation (DFG) under grants SM 281/17-1 and SM 281/20-1 as well as by the German Federal Ministry for Digital and Transport (BMDV) within the mFUND program under grant 19FS2013B. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of DFG or BMDV.

References

Al-Zuriqat, T., Chillón Geck, C., Dragos, K., & Smarsly, K. (2023) Adaptive fault diagnosis for simultaneous sensor faults in structural health monitoring systems. Infrastructures, 8(3), 39.

Cawley, P. (2018) Structural health monitoring: Closing the gap between research and industrial deployment. Structural health monitoring, 17(5), pp. 1225-1244.

Dragos, K. & Smarsly, K. (2016) Distributed adaptive diagnosis of sensor faults using structural response data. Smart Materials and Structures, 25(10), 105019.

Dragos, K., Makarios, T., Karetsu, I., Manolis, G. D., & Smarsly, K. (2020) Detection and correction of synchronization-induced errors in operational modal analysis. Archive of Applied Mechanics, 90(7), pp. 1547-1567.

Frank, P. M. (1990) Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. Automatica, 26(3), pp. 459-474.

Fritz, H., Peralta Abadía, J. J., Legatiuk, D., Steiner, M., Dragos, K., & Smarsly, K. (2022) Fault Diagnosis in Structural Health Monitoring Systems Using Signal Processing and Machine Learning Techniques. Structural Health Monitoring Based on Data Science Techniques, Cury, A., Ribeiro, D., Ubertini, F. & Todd, M. D., Eds. Cham: Springer International Publishing, pp. 143-164.

Isermann, R. & Balle, P. (1997) Trends in the application of model-based fault detection and diagnosis of technical processes. Control engineering practice, 5(5), pp. 709-719.

Kullaa, J. (2013) Detection, identification, and quantification of sensor fault in a sensor network. Mechanical Systems and Signal Processing, 40(1), pp. 208-221.

Kramer, M. A. (1992) Autoassociative neural networks. Computers & chemical engineering, 16(4), pp. 313-328.

Law, K. H., Smarsly, K., & Wang, Y. (2014) Sensor data management technologies for infrastructure asset management. Sensor Technologies for Civil Infrastructures. Wang, M. L., Lynch, J. P. & Sohn, H. Eds. Woodhead Publishing, pp. 3-32.

Li, L., Liu, G., Zhang, L., & Li, Q. (2019) Sensor fault detection with generalized likelihood ratio and correlation coefficient for bridge SHM. Journal of Sound and Vibration, 442(2019), pp. 445-458.

Li, L., Luo, H., Qi, H., & Wang, F. (2023) Sensor Fault Diagnosis Method of Bridge Monitoring System Based on FS-LSTM. Advances in Frontier Research on Engineering Structures, Yang Y., Eds. Singapore: Springer Nature Singapore. pp. 487-501.

Liu, B., Xu, Q., Chen, J., Li, J., & Wang, M. (2022) A New Framework for Isolating Sensor Failures and Structural Damage in Noisy Environments Based on Stacked Gated Recurrent Unit Neural Networks. Buildings, 12(8), 1286.

Luca A.V., Simon-Várhelyi, M., Mihály, N.B., & Cristea, V.M. (2023) Fault type diagnosis of the WWTP dissolved oxygen sensor based on Fisher discriminant analysis and assessment of associated environmental and economic impact. Applied sciences (Switzerland), 13(4), 2554.

Patton, R. J. (1990) Fault detection and diagnosis in aerospace systems using analytical redundancy. In: IEE Colloquium on Condition Monitoring and Fault Tolerance. London, UK, 11/06/1990.

Peralta, J., Fritz, H., Dragos, K., & Smarsly, K., (2022) Sensor fault diagnosis coupling deep learning and wavelet transforms. In: 13th International Workshop on Structural Health Monitoring (IWSHM). Stanford, CA, USA, 03/15/2022.

Qin, S. J. & Li, W. (1999) Detection, identification, and reconstruction of faulty sensors with maximized sensitivity. The Global Home of Chemical Engineers, 45(9), pp. 1963-1976.

Rao, A. R. M., Kasireddy, V., Gopalakrishnan, N., & Lakshmi, K. (2015) Sensor fault detection in structural health monitoring using null subspace–based approach. Journal of Intelligent Material Systems and Structures, 26(2), pp. 172-185.

Shen, D., Yang, D., Lyo, C., Ma, J., Hinds, G., Sun, Q., Du, L., & Wang, L. (2024) Multi-sensor multi-mode fault diagnosis for lithium-ion battery packs with time series and discriminative features. Energy, 290(2024), 130151.

Smarsly, K. & Petryna, Y. (2014) A Decentralized Approach towards Autonomous Fault Detection in Wireless Structural Health Monitoring Systems. In: Proceedings of the 7th European Workshop on Structural Health Monitoring (EWSHM). Nantes, France, 07/08/2014.

Smarsly, K. & Law, K. H. (2014) Decentralized Fault Detection and Isolation in Wireless Structural Health Monitoring Systems using Analytical Redundancy. Advances in Engineering Software, 73(2014), pp. 1-10.

Smarsly, K., Dragos, K., & Kölzer, T. (2022a) Sensor-integrated digital twins for wireless structural health monitoring of civil infrastructure [Sensorintegrierte digitale Zwillinge für das automatisierte Monitoring von Infrastrukturbauwerken]. Bautechnik 99(6), pp. 471-476.

Smarsly, K., Worm, M., & Dragos, K. (2022b) Design and validation of a mobile structural health monitoring system based on legged robots. In: Proceedings of the 29th International Workshop on Intelligent Computing in Engineering (EG-ICE). Aarhus, Denmark, 07/08/2022.

Steiner, M., Legatiuk, D. & Smarsly, K. (2019) A support vector regression-based approach towards decentralized fault diagnosis in wireless structural health monitoring systems. In: Proceedings of the 12th International Workshop on Structural Health Monitoring (IWSHM). Stanford, CA, USA, 09/10/2019.

Yan, K., Zhang, Y., Yan, Y., Xu, C., & Zhang, S. (2020) Fault diagnosis method of sensors in building structural health monitoring system based on communication load optimization. Computer Communications, 159(2020), pp. 310-316.

Yu, C. B., Hu, J. J., Li, R., Deng, S. H., & Yang, R. M. (2014) Node fault diagnosis in WSN based on RS and SVM. In: 2014 International Conference on Wireless Communication and Sensor Network (IEEE). Wuhan, China, 12/14/2014.

Zhang, Z., Mehmood, A., Shu, L., Huo, Z., Zhang, Y., & Mukherjee, M. (2018) A survey on fault diagnosis in wireless sensor networks. IEEE Access, 6(2018), pp. 11349-11364.

ESTIMATE ROAD ROUGHNESS USING SAMRTPHONE RESPONSE DATA - A SEMI-SUPERVISED LEARNING APPROACH

Qiqin Yu[1], Ye Sang[1], Yihai Fang[1], Viet Vo[2], and Richard Wix[3]

[1]Department of Civil Engineering, Monash University, Melbourne, Australia
[2]Faculty of IT, Monash University, Melbourne, Australia
[3]National Transport Research Organisation, Melbourne, Australia

Abstract

The smartphone-collected vehicle response data is being used to estimate the International Roughness Index (IRI). Among the existing smartphone-based methods, the data-driven approach, which involves training a machine learning model, is drawing more attention. However, surveying the IRI using conventional methods is expensive and there is limited labelled response data. In contrast, there exists a wealth of unlabelled response data from road users. This scenario presents opportunities for exploring semi-supervised learning (SSL) algorithms, which have been insufficiently researched in this domain. This study addresses this gap by applying an SSL framework to refine an IRI estimation model. Our results show that the SSL-trained model achieves a lower RMSE than the fully supervised trained model.

Introduction

Pavement deteriorates under the effects of traffic loading and environmental conditions, resulting in defects such as potholes and cracks. These failures affect ride quality and increase the likelihood of vehicle damage and safety risks. A road roughness index indicates the condition of the road, and the most prevalently used one is the International Roughness Index (IRI) (Yu et al., 2022). Conventional roughness index estimation (RIE) systems are labour-intensive, and require a complex set-up Qiqin Yu and Wix (2023). As a result, these systems are limited in surveying frequency and spatial coverage. Studies have proposed utilising a crowdsourcing-based approach to tackle these issues, since the proliferation of smartphones presents an opportunity to leverage data from road users (Staniek, 2021). The embedded sensors, such as the accelerometer and gyroscope, could be exploited to characterise road pavement based on the traversing vehicles' responses. Such a collaborative assessment approach could potentially supplement conventional instruments by locating distressed road segments more responsively and cost-effectively.

Existing studies have focused on increasing the accuracy of smartphone-based systems by applying statistical-based methods, vehicle mechanistic model-considered and machine learning methods (Yu et al., 2022). However, a dominant difference between the smartphone-based method and the conventional approaches is the amount of data generated. The data generated from the smartphone-based approach is thousands of folds to that of the conventional approaches. Many vehicles are driving on relatively fixed routes (e.g., from home to work), and the response data from these vehicles could effectively indicate the pavement roughness condition. In practice, collecting data in road segments labelled by ground-truth IRI is expensive and likely unavailable in remote regions. As a result, only a small amount of smartphone response data could be mapped with the ground-truth IRI. Meanwhile, there is a large amount of useful response data driving on road networks whose IRI is unknown and such data becomes unlabelled. It is believed that such response data should also be exploited to contribute to the IRI estimation model. Essentially, only a small percentage of samples are labelled while the majority remains unlabelled, and this scenario presents opportunities for exploring semi-supervised learning (SSL) algorithms and little research has touched this area. To fill the gap, this study proposes an SSL-based framework for estimating IRI by considering both labelled and unlabelled vehicle response data.

The rest of the paper is organised as follows. The Literature Review section summarises relevant studies in SSL and smartphone-based roughness index estimation (sRIE) methods. Next, the Methodology section elaborates on the proposed framework for IRI estimation using SSL. The Data Collection and Data Pre-processing sections provide information on obtaining and preparing data for training. The Experiment section details the training of the initial IRI-estimation model and the implementation of the SSL technique. Lastly, the Discussion section breaks down the contributions and limitations of this study and sheds light on directions of future research.

Literature review

Semi-supervised learning

As the quantity of available data grows and the hardware computational capabilities grow, deep learning (DL) has been making significant contributions in various applications (Pouyanfar et al., 2018). However, the prerequisite of these encouraging results usually relies on a large amount of labelled data, which is often expensive and difficult to obtain compared with unlabelled data (Ouali et al., 2020). Hence, semi-supervised learning (SSL) was proposed to leverage a volume of unlabelled data to improve the DL model performance based on a small number of labelled data. Specifically, with respect to loss function and model design, the SSL methods could be categorised into 5 groups, deep generative methods, consistency regularization methods, graph-based methods, pseudo-labeling methods, and hybrid methods (Yang et al., 2022). For example, given labelled data, the typical Generative Adversarial Networks (GAN), as described in

Ian et al. (2020), consist of two neural networks, a generator, and a discriminator, and were supposed to generate new samples through adversarial training. As a consistency regularization method, Mean Teacher (Tarvainen and Valpola, 2017) leverages consistency regularization to train the model with both high accuracy and strong robustness by integrating Exponential Moving Average (EMA) during the training step. Meanwhile, the manifold assumption, a fundamental principle in semi-supervised learning (Learning, 2006), has been employed by GoodBadGNN to enhance the model's inference performance (Dai et al., 2017). Pseudo-label claimed a straightforward and effective approach for training a robust deep learning model by initially training it on labelled data and subsequently retraining the model with high-confidence predictions on unlabelled data (Lee et al., 2013). Moreover, since most semi-supervised learning methods rely on unique structure in the domain data, a hybrid method, the VIME (Value Imputation and Mask Estimation) was proposed for tabular data that does not have explicit structure as image data (Yoon et al., 2020). Overall, SSL methods are used to learn the underlying pattern from unlabelled datasets and hence to augment the labelled dataset. Specifically, Consistency regularization methods directly improve the performance and robustness of the DL model by contrasting learning. Pseudo-labeling enhances model performance by initially training on a limited labeled dataset, followed by generating predicted labels with high confidence from an unlabeled dataset, which are subsequently integrated into the training process to retrain the initial deep learning model.

In terms of using semi-supervised learning in data-driven approach, it didn't apply any SSL algorithm to leverage the unlabelled response data to improve the performance of an IRI estimation model. To the author's knowledge, no other studies have applied SSL algorithms in estimating the IRI.

Crowdsourcing-based infrastructure health monitoring

Estimations made by a single smartphone may be inadequate in accuracy and consistency. Crowdsourcing smartphone measurements aim to aggregate measurements from a large number of participants. Big data enables analysis of the impact of practical factors, including speed, vehicle types, smartphone models, and mounting configuration/position, on the response-based IRI estimation method. Li and Goldberg (2018) proposed crowdsourcing systems to evaluate pavement roughness. Recently, Jeong and Jo (2023) implemented a validated IRI-estimation approach using response data collected from 29 vehicles, 8 smartphone models, and 5 mounting types. By considering various practical settings, their study was conducted in a controlled environment and the participants surveyed a single route of 18.2 km.

Methodology

This section provides an overview of the proposed framework for estimating the IRI by incorporating the profiler, sRIE systems and the crowdsourcing participants. The framework contains two modules. The first module features a multilayer perceptron (MLP) model for estimating the IRI, and the second module engages crowdsourcing data collection and leverages an SSL framework to improve the accuracy of the trained model. A schematic

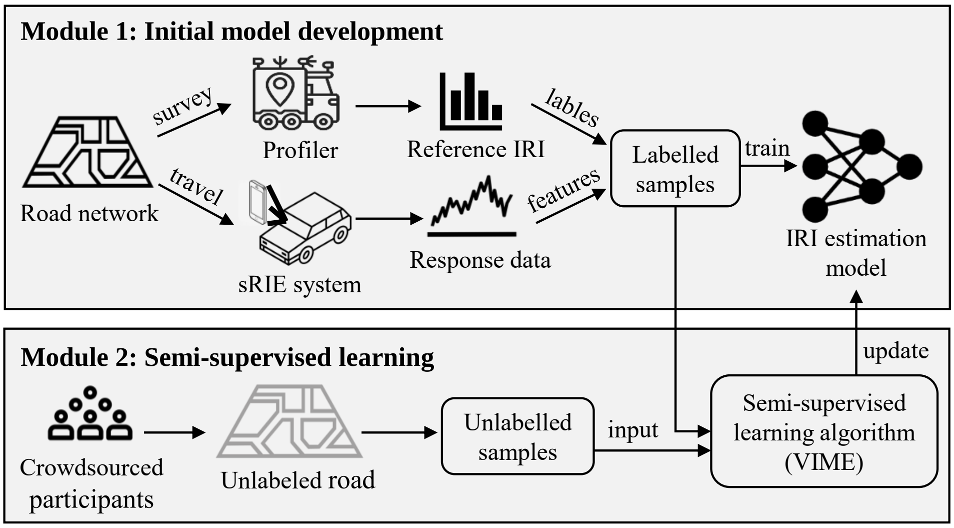


Figure 1: Semi-supervised learning-based IRI estimation schematic framework

dynamic response data was recorded using win
mounted smartphones. A smartphone App was d
to collect the Inertial measurement unit (IMU) a
data. The obtained response data is essentially
series of triaxial acceleration and is delineated in
ent road segments based on the GPS. Features fr
statistical and frequency domains were then ext
train an IRI estimation model using MLP.
In Module 2, since not the entire road networks w
veyed with the profiler, a large amount of unlabe
is generated when the participants travel on a roa
ground-truth IRI is unknown. Therefore, route
vided into two groups. Road segments that were
by the profiler are called labelled roads since their
truth IRI is known and the vehicle's response data
routes are labelled data. In contrast, other roads a
unlabelled roads and their vehicle response da
labelled. The SSL makes use of a small amou
belled data with a large amount of unlabelled d
et al., 2021), and increases the accuracy of the MI
developed in Module 1. This study adopts the
framework "VIME (Value Imputation and Mask Estimation)" that employs a data augmentation method, generating augmented samples for each feature set (Yoon et al., 2020). Specifically, unlabelled data are fed into VIME, which learns the underlying patterns within features. To achieve this, the original feature matrix X is shuffled to create a new feature matrix $\hat{X}$. A mask matrix M, with the same dimensions as the feature matrix X, is generated based on the Bernoulli distribution, and its counter-

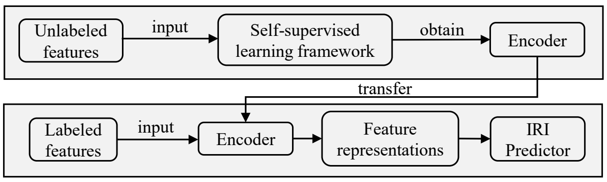


Figure 2: Semi-supervised learning method (VIME)

Data collection

This section details the experiment routes and set-up of the data collection instruments, including the profiler and the smartphone-based response data collection system.

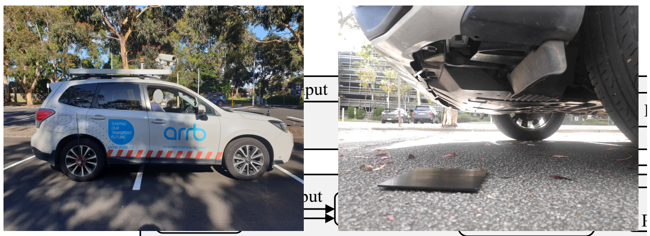
Survey routes

As shown in Figure 3, six routes around the Clayton campus of Monash University, in Melbourne, Australia, were selected. The routes collectively span a total length

Figure 3: Experiment survey routes

Ground-truth IRI survey instrument

The ground-truth IRI was surveyed using an inertial profiler system from the National Transport Research Organisation (NTRO). This instrument was certified in accordance with the Austroads standards on profiler validation (Austroads, 2016). The inertial profiler consists of a laser profilometer and accelerometers. The laser profilometer measures the distance to the road surface while the self-movement of the vehicle body is integrated from the accelerometer measurements. Knowing the initial height of the laser profilometer, the roughness elevation profile can then be calculated. The IRI is then computed by running the Golden-Car simulation on the measured profile. Two repetitive runs were conducted on all routes to obtain robust ground-truth IRI measurements which were sampled at a 10m interval. The distribution of the ground-truth IRI values in five survey routes is illustrated in Figure 3.

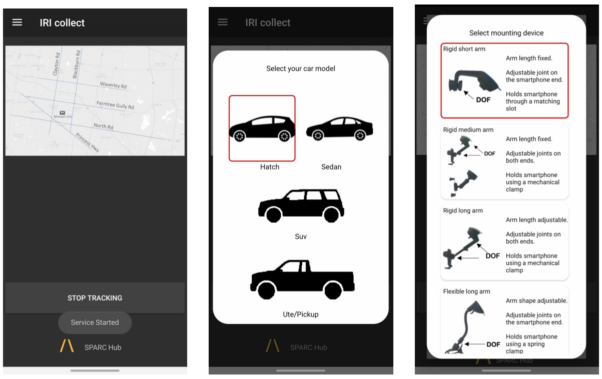


(a) Survey vehicle (b) Profiler (under front bumper)
Figure 4: Ground-truth IRI instrument

Table 1: Ground-truth IRI distribution

IRI Band	Quantity	Percentage
0-2	44,665	45.8%
2-4	44,213	45.4%
4-6	6,823	7%
6-8	1,445	1.5%
8-10	216	0.2%
>10	123	<0.1%

Figure 5: Smartphone IMU sensor axial orientations.



(a) Home (b) Vehicle select (c) Mount select

Figure 6: Data collection App user interface.

Smartphone-based response data collection system

The vehicle response data is collected from common passenger car models of different body types. Four smartphones of the same model, each held using a different mount, were deployed to collect the vehicle response data, as demonstrated in Figure 5. The four mountings were different in their arm length and mechanism type and represented the range of commonly available models in the market. Meanwhile, an Android application was developed to collect the smartphone's IMU and GPS data. The collected data includes the timestamp, tri-axial accelerometer, tri-axial gyroscope, rotational sensor, GPS and speed. The App interfaces are shown in Figure 6. The user's vehicle body type and mounting type were also recorded in the App.

Data preprocessing

Data preprocessing contains two tasks, data mapping and data cleaning. Data mapping assigns the smartphone measurements with the ground-truth IRI. Each acceleration measurement now has its own latitude and longitude records. Meanwhile, the latitude and longitude of the start and end points of each 10m ground-truth IRI segment are known. By selecting the smartphone data points that are spatially closest to the reference start and end points of each segment and knowing their timestamp, the time window that corresponds to the reference segment could be determined. The mapping was completed by assigning the smartphone data points in this time window with their corresponding ground-truth IRI.
The next step is data cleaning. The raw smartphone sensor data contains invalid records that must be removed. Among the collected 1.82×10^7 sensor data, the number of "NaN" in the accelerometer, gyroscope and rotation sensors are 6.06×10^3, 1.02×10^3 and 1.38×10^6, respectively. These "NaN" records were imputed with the local mean, which was calculated from data points in the same segment where the "NaN" record sits. Moreover, it is worth noting that the profiler does not record an IRI when the travel speed is less than 20km/h. Hence, there was no ground-truth IRI (label) for some segments and as a result, the corresponding smartphone data was also removed. In addition, road roughness may not excite movement of the vehicle body when the travel speed is low. In this study, the minimum speed was set as 20km/h and response data recorded below this speed was also excluded.

Experiment

Multilayer perceptron training - Module 1

This section elaborates on the training of the MLP-based IRI-estimation model. Statistical and frequency features were extracted from the signals collected from the tri-axial accelerometer, the gyroscope, and the rotational sensor. Statistical features were chosen according to related empirical work (Souza et al., 2018), including mean, median, average magnitude value, standard deviation, zero-crossing rate, skewness, kurtosis, entropy, maximum-to-minimum difference, root mean square level, peak magnitude to RMS ratio, the mean number of peaks, crest factor and IQR (Interquartile Range). On the other hand, after applying Fast Fourier Transform (FFT) to the time-domain signal, 6 frequency features including spectral entropy, spectral irregularity, flux, bandpower, max peak and max peak location were computed. Additionally, categorical features such as car model and mounting device type were also included in the feature set. Consequently, when provided with a label (IRI value), the input to the DL model comprises a vector with a dimension of 258, encompassing 256 numerical features and 2 categorical features.

In the training process, the Adam optimiser was employed with a parameter learning rate of 1e-3. The networks underwent training with a batch size of 128, a seed value of 777 for reproducibility, and a total of 50 epochs. The framework used to construct the model in this study was TensorFlow (TF). In this study, 80% of the pre-processed data was allocated for training the model, while 10% was reserved for validation to fine-tune the model and optimise hyperparameters. The remaining 10% of the data was kept for testing. Different model structures were trialled and the fully connected layers with neurons of 128-256-512 performed the best. The model achieved an root-mean-square error (RMSE) of 0.57, as shown in Figure 7. The R-square of the linear regression between the estimated and ground-truth IRI was 0.79, as demonstrated in Figure 8. The calculation of RMSE is shown in the below formula.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

$\boldsymbol{n}$: The number of samples,
$\boldsymbol{y_i}$: Ground-truch IRI ,
$\boldsymbol{\hat{y}_i}$: Predicted IRI.

Semi-supervised learning - Module 2

Besides the labelled data, unlabelled data is generated when vehicles drive on road networks whose ground-truth IRI is unknown. Based on these data, VIME was adopted as the semi-supervised learning methodology. The hyper-parameters (learning rate, p_m, and alpha) and output dimension were subsequently refined to align more suitably with the objectives of our work. After applying the SSL algorithm, the performance of the original DL model developed in Multilayer perceptron training - Module 1 improved, as evidenced by the lower RMSE, indicated in Figure 9.

A systematic validation procedure was undertaken to evaluate the efficacy of VIME. A predetermined label rate, ranging from 30% to 90%, was set to extract a labelled subset of the initial entire labelled dataset. The selected labelled subset was then employed to train the supervised model elaborated upon in Section Multilayer perceptron training - Module 1. Simultaneously, the remaining data was regarded as unlabelled data, and the encoder in the VIME algorithm was trained using this unlabelled data to learn the underlying feature patterns and generate a novel representation for each sample. Then, the labelled subset data along with the representations of the original unlabelled subset were both used to retrain the model. This semi-supervised trained model was then compared with the model trained exclusively using the labelled data subset. This study trialled different label rates ranging from 30% to 90% with a step of 10%, and the RMSE values of the two streams of models are plotted in Figure 9. It is shown that both lines go downwards as the label rate increases, indicating that the performance of both models improves with more labelled data. However, it should be noted the SSL-trained model consistently sits marginally below the fully supervised-trained model, suggesting that the SSL-trained model achieves a lower RMSE, and therefore shows a better performance across the entire label rate range. It should be noted that the improved performance is not solely attributed to using more data, and the implication is that leveraging both labelled and unlabelled data to capture underlying patterns more effectively improves the model's generalisation and performance on unseen data.

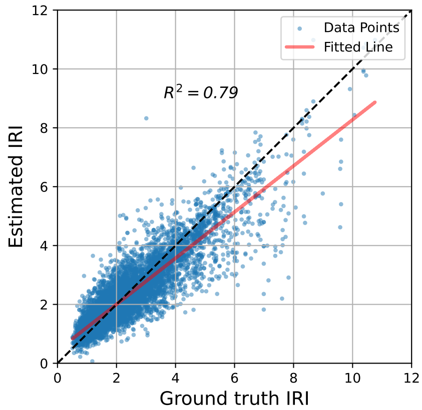


Figure 8: Estimated IRI vs Ground truth IRI datapoints and regression line.

Discussion

This section summarises the contributions of the proposed SSL framework, discusses the limitations of the methods developed, and suggests the areas of future research that promote the application of data-driven sRIE systems.

Contributions

Our study advances the field of data-driven road roughness evaluation, by introducing an SSL framework that leverages real-world data. Key contributions include:

Considering the real-world practical set-up challenges.

This is the first study to simultaneously utilise real-world data while accounting for variations in car models and

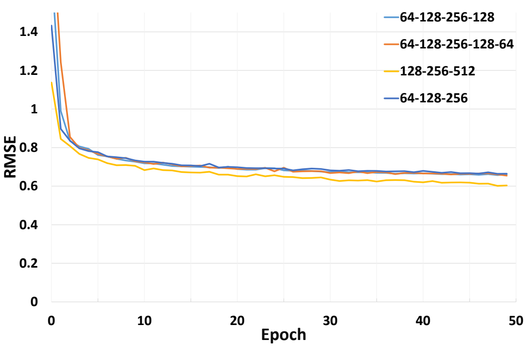


Figure 7: The MSE vs Epoch plot between different MLP model structures using validation dataset.

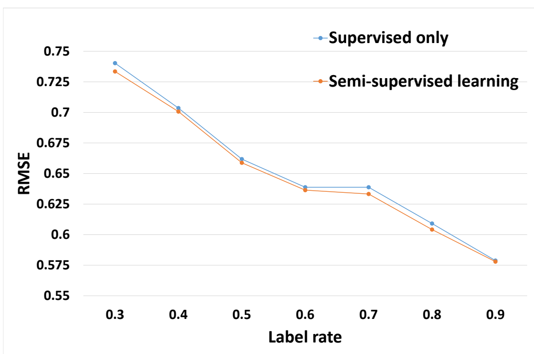


Figure 9: RMSE of the fully-supervised and semi-supervised models in different label rates.

mounting devices, unlike previous studies that worked on simulated vehicle responses or involved single vehicle and mounting configuration.

Adopting SSL to leverage unlabelled data.

Extending the application of SSL in data-driven infrastructure health monitoring, our study validates the idea of using limited labelled data while leveraging the crowdsourced unlabelled data to improve the model's performance, as demonstrated by the lower RMSE values. It's important to note that this improvement partially results from the inclusion of additional unlabelled data, which results in a more comprehensive dataset than what is available to fully-supervised models. Our findings underscore SSL's potential for scalable and cost-effective infrastructure monitoring, especially in contexts where labelled data is limited.

Setting a benchmark for SSL-based pavement roughness evaluation.

By integrating semi-supervised learning with real-world data collection, our proposed framework utilises both specialised survey instruments and public vehicles, and the SSL method yielded a lower RMSE than the model trained with full supervision. This study sets a benchmark for future studies exploring this field and contributes to the knowledge of a scalable and cost-effective approach for real-time pavement roughness monitoring.

Limitations

Limitations of this study are acknowledged:

Ground-truth data imbalance.

As an inherent property of the experiment route's roughness, the ground-truth IRI data is imbalanced, as illustrated in Table 1. Specifically, the ground-truth data sit mostly in the range of 0-4 mm/m and lacks in high IRI ranges (above 6 mm/m) and such imbalance limits the SSL algorithm's performance. Since the SSL algorithms involve learning the patterns of the unlabelled features, and due to a lack of response data for high-IRI road segments, there is not enough data for the algorithm to learn the patterns from. As a result, the SSL can exhaustively learn the complete patterns of features in the low-IRI range, and improves the model's estimation accuracy in this range better than that on the high-IRI range.

Limited semi-supervised learning method.

This study only explored one SSL method, namely VIME. While VIME has shown promising results in our context, our study struggled to find other open-sourced SSL methods that work on tabular data features, instead of the image data, which is predominantly studied in SSL. The improvement in the model's accuracy achieved in this study using VIME does not represent the benefits that generic SSL methods could provide in this application context.

Resource intensity of Semi-supervised learning.

Under the context of crowdsourcing-based road roughness monitoring, crowdsourced smartphone sensor data could become resource-intensive, in terms of computational power and data storage. Moreover, SSL techniques involve iterative mutation of the raw data, posing challenges for scalability and real-time data processing, which could be crucial for data-driven monitoring approaches.

Future works

Implementation using crowdsourced data

Our research has validated the effectiveness of using SSL techniques to leverage unlabelled data in training the IRI estimation model. In the next stage, the team plans to engage crowdsourcing participants to collect unlabelled data in real-world scenarios. It is expected that more variations exist in the public participants' practical set-ups and the response data collected under such conditions contains underlying patterns that could be learned by self-supervised and semi-supervised learning algorithms. Consequently, by accounting for these variations, a more robust and adaptable estimation model could be developed.

Dependency on participant compliance

Crowdsourcing relies on the active and correct usage of our data collection app by participants and the robustness of collected data depends on the participant's compliance with collection rules. Such an issue, which exists in crowdsourcing-based research studies, becomes more critical when the scale of the data collection grows. Hence, strategies to enhance participant compliance and retention become necessary. These strategies could may more engaging App user interfaces, regular reminders, incentives aligned with long-term participation, etc., and can be explored in future research.

Incorporating connected vehicles in RIE

While this study adopted smartphones as the roughness data collector, the prevalence of connected vehicles has drawn research focus. These vehicles are equipped with GPS and motion sensors and the data is uploaded to the manufacturer's servers, providing a rich data source. However, the integration of data from connected vehicles in the proposed SSL framework poses many challenges. One of the primary concerns is addressing inaccuracies that may arise from differences in vehicle types. Identifying and rectifying these inaccuracies is crucial to ensure the validity of roughness monitoring using this approach.

Applying other SSL methods

A comparative analysis of other SSL techniques could be explored in future research. Such exploration would also enable a more comprehensive understanding of the strengths and limitations of various semi-supervised learning strategies in this application context.

Conclusion

This study proposed an SSL-based framework for IRI estimation and validated its effectiveness using real-world data. Smartphone-collected vehicle response data was collected from different mounting configurations and vehicle models. An MLP-based IRI estimation model was first trained and an SSL algorithm (VIME) was adopted to learn patterns from the unlabelled data and to generate new representations of the data features used to predict the IRI. The results suggest that the SSL-trained model achieves a lower RMSE than the fully supervised trained model, when the same amount of labelled data was provided. Our study adds to the body of knowledge in smartphone-based road roughness monitoring by 1) developing an IRI-estimation model trained using real-world vehicle responses, considering the vehicle and mounting variations, 2) proposing an SSL framework that leverages both labelled and unlabelled response data, 3) improving the accuracy of an existing model by using SSL techniques, and 4) validating the benefits of incorporating SSL in crowdsourced pavement roughness monitoring.

Acknowledgments

This research work is part of research projects (Project No. 2.7 and No. 3.4) sponsored by the SPARC Hub (https://sparchub.org.au) at the Department of Civil Engineering, Monash University funded by the Australian Research Council (ARC) Industrial Transformation Research Hub (ITRH) Scheme (Grant number: IH180100010). The financial and in-kind support from the National Transport Research Organisation (NTRO) and Monash University is gratefully acknowledged. Also, the financial support from ARC is highly acknowledged.

References

Austroads (2016). Validation validation of an inertial profilometer for measuring pavement roughness (loop device method).

Dai, Z., Yang, Z., Yang, F., Cohen, W. W., and Salakhutdinov, R. R. (2017). Good semi-supervised learning that requires a bad gan. Advances in neural information processing systems, 30.

Ian, Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., and Bengio, Y. (2020). Generative adversarial networks. Communications of the ACM, 63(11):139–144.

Jeong, J.-H. and Jo, H. (2023). Toward real-world implementation of deep learning for smartphone-crowdsourced pavement condition assessment. IEEE Internet of Things Journal, pages 1–1.

Learning, S.-S. (2006). Semi-supervised learning. CSZ2006. html.

Lee, D.-H. et al. (2013). Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks. In Workshop on challenges in representation learning, ICML, volume 3, page 896. Atlanta.

Li, X. and Goldberg, D. W. (2018). Toward a mobile crowdsensing system for road surface assessment. Computers, Environment and Urban Systems, 69:51–62.

Liu, C., Wu, D., Li, Y., and Du, Y. (2021). Large-scale pavement roughness measurements with vehicle crowdsourced data using semi-supervised learning. Transportation Research Part C: Emerging Technologies, 125:103048.

Ouali, Y., Hudelot, C., and Tami, M. (2020). An overview of deep semi-supervised learning. arXiv preprint arXiv:2006.05278.

Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M. P., Shyu, M.-L., Chen, S.-C., and Iyengar, S. S. (2018). A survey on deep learning: Algorithms, techniques, and applications. ACM Computing Surveys (CSUR), 51(5):1–36.

Qiqin Yu, Y. F. and Wix, R. (2023). Evaluation framework for smartphone-based road roughness index estimation systems. International Journal of Pavement Engineering, 24(1):2183402.

Souza, V. M., Giusti, R., and Batista, A. J. (2018). Asfault: A low-cost system to evaluate pavement conditions in real-time using smartphones and machine learning. Pervasive and Mobile Computing, 51:121–137.

Staniek, M. (2021). Road pavement condition diagnostics using smartphone-based data crowdsourcing in smart cities. Journal of Traffic and Transportation Engineering (English Edition), 8.

Tarvainen, A. and Valpola, H. (2017). Mean teachers are better role models: Weight-averaged consistency targets improve semi-supervised deep learning results. Advances in neural information processing systems, 30.

Yang, X., Song, Z., King, I., and Xu, Z. (2022). A survey on deep semi-supervised learning. IEEE Transactions on Knowledge and Data Engineering.

Yoon, J., Zhang, Y., Jordon, J., and van der Schaar, M. (2020). Vime: Extending the success of self-and semi-supervised learning to tabular domain. Advances in Neural Information Processing Systems, 33:11033–11043.

Yu, Q., Fang, Y., and Wix, R. (2022). Pavement roughness index estimation and anomaly detection using smartphones. Automation in Construction, 141:104409.

MONITORING CONCRETE POURING WITH KNOWLEDGE GRAPH-ENHANCED COMPUTER VISION: A CASE STUDY FROM MUNICH, GERMANY

Fabian Pfitzner[1], Songbo Hu[2], Alex Braun[1], André Borrmann[1], and Yihai Fang[2]
[1]Chair of Computational Modeling and Simulation, Technical University of Munich, Germany
[2]Department of Civil Engineering, Monash University, Australia

Abstract

This paper introduces a novel approach for monitoring concrete pouring. Traditional manual tracking methods are tedious, while automated solutions, such as Computer Vision (CV)-enabled methods, are challenged with occulted data and limited adaptability to diverse crane behaviour patterns. We propose a knowledge graph-enhanced CV method that combines context knowledge with object recognition. This approach analyses tower crane behaviours and their interactions with workers, truck mixers, and building elements, providing a detailed and resilient interpretation of concrete pouring progress. Preliminary findings reveal the method's capacity to interpret incomplete data and comprehend complex site dynamics, demonstrating promising potential in a real-world scenario.

Introduction

Concrete pouring is a common and critical construction activity, significantly impacting both the completion time and cost of construction projects (Wang et al., 2022). The crane plays a pivotal role in this activity, as the "crane-and-skip" method emerging as one of the most prevalent techniques for concrete pouring (Lu et al., 2003). In this process, concrete is mixed and then poured into the skip on the ground by workers; then the crane lifts the skip to one or more locations requesting concrete; once in position, the skip is tilted or manipulated to pour, before the empty skip is lowered back to the mixer for refill.

Traditionally, monitoring the progress of concrete pouring processes has been manual and approximate, where the total volume of concrete delivered to the site is used as an indirect indicator of pouring progress (Lu & Anson, 2004). This method merely provides a rough estimate of the progress, and cannot capture nuances related to the pouring process, such as the cycle time of crane lifts and the waiting time of truck mixers. Consequently, it provides limited insight into the bottlenecks of critical site resources (e.g., crane availability) and yields minimal contributions to crucial decisions that impact site productivity (e.g., hiring extra cranes) (Hu et al., 2021).

To gain insights into the pouring process, which requires a large volume of data, Computer Vision (CV) has emerged for automated data acquisition and analysis. For instance, Gong and Caldas (2010) developed a CV-based method to track crane hooks and concrete buckets (i.e., skips), enabling the analysis of concrete pouring states (e.g., bucket readiness, pouring into specific columns) and cycle times. Nevertheless, these CV-based methodologies often confront adaptability challenges. They rely on the assumption of continuous and uninterrupted concrete pouring, following an overly rigid crane behaviour pattern in each pouring operation (e.g., the buckets have to return to the mixer after each pouring). Meanwhile, these methods often necessitate project-specific parameters, such as designated loading and unloading zones, to streamline data interpretation (Yang et al., 2014).

These rigid rules in data analyses often misalign with the dynamic environment of construction sites, leading to inaccurate data interpretations and limited generalisability in different construction projects. Compounding this issue are hardware limitations, such as cameras missing crucial frames or occlusions blocking the view. These limitations will also disrupt the crane behaviour pattern recognition, leading to inaccuracies or frequent relocating of cameras.

Therefore, an adaptable reasoning mechanism is required that ensures enhanced robustness and greater generalisability for monitoring concrete pouring. Knowledge Graphs (KGs) offer significant potential to elevate the reasoning capabilities of CV systems (Fang et al., 2020). KGs enable machines to comprehend complex relationships and enriched contexts within data, going beyond mere visual recognition (Pfitzner et al., 2023a).

By integrating KG and CV, the authors propose a universally adaptable interpretation of concrete pouring processes. In later sections, the paper includes a literature review on concrete pouring monitoring and KG-enhanced CV systems in the construction industry, followed by a detailed description of our methodology and a case study from an active construction site in Munich. It concludes with a discussion of the method's contributions, limitations and potential improvements.

Literature review

Methods of Monitoring Concrete Pouring

Over the past decades, continuous research has focused on monitoring concrete pouring, with relevant studies summarised in Table 1. Initially, this monitoring was mainly for quality control, aiming to predict concrete curing conditions and strength development (Moon & Yang, 2017). The process involved manual documentation of the concrete pouring time into formwork and environmental factors, such as humidity, to ensure compliance with curing standards.

Table 1: Literature on Monitoring Concrete Pouring

Reference	Year	Tracked item	Tracking Mechanism	Collected data	Outputs	Limitation
(Lu & Anson, 2004)	2004	Concrete Mixer	Retrieving information from manually prepared documents	Site arrival /departure time; Location and element type for pouring; Pouring method.	A productivity evaluation (i.e., volume/time) for different pouring methods	A coarse granularity of tracking at a dozen of truck trips level; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Moon & Yang, 2010)	2010	Concrete Mixer and concrete pump	Radio-frequency identification device (RFID)	Site arrival time, mixer entering/ leaving work zone time.	Total amount of poured/remaining concrete; Average pouring speed; Estimated time to complete.	A coarse granularity of tracking at a truck trip level; Manual installation of RFID tags; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Gong & Caldas, 2010)	2010	Bucket	Computer Vision (CV)	Buckets' presence in user-specified waiting zone and column zones; Buckets' absence (indicating their loading).	Durations of each user-defined pouring states for a particular pouring activity.	Assuming uninterrupted pouring; The case study merely focuses on the pouring of two columns; Relying on fixed states in the column pouring and site-specific spatial context; Requiring meticulous/dynamic planning of camera locations.
(Yang et al., 2014)	2014	Crane jib	Computer Vision (CV)	Crane jib trajectory and its overlap with site layout.	State durations of concreting pouring and non-concrete pouring activities.	Assuming concrete mixers' locations are fixed; Replying on site-specific spatial context and observed probability when differentiating concrete pouring states.
(Danel et al., 2022)	2022	Crane	Crane inbuilt motion trackers and load sensor	Slewing angle, trolley distance, hoist height, and load weight.	State durations of concreting pouring activities.	Assuming uninterrupted concrete pouring process; Relying on a standard crane behaviour pattern.
(Wang et al., 2022)	2022	Crane cable and bucket	Computer Vision (CV)	Concrete buckets' presence in a meticulously planned imaged area (e.g., on the top of the dam to be poured).	Cycle time of crane lifts.	Assuming uninterrupted concrete pouring processes; Relying on a standard crane behaviour pattern; Requiring meticulous locating of camera; Not revealing delays in concrete pouring states and related resource bottlenecks.
(Kim et al., 2023)	2023	Sound	Classifying site sound with deep-learning algorithms to recognise concrete pouring	Acoustic signal from construction sites.	Start and finish time of concrete pouring; Abnormal concrete pouring such as the impact sound of the vibrator's formwork and the sound of concrete slowly leaking.	Focusing on the safety of concrete pouring; Not revealing productivity-related insights (e.g., cycle times).

Lu & Anson (2004) analysed hundreds of digital quality control records to explore the link between concreting speed and pouring methods. Their work provided significant insights for enhancing productivity through method selection. However, their reliance on manual data collection leads to a coarse tracking granularity (e.g., dozens of truck mixer trips) and restricts insights for site management.

The introduction of IoT devices creates a shift towards automated, sensor-based tracking methods. Moon & Yang (2010), for instance, employed RFID technology to monitor the movement of truck mixers on-site, enhancing tracking to a granularity of individual trips. Kim et al. (2023) used acoustic sensors combined with deep learning algorithms to pinpoint the state when concrete is poured into formwork. While it does not monitor every state of the concrete pouring process, such as loading the concrete, this method further refines the granularity for tracking to specific pouring states. Danel et al. (2022) further advanced the tracking by using in-built crane sensors to monitor every state of the concrete pouring process, thereby uncovering prolonged delays caused by resource miscoordination. However, IoT methods often struggle to gather context information efficiently, limiting their ability to correlate lift operations with overall construction progress.

Recent research has turned to CV for more nuanced monitoring. CV methods typically consist of two steps: object recognition and reasoning. Object recognition identifies concrete-related resources in images, while reasoning interprets the actions of these objects. Gong & Caldas (2010) were among the first to apply CV in this field. They detected concrete buckets and mapped them onto predefined work zones in the images to identify specific pouring states with an accuracy of 84.7%. However, the need for user input limited their method's applicability. In contrast, Wang et al. (2022) applied CV to dam construction, tracking concrete buckets and crane cable interactions to indicate pouring cycles, achieving an even higher accuracy (>99%). However, the method assumed an uninterrupted pouring process, which is unrealistic for building construction. Yang et al. (2014) focused on crane jib recognition to track crane movements in relation to site layout plans, allowing for the recognition of mixed crane operations (i.e., concrete pouring and other lifting activities). Yet, this required distinct separation of concrete and other materials' loading zones and relied on observed site-specific probabilities in the reasoning process, potentially limiting adaptability to other sites.

In addition to the generalisability issue, one of the main challenges with these CV methods is their struggle to correlate crane operations with specific building elements and to detect the particular causes of idling. For example, they were often unable to identify scenarios where concrete buckets were in position but left unattended. These limitations highlight the imperative need to develop a reasoning mechanism that can adapt to varied site-specific parameters (e.g., site layouts) and effectively highlight the interactions between concrete pouring resources (e.g., cranes, buckets, workers, and mixers) and construction products (e.g., building elements).

Leveraging the Knowledge Graph to Enhance CV in Construction Applications

Knowledge Graphs (KGs) are poised to significantly augment the reasoning capabilities of CV, providing a universally applicable, generalisable context. By focusing on universal heuristics, such as the spatial relationships among objects (e.g., workers), KGs actively process and interpret CV outcomes, transforming raw data into information and enabling the derivation of new insights. In the construction industry, KG-enhanced CV systems have shown particular promise in safety management. Fang et al. (2020) demonstrated this by developing a KG-based CV system that identifies potential hazards through the spatial relationships among workers, PPE, and heavy machinery. Lee and Yu (2023) employed a KG to standardise and identify safety hazards in mobile scaffolding use, effectively pinpointing common misuses, including absent outriggers or missing guardrails. Both KGs perform spatial analysis of relevant objects to identify unsafe conditions or behaviours.

Unlike unsafe behaviours that often occur in isolated instances, construction activities, such as concrete pouring, typically consist of multiple states following a specific sequence. Therefore, solely specifying spatial relationships may only allow for a rough productivity evaluation at a broader level (e.g., site level) rather than recognising specific construction activities (Pfitzner et al., 2023a). To link planned sequences to actual site conditions, Braun et al. (2020) proposed a graph-based method to derive construction sequences from BIM models, overlay the sequence with detected objects, and thereby compensate for data gaps caused by occlusions. This application underscores the capability of KGs to embed sequence information.

However, the sequence of concrete pouring states (i.e., lifting states) is more dynamic. Addressing this challenge, Hu et al. (2023) utilised a KG to organise heterogeneous data from motion trackers, weight sensors and images for recognising crane lifting states (e.g., pick-up, suspend). This approach embeds dependencies of states instead of a standard sequence to allow flexibility. Thus, it is able to understand complex operations, such as those with multiple unloads in one lift. Despite its effectiveness in crane operation recognition, Hu et al.'s approach falls short in differentiating concrete pouring from other lifting activities and in linking crane operations to specific building elements. Moreover, their method is not tailored for enriching CV data. This underscores the necessity for a specialised KG tailored for concrete pouring processes and CV data formats. Such a KG should be proficient in interpreting concrete pouring operations, distinguishing between various lifting activities and linking pouring operations to building elements seamlessly.

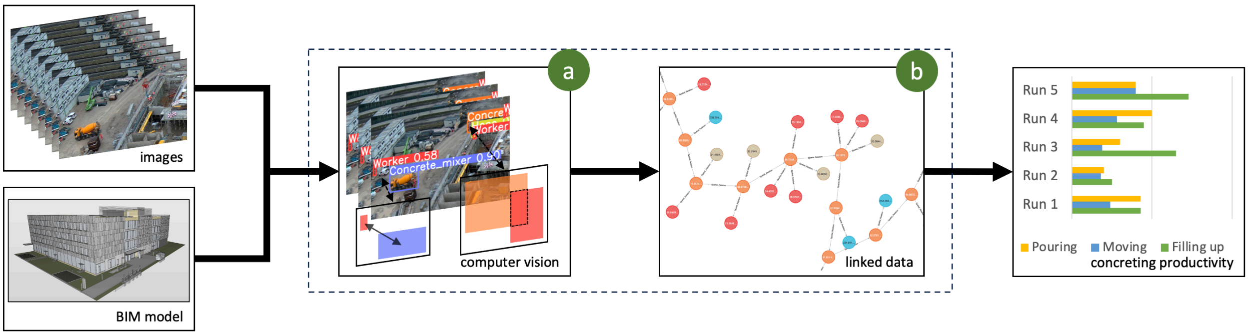


Figure 1: Overview of pipeline

Methodology

Overview of the approach

The scope of this study is to determine the utility of the crane in concrete pouring activities and associate crane operations with building progress. The underlying purpose of activity monitoring is to optimize the coordination of concrete delivery schedules, thereby mitigating overproduction and minimizing material waste during concrete pouring. Streamlining the concrete pouring process could condense construction timelines and elevate resource utilization.

We use semantic knowledge derived from the state dependencies in concrete pouring and BIM model to fuse CV object recognition results. As a result, we automatically monitor the pouring process in fine granularity (i.e., states), with the elapsed time of distinct process states recorded to provide precise insight into the required time for pouring diverse building zones. The results are anticipated to facilitate detailed control of construction progress.

The CV-based pipeline, shown in Figure 1, is developed to convert raw image and geometry data to precise process-level information. Instead of using data-heavy video streams, a low frequency of input images is chosen to avoid extensive computation times for processing long construction periods.

CV methods (a) are applied to extract information from the images. This step includes interpreting spatial dependencies in the context of the concrete pouring phase. In addition, the information is mapped to the BIM model using grid-based zones.

The knowledge graph (b) is used to determine unseen states and to link the as-planned geometry data. Unknown states are predicted based on their predecessors and successors. This predictive capability is particularly valuable in addressing data acquisition limitations, such as missing frames or occlusion, ensuring a robust analysis of individual pouring states while considering the as-planned quantities.

CV-based state recognition

The computer vision part contains three steps: Object detection, spatial-temporal reasoning, and geometric projection.

A YOLOv8 model is used for object detection, covering the following classes: concrete_bucket, hook, concrete_mixer, hose, and worker. The model is deployed on image sequences representing the concrete pouring states. Figure 2 illustrates the distinct states of the concrete pouring process: Loading the concrete bucket at the concrete mixer, moving the bucket to the building element, and filling the building element's formwork with concrete. To derive these instant states from the images, spatial reasoning is necessary.

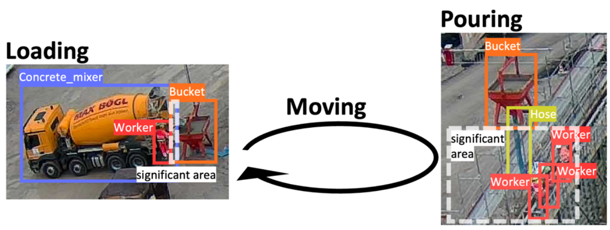


Figure 2: Spatial relationships of concreting cycle

In the case of crane-and-skip concrete pouring, the location of the concrete bucket and its relation to other elements defines the current action. Fig. 2 shows the essential scenarios of the concrete bucket to determine the process state: "Loading", "Moving", and "Pouring". At the start of the process, the concrete bucket moves close to the concrete mixer to be filled by a worker. The bucket stays for a period of time (e.g., 30s) and then gets lifted by the crane. Intersection over Union (IoU) is used to determine the bounding boxes' spatial dependency.

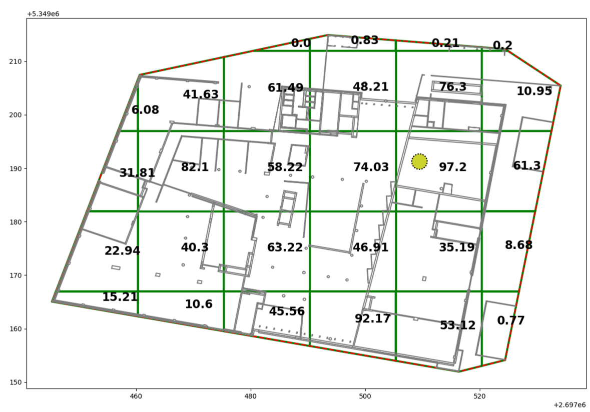


Figure 3: As-planned zones and volumes [m³] derived from the BIM model and projected base point of the hose (light green).

Subsequently, the bucket moves to the building element, and the pouring state starts. During the pouring state, a hose is rolled down, and a worker controls the nozzle of the hose to fill the formwork. Like in the "Loading" scenario, the activity is derived by interpreting the spatial relationships of the bounding boxes using IoU. During the "Pouring" state, workers and the hose are within the area underneath the bucket. The "Moving" state is computed based on the known order of process steps, discussed in the following section.

To compare the as-performed pouring volumes to the as-planned volumes, the image information is enriched. First, the as-performed quantities are computed using known volumes of the concrete mixer and bucket. Second, the as-planned volumes are generated using the BIM model, illustrated in Fig 3. The quantities of the BIM model are computed storey-wise to specific zones of a grid. The grid contains the quantity of all building elements within one zone. The location of the hose is projected to the BIM model using a perspective transformation approach presented by the authors in previous publications (Pfitzner et al., 2023b). As the hose traverses a specific zone, the as-performed data gets linked with the as-planned data. Based on the expected concrete amount of a specific zone, the amount of poured concrete is validated. Moreover, the amount of wasted concrete can be detected. A detailed investigation of geometry mapping is outside this paper's scope and will be discussed by the authors in future publications.

Ontological model for process reasoning

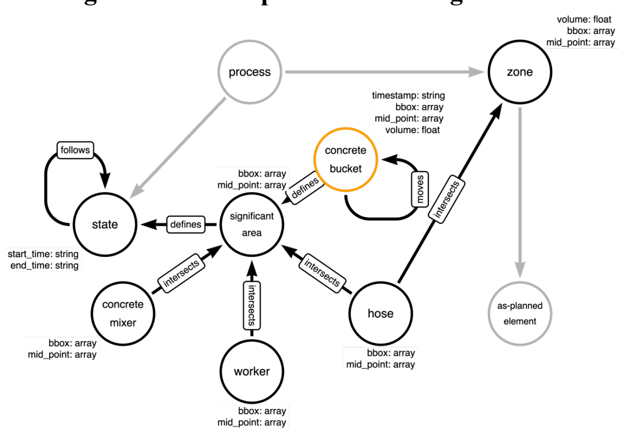


Figure 4: Ontological model

The ontological model, shown in Fig. 4, is designed according to the states representing the concrete pouring process. The process's states are defined by the time-dependent actions of the concrete bucket. As such, the bucket nodes have a timestamp property and a relationship to their predecessors and successors. This ensures that the bucket's current, past, and future actions can be determined based on the bucket's significant area. The significant area is defined by typical construction-related dimensions; it encompasses the space beneath the bucket and incorporates a buffer, considering the objects utilising the bucket manifest beneath it.

Unseen states are parts of the process that cannot be detected on the frames. This is the case when the concrete bucket moves. The predecessor and successor relationships are utilised to compute the missing information. The unseen states of the concrete pouring cycle are determined based on the four different relationship patterns shown in Fig. 5. If the detected state changes, the bucket's state is considered to be "Moving" in between, encompassing both scenarios "moving from mixer" and "moving from building elements". In addition to that, the appearance of the concrete mixer is an indicator of pouring. If there is no mixer on-site, the crane is not involved in the pouring process.

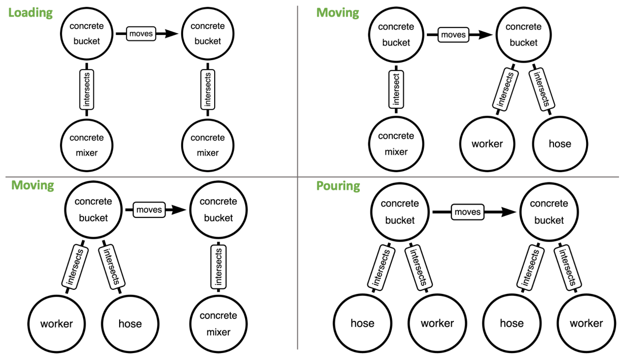


Figure 5: Relationship patterns for process reasoning

Once the relationship patterns are detected in the graph, the state information is stored in the equivalent state nodes. Subsequently, the process chain is created by sequential states containing start and end times. This has the advantage of querying the elapsed times straightforwardly. The start of the process chain is defined by the first time the concrete bucket gets loaded by the concrete mixer. The process chain ends once the bucket no longer returns to the concrete mixer and has moved away from the element. The link to the as-planned zones is created through the projected tip of the hose once it intersects with the corresponding zone. The process chain is used to investigate the elapsed time of individual states, the number of cycles, and the amount of concrete used during the process.

Case Study

The case study, shown in Fig. 6, was conducted on a building construction project near Munich, Germany. The building was partially constructed using cast-in-place concrete pouring.

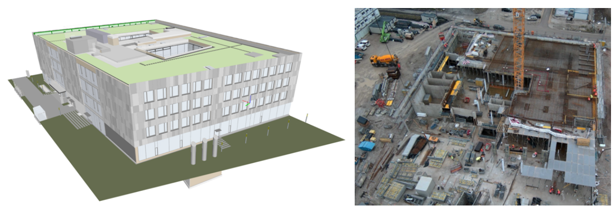


Figure 6: BIM model and crane camera perspective from the construction project

Data and setup

The images were taken every 30 seconds from fixed crane cameras at diverse heights and perspectives. The dataset consists of a total number of 270k images. The annotated image dataset containing 325 images was split by 80/20. An additional model trained on the MOCS dataset (Xuehui et al., 2021) was included in the pipeline to detect workers. The model training on a Nvidia RTX 8000 GPU took 0.98 hours. To receive the as-planned geometry, we used an existing IFC model of the building. A Neo4j server was set up in a docker environment to host the labeled property graph. Four concrete pouring examples were investigated with a total number of 623 images.

Results

The YOLOv8 model was trained on 79 epochs and reached an mAP score of 92.1 %, as shown in Table 2. The hose class performed weaker than the other classes due to fewer occurrences in the dataset.

Table 2: Object detection results

class	Precision	Recall	$mAP^{0.5}$
bucket	92.6%	98.3%	97.7%
hook	97.6%	96.9%	99.2%
concrete mixer	98.1%	97.4%	97.8%
hose	87.3%	77.8%	73.5%
all	93.9%	92.6%	92.1%

The model was deployed batch-wise using a self-implemented PyTorch-based environment. The extracted image information was inserted into the Neo4j graph database using the Python library neomodel and the ontological model, illustrated in Figure 4. Using the distinct relationship patterns, the state nodes were generated and integrated into the graph. Based on the start- and end-time attributes of the state nodes, the elapsed time was queried. The buckets' bounding boxes, timestamps, and the particular concrete pouring states were annotated across two datasets to evaluate the ground truth of all states. The first dataset contained 119 samples; the second dataset contained 184 samples.

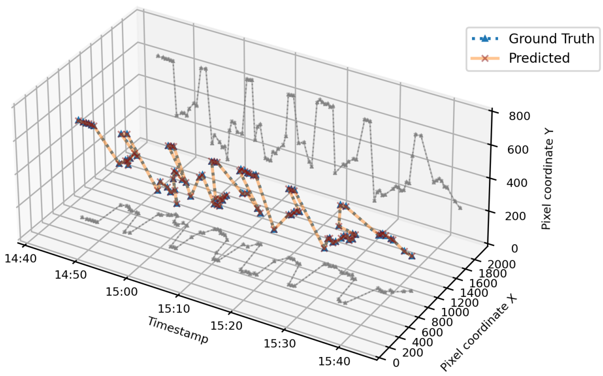


Figure 7: Comparison of the bucket's pixel coordinates between ground truth and detection result in dataset 1

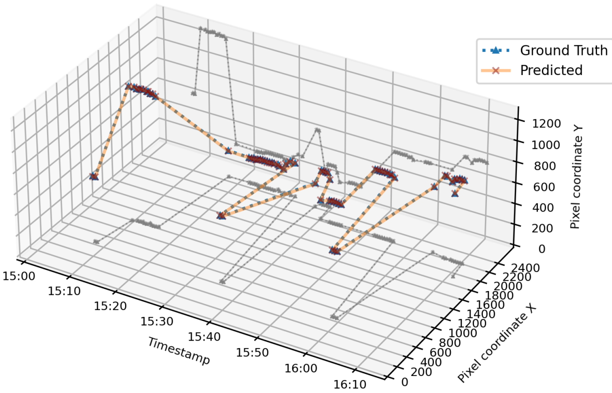


Figure 8: Comparison of the bucket's pixel coordinates between ground truth and detection result in dataset 2

Figures 7 and 8 illustrate the bucket detection accuracy by comparing the coordinates of the buckets' bounding boxes against ground truth data. The comparison is conducted over all timestamps across the two datasets, providing insights into the accuracy of the detection algorithm relative to the actual positions of the buckets. The predicted pouring states compared against the ground truth are shown in Figures 9 and 10. A state accuracy of 92.44% was achieved for the first dataset and 95.11% for the second dataset. The consistency and precision depicted in Fig. 7 and 8 underscore the robustness of the bucket tracking mechanism. However, Fig. 9 and 10 show some challenges in the state reasoning. Our method performed better on the second dataset due to the lower frequency of cycles.

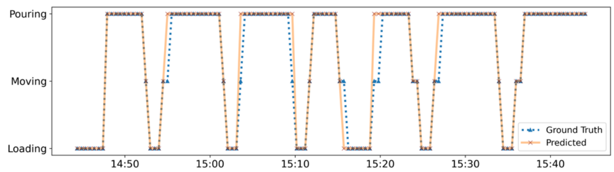


Figure 9: States of the pouring phase in dataset 1

In particular, states with shorter durations are more challenging to detect. It is important to note that exact moving times cannot be computed for the rare cases of moving time below the frames' interval (30 seconds).

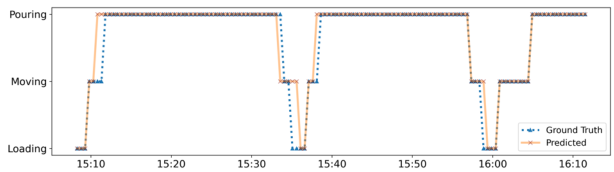


Figure 10: States of the pouring phase in dataset 2

We investigated the elapsed time of four different concrete pouring samples with the introduced pipeline. The times of the individual phases and the number of cycles are shown in Table 3. In general, pouring consumes most of the time. The results show that the process's time can significantly vary depending on the construction scenario.

In certain situations, precisely when waiting for the following concrete mixer's arrival, the moving time of the bucket substantially differs. Moreover, factors like location and size of building elements play a substantial role. To allow further investigations, the amount of concrete volume was calculated based on the IFC model and included in the approach. For this step, the geometry from the IFC model was derived and processed using IfcOpenShell. Based on the building floor, the building elements' vertices were projected to the corresponding 2D floor, and polygons were generated to define the building elements' region, shown in Fig. 3. The grid was created using 15x15 metre cells. All building elements within a cell were considered and summed up to get the total volume amount of the cell. Finally, the correlating grid cell was determined based on the projected basepoint of the hose's bounding box and linked within the graph. Further studies of this topic will be included in future publications of the authors.

Table 3: Elapsed time of the specific cast-in-place states

Sample	No. 1	No. 2	No. 3	No. 4
No. cycles	7	2	7	3
Loading	0:13:06	0:05:04	0:20:44	0:03:31
Moving	0:04:12	0:05:05	0:23:24	0:11:11
Pouring	0:34:52	0:10:06	0:54:04	0:45:31

Discussion and Future Work

The case study results validate the feasibility of our proposed KG-enhanced CV approach in monitoring and analysing the concrete pouring process. Capturing images at 30-second intervals ensures sufficient granularity, achieving 92.44% and 95.11% accuracy in identifying concrete pouring states. This accuracy surpasses previous efforts, such as the 84.7% reported by Gong & Caldas (2010).

With satisfactory accuracy, the case study also demonstrates our methods' superiority in flexibility and extensibility through our innovative adoption of KG in enriching CV systems. Traditional CV methods have a limited application scope, often requiring site-specific model training and manual annotations for critical areas, with a lack of correlation between crane operations and construction progress. Our approach uses CV to identify basic construction objects like workers, concrete mixers, and buckets, enhancing detection success.

The KG then streamlines their relationships with semantic context, enabling a generalisable data interpretation that adapts to changing construction contexts (e.g., concrete loading areas) without reprogramming. This facilitates efficient and resilient processing of diverse data across various construction stages, even different projects. Additionally, this method employs grid-based mapping to indirectly link crane operations with building elements. Although this connection is approximate, it can be further refined to recognise the specific elements being poured based on the positions of workers in a particular zone.

As a result of this method, the data on concrete pouring states offers a dynamic, real-time view of on-site resource coordination, uncovering subtleties often missed in traditional monitoring. For instance, KG analysis reveals that prolonged "moving states" usually miss workers who operate the nozzle. This indicates time wasted mobilising workers, suggesting potential optimisations in pouring sequences. For instance, pouring activities should prioritise the occupied zones to minimise workers' travel. It also informs future development of our ontological model, where tracking worker movements could be valuable as it can highlight unnecessary movements.

Despite its innovative aspects, our approach has limitations. Currently, it focuses on crane-and-skip methods, with pump-based pouring scenarios outside our scope. Meanwhile, its success hinges on accurately detecting discrete states based on spatial relationships between construction elements, which can fail in cases of occlusion or distance. In particular, the hoses have a lower detection accuracy compared with other classes, suggesting the vulnerability of current reasoning rules. To address this, additional parameters and reasoning rules could be used to infer pouring states. For example, the distances between concrete buckets and workers, along with buckets' moving speed, can be used for reasoning. Future work will also explore graph-based machine learning, such as graph neural networks, for automated state classification.

Conclusion

This paper introduces a KG-enhanced CV approach to monitoring concrete pouring. By integrating KG with CV, this method offers a dynamic and adaptable system capable of interpreting complex site dynamics and managing incomplete data. Achieving high accuracy rates of 92.44% and 95.11% in identifying concrete pouring states, our approach matches or surpasses earlier efforts in terms of data interpretation accuracy.

The strength of our method also lies in its ability to efficiently process big data with diverse quality and adapt to changing construction contexts, such as varying concrete loading areas, without needing reprogramming. This adaptability extends across different construction stages and projects, showcasing its potential for broad applicability. Additionally, the use of grid-based mapping to correlate crane operations with as-planned BIM models, although approximate at the moment, opens avenues for more precise recognition of specific pouring elements, even when cameras are positioned at a distance.

In general, our approach's real-time data analysis capability offers a nuanced view of concrete pouring-related resources, effectively highlighting miscoordinations. These insights pave the way for optimising resource allocation and improving site productivity. Anticipated further research aims to integrate this KG with semantic knowledge related to the pump-based concrete pouring method and enhance independence from the detection of site-specific or hard-to-recognise objects (e.g., hoses) by analysing the spatial relationships of easier-to-detect objects (e.g., buckets and workers) or employing machine learning algorithms to discover patterns in KG's topology for classification.

In summary, our KG-enhanced CV method represents a significant stride in concrete pouring monitoring, offering improved accuracy, adaptability, and insight into construction processes. It promises to reshape traditional practices, leading to more efficient and effective construction project management.

Acknowledgements

We thankfully acknowledge Innovation Management Bau GmbH for their financial support and for providing us access to multiple construction sites to collect valuable data. In addition, we would like to thank Siemens Real Estate AG and Max Bögl.

References

Braun, A., Tuttas, S., Borrmann, A., & Stilla, U. (2020). Improving progress monitoring by fusing point clouds, semantic data and computer vision. Automation in Construction, 116. https://doi.org/10.1016/j.autcon.2020.103210

Danel, T., Lafhaj, Z., Puppala, A., BuHamdan, S., Lienard, S., & Richard, P. (2022). Identifying tower crane activities with data: the case of the concrete pouring. Engineering, Construction and Architectural Management. https://doi.org/10.1108/ECAM-10-2021-0936

Fang, W., Ma, L., Love, P. E. D., Luo, H., Ding, L., & Zhou, A. (2020). Knowledge graph for identifying hazards on construction sites: Integrating computer vision with ontology. Automation in Construction, 119. https://doi.org/10.1016/j.autcon.2020.103310

Gong, J., & Caldas, C. H. (2010). Computer Vision-Based Video Interpretation Model for Automated Productivity Analysis of Construction Operations. Journal of Computing in Civil Engineering, 24(3). https://doi.org/10.1061/(asce)cp.1943-5487.0000027

Hu, S., Fang, Y., & Bai, Y. (2021). Automation and optimization in crane lift planning: A critical review. Advanced Engineering Informatics, 49. https://doi.org/10.1016/j.aei.2021.101346

Hu, S., Wang, J., & Fang, Y. (2023). Semantic web-assisted progress monitoring of tower crane operations in construction projects. The 30th EG-ICE: International Conference on Intelligent Computing in Engineering.

Kim, I., Kim, Y., & Chin, S. (2023). Deep-Learning-Based Sound Classification Model for Concrete Pouring Work Monitoring at a Construction Site. Applied Sciences (Switzerland), 13(8). https://doi.org/10.3390/app13084789

Lee, S. K., & Yu, J. H. (2023). Ontological inference process using AI-based object recognition for hazard awareness in construction sites. Automation in Construction, 153. https://doi.org/10.1016/j.autcon.2023.104961

Lu, M., Anson, M., Tang, S. L., & Ying, Y. C. (2003). HKCONSIM: A Practical Simulation Solution to Planning Concrete Plant Operations in Hong Kong. Journal of Construction Engineering and Management, 129(5), 473–585. https://doi.org/10.1061/ASCE0733-93642003129:5547

Lu, M., & Anson, M. (2004). Establish Concrete Placing Rates Using Quality Control Records from Hong Kong Building Construction Projects. Journal of Construction Engineering and Management. https://doi.org/10.1061/ASCE0733-93642004130:2216

Moon, S., & Yang, B. (2010). Effective Monitoring of the Concrete Pouring Operation in an RFID-Based Environment. Journal of Computing in Civil Engineering, 24(1), 127. https://doi.org/10.1061/ASCECP.1943-5487.0000004

Pfitzner, F., Braun, A., & Borrmann, A. (2023a). Object Detection Based Knowledge Graph Creation: Enabling Insight into Construction Processes. ASCE International Conference on Computing in Civil Engineering 2023. https://doi.org/10.1061/9780784485224.023

Pfitzner, F., Braun, A., & Borrmann, A. (2023b). From Data to Knowledge: Construction Process Analysis Through Continuous Image Capturing, Object Detection, and Knowledge Graph Creation. Automation in Construction, under revision

Wang, D., Wang, X., Ren, B., Wang, J., Zeng, T., Kang, D., & Wang, G. (2022). Vision-Based Productivity Analysis of Cable Crane Transportation Using Augmented Reality–Based Synthetic Image. Journal of Computing in Civil Engineering, 36(1). https://doi.org/10.1061/(asce)cp.1943-5487.0000994

Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., & Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. Automation in Construction, 122. https://doi.org/10.1016/j.autcon.2020.103482

Yang, J., Vela, P., Teizer, J., & Shi, Z. (2014). Vision-Based Tower Crane Tracking for Understanding Construction Activity. Journal of Computing in Civil Engineering, 28(1). https://doi.org/10.1061/(asce)cp.1943-5487.0000242

INTERACTIVE AI FOR GENERATIVE HOUSING DESIGN BASED ON GRAPH NEURAL NETWORKS AND DEEP GENERATIVE MODELS

Tian Xia[1], Alex Ledbetter[1], Alexandru Bobe[1], Jeroen Hofland[1], Berend Krouwels[1], ,Tong Wang[1], Luciano Cavalcante Siebert[1], Paul Chan[1], and Jian Yang[2]
[1]Delft University of Technology, Delft, Netherlands
[2]Shanghai Jiao Tong University, Shanghai, China

Abstract

Nowadays, the complexity of housing design has grown substantially to meet multitudinous requirements (e.g., cost, space, esthetics, sustainability, circularity, and modularity) of diverse stakeholders, which poses great challenges to traditional labor-intensive design processes. Automatic design tools are being developed to assist designers handle tedious work at scale. However, knowledge gaps still exist in harnessing deep learning models to learn from human experience for more efficient design generation while keeping the data understandable and interoperable. Moreover, human-in-the-loop approach is largely neglected in the automatic design tools, which are essential for more customized and user-centered design. This research utilizes graph data to parametrically represent housing designs and graph-representative deep generative models for design generation, which provides an interactive design approach for the users at every step. This method enables deep learning models to semantically understand hidden patterns and knowledge in housing designs and facilitate the human-centered design process by returning feasible and parametric housing design alternatives. All codes can be found at: https://github.com/jlhofland/housing-design.

Introduction

The design process has been identified as one of the most significant elements influencing housing quality (Hamzah et al., 2011). Traditionally, the creation of housing designs has involved labor-intensive, manual processes, relying on the expertise of designers. This conventional approach has been characterized by iterative design revisions and historical references. However, it faces challenges related to scalability and adaptability to evolving societal needs (Williamson and Wong, 2022). In response to these challenges, recent advancements in technology and more specifically in the area of Generative AI have introduced the possibility of partially automating the housing design process (Ko et al., 2023). These advancements have the potential to accelerate design iterations and augment the creativity of the proposed designs (As and Basu, 2021).

While various techniques for automatic housing design generation offer distinct advantages, they also come with some limitations. Generative Adversarial Networks (GANs) have the ability to produce realistic floorplan images (He et al., 2022; Wu et al., 2019). However, the image-based approaches lack interoperability with common architectural applications (e.g., CAD and BIM) (Ko et al., 2023). Moreover, it is hard to incorporate various constraints and user requirements into the image-based floorplan generation models. Graph-based methods like Graph Neural Networks (GNNs) can be a complementary approach which has good interoperability due to its parametric nature and compatibility with state-of-the-art generative models (Guo and Zhao, 2023). Therefore, it is worth noting that a promising approach lies in combining GANs with graph-based methods for their generative ability and interoperability.

Numerous studies have incorporated graphs as semantic representation to guide the floorplan generation with GAN, which could be referred to as graph-constrained Relational Generative Adversarial Networks (RaGAN) (Nauata et al., 2021, 2020; Shabani et al., 2022). By using graphs, the users can more easily input and modify the layout at the semantic level, which provides a user-in-the-loop approach and ensures that the model not only produces functional designs but also aligns closely with the user requirements. Bubble diagrams have long been used as a conceptual tool in housing design for visualizing spatial relationships and zoning (Zheng and Petzold, 2023), which has been widely incorporated as graph input to guide floorplan generation with GANs in existing models (Nauata et al., 2021, 2020; Shabani et al., 2022). However, most of the models only generate floorplans based on bubble diagrams given by users and neglect the room layout arrangement task in the conceptual design stage (Nauata et al., 2021, 2020; Shabani et al., 2022). Moreover, bubble diagrams primarily serve as a conceptual tool in housing design for visualizing numbers and spatial relationships among rooms, which can not incorporate constraints from building boundaries (Zheng and Petzold, 2023). Graph2Plan (Hu et al., 2020) allows for building boundaries as user input and then extracts the layout bubble diagrams and generates floorplans. However, Graph2Plan does not generate the graph but rather retrieves it from a set of similar floorplans from a database. This makes it less flexible as certain input constraints might not coincide with the floorplans in the database. Moreover, the adjacencies between rooms and exterior walls cannot be defined. It is of vital importance to incorporate such conditions if the housing design needs to fit into constraints from existing walls and rooms, which happens a lot in renovation cases. Also, the edge features of the graph are non-configurable, which means it does not allow for interior doors to be specified by the user.

Heterogeneous graphs could offer a more expressive approach for housing design generation by providing a data-driven framework that explicitly encodes the attributes and relationships among different rooms and components (e.g., walls, windows, and doors) (Gan, 2022). Therefore, instead of using bubble diagrams, we develop heterogeneous graphs to capture complex constraints and diverse attributes for more powerful and versatile housing design generation. To achieve higher flexibility, we generate both layout graphs and floorplans by using generative models based on graph representation (Figure 1).

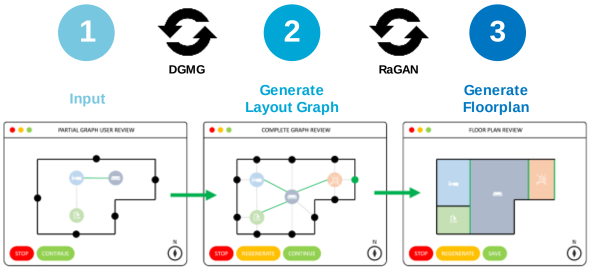


Figure 1: Interactive design generation process where the dots are walls and rooms, and edges represent adjacency. For envisioned graphical interfaces see Figure A4.

Specifically, heterogeneous graph representation will be developed to explicitly encode the attributes and relationships among the exterior walls and rooms which would enable users to input their requirements and constraints (Step 1). Starting from the input, our generative housing design pipeline encompasses layout graph generation model (Step 2) and floorplan generation model (Step 3), which are developed based on DGMG (Li et al., 2018) and RaGAN from Housegan++ (Nauata et al., 2021) because of its wide adaptability and state-of-the-art performance. We made improvements to the models by incorporating edge features in the message passing over heterogeneous layout graphs. On top of this, the user can intervene at each step in the pipeline to make changes to all interim results. By doing this we create a model that proactively keeps the user in the loop and achieves more user-centered design through a collaborative approach compared with end-to-end models (e.g., FloorGAN (Upadhyay et al., 2023)).

Research Methods

Overview

An overview of the pipeline is illustrated in Figure 2, where the top depicts user interaction and the bottom is model training. Starting with the user-input file containing user constraints, augmented DGMG (see next sections) will generate nodes and edges to finish the layout graph. User interaction is available to evaluate the generated layout and make changes along the way. When they are satisfied, the pipeline will pass the layout graph to Hetero-HouseGAN++ (HHGPP, see next sections) for floorplan generation by representing the spatial configuration of rooms through masks, which would then be saved in parametric format. A final user-input step will occur offering to regenerate the plan as necessary.

Initial user input: The user can input the building boundary and specify the location of entrance doors. It is also allowed to provide initial constraints such as the number of rooms for each type. More importantly, the location and adjacency of certain rooms can be specified in the input as partial layout graphs, which is an essential feature for the renovation housing design. It should be noted that only a text-based interface is available for initial user input now, which, however, can be transformed into a graphical interface and embedded into domain applications (e.g., CAD and BIM) in future work. The envisioned graphical interface is depicted in Figure A4.

Generation of layout graph: To generate the customized layout graph based on user input, we adopt a graph generation model based on augmented DGMG (Li et al., 2018). The generated layout graph will explicitly formulate the adjacency relationships of rooms and walls. For instance, a bedroom could be adjacent to the south side of an exterior wall and connected to the north side of the living room through an interior wall with a door. The model will be trained on the large-scale floorplan dataset LIFULL, which was transformed into parametric representation (Nauata et al., 2020). The users can evaluate and modify the generated layout graph or regenerate it based on their preferences.

Generation of floorplan: To acquire the floorplan with precise spatial configurations that fits the layout, we adopt a graph-constraint RaGAN based on Housegan++ (Nauata et al., 2021), which is referred to as HHGPP. The model could aggregate the features of the layout graph through message passing and generate the spatial configuration of rooms as volume masks. The output of

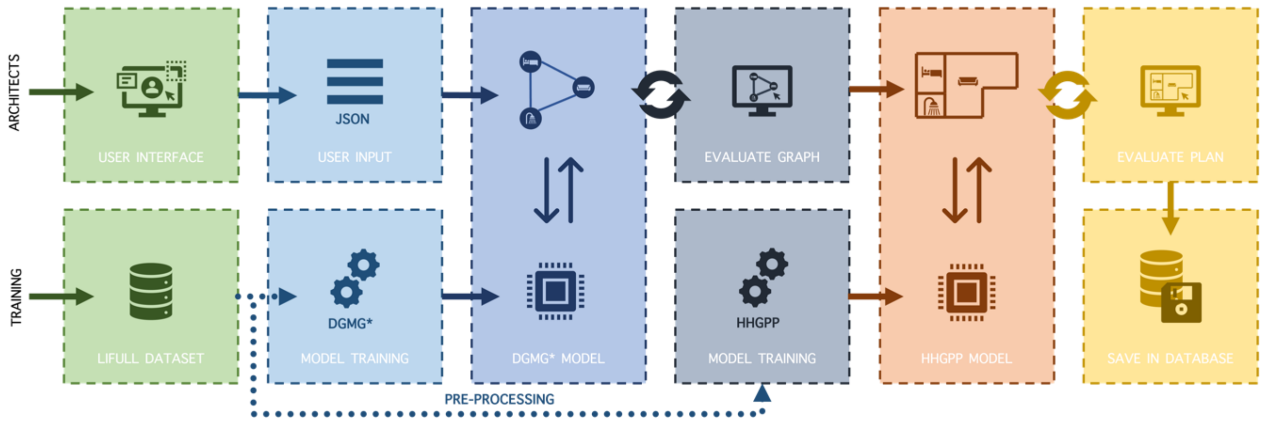


Figure 2: Pipeline for housing design, top: user interaction, bottom: model training

the pipeline will be a parametric floorplan and a graph that depicts the layout arrangement, which can be fed into domain applications for further editing and optimization.

The following sections explain these steps in more detail.

Initial user input

The user could input the requirements and constraints for the housing design(Step 1, Figure 3). The requirements would be represented as a conditioning vector. The constraints would be transformed into a partial graph (Step 2, Figure 3).

User input:

- A vector representing user requirements: the number of various room types, and additional entries indicating if the room quantity is upwards flexible (3 vs 3+ for example).
- The constraints from building boundary and predefined location of rooms (i.e., preferred location of specific rooms or fixed location of existing rooms), transformed into heterogeneous-graph representation: exterior walls as nodes (node features with starting and ending vertices, final element indicating with entrance door or not); corners which connect the exterior walls as edges (edge feature represent the corner type); predefine rooms as nodes (node feature represent room types); adjacency between rooms and exterior walls as edges (edge features represent relative direction); adjacency between rooms as edges (edge features represent the relative direction and connection type (i.e. through a wall or a door, or directly connected)).

Generation of layout graph

In this section, we develop the model to generate layout graphs conditioned on user input by completing the input partial graph from user constraints. User requirements such as the number of bedrooms and bathrooms are fed into the model as a conditioning vector (Li et al., 2018) to generate graphs that meet a client's needs. The graph generation model is based on augmented DGMG, which could incorporate edge features in the message passing for heterogeneous layout graph generation. Through training over a large set of real architect-created home layout graphs (Nauata et al., 2020), the model learns to generate graph layouts that respect various implicitly learned architectural design rules and meet user requirements and constraints.

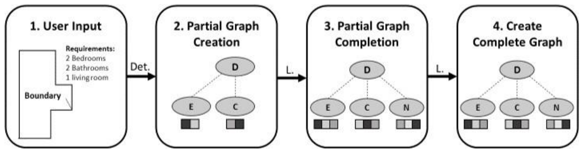


Figure 3 Layout Graph Generation Model. D(decision), E(Add Edge), C(Choose connection), N(Add Node). Det. specifies that the process is deterministic, while, L. states that the step is learned from data.

The layout graph generation process based on augmented DGMG follows an autoregressive generation mode (Li et al., 2018), which is depicted in Steps 3 and 4, Figure 3. The general procedure and our modification are described as follows, and a more detailed illustration can be found in (Li et al., 2018).

Layout graph generation:

(a) Choose to add a new node A (yes/no).

(b) Choose to add a new edge from source node A (yes/no).

(c) If yes:

  i. Choose a destination node B from pre-existing nodes in the graph. In this research, we also predict the feature of the added edge in this step.

  ii. Perform GNN message-passing to update node representations. In particular, we augment DGMG by incorporating edge features in the message passing.

When the model chooses not to add another node, the graph is complete. The output of the model would be a complete layout graph, which could be evaluated, edited, and regenerated through the envisioned user interface.

Generation of floorplan

To generate the floorplans with precise spatial configuration of rooms that respect the constraints specified in this heterogenous layout graph, a graph-constraint RaGAN is developed based on augmentation of Housegan++ (Nauata et al., 2021) (Figure 4).

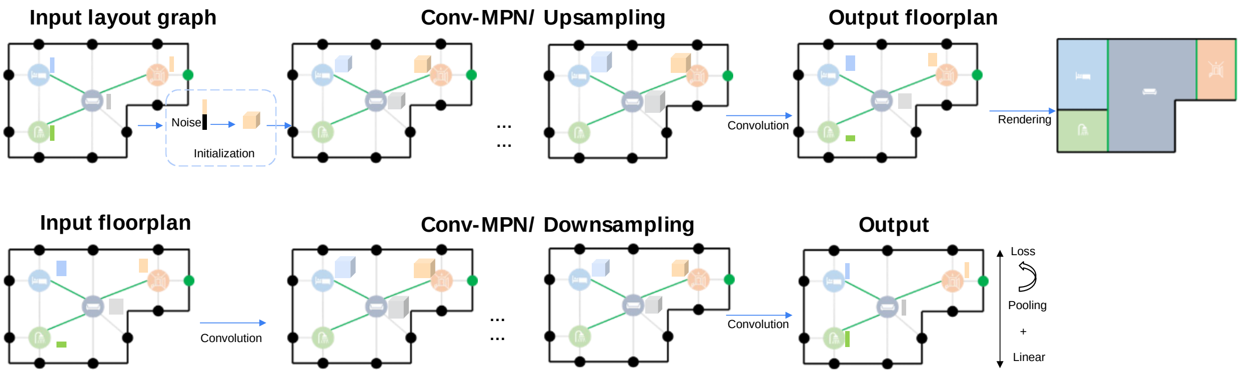


Figure 4 Floorplan generator (top) and discriminator (bottom). ConvMPN is our backbone architecture [9]. The input graph constraint is encoded into the graph structure of their relational networks.

Hetero-HouseGAN++ (HHGPP) is a generative adversarial network whose generator transforms input data (lists of data representing layout graph, e.g. nodes, edges, features, etc.) through Convolutional Message-Passing Network (ConvMPN) into pixel-based room masks. These room masks are later vectorized into parametric representations of floorplans. The discriminator then performs the reverse process, starting with room masks, and resulting in a binary response as to the validity of the room masks it was provided, given the graph data also fed in alongside.

Our improvement in HHGPP compared with Housegan++ includes incorporating new node types, edge types, and node and edge features for graph-constrained floorplan generation. First, GNN message-passing occurs as convolutions over a large data block constructed as the concatenation of the graph node feature volumes, the pooled features volumes of their neighbours which get encoded with the edge type and edge features across these connections, and the pooled feature volumes of the disconnected nodes. These convolutions happen three times, downscaling as we go to result in the final encoded nodal feature volume representations. The second innovation is to use conditional generation whereby some room masks are predefined during mask generation such that the generator learns to paint the remaining masks for the floorplan.

Results and Evaluation

Training data preparation

All the data used for this project are extracted and augmented from the LIFULL dataset (available at: https://www.nii.ac.jp/dsc/idr/en/lifull/), a database of real floor plan designs. Each floor plan is represented as a list of values for room ID and type, room bounding boxes, floorplan edges, edge adjacencies, and IDs of edges with doors. Note that edges here refer to exterior and interior walls, not graph edges.

We filter out floorplans with either completely disconnected rooms or rooms that are only connected at a corner. From this raw data, we create two types of files used for training the layout graph generation model: user-input files and sequences of graph generation. For training the floorplan generation model, we transform the bounding boxes of rooms into the masks (3D volumes) of room nodes to represent their spatial configurations. The pairs of layout graphs and floorplans (graphs with masks for room nodes) will be used to train the generator and discriminator in the HHGPP.

Network training

The layout graph generation model will be trained through teacher forcing throughout the autoregressive generation process of DGMG (Li et al., 2018). The sequences-type files are lists of the ground-truth answers to the sequential generation questions, "should another node be added ?", "what node should this edge be connected to?," such that, when followed, producing the precise home layout graph corresponding to the LIFULL homes. Some decisions in the sequence are binary yes/no responses, while others are multi-class prediction-type responses. In either case, decisions are of a discrete number of choices and proceed by the model calculating a prediction for the logits of each response type. This is a numerical value where a highly positive value represents high confidence in that prediction value being the correct one, while a low or highly negative value indicates low confidence in that option.

We then convert these predicted logits to probability mass values (or likelihoods) via the sigmoid (binary) or softmax (multi-class) equations. Finally, our loss accumulated at each decision is equal to the negative log of the predicted likelihood for the correct response. Minimizing this value then seeks to maximize the likelihood of the correct response. The corresponding loss function is illustrated in (1):

$$loss = -log(softmax(predicted-logits))|correct-response \quad (1)$$

During training, the model will make a predicted answer, and loss will be accumulated if the answers are wrong. After each batch, the loss is backpropagated through the various network layers, and the model parameters are updated to minimize this loss. The loss and generated graphs can be found in Figure A1 and A2.

To train the floorplan generation model, we iterate through the dataset in batches, following a typical training process of GAN models by iteratively training the discriminator, and then the generator, each per batch. Training the discriminator involves generating a set of room masks, and passing these into the discriminator to get a "fake validity" score (binary response, real or fake), then passing in the corresponding real masks to get a "real validity score". These are summed with fake scores negated, along with a gradient penalty term. This loss is then used to backpropagate and update weights. Training the generator then proceeds by the generator producing masks, the discriminator scoring them, and any predicted as fake are counted up as the loss. Additionally, to train the generator to not change the conditioned masks, an L1 loss between the conditioned and generated masks is added to the discriminator loss. The detailed description of the loss function can be found in (Nauata et al., 2021, 2020). This loss is then backpropagated over the generator and weights are updated (Figure A3).

Evaluation

In this section, the evaluation methods are discussed at different points in the pipeline.

The validity of the generated layout graph is evaluated based on compliance with user input and basic design rules. It should be noted that the quality of the generated layout should be evaluated from a more comprehensive architectural design perspective and with potential human expert evaluation for benchmarking against other models, which will be future works and will be discussed in the next chapter. The validity evaluation criteria for the current research are as follows:

R1. the user input constraints are all fully met.

R2. all rooms have at least one door.

R3. exterior walls connect to two other walls and one room, which ensures no solitary walls exist.

R4. each room connects to at least one other room (or wall), which ensures no solitary rooms exist.

R5. each room has outgoing room adjacencies that cover at least two cardinal directions. This attempts to ensure that a room is bounded in all directions. Ground truth homes show that 99.97% of homes meet this criteria.

Table 1: Evaluation Validity Results for DGMG, 10x100 Graph Runs

	R1	R2	R3	R4	R5
Mean	0.13	0	0.061	0.091	0.337
STD	0.126	0	0.167	0.201	0.192

Table 1 shows the ratio of generated home layout graphs that fail each of the validity criteria. These results come from generating 100 layout graphs each from ten different user-input files, with results averaged over the ten runs. Finally, of the 1000 total graphs generated, they are valid 54.1% of the time with a standard deviation of 25.5% between the ten runs. It should be noted that the failure rate to meet R5 is significantly higher, which implies that the graph-representative model finds it more challenging to capture the directional relationships among rooms. It will be discussed in the next section.

The floorplan generation model is evaluated using the Fre´chet Inception Distance (FID). FID is a metric for image generators that evaluates the realism of the generated images by comparing them to a set of real, given images (Heusel et al., 2017). In our evaluation, the FID is used to assess the quality of the floorplans generated by the model using the floorplans in the LIFULL database.

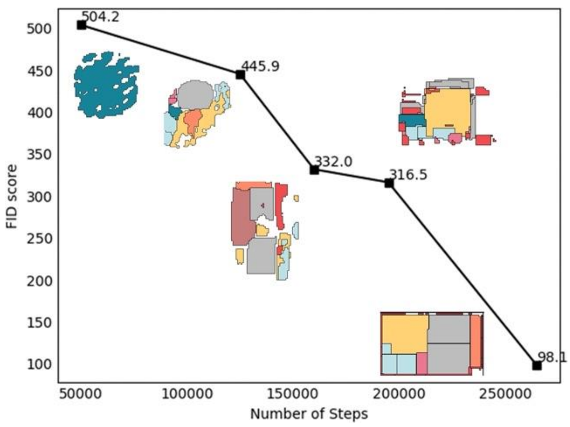


Figure 5 FID Score for HHGPP

Our FID score, Figure 5, is calculated after roughly every 50000 training steps. As the model learned, the floorplan generation model decreased the FID score from 504.2 to a final score of 98.1, indicating more realistic generated images. The score of 98.1 indicates a high similarity between the two distributions of real and generated images. Given enough computing power and training steps, the results could be further improved.

Discussion

We set out to show that housing design assisted by interactive AI based on graph representation is possible. Through the design of the envisioned user interface, review of the workflow, and evaluation of the supporting AI models, we have shown that architects have a new tool at their disposal to generate both home layout graphs and corresponding floorplans pertaining to their needs at higher efficiency. Currently, only a text-based interface is available for users. More advanced graphical interfaces could be developed and embedded into domain applications (e.g., CAD and BIM) based on the graph-based approach.

Both our augmented DGMG framework is a sufficiently capable architecture to learn to generate valid layout graphs that respect the user's input constraints, and the graph-constraint RaGAN architecture present in our HHGPP model is sufficient to learn to translate layout graphs into floorplans that maintain the constraints. This is important because it shows that the user input is not too complex to be learned by generative models in general.

However, the prediction of edge direction in the layout generation model is still imperfect. When the model adds a new node and a new edge from that node, it must then decide in what direction the connection will point. It is important to avoid contradictory constraints. The graph generation model still finds it hard to capture and fulfill all the hard constraints through learning. The reasons for the imperfection can be representation-related, model-related, or training-related. We believe the training process is adequately done considering the loss has dropped to a reasonably low level in different training parameters. The graph representation of the floorplan is quite complex in our current approach, which contains different types of edges and nodes with different attributes and semantic meanings. The tradeoff here is that more expressive graph representation of the housing designs would leave us more possibilities in the interactive design process and in combining computational design approaches. But it could also be more challenging for the models to capture the complex semantic information in the representation of housing designs since the generative models are mostly developed for applications with simple representations and big data (e.g., predictions of proteins or monocular structures).

To tackle this challenge, we will test different graph generation models (e.g., GraphRNN) to potentially identify the one with more expressive power. We would also try other graph representations for the layout to help the model better capture the embedded semantic and spatial relationships. In the meantime, the graph representation enables us to fix the generated layout through hardcoded rules. Also, different computational design methods can be implemented as complementary approach to the generative design by optimizing the results.

Another interesting finding is that small errors in generated layout graphs could result in noticeable contradictions in the corresponding generated floorplans. Even though the graph-based generative housing design pipelines allow more user interaction and control in every intermediate step, it could result in higher propagated error compared with end-to-end models.

Due to time and resource limitations, we only validate our models regarding compliance with user input and basic design rules, and the realism of floorplans. To bring our approach forward, we would scale up the testing and evaluation in two ways, including developing the qualitative matrices to evaluate the architectural quality of the generative housing design through a computational approach and developing a graphical user interface for larger-scale user evaluation. We would also benchmark the performance of our model against other state-of-the-art models in future research.

Since the graph representation is parametric, our graph-based approach could be compatible with various design tasks such as building equipment arrangement (Wang et al., 2019) and structural design (Zhao et al., 2024). It would be a promising direction to investigate how to integrate our graph-based method into a more generic design approach so as to ensure a more realistic and functional design solution.

Conclusions

In this paper, we presented an interactive AI framework for generative housing design, leveraging Graph Neural Networks (GNNs) and Deep Generative Models. Our approach, exemplified by the augmented DGMG framework for layout graph generation and Hetero-HouseGAN++ (HHGPP) for floorplan generation, demonstrates the viability of incorporating user constraints and requirements into the generative design process. Through an interactive interface, architects and stakeholders could actively participate in the design workflow, ensuring the production of feasible and user-centered housing design alternatives.

Moving forward, several avenues for improvement and exploration emerge. Firstly, the model is still insufficient in predicting the edge features and meeting all hard constraints in layout graph generation. We will test different models and representations for more expressive power in the learning for generative design. Future research could also look into computational methods as a complementary approach to fix and optimize the generated design alternatives based on graph representation. Moreover, user feedback and architectural quality metrics will be integrated into the evaluation process, providing a more comprehensive assessment of the generated designs.

This research would facilitate the housing design process through a collaborative and user-centered approach. The graph-based approach can also be integrated into more advanced design tasks (e.g., generative design of modular buildings based on BIM components library) due to its interoperability with the parametric data (e.g., IFC and IFCOWL), which would greatly facilitate the complex design process by harnessing the power of learning-based models and computational design approach. Despite only text-based interface is available for users currently, more advanced graphical interface could be further developed and easily embedded into domain applications (e.g., CAD and BIM) to facilitate the real-world design process.

Acknowledgements

This work was supported by TU Delft AI Labs. All codes can be found at: https://github.com/jlhofland/housing-design.

References

As, I., Basu, P., 2021. AI & architecture. Routledge.

Gan, V.J.L., 2022. BIM-based graph data model for automatic generative design of modular buildings. Autom. Constr. 134, 104062. https://doi.org/10.1016/j.autcon.2021.104062

Guo, X., Zhao, L., 2023. A Systematic Survey on Deep Generative Models for Graph Generation. IEEE Trans. Pattern Anal. Mach. Intell. 45, 5370–5390. https://doi.org/10.1109/TPAMI.2022.3214832

Hamzah, N., Ramly, A., Salleh, H., Tawil, N.M., Khoiry, M.A., Che Ani, A.I., 2011. The Importance of Design Process in Housing Quality. Procedia Eng., 2nd International Building Control Conference 20, 483–489. https://doi.org/10.1016/j.proeng.2011.11.191

He, F., Huang, Y., Wang, H., 2022. iPLAN: Interactive and Procedural Layout Planning. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7793–7802.

Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S., 2017. GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium, in: Advances in Neural Information Processing Systems. Curran Associates, Inc.

Hu, R., Huang, Z., Tang, Y., Van Kaick, O., Zhang, H., Huang, H., 2020. Graph2Plan: learning floorplan generation from layout graphs. ACM Trans. Graph. 39, 118:118:1-118:118:14. https://doi.org/10.1145/3386569.3392391

Ko, J., Ennemoser, B., Yoo, W., Yan, W., Clayton, M.J., 2023. Architectural spatial layout planning using artificial intelligence. Autom. Constr. 154, 105019. https://doi.org/10.1016/j.autcon.2023.105019

Li, Y., Vinyals, O., Dyer, C., Pascanu, R., Battaglia, P., 2018. Learning Deep Generative Models of Graphs. https://doi.org/10.48550/arXiv.1803.03324

Nauata, N., Chang, K.-H., Cheng, C.-Y., Mori, G., Furukawa, Y., 2020. House-GAN: Relational Generative Adversarial Networks for Graph-Constrained House Layout Generation, in: Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (Eds.), Computer

Vision – ECCV 2020, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 162–177. https://doi.org/10.1007/978-3-030-58452-8_10

Nauata, N., Hosseini, S., Chang, K.-H., Chu, H., Cheng, C.-Y., Furukawa, Y., 2021. House-GAN++: Generative Adversarial Layout Refinement Network towards Intelligent Computational Agent for Professional Architects. Presented at the Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13632–13641.

Shabani, M.A., Hosseini, S., Furukawa, Y., 2022. HouseDiffusion: Vector Floorplan Generation via a Diffusion Model with Discrete and Continuous Denoising. https://doi.org/10.48550/arXiv.2211.13287

Upadhyay, A., Dubey, A., Mani Kuriakose, S., Agarawal, S., 2023. FloorGAN: Generative Network for Automated Floor Layout Generation, in: Proceedings of the 6th Joint International Conference on Data Science & Management of Data (10th ACM IKDD CODS and 28th COMAD), CODS-COMAD '23. Association for Computing Machinery, New York, NY, USA, pp. 140–148. https://doi.org/10.1145/3570991.3571057

Wang, K., Lin, Y.-A., Weissmann, B., Savva, M., Chang, A.X., Ritchie, D., 2019. PlanIT: planning and instantiating indoor scenes with relation graph and spatial prior networks. ACM Trans. Graph. 38, 132:1-132:15. https://doi.org/10.1145/3306346.3322941

Williamson, E., Wong, K., 2022. Valuing architectural work: The human effects. Archit. Aust. 111, 58–59. https://doi.org/10.3316/informit.604522451937359

Wu, W., Fu, X.-M., Tang, R., Wang, Y., Qi, Y.-H., Liu, L., 2019. Data-driven interior plan generation for residential buildings. ACM Trans. Graph. 38, 234:1-234:12. https://doi.org/10.1145/3355089.3356556

Zhao, P., Liao, W., Huang, Y., Lu, X., 2024. Beam layout design of shear wall structures based on graph neural networks. Autom. Constr. 158, 105223. https://doi.org/10.1016/j.autcon.2023.105223

Zheng, Z., Petzold, F., 2023. Neural-guided room layout generation with bubble diagram constraints. Autom. Constr. 154, 104962. https://doi.org/10.1016/j.autcon.2023.104962

Appendix

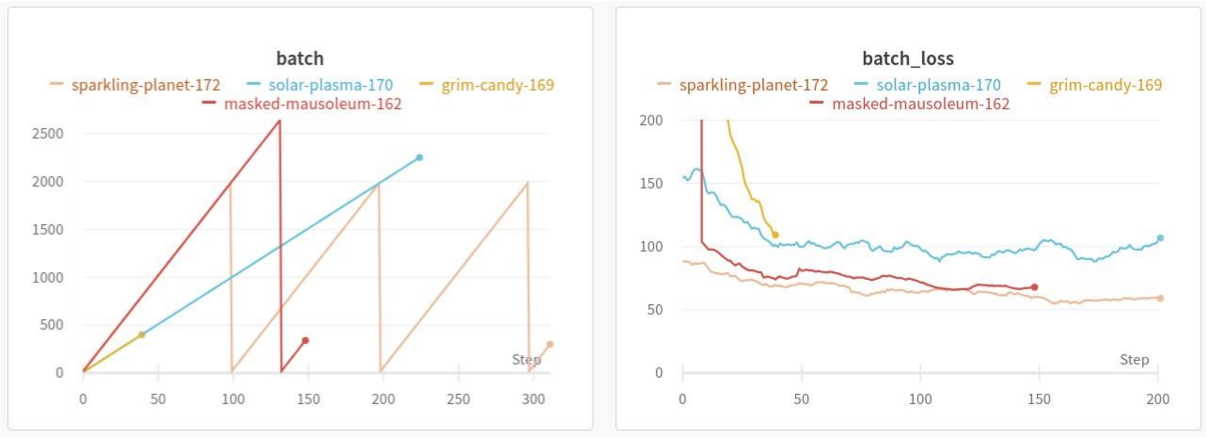


Figure A1: Training of layout graph generation model

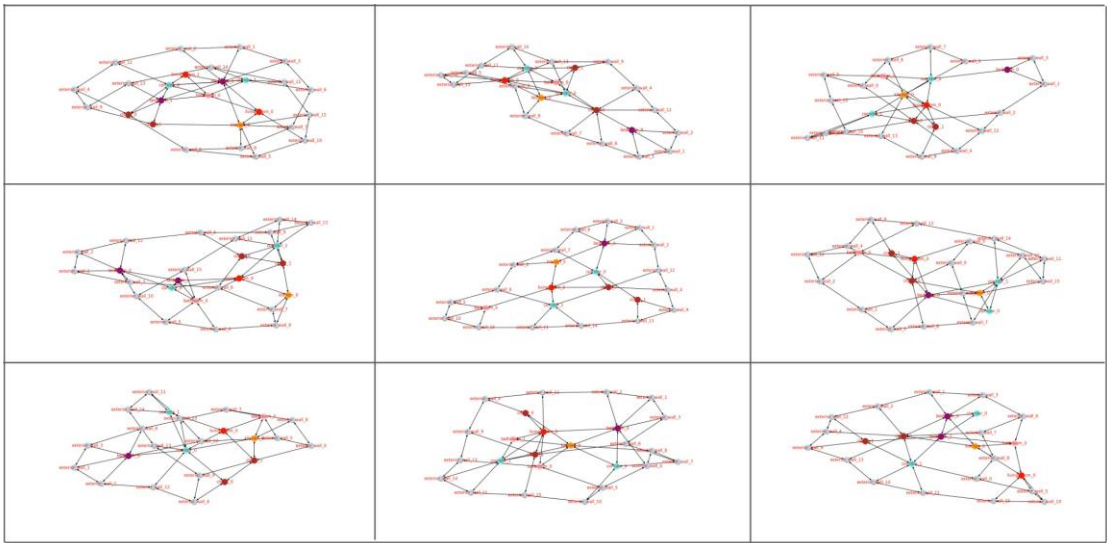
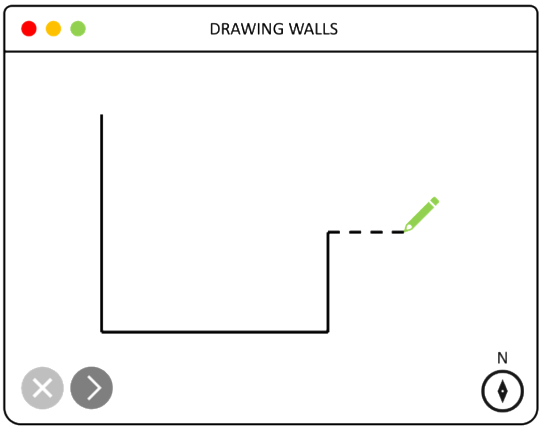
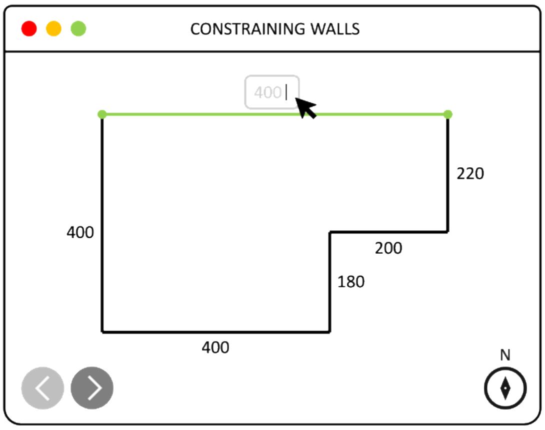


Figure A2: Sample of generated layout graphs

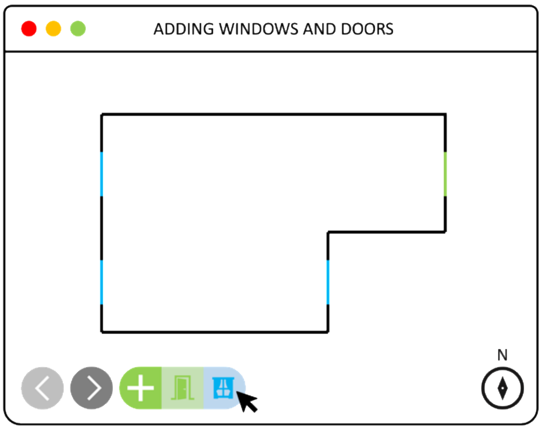
Figure A3: Training of floorplan generation model



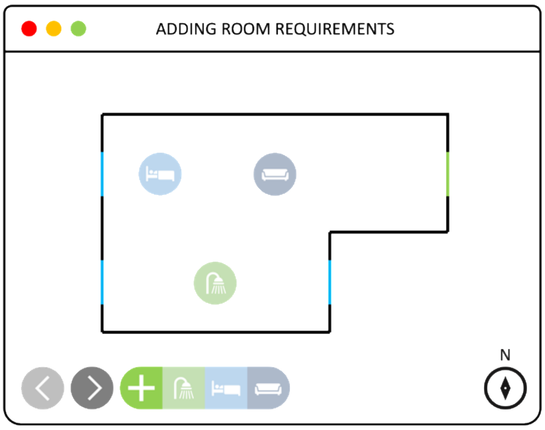
(a) Drawing outline of building using mouse. At each step we can use the buttons in the bottom to navigate through the user input steps.



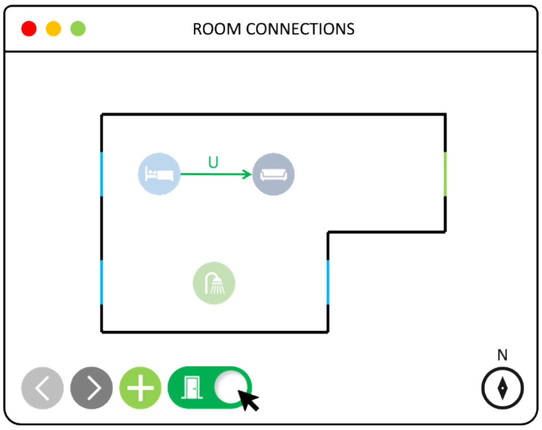
(b) Add length and angular constraints to each wall and corner. The interface could allow for natural angles but currently our model is only trained on angles of 90 degrees.



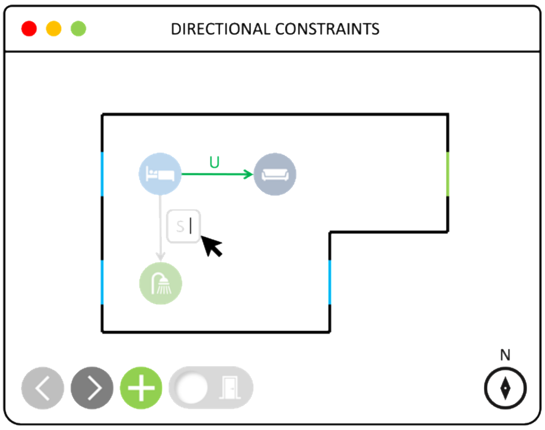
(c) Add facade features like doors and windows. In our case we only allow for the placement of doors as our model does not include window information.



(d) We can now add rooms, notice that the position of the nodes does not represent the spatial location. Our model only constrains the amount of baths, beds and living rooms but can be expanded.



(e) Adding edges (connections) between rooms. We can define wether or not the connection is an adjacency or door connection.



(f) We can also add an adjacency/door between rooms with a constrained relative direction. In the example, we say the bathroom should be south (S) of the bedroom without a door.

Figure A4: Envisioned graphical user interface

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

DEVELOPMENT OF A DIGITAL TWIN-BASED SIMULATION SYSTEM AND A NOVEL SYNTHETIC VIDEO DATASET FOR ENHANCING COMPUTER VISION IN CONSTRUCTION SITE SAFETY

Zhengyu Wu, Yuxiang Feng, Yiannis Demiris, Panagiotis Angeloudis
Imperial College London, London, United Kingdom

Abstract

In this work, we introduce a novel construction simulation system and the first photo-realistic synthetic video dataset for the construction industry, named ConSynth. The system simulates vehicle dynamics, worker behaviours, and various environmental conditions (Figure 1), generating a dataset of 200 video sequences and 24,000 images with automated labels. Our experiments have shown that models trained on ConSynth achieve robust detection with less reliance on real-world data, indicating enhanced diversity, improved performance in underrepresented scenarios, and substantial generalisability. Our work promises significant advancements in construction safety monitoring through computer vision.

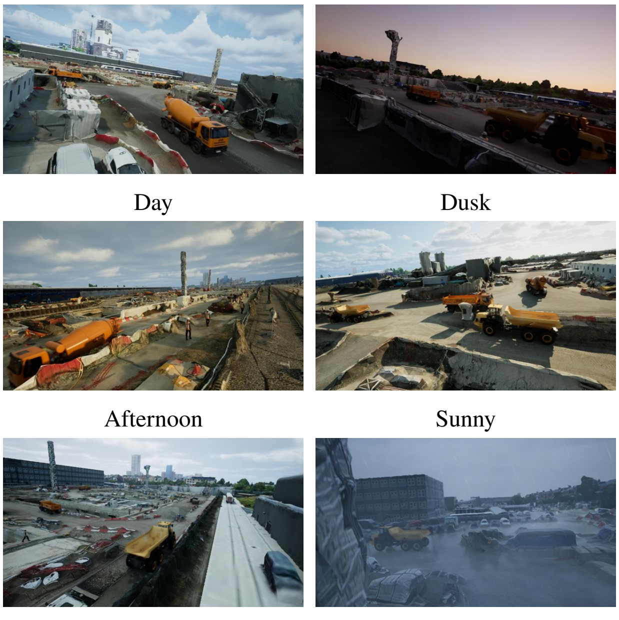


Figure 1: ConSynth contains video sequences initialised and rendered with different domain parameters

Introduction

Background

Construction sites are inherently hazardous. With the increasing application of deep learning techniques in safety monitoring on construction sites, accurate detection of construction equipment and workers has grown in importance. However, a significant barrier to this progression is the limited availability of labelled data, such as images or video frames, where each object (e.g. construction equipment or workers) is identified and annotated with a bounding box and a category. Currently, the two largest public datasets in this domain are the Detecting Moving Objects in Construction Sites (MOCS) dataset (Xuehui et al., 2021) and the Alberta Construction Image Dataset (ACID) (Xiao and Kang, 2021). Although models trained on such datasets have demonstrated the ability to effectively detect a variety of construction equipment or workers, a notable 'domain gap' often exists in their applications on real construction sites. This 'domain gap', a common issue in deep learning, signifies the discrepancies in sample distributions between the data that models are being trained on and the data in real-world applications. One common solution is to collect and annotate extra data from target construction sites and fine-tune the models. However, the manual annotation is time-consuming and error-prone, introducing noise and potential privacy concerns. Furthermore, the current focus of construction site datasets predominantly lies in 2D object detection, tracking, and instance segmentation. Other critical perception tasks witnessed in autonomous driving research, such as depth estimation or 3D object detection, remain largely unexplored in construction sites. This is mainly constrained by the significantly increased data collection and annotation cost associated with these tasks. To tackle these challenges, exploring synthetic datasets has emerged as a promising and vital research direction.

Research gaps

While synthetic construction datasets provide significant potential, investigations in current literature have revealed several limitations. One common method involves rendering 3D models of construction equipment or personnel and superimposing them onto backgrounds of construction sites, known as the crop and paste method (Lee et al., 2022). However, the generated images lack realism due to inconsistencies in camera perspective and lighting with the background images. Alternative methods attempt to craft a construction site manually using 3d modelling software and then render the models of construction objects within it (Soltani et al., 2016; Neuhausen et al., 2020; Lee et al., 2023). Yet, hand-crafting scenes fail to capture the details and complexity of real-world sites. Moreover, manual creation is often limited to specific research contexts and is difficult to adapt to general construction projects. Consequently, models trained on these synthetic datasets tend to be constrained by substantial domain gaps, undermining their performance in actual application scenarios. In the field of autonomous driving, strategies such as transfer learning, domain randomisation, and domain adaptation are typically employed to bridge this gap. However, their application in construction site monitoring is still rare. Lastly, most datasets for the construction industry contain only images. These discontinuous images, sampled at different intervals, restrict the extraction of spatio-temporal features of the scene.

Contributions

The contributions of this paper are two-fold. First, we propose a digital twin-based simulator that simulates realistic vehicle dynamics and worker behaviours. The simulator also adapts domain randomisation with the weather and lighting parameters to minimise domain gaps between reality and simulation. Second, our proposed synthetic dataset, 'ConSynth', is the first synthetic video dataset in the construction industry with the aim to broaden the scope of task applicability. This novel dataset provides continuous image sequences with precise labels for traditional computer vision tasks, such as 2D and 3D object detection, tracking, and instance segmentation, and enables additional tasks such as depth estimation, action recognition and behavioural analysis. Our research showcases that combining photo-realistic synthetic and real-world data significantly improves model robustness and accuracy in construction site scenarios, effectively reducing the reliance on extensive real-world data collection and manual annotation. Our findings highlight the potential of synthetic data in enhancing model robustness, offering a solution for deploying computer vision in dynamic and challenging construction environments.

Related Works

In this section, we review existing research on developing datasets for deep learning applications. Initially, it outlines the current state of real-world datasets in autonomous driving, underscoring the emergence of synthetic datasets as a promising alternative to manual data collection. Following this, the section discusses the development of synthetic datasets in the construction industry.

Real-world Datasets

Deep learning models are highly data-dependent. Over the past decade, the rapid advancement in autonomous driving research has been significantly fuelled by vast amounts of data from real-world sensors (Feng et al., 2021). The collaboration between the industrial and academic sectors has led to the annotation and release of several mature datasets, such as KITTI (Geiger et al., 2013), nuScenes (Caesar et al., 2019), and Argoverse (Chang et al., 2019). These datasets have catalysed the progress in autonomous driving research by supporting various tasks, notably 2D/3D detection and tracking, depth estimation, and stereo vision.

In the construction industry, researchers are also developing datasets to train deep-learning models. Xuehui et al. (2021) released the first large-scale construction site dataset, MOCS, which offers around 42,000 images covering 13 common construction equipment categories, supporting 2D object detection and segmentation. Similarly, the ACID dataset (Xiao and Kang, 2021) provides 10,000 images covering ten types of construction equipment, targeted primarily for 2D object detection. They have paved the way for deploying 2D detectors like Faster-RCNN, YOLO, and Mask RCNN in the industrial sector. However, compared with the autonomous driving sector, datasets for the construction industry are still relatively scarce and limited to a narrow range of tasks.

Synthetic Datasets in Autonomous Driving

The development of synthetic datasets is emerging as a solution to the challenges of data collection. In autonomous driving, a pioneering example is the dataset introduced by Richter et al. (2016), which utilises highly realistic imagery from commercial video games and intricate simulations of interactions between pedestrians and vehicles to provide pixel-level semantic segmentation labels. Combining synthetic with real datasets such as KITTI (Geiger et al., 2013) has enhanced model training accuracy while reducing reliance on expensive manual annotation of real-world data. Similarly, SYNTHIA (Ros et al., 2016) generates semantic segmentation labels for driving scenarios using a simulator built with the Unity game engine. Experiments in this work have demonstrated that models trained on SYNTHIA and fine-tuned with real-world datasets show improved accuracy. VIPER (Richter et al., 2017) extends the scope by providing annotations for multiple tasks, including 2D/3D object detection, multi-object tracking, and optical flow estimation. Additionally, it includes data from five different environmental conditions and diverse urban landscapes. AIODrive dataset (Weng et al., 2020) provides point cloud data from LiDAR sensors in varying densities, in addition to complete annotations for object detection, and covers diverse scenarios and weather conditions. CarlaScenes (Kloukiniotis et al., 2022) provides a practical solution for annotating real-world scene data for multiple challenging scenarios using the CARLA simulator (Dosovitskiy et al., 2017). SHIFT (Sun et al., 2022), another dataset based on the CARLA simulator (Dosovitskiy et al., 2017), encompasses the 13 most significant perception tasks in the domain of autonomous driving. The continuous domain shifts, a distinctive feature of SHIFT, enable the researchers to evaluate the 'domain gaps' and benchmark domain transfer algorithms.

Synthetic Datasets in Construction

To date, the synthetic approach has not been thoroughly explored in the research domain of construction. Most studies used synthetic datasets for specific tasks where manual data annotation is challenging. For example, stud-

(a) Crop and paste (Lee et al., 2022) (b) 3D rendering (Barrera-Animas and Delgado, 2023) (c) Digital twin-based simulation **(Ours)**

Figure 2: Comparison of methods used in generating construction synthetic dataset. Our simulation system generates more realistic images in terms of texture and environment.

ies by Soltani et al. (2016); Lee et al. (2022) rendered 3D models of excavators in various poses and superimposed them onto diverse construction site backgrounds to enhance the accuracy of detecting excavators in images, particularly smaller objects (Figure 2a). Similarly, Assadzadeh et al. (2022) synthesised images of excavators for training object detection models and extended the research to quantitatively analyse the role of domain randomisation in bridging the gap between synthetic and real-world images. Neuhausen et al. (2020) used manually created scenes and pre-captured motion animations to render 3D worker models under varying lighting and weather conditions, aiming to improve the tracking accuracy of construction workers. Additionally, Lee et al. (2023) employed a game engine to create a synthetic dataset to enhance the detection precision of workers' small personal protective equipment for safety monitoring. Barrera-Animas and Delgado (2023) expanded the research scope by covering multiple equipment categories, combining a series of 3D models, and rendering them under diverse lighting scenarios. However, the effectiveness of these datasets is still in question due to their severe lack of photorealism. Results indicated that models trained on these datasets performed poorly when tested on real-world data Barrera-Animas and Delgado (2023). Despite these limitations, these pioneering datasets have laid the groundwork for developing more complex, multi-task synthetic datasets in construction safety research.

Developing Construction Simulation System and Synthetic Video Dataset

In this paper, we created a novel construction simulation system to simulate the behaviours of various agents under different lighting and weather conditions. Within the system, we generated a novel synthetic video dataset, ConSynth, featuring annotations for various tasks. The overall process can be found in Figure 3.

Generation of the Digital Twin

As discussed in the previous section, existing methods generated images with a lack of realism. In this work, we propose using a digital twin model to preserve the details of actual sites. First, we selected an actual construction site in northwest London, UK, as the case study area. Then, a UAV was scheduled to follow a pre-designed route over the site, capturing images at a frequency of once per week. We estimated the three-dimensional structure using the Structure from Motion (SfM) algorithms based on images and the known camera parameters during capture. This process involved the use of Bentley ContextCapture software.

Although our digital twin model was regularly updated to reflect the continuous changes at the construction site, it only includes the case study site and a small surrounding area. A large proportion of the background was missing when we used cameras within the simulator to capture images at a flat angle. Tremblay et al. (2018) found that diversity in the background of synthetic datasets is crucial, as it encourages the model to learn the most representative features of the target objects, thereby distinguishing them from the complicated background. Expanding the UAV capturing area would be computationally costly and time-consuming. As an alternative, we integrated the 3D model from Google Earth to fill the distant scenery. After localising the scanned model in a geographic coordinate system, we removed the scanned region from Google's 3D model and overlaid our model. This approach ensures the most recent scanned model is used for the construction site area, while Google Earth, with less frequent changes, is used as the distant scenery.

Simulation of Site Logistics

The next step involves simulating the agents, such as the construction equipment and workers, with realistic behaviours. This simulation comprises three parts: (1) site road configuration, (2) generation of site agents, and (3) movement of each agent.

To accurately represent the site road configuration, our approach involves creating a vector map in alignment with the site's logistic plan (Figure 5). We utilised the map handling framework Lanelet2 (Poggenhans et al., 2018), a tool prevalent in the autonomous driving sector for map reading and trajectory planning. These vector maps, essential

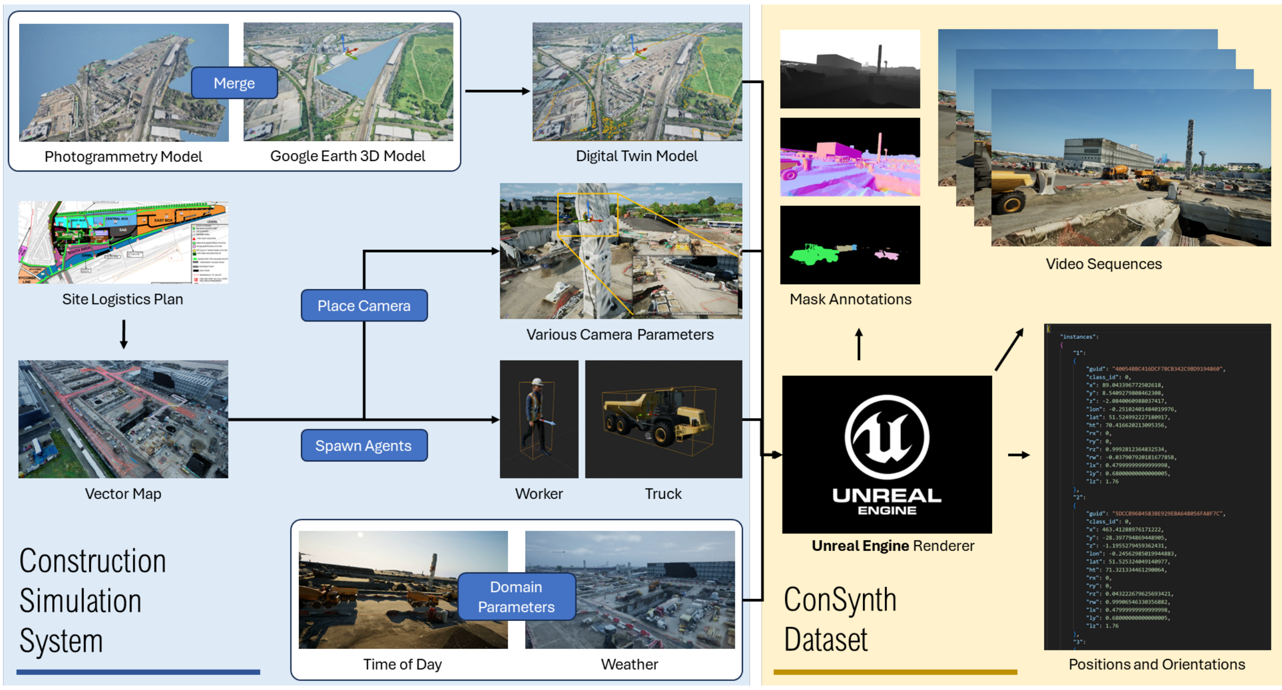


Figure 3: Diagram of the overall process to generate the synthetic video dataset within the construction simulation system

for autonomous driving, offer detailed vector-based road information, including lanes, crossings, and intersections. Leveraging Lanelet2, we implemented the behavioural logic for agent movements. For instance, trucks are programmed to travel on vehicle lanes and navigate through intersections, while workers are set to follow pedestrian pathways and crossings. These movements are governed by the underlying logic provided by the Lanelet2 framework to ensure a more realistic simulation.

We then developed a rule-based system for generating agents based on the vector map. Initially, we tagged haul roads and one-way roads with a 'vehicle' tag and intersections with both 'vehicle' and 'intersection' tags. Pedestrian footways and crossings received a 'pedestrian' tag. The generation follows these rules:

1. Vehicles, such as dump trucks and mixer trucks, are only generated on lanes tagged with 'vehicle' and not 'intersection', with a minimum spacing of $S_{min,veh}$ and a maximum of N_{veh} vehicles on the map;
2. Pedestrians are generated only on lanes tagged 'pedestrian', with a minimum spacing of $S_{min,ped}$ and a maximum of N_{ped} pedestrians on the map;
3. For candidate lanes tagged 'pedestrian', we check if the lane intersects with any lane tagged with 'vehicle', identifying it as a pedestrian crossing.
4. For each lane tagged 'intersection', we use the $k-NN$ algorithm to cluster lanes belonging to the same intersection and generate a special agent as the controller: the intersection agent.

All agent movements are based on their respective lanes. Vehicles move on lanes tagged with 'vehicle', randomly choosing a connected lane at the end of their current lane. If the upcoming lane belongs to an intersection, the corresponding intersection agent will take over the control when a vehicle approaches the stop line. Similarly, pedestrians move only on pedestrian-tagged lanes and randomly select the next lane. If a pedestrian crossing is found to be chosen as the next lane, the agents check if the crossing holds a 'walkable' tag. If the tag is absent, they wait at the end of the crossing. Otherwise, they continue to walk through the crossing. Each intersection agent, simulating a traffic light, sequentially checks its managed lanes with a 'vehicle' tag. For each lane, if a vehicle is found waiting at a stop line, the agent allows it to pass through the intersection and assigns it to a lane at the other end. After a round of checks, a 'walkable' tag is added to all managed pedestrian crossings for a duration of t, allowing all pedestrian agents to cross the vehicle lanes freely.

Domain Randomisation

One major drawback of the current public datasets in construction is that most samples were collected under good illumination conditions. Consequently, models trained on these datasets struggle to handle adverse weather and lighting conditions effectively. To improve model robustness under varying conditions, such as night and rainy and foggy weather, we propose the use of domain randomisation. Specifically, we randomised the weather and time of day in the game engine to create scenes with diverse appearances (Table 1).

Camera Placement

We developed an algorithm for camera placement to ensure that they capture agent movements in the camera view rather than just the background. To begin with, we ran-

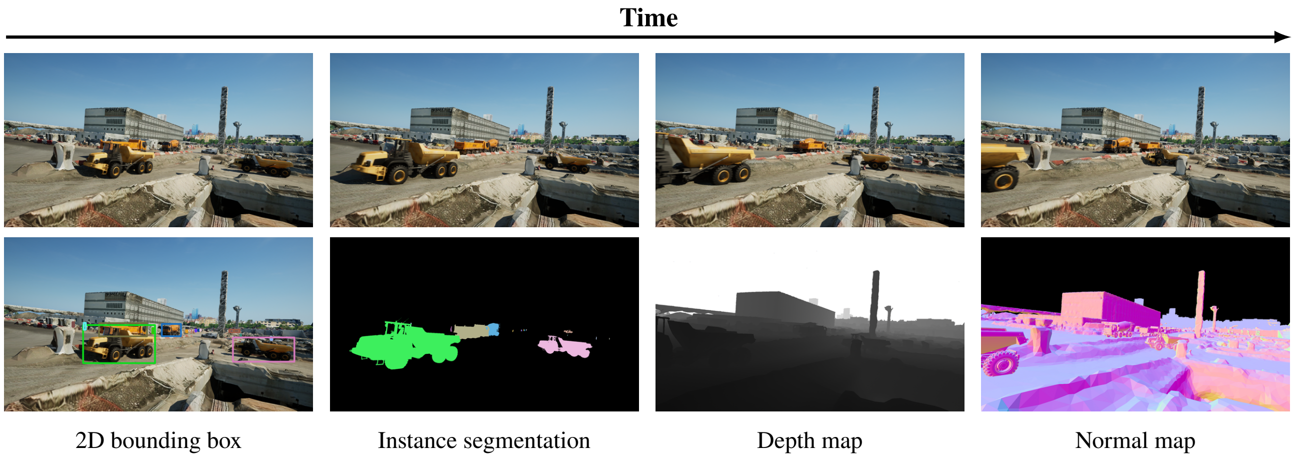


Figure 4: ConSynth provides continuous image sequences with precise labels of 2D bounding boxes, instance segmentation, depth map and normal map. The dataset supports both image-based and video-based computer vision tasks including object detection and tracking, instance segmentation, depth estimation, action recognition and behaviour analysis.

Figure 5: Vector map for the case study construction site

Table 1: Distribution of domain randomisation parameters

Parameter	Variations	Probability
Weather	Clear sky	0.20
	Cloudy (Regular Cloud Cover)	0.10
	Cloudy (Partial Cloud Cover)	0.10
	Foggy	0.20
	Rainy (Light Rain)	0.10
	Rainy (Thunderstorm)	0.10
	Overcast	0.20
Time of Day	Daylight	0.75
	Night-time	0.25

domly selected a point on one of the lanes from the vector map. Then, a centre was set between z_{min} and z_{max} above this point, generating a semi-sphere with a radius between r_{min} and r_{max}. A point on the surface of this semi-sphere was chosen randomly as the initial camera position. We then performed ray tracing from the centre to this point to avoid obstructions and repeated the selection process until the view was unobstructed. A camera was then positioned at this point, oriented towards the centre. Finally, we randomised the camera's intrinsic and extrinsic parameters, including the field of view and orientation, to ensure our dataset covers a variety of perspectives.

Automated Annotation

We leveraged Unreal Engine 5 for rendering photo-realistic RGB images and mask annotations. In addition, we developed a C++-based tool for automated annotation that, for each rendered frame, collects 3D transformations of all agents in the scene, including relative coordinates (x, y, z), world coordinates $(latitude, longitude, height)$, rotation angles (r_y), and dimensions of 3D bounding boxes (l_x, l_y, l_z). It also gathers camera intrinsic and extrinsic parameters, exporting them along with the agent transformations. The segmentation annotations are initially exported as mask images, where a unique colour represents each agent. Following the generation, we utilised OpenCV (Bradski, 2000), a public computer vision library for image processing, to identify these colours and extract polygon-based segmentation data. Finally, the tool calculates the smallest enclosing rectangle from the segmentation as the 2D bounding box for each agent. The automated annotation system, illustrated in Figure 4, is of significance in reducing the time and resources required for the generation of extensive, high-quality and accurate labelled datasets.

Dataset Statistics

Our ConSynth dataset encompasses 200 video sequences, each uniquely produced within the digital twin environment using randomly varied parameters and captured from different locations (see Figure 1). Every sequence consists of 120 frames, recorded at 12 frames per second. Within the dataset, we focus on three primary categories: workers, dump trucks, and mixer trucks. The most populous category is the workers, with $237,215$ instances, which is twice the number of dump truck instances. Mixer trucks, being the least common type, have $31,800$ instances. Figure 6 details the distribution across these categories. Notably, the dataset totals 24,000 images, which results in an average of around 15 instances per image. This density of instances is vital for training robust models, offering a diverse and comprehensive range of data for each category.

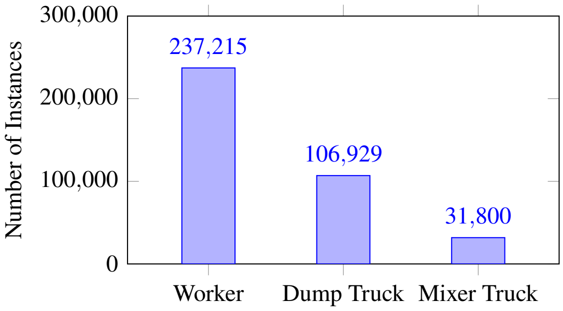


Figure 6: Distribution of different categories (workers, dump trucks, and mixer trucks) in the ConSynth dataset.

The images generated in the ConSynth dataset have a resolution of 1920×1080 pixels. We referred to the standards for object sizes set by the COCO dataset to analyse the distribution of objects of various sizes. According to COCO, in images of 640×480 resolution, objects with an area less than 32×32 pixels are classified as small objects, and those less than 96×96 pixels as medium objects, with the remainder being large objects. We scaled this standard proportionally to match our resolution, resulting in the distribution of small, medium, and large objects in our dataset being 63%, 24%, and 13%, respectively.
This significant presence of small objects in the dataset is a critical feature, as it aligns with real-world scenarios where computer vision models need to identify small yet significant objects on construction sites, commonly workers. To empirically evaluate the effectiveness of the ConSynth dataset in addressing real-world challenges, the next section will present a series of experiments. These experiments are designed to test the dataset's performance in enhancing object detection in real-world construction sites, particularly of small objects.

Experiments

In this section, we explore the effectiveness of using the synthetic ConSynth dataset in the field of object detection, with three focuses: enhancing real data with ConSynth to improve model performance, reducing the reliance on real data by leveraging ConSynth, and evaluating the performance of ConSynth in a case-study construction site.

Training and Validation Setup

We employed the YOLOv8 model, a commonly used open-source one-stage object detector known for its anchor-free design, fast speed, and high accuracy. We used two datasets for training: a real dataset and a synthetic dataset (i.e. ConSynth). The real dataset combines the largest available construction site datasets, MOCS (Xuehui et al., 2021) and ACID (Xiao and Kang, 2021). For ACID, we split it into 80% for training and 20% for validation. The 80% from ACID was merged with the MOCS training set

(a) Real (MOCS + ACID) (b) Synthetic (ConSynth) (c) Case study

Figure 7: Examples from the experiment datasets. ConSynth dataset provides images similar to images from the case study dataset but targets adverse lighting and weather conditions.

to form our training dataset, and the 20% was combined with the MOCS validation set for our validation dataset. The final composition of the real dataset included 24,233 images in the training set and 5,219 images in the validation set. For the ConSynth dataset, encompassing 24,000 images, we similarly allocated 80% (19,200 images) for training and 20% (4,800 images) for validation.
We also created another dataset, the case study dataset, for validation independent from any of the training datasets. We collected 500,000 images from real cameras at the case study construction site, which covers various construction equipment, personnel, weather, and lighting conditions. From these, 10,000 images were randomly selected following a normal distribution. Using a MobileNetv2 model (Sandler et al., 2019) trained with the ImageNet dataset (Deng et al., 2009), we computed the visual embeddings for each image and selected the 100 most unique images based on these embeddings. These images were then manually annotated and reviewed by experienced engineers. Examples from these datasets can be seen in Figure 7.
The training was conducted on four RTX 4090 24G GPUs with a batch size of 16 over 100 epochs. The base learning rate was set to $lr_0 = 0.002$, employing an SGD optimiser with a learning rate factor of $lr_f = 0.01$. The size of the input image was maintained at 640×640. After the training, a key performance metric, mean Average Precision (mAP), was employed to evaluate each model's performance. The mAP is defined as the percentage of correct predictions, algebraically represented as:

$$mAP = \frac{TP}{TP + FP} \quad (1)$$

where TP denotes the number of true positives, FP represents the number of false positives.

Enhancing Model Performance with Synthetic Data

This section explores the impact of using the synthetic dataset ConSynth on model performance in three scenarios: (1) training separately on real and synthetic train sets and validating on their respective validation sets; (2) training separately on real and synthetic train sets and validating on the other validation set; (3) training on combined train sets and validating on combined validation sets. The results are detailed in Table 2.
Our research found that models trained on both the real and synthetic datasets performed well on their respective validation sets, indicating the effectiveness and domain-

Table 2: Performance Comparison of Models. Notably, models trained solely on the Consynth dataset improved the accuracy of detecting dump trucks by 20%. *Models trained on the combined datasets improved the accuracy of detecting dump trucks by* 40% *and all classes by* 13%.

Train	Validation	Worker	Dump Truck	Mixer Truck	All Classes
Real	Real	62.6	67.2	78.3	69.3
Synthetic	Synthetic	31.3	69.8	68.1	56.4
Real	Synthetic	6.1	11.4	10.2	9.3
Synthetic	Real	9.0	3.8	4.3	5.7
Real + Synthetic	Real	62.8	65.1	75.6	67.8
Real + Synthetic	Synthetic	25.9	64.7	61.9	50.8
Real	Case Study	10.4	6.4	13.1	9.9
Synthetic	Case Study	1.9	7.7 (+1.3)	3.1	4.2
Real + Synthetic	Case Study	**10.5** (+0.1)	**9.0** (+2.6)	**14.0** (+0.9)	**11.2** (+1.3)

specific robustness of the datasets. However, the performance of these models on the opposite validation sets was poorer. This performance drop is primarily attributable to the domain gap between the real and synthetic datasets. Despite our efforts to minimise this gap through photo-realistic rendering and domain randomisation during the data generation, a discernible difference remains. Another reason for the performance discrepancy is that the ConSynth dataset more closely resembles the data collected from the specific construction site used in our case study. The distribution of object types, appearances, and sizes in ConSynth differ from those in the public construction datasets, creating an inherent domain gap that will be further analysed in subsequent sections.

We also combined the real and synthetic datasets to bridge the gap between domains. Interestingly, the model trained on the combined dataset showed a slight decrease in performance metrics in each domain. However, this does not necessarily indicate a deterioration in model performance. Instead, it suggests that the model has adapted to data from multiple domains, leading to a more generalised model. This generalisation is a positive attribute, especially for practical applications in complex construction sites where conditions vary significantly.

To validate this generalisation, we evaluated all trained models on the dataset created and annotated from the case study construction site. As shown in Table 2, the model trained on the synthetic dataset has shown promising improvement in dump truck detection as the Mean Average Precision (*mAP*) improved by 20%. The model trained on the combined dataset outperformed the other two models in all aspects. Notably, the improvement in dump truck detection was significant, jumping from 6.4 to 9.0 by 40% while the *mAP* for all classes was improved by 13%. However, despite these improvements, all three models' *mAP* was generally low. This is attributed to the high resolution of site cameras and the distance of these cameras from the targets, resulting in small object sizes in the images. The down-sampling of high-resolution images from 1920×1080 pixels to 640×360 pixels for convolution operations in the backbone likely leads to a loss of critical features, intensifying the challenge of detecting small objects. Yet, with the presence of a large number of small objects in the dataset, the ConSynth dataset helps enhance the model's performance in detecting small objects on construction sites.

Finally, we conducted a quantitative analysis of video footage from our case study construction site to investigate how the ConSynth dataset enhances model performance in specific construction scenarios. The results, depicted in Figure 8, compare the performance of models: (a) trained solely on the real dataset and (b) trained exclusively on the ConSynth dataset, with the latter having no exposure to real dataset images. The findings demonstrate that the model can effectively learn features of different classes from the ConSynth dataset and make accurate predictions, even without real data. In this comparison, the model trained with ConSynth data notably reduced false positives in the worker category and improved the detection of overlapping dump trucks, thereby reducing false negatives. However, it struggled to detect workers partially obscured by railings, a limitation likely due to a lack of representative samples in the synthetic dataset. This issue can be mitigated by combining the ConSynth dataset with real data. Overall, the ConSynth dataset showed promising results, effectively reducing reliance on real data and demonstrating significant potential.

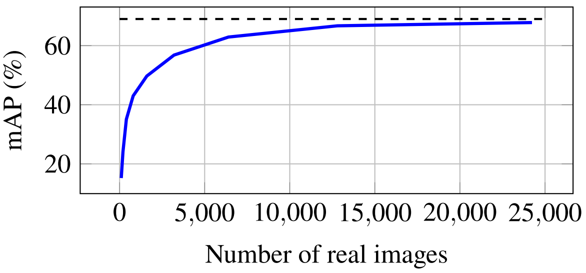
(a) Model trained on real dataset

(b) Model trained on ConSynth dataset

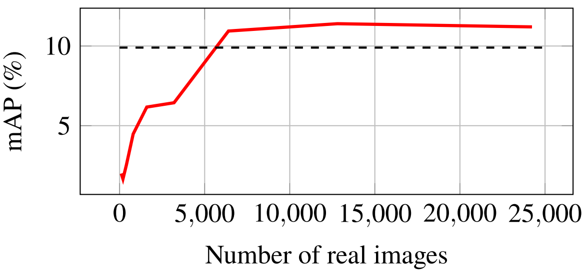
Figure 8: Models trained solely on the ConSynth dataset yield better accuracy in detecting dump trucks and comparable accuracy in detecting workers.

Relieving the Burden of Manual Labelling

In this section, we primarily explore how synthetic datasets can alleviate the reliance on real-world data. We randomly selected k samples from a real training set comprising 24,233 images. The sampling followed a normal distribution, where $k \in [100, 200, 400, 800, 1600, 3200, 6400, 12800]$. We combined the selected samples with the entire synthetic dataset to form eight distinct training sets. These datasets were then used for training with the same hyper-parameters from the previous section.



(a) Model performance (mAP) on the real dataset



(b) Model performance (*mAP*) on the case study dataset

Figure 9: Comparison of model performance in response to varying proportions of real data in training. Reliance on real data can be significantly reduced with synthetic data, especially in underrepresented scenarios targeted by synthetic data.

Validation was performed on the real dataset's validation set, and the results are depicted in Figure 9a as a solid blue line. The dashed line represents the baseline ($mAP = 69.0\%$), which is the performance using only the real dataset for training and validation. The graph indicates that as the proportion of real data in the mixed dataset increases, the model's performance on the real training set improves, signifying the learning of domain-specific features. While the mixed dataset did not surpass the baseline, it is notable that when the real data was reduced by 47.2% (from 24,233 to 12,800), the model's performance remained close, with *mAP*s of 67.8 and 66.7, respectively. Further reduction of real data to 26.4% (from $24,233$ to $6,400$) still resulted in a significant *mAP* (62.9%). This demonstrates that mixing synthetic with real data for training substantially decreases the dependency on real data.

We also verified the performance of models trained on the mixed datasets using data collected from a case study construction site, as shown in Figure 9b. The dashed line represents the performance of models trained on real data for this case study dataset. The trend observed was similar to the previous experiment, where reducing the real data in the dataset by approximately 75% still yielded comparable performance. Notably, the model exceeded the baseline performance on this dataset, indicating the use of synthetic data in improving the model's robustness. This can be attributed to using a digital twin model closely resembling the actual construction site in the synthetic data generation process, which aids in improving the model's performance in underrepresented site scenarios.

Conclusions

In this work, we have introduced a digital twin-based simulation system for generating synthetic datasets for construction site scenarios. Our method outperforms previous approaches by producing photo-realistic images to enhance the overall model performance in various construction scenarios. Moreover, we have introduced ConSynth, a novel synthetic dataset. The dataset is the first video dataset in construction that supports common perception tasks and provides valuable annotations for potential tasks which have not been extensively applied in the industry.

Currently, ConSynth includes three different categories of construction personnel equipment. Future work could expand the dataset's scope by incorporating additional construction equipment, such as excavators, loaders, and cranes, into the simulation. This expansion would further enhance the dataset's applicability and relevance to real-world construction scenarios. In addition, the models evaluated in this study exhibit a relatively low mean Average Precision (mAP) for detecting small objects in high-resolution images. Future work should conduct extensive experiments to explore improvements in mAP, potentially through methods like data augmentation and advanced algorithms, including the Slicing Aided Hyper Inference (SAHI) (Akyon et al., 2022).

We anticipate that our approach will not only improve the accuracy and robustness of computer vision applications in dynamic and challenging construction environments but also significantly reduce the costs associated with data collection and labelling. With a richer dataset, computer vision models are expected to not only accurately locate key targets and detect potential hazards but also contribute to the proactive enhancement of safety monitoring on construction sites, thereby mitigating risk and ensuring the well-being of personnel. In light of this, the ConSynth dataset will be made available at `https://github.com/tsl-imperial/ConSynth`.

Acknowledgments

The authors wish to extend their sincere gratitude to their industry collaborator, Balfour Beatty VINCI SYSTRA Joint Venture (BBVS JV), for their invaluable support and collaboration in this research. The provision of funding and access to the case study construction site for data collection were integral to the success of this study. Their contributions have been fundamental in advancing the research presented in this paper.

References

Akyon, F. C., Altinuc, S. O., and Temizel, A. (2022). Slicing aided hyper inference and fine-tuning for small object detection. 2022 IEEE International Conference on Image Processing (ICIP), pages 966–970.

Assadzadeh, A., Arashpour, M., Brilakis, I., Ngo, T., and Konstantinou, E. (2022). Vision-based excavator pose estimation using synthetically generated datasets with domain randomization. Automation in Construction, 134:104089.

Barrera-Animas, A. Y. and Delgado, J. M. D. (2023). Generating real-world-like labelled synthetic datasets for construction site applications. Automation in Construction, 151:104850.

Bradski, G. (2000). The OpenCV Library. Dr. Dobb's Journal of Software Tools.

Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., and Beijbom, O. (2019). nuscenes: A multimodal dataset for autonomous driving. arXiv preprint arXiv:1903.11027.

Chang, M.-F., Lambert, J., Sangkloy, P., Singh, J., Bak, S., Hartnett, A., Wang, D., Carr, P., Lucey, S., Ramanan, D., and Hays, J. (2019). Argoverse: 3d tracking and forecasting with rich maps.

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, pages 248–255.

Dosovitskiy, A., Ros, G., Codevilla, F., Lopez, A., and Koltun, V. (2017). CARLA: An open urban driving simulator. In Levine, S., Vanhoucke, V., and Goldberg, K., editors, Proceedings of the 1st Annual Conference on Robot Learning, volume 78 of Proceedings of Machine Learning Research, pages 1–16. PMLR.

Feng, D., Haase-Schütz, C., Rosenbaum, L., Hertlein, H., Gläser, C., Timm, F., Wiesbeck, W., and Dietmayer, K. (2021). Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. IEEE Transactions on Intelligent Transportation Systems, 22(3):1341–1360.

Geiger, A., Lenz, P., Stiller, C., and Urtasun, R. (2013). Vision meets robotics: The kitti dataset. International Journal of Robotics Research (IJRR).

Kloukiniotis, A., Papandreou, A., Anagnostopoulos, C., Lalos, A., Kapsalas, P., Nguyen, D.-V., and Moustakas, K. (2022). CarlaScenes: A synthetic dataset for odometry in autonomous driving. In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE.

Lee, H., Jeon, J., Lee, D., Park, C., Kim, J., and Lee, D. (2023). Game engine-driven synthetic data generation for computer vision-based safety monitoring of construction workers. Automation in Construction, 155:105060.

Lee, J. G., Hwang, J., Chi, S., and Seo, J. (2022). Synthetic image dataset development for vision-based construction equipment detection. Journal of Computing in Civil Engineering, 36(5).

Neuhausen, M., Herbers, P., and König, M. (2020). Using synthetic data to improve and evaluate the tracking performance of construction workers on site. Applied Sciences, 10(14):4948.

Poggenhans, F., Pauls, J.-H., Janosovits, J., Orf, S., Naumann, M., Kuhnt, F., and Mayr, M. (2018). Lanelet2: A high-definition map framework for the future of automated driving. In Proc. IEEE Intell. Trans. Syst. Conf., Hawaii, USA.

Richter, S. R., Hayder, Z., and Koltun, V. (2017). Playing for benchmarks. In 2017 IEEE International Conference on Computer Vision (ICCV). IEEE.

Richter, S. R., Vineet, V., Roth, S., and Koltun, V. (2016). Playing for data: Ground truth from computer games.

Ros, G., Sellart, L., Materzynska, J., Vazquez, D., and Lopez, A. M. (2016). The SYNTHIA dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., and Chen, L.-C. (2019). Mobilenetv2: Inverted residuals and linear bottlenecks.

Soltani, M. M., Zhu, Z., and Hammad, A. (2016). Automated annotation for visual recognition of construction resources using synthetic images. Automation in Construction, 62:14–23.

Sun, T., Segu, M., Postels, J., Wang, Y., Van Gool, L., Schiele, B., Tombari, F., and Yu, F. (2022). Shift: A synthetic driving dataset for continuous multi-task domain adaptation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pages 21371–21382.

Tremblay, J., Prakash, A., Acuna, D., Brophy, M., Jampani, V., Anil, C., To, T., Cameracci, E., Boochoon, S., and Birchfield, S. (2018). Training deep networks with synthetic data: Bridging the reality gap by domain randomization.

Weng, X., Man, Y., Cheng, D., Park, J., O'Toole, M., and Kitani, K. (2020). All-In-One Drive: A Large-Scale Comprehensive Perception Dataset with High-Density Long-Range Point Clouds. arXiv.

Xiao, B. and Kang, S.-C. (2021). Development of an image data set of construction machines for deep learning object detection. Journal of Computing in Civil Engineering, 35(2).

Xuehui, A., Li, Z., Zuguang, L., Chengzhi, W., Pengfei, L., and Zhiwei, L. (2021). Dataset and benchmark for detecting moving objects in construction sites. Automation in Construction, 122:103482.

INDUCTIVE THEORIZATION OF CONSTRUCTION DIGITAL TWINS: DERIVING A TAXONOMY OF APPLICATIONS AND ENABLING TECHNOLOGIES

Wahib Saif [1,2], SeyedReza RazaviAlavi[1], Kay Rogage[1], Xiang Xie[3], and Mohamad Kassem[3]
[1]Northumbria University, Newcastle, UK
[2]BIM Academy, Newcastle, UK
[3]Newcastle University, Newcastle, UK

Abstract

Digital twin in the construction phase—termed Construction Digital Twin (CDT)—faces more significant challenges than in other phases, such as operation or maintained phase, due to the dynamic and evolving nature of construction sites and their broad spectrum of applications. To enhance our understanding of the CDT domain, it is crucial to clearly define it, establish a detailed taxonomy of its current applications and enabling technologies, and elucidate how attributes, data requirements, and technology choices within CDT vary across different construction applications. This objective remains unmet in existing literature, a gap this paper addresses through an approach combining systematic review, thematic coding, and conceptualization of CDT architecture comprising of five layers: sensing, communication, storage, analytics, and visualization. The study identifies seven major applications of CDT and maps them to the five architectural layers and their enabling technologies, providing insights into the suitability and prevalence of these technologies for specific applications of interest.

Introduction

Digital twin (DT) is an emerging and promising advancement that has the potential to transform the construction industry and drive its productivity (Boje *et al.*, 2020; Opoku *et al.*, 2021). DT is generally perceived as an up-to-date digital representation of a physical system and its functional properties leveraging data streaming from a wide range of data capturing and communication technologies (Sacks *et al.*, 2020). It can provide data-driven insights about the state and performance of an object or a process in real time thus enabling informed decision-making capabilities. DT generally has three main elements: the physical environment, its digital representation, and the streaming of data/feedback that connects the two, enabled by data sensing technologies, artificial intelligence functions, and high-speed networking (Sacks *et al.*, 2020).

Although DT is still regarded as a nascent concept within the construction industry, it has attracted substantial attention from researchers over the past five years investigating its applications during the different phases of built asset projects, with a particular focus on the design and operational phases (Opoku *et al.*, 2021). DT applications during the construction phase have been emerging recently and still require extensive research.

Hence, this paper aims to present a systematic approach to reviewing, mapping, and classifying the different applications of DT exclusively during the construction phase of built assets, namely Construction Digital Twin (CDT). Doing so provides transparency to readers on the most comprehensive state-of-the-art applications of DT and further supports its adoption during the construction phase. This paper also aims to thoroughly examine the different DT systems presented within the literature to give a generalized anatomy of DT architecture and the different technologies integrated within its identified layers. This will provide knowledge of the concepts, technologies, and processes required to replicate or build upon existing systems thus enhancing the scalability and adaptability of future-developed DTs.

Research Methodology

A systematic review was performed by following the reporting checklist of PRISMA to present a transparent, replicable, updateable, and comprehensive summary of previous studies related to CDT domain. Three databases were selected (i.e., Scopus, Web of Science, and Google Scholar) for carrying out the preliminary literature search. For Google Scholar, the search string included keywords of "Digital twin" and "Construction site" or "Construction Phase" and only results within the first 30 pages were included as results were not relevant anymore to the scope of this search. For Scopus and Web of Science, these keywords were expanded to form a broad search string presented in Table 1. The search, unrestricted by publication date up to July 2023, retrieved 504 DT publications from three databases, which were then systematically filtered using the inclusion and exclusion criteria detailed in Table 2. Figure 1 summarizes the literature search and selection procedure. The final set of 112 papers was examined to provide a thematic analysis that highlights the different applications of CDT and enabling technologies employed within each layer of a proposed conceptualized architecture.

Table 1. Scopus search string

(TITLE-ABS-KEY ("digital" OR "virtual" OR "cyber") AND TITLE-ABS-KEY ("twin*" OR "replica*" OR "shadow" OR "counterpart*") AND TITLE-ABS-KEY ("construction site*" OR "building site" OR "infrastructure site" OR "*project site*" OR "construction phase*" OR "building phase" OR "construction stage" OR "building stage" OR "engineering projects" OR "engineering site*" OR "construction operation*" OR "construction task*" OR "construction activity*" OR "building operation*" OR "building task*" OR "building activity*" OR "building project*" OR "construction management*"))

Table 2. *Inclusion and exclusion criteria for the literature search*

Inclusion criteria	Exclusion criteria
- Journal Articles - Conference papers - Review papers - Final publication stage - Any publication date - English language documents - Literature covering the scope of this research (i.e., DT for construction phase applications). - Literature that covers DT for the whole construction industry and includes sections related to the construction phase or construction management.	- Non-English language. - In press documents - Book chapters - Duplicate studies reported on research that had been previously published (e.g., a conference paper gets extended into a journal article or vice versa) - Studies that are irrelevant to DT concept - Research Studies from other industries (e.g., manufacturing, aerospace, oil & gas, naval, marine, aviation, urban/city planning, etc.) - Studies that are out of scope (e.g., DT implementations during other phases like the design, operational, or maintenance phases.

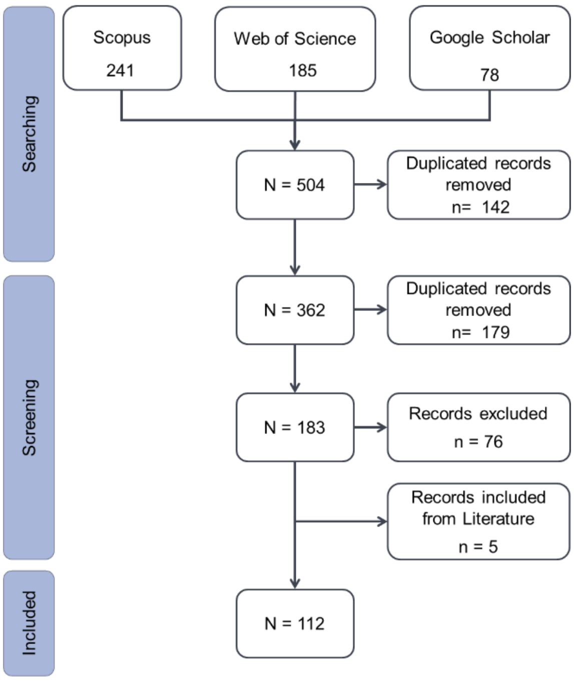


Figure 1. Study selection process based on PRISMA.

Taxonomy of CDT Applications

Seven major CDT applications and 17 sub applications emerged from the conducted thematic analysis. This includes 1- Safety and Risk Management (SRM), 2- Progress Monitoring and Control (PMC), 3- Data Integration and Management (DIM), 4- Construction Robotics and Automation (CRA), 5- Supply Chain and Logistics (SCL), 6- Quality Control and Assurance (QCA), 7- Sustainability and Circular Construction (SCC). Figure 3 demonstrates the distribution of these major applications within the literature, providing insights into the level of research corresponding to each application. For instance, "Safety and Risk Management" emerged as the most investigated area of CDT application while "Sustainability and Circular Construction" attracted the least research attention, thus offering a promising area of CDT for future research. Table 3 summarizes the seven CDT applications, their sub-implementations, and the corresponding key technologies and processes.

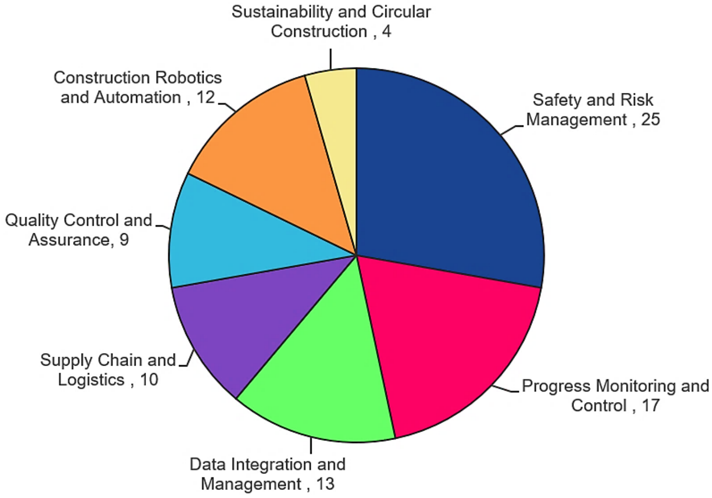


Figure 2. Distribution of the identified CDT applications across the reviewed studies (total number of 90).

CDT Architecture and Enabling Technologies

This section introduces a conceptualized architecture of CDT to better explore and categorize the different enabling technologies and processes integrated into a CDT to serve its intended application, (see Figure 3). The proposed architecture presents five distinct layers constructing a CDT: sensing, communication, storage, analytics, and visualization layer, their data interaction, and the feedback-decision making- loop mechanism. In this context, a "layer" denotes the logical combination of the hardware and/or software components sharing common functionality and purpose. Systematically, a "CDT architecture" is an enclosed combination of the required layers constructed to interact in a way that enables data collection, transmission, processing, and visualization to support informed decision making.

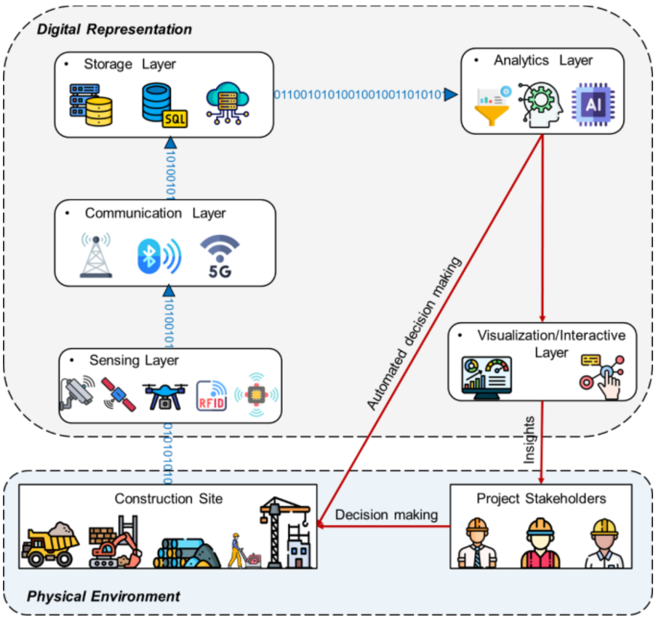


Figure 3. The proposed conceptualized CDT architecture

Table 3. A summary of the identified thematic applications of CDT and their corresponding technologies and processes

Application	Sub-application	Implementations Summary	Key technologies and processes	Reference
SRM	Workers' Health and Safety Management	Tracking workers' locations and issuing proactive warnings for trespassing hazardous zones; Monitoring Worker's physical health and posture; and Interactive site visualization for safety training.	RFID, LoRa, GPS, Insole pressure sensors, YOLOv4, Deep SORT, Deep learning, VR, MR headsets	(Lee and Lee, 2023; Osti et al., 2021; Wu et al., 2022; Zhang et al., 2023)
	Hazards Identification and Prevention	Identify and proactively prevent safety threats in a job site (e.g., tunnelling, crane operations, and drilling tasks).	Photogrammetry, kinematic sensors, spatial/orientation sensors, BIM, RFID, DNN, LSTM, XGBoost, BPNN, and AdaBoost.	(Jiang, Ding, et al., 2021; Kamari and Ham, 2022; Shariatfar et al., 2022; Zhao et al., 2022)
	Equipment Health Monitoring	Monitoring the operating status of construction equipment, mainly tower cranes, against mechanical and structural failure.	Mechanical and motion sensors, inclination and load sensors, RFID, IMU, Finite Element analysis (FEM), Multi Fidelity Surrogate model, Deep learning, LoRa, Visualization interface	(Jiang, Ding, et al., 2022; Lai et al., 2022; Li et al., 2021; Liu, Li, et al., 2022)
	Human-Machinery Interaction	Enhancing workers-machinery interactions on site to prevent accidents and fatalities in limited space. Earthwork is the main use case	Bluetooth Low Energy (BLE), Real-time location system (RTLS), Bayesian network, cameras, YOLOv5	(Assadzadeh et al., 2023; Huang et al., 2021; Kojima et al., 2020; Leonardo et al., 2020)
PMC	Productivity Estimation and Prediction	Equipment tracking for estimating and predicting productivity and utilization rates (i.e., mainly plant equipment).	IMU, GPS, 3D Accelerometer, kinematic sensors, Discrete Event Simulation, ANN, cloud computing, MySQL database, Supervised learning, dashboards, CNN, RNN, Cellular Network	(Fischer et al., 2023; Rogage et al., 2022; Salem and Moselhi, 2021)
	Progress Tracking and Control	Monitoring and most importantly visualizing the progress of certain construction activities (e.g., prefabricated element installation)	Onsite cameras, sound sensors, GPS, RFID, 4D BIM, YOLOv4, DNN, 3D visualization model, Web-based platform, Wireless communication, Cloud-based computing and storage	(Hasan and Sacks, 2023; Huang et al., 2021; Shariatfar et al., 2022)
	As-Built vs As-Planned Comparison	Capturing as-built progress (mainly point cloud data) and comparing it with as-planned data (mainly BIM data) for detecting progress deviations.	Laser scanning, LiDAR, Photogrammetry, Point clouds, 4D BIM, Reality capture, Object recognition, Semantic segmentation.	(Alizadehsalehi and Yitmen, 2023; Pan and Zhang, 2021; Pour Rahimian et al., 2020; Rausch and Haas, 2021)
DIM	Blockchain-enabled DT	Integrating DT with blockchain to enable traceable, immutable, and secure data sharing.	Blockchain, Smart contracts, SHA-256 algorithm, Data querying	(Jiang, Liu, et al., 2021; Lee et al., 2021; Zhao et al., 2023),
	Data Schemas-enabled DT	Utilizing semantic web technologies and data schemas to address the integration of heterogenous construction data.	Ontologies, Neo4j graphs, Semantic web, IFC-schema, SPARQL.	(Ayinla et al., 2021; Dong et al., 2021; Schlenger et al., 2022)
CRA	Machinery Remote Control	Integrating interactive and real-time reality capturing for remotely controlling equipment, mainly plant equipment.	VR, MR, LiDAR, Cameras, 4G/5G networks. WebSocket, node-red UI.	(Hasan et al., 2022; Heikkilä et al., 2022; Hoffmann et al., 2022; Schönböck et al., 2022)
	Robotics Tasks Management	Planning and supervising robotic tasks by controlling the robot's joint angles and/or its cartesian path using a simulator platform.	VR, MR, Gazebo simulator, MQTT, ADS protocol, KUKA KR120, IFC-based schema, SLAM, Context-awareness, Ethernet	(Asadi et al., 2018; Ding et al., 2020; Liang et al., 2022; Wang et al., 2021)
SCL	Material Logistics	Integrating and visualizing logistics data including material inventory, incoming deliveries, and material consumption rates.	RFID, GPS, BIM data, Deep reinforcement learning (DRL)	(Deria et al., 2022; Gehring and Rüppel, 2023; Jiang, Li, et al., 2022; Lee and Lee, 2021)
	Resource Allocation	Managing and optimizing the allocation of construction resources	RFID, DRL, Dijkstra algorithm, Queuing model, Genetic algorithm, 4D BIM.	(Esmaeili and Simeone, 2023; Lee et al., 2022; Liu, Shi, et al., 2022)
QCA	As-Built vs As-Designed Verification	Ensuring the geometric quality of as-built products by verifying them with as-designed models; common with prefabricated elements.	Laser scanning, LiDAR, photogrammetry, MR, AR, Scan-vs-BIM, Scan-to-BIM, CNN.	(Rausch and Haas, 2021; To et al., 2021; Tran et al., 2021)
	Regulatory Standards Compliance	Checking the quality of construction products and processes based on their compliance with relevant regulations and standards (e.g., concrete maturity properties and road compaction).	Piezoelectric sensor, Velocity sensor, BIM, Thermocouple sensor, FEM.	(Han et al., 2022; Kosse et al., 2022; Posada et al., 2022)
SCC	Embodied Carbon Monitoring	Estimating and monitoring the emission of embodied carbon dioxide during the construction process of a built asset.	QR code, UWB, IFC-based data schema, RFID, Accelerate sensor, BIM model, Wi-Fi	(Chen et al., 2021; Mao et al., 2018; Shen et al., 2022)
	Circular Construction	Facilitating the reuse and recycling of construction materials to reduce waste and enhance the efficiency of resource utilization	GPS, GIS data, Monte Carlo-based simulation	(Züst et al., 2021)

Sensing layer

The sensing layer is the first and most fundamental layer within a DT architecture that employs one or more data-sensing devices to serve as the nervous system responsible for gathering raw data from the physical environment in which they were deployed. While only devices that stream data in a timely manner should technically be considered part of the sensing layer, static data sources (e.g., BIM models, GIS, schedules, and regulations) are also included to provide a more comprehensive overview of all data sources available for supporting a DT system. The selection of data sources and sensing technologies significantly affects the accuracy, reliability, and acquisition speed of raw data, and consequently, the efficiency and overall performance of the DT system. Thus, understanding the data requirements that align with the DT's primary purpose is crucial. Figure 4a maps the different sensing technologies employed within the CDT systems reviewed from the literature to the identified applications. The size of each node represents how frequently a technology was used. RFID was the most common sensing technology for safety and risk management implementations. GPS was widely used for progress monitoring, robotics, and logistics due to its real-time, precise location data for equipment, robots, and vehicles. IMUs followed GPS, mainly capturing motion data for progress monitoring and construction robotics.

Communication layer

The communication layer plays a key role in receiving the raw data acquired by the sensing layer and transmitting it to the CDT hub to be stored, processed, or visualized. Typically, the communication layer utilizes either a wireless network (e.g., Wi-Fi, Bluetooth, LoRa, 5G/4G cellular, etc.), a wired network (e.g., Ethernet, USB, and Fieldbus), or occasionally a hybrid of both. Figure 4b maps the identified applications to the different employed communication networks and protocols to give insights into their suitability and popularity. For example, Ethernet is featured in most studies related to controlling scale model robotics (Asadi *et al.*, 2018; Liang *et al.*, 2022). By employing the MQQTT protocol, Ethernet allows for bi-directional communication between the physical robot and its digital model enabling smooth robotic control (Han *et al.*, 2022; Liang *et al.*, 2022). Short-range wireless technologies (i.e., Bluetooth and Wi-Fi) are commonly utilized in various applications due to their higher transfer rates. Bluetooth Low Energy (BLE), a more energy-efficient choice compared to traditional Bluetooth, is also used in several studies as it is suitable for battery-powered devices and when long-period data collection with short-range coverage is required. Within long-range wireless technologies, LoRa is found to be a popular option for safety and risk management applications, specifically with hoisting operations owing to its long-range coverage (~10 km) and low-power requirement personnel (Zhang *et al.*, 2023). In general, selecting a communication network and protocols depends on several factors determined by the overall purpose of CDT including the required transmission rate, latency, range of coverage, security, and power consumption.

Storage layer

This layer provides a repository for hosting and managing the vast amount of data including the sensed raw data, data from other sources, and historical data. As shown in Figure 4c, cloud-based storage emerged as the predominant choice for data storage within CDT on account of its capacity to accommodate extensive data volumes and streamline their retrieval over the Internet (Pan and Zhang, 2021). Both SQL, relational databases, and NoSQL, non-relational databases, have been adopted in various construction applications. The selection of either relies on the types of data, its structure, and volume. For unstructured and variant data in substantial volumes (e.g., visual data), NoSQL is best due to its dynamic schema with higher scalability and adaptability. On the other hand, SQL is fit to manage well-defined data that has a structured schema (e.g., workers' information) as it gets arranged in structured tables with predefined relationships.

Analytics layer

This layer functions as the CDT brain in charge of processing, analyzing, and deriving insights from the collected data to support informed decision-making. Figure 4d maps the various data analytics techniques used for each CDT application. These techniques can be streamed into three main categories: a) ML-based methods that employ deep neural networks and object recognition and tracking algorithms (e.g., CNN, RNN, YOLO, and Deep SORT). b) Numerical/Simulation-based methods (e.g., Finite element method, discrete event simulation, queuing theory, Dijkstra algorithm, and genetic algorithm) that are mostly common with applications requiring simulation and optimization (e.g., equipment health monitoring, hazards identification, resource allocation, and material logistics). c) Human interpretation which is mostly featured in quality inspection applications where defects get visually detected in as-built data (e.g., point clouds). Construction robotic applications also highlighted human supervision for planning and remotely controlling robotic tasks.

Visualization layer

This layer enables the visual representation of the processed data and insights in a dynamic and interactive environment. Different visualizations can be employed within the CDT architecture depending on the data type intended to be visualized. For example, 3D data (e.g., point clouds and BIM models) are commonly utilized to visualize progress data for quality control and progress monitoring purposes. 4D BIM models are also employed to identify and visualize hazardous activities and locations and create safety precautions. Extended reality (XR) technologies are employed in various CDT applications as they allow stakeholders to intuitively interact with visualized data in an easily accessible way.

Several game engines and modelling software are featured within the literature to create 3D immersive environments. As presented in Figure 8, the “Unity engine” is the most used game engine for this purpose as it supports modeling and hosting DT entities, data integration from multiple sources, and running data analytics within its environment. Such features make it a favorable platform for visualizing DT data. Other game engines and modelling software are also highlighted, including Unreal Engine, and Blender platform. 3D visualization environments were not included in many applications (e.g., material logistics, embodied carbon estimations, and productivity measurements) as the type of data is mainly numerical and can be visualized using graphical representations within customized dashboards.

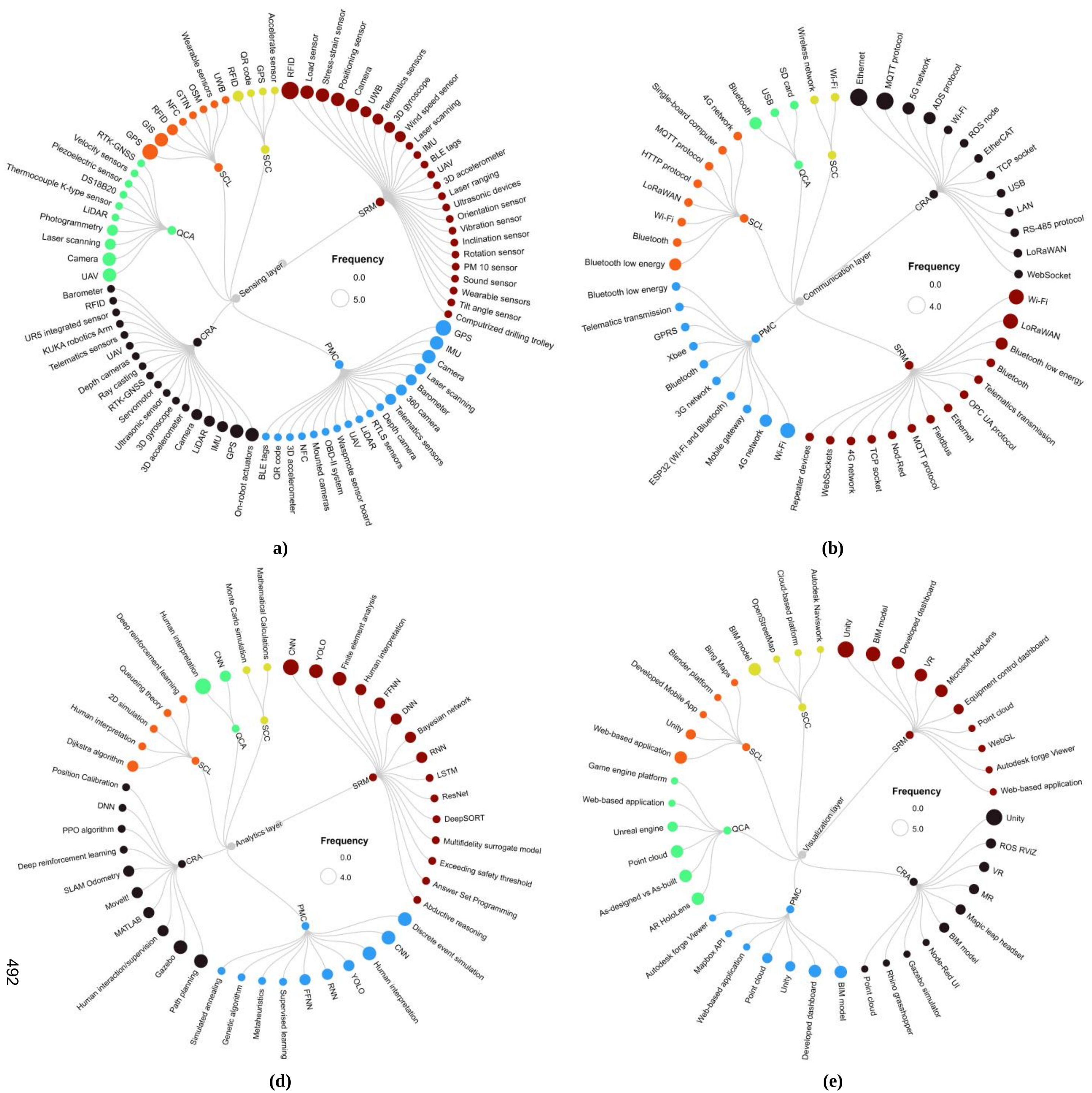


Figure 4: The different enabling technologies employed within each DT layer mapped to the identified applications: (a) Sensing, (b) Communication, (c) Storage, (d) Analytics, (e) Visualization.

SRM: Safety and Risk; Management; PMC: Progress Monitoring and Control; CRA: Construction Robotics and Automation; QCA: Quality Control and Assurance; SCL: Supply Chain and Logistics; and SCC: Sustainability and Circular Construction.

Discussion

Evaluation of CDT applications

Comparing the findings of this study with previous review studies highlights the significant progression of DT implementations during the construction phase of built projects as most prior implementations are mainly focused on the operational and maintenance stage. Additionally, most construction applications highlighted in prior review studies are mostly nD BIM applications (e.g., clash detections, cost estimation, scheduling, and communication) (Boje *et al.*, 2020) or implementation of the integration of BIM with other technologies (e.g., XR, Simulation, AI). However, the thematic applications that emerged in this study are mainly DT-related implementations. Some of these applications align with DT research clusters categorized by (Jiang, Ma, *et al.*, 2021) in early 2021including safety, quality, and progress management. These applications have been expanded since then and new ones have emerged including construction robotics, logistics, and sustainability. Promising CDT research areas are recommended that are related to applications such as "circular construction", "embodied carbon estimations", "full-scale robotics", and "graph-based DT". New construction applications are anticipated to emerge, contributing to the increasing adoption of DT in both academia and industry.

What is CDT?

This section provides a discussion on the key attributes of a CDT system including the data acquisition mechanism, the feedback-decision making-loop dynamics, and the role of 3D modeling and BIM within CDT systems. Ultimately, synthesizing these discussions to propose a definition capturing the essence of CDT.

Data streaming mechanism

Almost all reviewed studies integrated at least one form of data sensing mechanism to provide DT with a dynamic stream of data without which a system cannot be considered as DT. Nonetheless, the level of automation, frequency, and latency parameters of data acquisition can all vary across different DTs based on several factors including the main purpose of the DT, the type of data that needs to be collected, and the rate at which is being updated. Several studies labelled their data-collecting mechanism as *"real-time"* or *"near-real-time"* without numerically defining these abilities in terms of latency and frequency. Only a few studies provided statistical information about these parameters to accurately reflect the capability of their data acquisition system (Deria *et al.*, 2022; Liang *et al.*, 2022; Posada *et al.*, 2022). Although in many articles especially within the manufacturing industry, being able to provide real-time data with high frequency and low latency is a requirement for a system to be a DT, in many construction applications that might not be achievable or even valuable. For example, in some applications, it was adequate to collect data, mostly visual, at a frequency of once a day (Hasan and Sacks, 2023; Pour Rahimian *et al.*, 2020; Zhao *et al.*, 2022). Yet, in other applications where data gets updated more dynamically, a higher frequency is essential. For instance, for tracking the angle movement of a robotic arm, a frequency of up to 250 Hz might be required to synchronize with the rapid rate of update of the moving arm (Liang *et al.*, 2022).

In general, acquiring vast amounts of construction data with high frequency might not be needed and comes with high costs and significant environmental impact associated with collecting, transmitting, storing, and processing this data. Therefore, when designing a DT, these attributes should be optimized to create a balance of being efficient and meeting its overall purpose requirements. Similarly, the degree of autonomy in data acquisition varies widely, with some cases featuring semi-automated approaches, particularly in laser scanning and photogrammetry where human intervention is required for capturing and processing the visual data. In contrast, many studies employ fully automated sensing sources, like positioning sensors capable of transmitting data without any human involvement.

This is to emphasize that the spectrum of data acquisition capabilities for DTs is broad, and achieving a fully automated data stream within specific latency and frequency parameters is not always a must-met criterion for classifying a system as a DT.

Feedback loop—decision making—dynamics

The majority of the DT research in digitally advanced fields (e.g., manufacturing) emphasizes automated bi-directional communication between the physical entity and its digital counterpart (Ladj *et al.*, 2021). In manufacturing, its typically well-designed and controlled environments (e.g., an assembly line) enable the establishment of an automated feedback loop. However, with dynamic, variant, and unpredictable construction operations, establishing an automated feedback channel from the DT to the site can be challenging. Instead, the feedback in some construction applications often involves visualized insights presented to stakeholders for decision-making, highlighting a human-oriented feedback channel, (see Figure 3). For instance, within progress monitoring applications, site stakeholders are presented with performance metrics, which they then align with baseline targets to inform their decision-making.

Almost all construction robotics and automation applications feature fully automated feedback channels where control instructions are given back to a prototype machinery or a robot to conduct a desired task based on the acquired data. Nonetheless, most of these studies used scaled machinery or robotic assembly to validate their systems, raising concerns about their scalability and real-world construction site applicability. Automated feedback was also employed in some safety applications in the form of proactive visual or audible warnings automatically triggered when a potential hazard is detected (Jiang, Ding, *et al.*, 2022; Zhang *et al.*, 2023).

BIM-CDT relationship

The relationship between DT and BIM has been a subject of debate among researchers, with some viewing DT as an extension of BIM with the incorporation of new technologies (e.g., IoT, AI, and simulation), while others perceive them as two distinct concepts with significant differences in their characteristics and applications. In many applications where objects' non-geometric information (e.g., attributes and relationships to other objects) are the main data of interest rather than the geometric information, BIM may not always be necessary for DT functionality and simple 3D models with low Level of Detail (LoD) would be enough for visual representation.

Examples of such applications are workers' health monitoring where workers' information and status are of interest or in material logistics where material quantities or locations are tracked. Nevertheless, in several applications, BIM models were still integrated into CDT layers, serving different purposes as follows: a) within the sensing layer as a source of static data (e.g., schedule, quantities, semantic and geometric data); b) within a storage layer as a unified storage model that integrates and stores various relevant data; c) provides a 3D visualization and interactive environment, possibly integrated with dashboards, within the visualization layer.

A proposed CDT definition

Several definitions of DT exist within the construction industry, often borrowed from advanced sectors like aerospace and manufacturing, where the term originated and developed. These definitions typically include terms like virtual replica, 3D model, digital counterpart, realistic model, and real-time representation of physical. However, a critical question arises: Is DT primarily about high LOD modeling and replicating physical objects, or is it fundamentally about data management and informed decision-making? Particularly in construction implementations, the focus is more on data capturing and analysis that enables informed decision making rather than 3D replication or modeling. Then, there is a need for a definition that emphasizes the data management aspect of DT in construction. Drawing insights from the extensive literature review and the conceptualized DT architecture in Figure 3, this study proposes the following definition:

> *Construction Digital Twin (CDT) is a system responsible for collecting and processing construction dynamic data about a physical entity within the intended application's temporal demand, translating this data into actionable knowledge to enable informed decision making.*

"*Dynamic data*" refers to data that is continuously changing and collected through automated sensing mechanism. It is important to note that CDT can also combine static data from BIM models or historical records to supplement the dynamic data when needed. A "*physical entity*" can be a tangible object (e.g., a structure, equipment, or person) or a physical process (e.g., material delivery or element installation). "*Temporal demand*" acknowledges the varied data latency and frequency requirements across different DT construction applications, as discussed earlier. Consequently, terms like "real-time" and "near real-time" are avoided to highlight the variability in data capturing and processing rates and the lack of precise numerical definitions for these terms. Informed decision-making can be either fully automated or human-assisted, as illustrated in Figure 3.

Literature gaps and recommendations for future research

- Many of the reviewed studies relied on a single sensing source for data acquisition mainly to avoid the challenges of integrating data streaming from multiple sources with different formats. Still, it is frequently stated that depending on a single source for onsite data falls short of providing accurate and comprehensive insights into scalable operations (Hasan and Sacks, 2023; Hoffmann et al., 2022; Pan and Zhang, 2021; Salem and Moselhi, 2021). Only by merging data from multiple sources, a better understanding of construction operations and better-informed decision-making is obtained. Hence, it is recommended to aim at integrating heterogeneous data across multiple domains, leveraging sensor fusion to bridge this gap. Moreover, publishing well-defined data schemas for construction applications can facilitate representing complex information and potentially contribute to developing industry standards.

- Another gap is the absence of standardisation in DT design and deployment including the configuration and specifications of key parameters including data acquisition frequency and latency, analytics accuracy, and its degree of autonomy. Most studies loosely labelled their DTs as "real-time" or "near-real-time" systems without numerically defining those parameters and linking their defined attributes to the DT's main purpose.

- Besides, many studies showcased a limited scope by carrying out a single data collection trial, failing to illustrate how the data and analyzed metrics are regularly updated in a timely manner. This raises questions about the usability, scalability, and versatility of their DT systems. Standardizing DT implementation is key for facilitating its adoption in the industry as it ensures consistency and compatibility.

- Several studies relied on human interpretations for analysing the collected data. This is especially the case with most of the quality inspection and progress tracking applications. Even when automated data analytics are integrated, they are limited to presenting descriptive analytics ("What has happened? As-is analysis"). Only a few studies employed predictive analytics (Lee and Lee, 2021; Li et al., 2021). It is recommended for future research to gear towards enabling CDT to provide predictive ("What will happen? -To-be analysis") and prescriptive analytics ("What should be done? -What-if analysis") by integrating advanced AI and simulation techniques. Doing so will unlock the complete potential of CDT as a proactive decision-support system.

Conclusion

This study serves as a valuable resource for researchers and practitioners in the construction industry by presenting a structured taxonomy of CDT applications and their interaction with different technologies through systematic review and thematic analysis. Seven main thematic applications of CDT are identified and discussed. Additionally, a conceptualized architecture of CDT, comprising five distinct layers, is proposed to aid in exploring and analyzing enabling technologies and processes, and mapping them to the identified applications. The paper also explores and discusses the characteristics that define the essence of CDT and proposes a definition emphasizing its focus on data management and decision-making. These findings not only consolidate existing knowledge in the CDT domain but also provide a foundation for guiding future research and development in this field.

References

Alizadehsalehi, S. and Yitmen, I. (2023), "Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX)", *Smart and Sustainable Built Environment*, Vol. 12 No. 1, pp. 200–236, doi: 10.1108/SASBE-01-2021-0016.

Asadi, K., Ramshankar, H., Pullagurla, H., Bhandare, A., Shanbhag, S., Mehta, P., Kundu, S., *et al.* (2018), "Vision-based integrated mobile robotic system for real-time applications in construction", *Automation in Construction*, doi: 10.1016/j.autcon.2018.10.009.

Assadzadeh, A., Arashpour, M., Li, H., Hosseini, R., Elghaish, F. and Baduge, S. (2023), “Excavator 3D pose estimation using deep learning and hybrid datasets”, *Advanced Engineering Informatics*, doi: 10.1016/j.aei.2023.101875.

Ayinla, K., Vakaj, E., Cheung, F. and Tawil, A.R.H. (2021), “A semantic offsite construction digital Twin-Offsite Manufacturing Production Workflow (OPW) ontology”, *CEUR Workshop Proceedings*, Vol. 2887.

Boje, C., Guerriero, A., Kubicki, S. and Rezgui, Y. (2020), “Towards a semantic Construction Digital Twin: Directions for future research”, *Automation in Construction*, doi: 10.1016/j.autcon.2020.103179.

Chen, C., Zhao, Z., Xiao, J. and Tiong, R. (2021), “A conceptual framework for estimating building embodied carbon based on digital twin technology and life cycle assessment”, *Sustainability (Switzerland)*, doi: 10.3390/su132413875.

Deria, A., Ghannad, P. and Lee, Y.C. (2022), “Integrating AI in an Audio-Based Digital Twin for Autonomous Management of Roadway Construction”, *Construction Research Congress 2022: Computer Applications, Automation, and Data Analytics - Selected Papers from Construction Research Congress 2022*, Vol. 2-B, pp. 530–540, doi: 10.1061/9780784483961.056.

Ding, L., Jiang, W., Zhou, Y., Zhou, C. and Liu, S. (2020), “BIM-based task-level planning for robotic brick assembly through image-based 3D modeling”, *Advanced Engineering Informatics*, doi: 10.1016/j.aei.2019.100993.

Dong, M., Yang, B., Liu, B., Wang, Z. and Zhang, B. (2021), “Realizing, Twinning, and Applying IFC-based 4D Construction Management Information Model of Prefabricated Buildings”, *Proceedings of the International Symposium on Automation and Robotics in Construction*, doi: 10.22260/isarc2021/0021.

Esmaeili, I. and Simeone, D. (2023), “A General Contractor’s Perspective on Construction Digital Twin: Implementation, Impacts and Challenges”, *Buildings*, Multidisciplinary Digital Publishing Institute, Vol. 13 No. 4, p. 978, doi: 10.3390/buildings13040978.

Fischer, A., Beiderwellen Bedrikow, A., Tommelein, I.D., Nübel, K. and Fottner, J. (2023), “From Activity Recognition to Simulation: The Impact of Granularity on Production Models in Heavy Civil Engineering”, *Algorithms*, Multidisciplinary Digital Publishing Institute, Vol. 16 No. 4, p. 212, doi: 10.3390/a16040212.

Gehring, M. and Rüppel, U. (2023), “Data fusion approach for a digital construction logistics twin”, *Frontiers in Built Environment*, Frontiers Media S.A., Vol. 9, p. 1145250, doi: 10.3389/fbuil.2023.1145250.

Han, T., Ma, T., Fang, Z., Zhang, Y. and Han, C. (2022), “A BIM-IoT and intelligent compaction integrated framework for advanced road compaction quality monitoring and management”, *Computers and Electrical Engineering*, doi: 10.1016/j.compeleceng.2022.107981.

Hasan, S. and Sacks, R. (2023), “Integrating BIM and Multiple Construction Monitoring Technologies for Acquisition of Project Status Information”, *Journal of Construction Engineering and Management*, doi: 10.1061/jcemd4.coeng-12826.

Hasan, S.M., Lee, K., Moon, D., Kwon, S., Jinwoo, S. and Lee, S. (2022), “Augmented reality and digital twin system for interaction with construction machinery”, *Journal of Asian Architecture and Building Engineering*, Vol. 21 No. 2, pp. 564–574, doi: 10.1080/13467581.2020.1869557.

Heikkilä, R., Immonen, M., Keränen, H., Liinamaa, O., Piri, E. and Kolli, T. (2022), “5G Wireless Communication for Autonomous Excavation”, *Proceedings of the International Symposium on Automation and Robotics in Construction*, doi: 10.22260/isarc2022/0076.

Hoffmann, N., Saunier, L., Prouten, S., Serroukh, O., Le Floch, J.C., Preda, M., Fetita, C., *et al.* (2022), “Industrial Use-Case : Digital Twin for Autonomous Earthwork in Virtual-Reality”, *Proceedings - Web3D 2022: 27th ACM Conference on 3D Web Technology*, doi: 10.1145/3564533.3565803.

Huang, Y., Hammad, A., Torabi, G., Ghelmani, A. and Guevremont, M. (2021), “Towards Near Real-time Digital Twins of Construction Sites: Developing High LOD 4D Simulation Based on Computer Vision and RTLS”, *Proceedings of the International Symposium on Automation and Robotics in Construction*, Vol. 2021-Novem, pp. 248–255, doi: 10.22260/isarc2021/0036.

Jiang, F., Ma, L., Broyd, T. and Chen, K. (2021), “Digital twin and its implementations in the civil engineering sector”, *Automation in Construction*, doi: 10.1016/j.autcon.2021.103838.

Jiang, W., Ding, L. and Zhou, C. (2021), “Cyber physical system for safety management in smart construction site”, *Engineering, Construction and Architectural Management*, doi: 10.1108/ECAM-10-2019-0578.

Jiang, W., Ding, L. and Zhou, C. (2022), “Digital twin: Stability analysis for tower crane hoisting safety with a scale model”, *Automation in Construction*, Vol. 138, doi: 10.1016/j.autcon.2022.104257.

Jiang, Y., Li, M., Li, M., Liu, X., Zhong, R.Y., Pan, W. and Huang, G.Q. (2022), “Digital twin-enabled real-time synchronization for planning, scheduling, and execution in precast on-site assembly”, *Automation in Construction*, doi: 10.1016/j.autcon.2022.104397.

Jiang, Y., Liu, X., Kang, K., Wang, Z., Zhong, R.Y. and Huang, G.Q. (2021), “Blockchain-enabled cyber-physical smart modular integrated construction”, *Computers in Industry*, Vol. 133, doi: 10.1016/j.compind.2021.103553.

Kamari, M. and Ham, Y. (2022), “AI-based risk assessment for construction site disaster preparedness through deep learning-based digital twinning”, *Automation in Construction*, Vol. 134, doi: 10.1016/j.autcon.2021.104091.

Kojima, H., Fujii, T., Mihara, Y. and Ihara, H. (2020), “Introduction of the new safety concept ‘safety2.0’ to reduce the risk of machinery accidents”, *Proceedings of the 37th International Symposium on Automation and Robotics in Construction, ISARC 2020: From Demonstration to Practical Use - To New Stage of Construction Robot*, doi: 10.22260/isarc2020/0197.

Kosse, S., Vogt, O., Wolf, M., König, M. and Gerhard, D. (2022), “Digital Twin Framework for Enabling Serial Construction”, *Frontiers in Built Environment*, Vol. 8, doi: 10.3389/fbuil.2022.864722.

Ladj, A., Wang, Z., Meski, O., Belkadi, F., Ritou, M. and Da Cunha, C. (2021), “A knowledge-based Digital Shadow for machining industry in a Digital Twin perspective”, *Journal of Manufacturing Systems*, doi: 10.1016/j.jmsy.2020.07.018.

Lai, X., He, X., Wang, S., Wang, X., Sun, W. and Song, X. (2022), “Building a Lightweight Digital Twin of a Crane Boom for Structural Safety Monitoring Based on a Multifidelity Surrogate Model”, *Journal of Mechanical Design, Transactions of the ASME*, doi: 10.1115/1.4053606.

Lee, D. and Lee, S. (2021), “Digital twin for supply chain coordination in modular construction”, *Applied Sciences (Switzerland)*, Vol. 11 No. 13, doi: 10.3390/app11135909.

Lee, D., Lee, S.H., Masoud, N., Krishnan, M.S. and Li, V.C. (2021), “Integrated digital twin and blockchain framework to support accountable information sharing in construction projects”, *Automation in Construction*, doi: 10.1016/j.autcon.2021.103688.

Lee, D., Lee, S.H., Masoud, N., Krishnan, M.S. and Li, V.C. (2022), “Digital twin-driven deep reinforcement learning for adaptive task allocation in robotic construction”, *Advanced Engineering Informatics*, Vol. 53, doi: 10.1016/j.aei.2022.101710.

Lee, J. and Lee, S. (2023), “Construction Site Safety Management: A Computer Vision and Deep Learning Approach”, *Sensors*, Multidisciplinary Digital Publishing Institute, Vol. 23 No. 2, p. 944, doi: 10.3390/s23020944.

Leonardo, M., Berardo, N., Alessandro, C., Luigi, R. and Giuseppe Martino, D.G. (2020), “Development of a Digital Twin Model for Real-Time Assessment of Collision Hazards”, pp. 14–19, doi: 10.3311/ccc2020-003.

Li, M., Lu, Q., Bai, S., Zhang, M., Tian, H. and Qin, L. (2021), "Digital twin-driven virtual sensor approach for safe construction operations of trailing suction hopper dredger", *Automation in Construction*, Vol. 132, doi: 10.1016/j.autcon.2021.103961.

Liang, C.-J., McGee, W., Menassa, C.C. and Kamat, V.R. (2022), "Real-time state synchronization between physical construction robots and process-level digital twins", *Construction Robotics*, Vol. 6 No. 1, pp. 57–73, doi: 10.1007/s41693-022-00068-1.

Liu, Z., Li, A., Sun, Z., Shi, G. and Meng, X. (2022), "Digital Twin-Based Risk Control during Prefabricated Building Hoisting Operations", *Sensors*, Vol. 22 No. 7, doi: 10.3390/s22072522.

Liu, Z., Shi, G., Qin, J., Wang, X. and Sun, J. (2022), "Prestressed Steel Material-Allocation Path and Construction Using Intelligent Digital Twins", *Metals*, doi: 10.3390/met12040631.

Mao, C., Tao, X., Yang, H., Chen, R. and Liu, G. (2018), "Real-Time Carbon Emissions Monitoring Tool for Prefabricated Construction: An IoT-Based System Framework", *ICCREM 2018: Sustainable Construction and Prefabrication - Proceedings of the International Conference on Construction and Real Estate Management 2018*, doi: 10.1061/9780784481738.015.

Opoku, D.G.J., Perera, S., Osei-Kyei, R. and Rashidi, M. (2021), "Digital twin application in the construction industry: A literature review", *Journal of Building Engineering*, doi: 10.1016/j.jobe.2021.102726.

Osti, F., de Amicis, R., Sanchez, C.A., Tilt, A.B., Prather, E. and Liverani, A. (2021), "A VR training system for learning and skills development for construction workers", *Virtual Reality*, doi: 10.1007/s10055-020-00470-6.

Pan, Y. and Zhang, L. (2021), "A BIM-data mining integrated digital twin framework for advanced project management", *Automation in Construction*, doi: 10.1016/j.autcon.2021.103564.

Posada, H., Chacón, R., Ungureanu, L.C. and García, D. (2022), "Closing the Gap between Concrete Maturity Monitoring and Nonlinear Time-dependent FEM Analysis through a Digital Twin. Case Study: Post-tensioned Concrete Slab of an Office Building, Barcelona, Spain", *Proceedings of the International Symposium on Automation and Robotics in Construction*, Vol. 2022-July, pp. 215–222, doi: 10.22260/isarc2022/0031.

Pour Rahimian, F., Seyedzadeh, S., Oliver, S., Rodriguez, S. and Dawood, N. (2020), "On-demand monitoring of construction projects through a game-like hybrid application of BIM and machine learning", *Automation in Construction*, Vol. 110, doi: 10.1016/j.autcon.2019.103012.

Rausch, C. and Haas, C. (2021), "Automated shape and pose updating of building information model elements from 3D point clouds", *Automation in Construction*, Vol. 124, doi: 10.1016/j.autcon.2021.103561.

Rogage, K., Mahamedi, E., Brilakis, I. and Kassem, M. (2022), "Beyond digital shadows: Digital Twin used for monitoring earthwork operation in large infrastructure projects", doi: 10.1007/s43503-022-00009-5.

Sacks, R., Brilakis, I., Pikas, E., Xie, H.S. and Girolami, M. (2020), "Construction with digital twin information systems", *Data-Centric Engineering*, doi: 10.1017/dce.2020.16.

Salem, A. and Moselhi, O. (2021), "Ai-based cloud computing application for smart earthmoving operations", *Canadian Journal of Civil Engineering*, doi: 10.1139/cjce-2019-0681.

Schlenger, J., Yeung, T., Vilgertshofer, S., Martinez, J., Sacks, R. and Borrmann, A. (2022), "A Comprehensive Data Schema for Digital Twin Construction", doi: 10.7146/aul.455.c194.

Schönböck, J., Kurschl, W., Augstein, M., Altmann, J., Fraundorfer, J., Freller, L.M., Steiner, L., *et al.* (2022), "From Remote-Controlled Excavators to Digitized Construction Sites", *Procedia Computer Science*, doi: 10.1016/j.procs.2022.01.315.

Shariatfar, M., Deria, A. and Lee, Y.C. (2022), "Digital Twin in Construction Safety and Its Implications for Automated Monitoring and Management", *Construction Research Congress 2022: Computer Applications, Automation, and Data Analytics - Selected Papers from Construction Research Congress 2022*, Vol. 2-B, pp. 591–600, doi: 10.1061/9780784483961.062.

Shen, K., Ding, L. and Wang, C.C. (2022), "Development of a Framework to Support Whole-Life-Cycle Net-Zero-Carbon Buildings through Integration of Building Information Modelling and Digital Twins", *Buildings*, Vol. 12 No. 10, doi: 10.3390/buildings12101747.

To, A., Liu, M., Hazeeq Bin Muhammad Hairul, M., Davis, J.G., Lee, J.S.A., Hesse, H. and Nguyen, H.D. (2021), "Drone-Based AI and 3D Reconstruction for Digital Twin Augmentation", *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Vol. 12774 LNCS, pp. 511–529, doi: 10.1007/978-3-030-77626-8_35.

Tran, H., Nguyen, T.N., Christopher, P., Bui, D.K., Khoshelham, K. and Ngo, T.D. (2021), "A digital twin approach for geometric quality assessment of as-built prefabricated façades", *Journal of Building Engineering*, doi: 10.1016/j.jobe.2021.102377.

Wang, X., Liang, C.-J., Menassa, C.C. and Kamat, V.R. (2021), "Interactive and Immersive Process-Level Digital Twin for Collaborative Human–Robot Construction Work", *Journal of Computing in Civil Engineering*, doi: 10.1061/(asce)cp.1943-5487.0000988.

Wu, S., Hou, L., Zhang, G. (Kevin) and Chen, H. (2022), "Real-time mixed reality-based visual warning for construction workforce safety", *Automation in Construction*, Elsevier B.V., Vol. 139, doi: 10.1016/J.AUTCON.2022.104252.

Zhang, M., Ghodrati, N., Poshdar, M., Seet, B.C. and Yongchareon, S. (2023), "A construction accident prevention system based on the Internet of Things (IoT)", *Safety Science*, doi: 10.1016/j.ssci.2022.106012.

Zhao, R., Chen, Z. and Xue, F. (2023), "A blockchain 3.0 paradigm for digital twins in construction project management", *Automation in Construction*, Vol. 145, doi: 10.1016/j.autcon.2022.104645.

Zhao, Y., Wang, N. and Liu, Z. (2022), "An Established Theory of Digital Twin Model for Tunnel Construction Safety Assessment", *Applied Sciences (Switzerland)*, Vol. 12 No. 23, doi: 10.3390/app122312256.

Züst, S., Züst, R., Züst, V., West, S., Stoll, O. and Minonne, C. (2021), "A graph based Monte Carlo simulation supporting a digital twin for the curatorial management of excavation and demolition material flows", *Journal of Cleaner Production*, doi: 10.1016/j.jclepro.2021.127453.

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

CAPTURING REALITY CHANGES FROM POINT CLOUDS FOR UPDATING ROAD GEOMETRIC DIGITAL TWINS

Diana Davletshina[1], Varun Kumar Reja[1], and Ioannis Brilakis[1]
[1]University of Cambridge, Cambridge, UK

Abstract

Proactive maintenance of roads enables increased asset lifespan with improved safety and minimised downtime. However, the absence of up-to-date structured information leads to expensive reactive maintenance. Geometric digital twins (GDT) provide digital replicas of object geometry, but no automatic tools exist to maintain them. This paper develops a method for detecting and applying geometric changes to road GDTs. Our solution performs a distance-based comparison of point clouds, estimates the changes and then applies them to GDT. It achieves a 77.87 F1-score in detecting changes and significantly saves time by automating the process, making such digital twins viable and practically applicable.

Introduction

The average travel delay in the United Kingdom has increased by 40% over the past three years due to roadwork disruptions, causing decreased journey time reliability and slower highway speeds (National Highways, 2023). These delays are not just inconvenient but also economically impactful, potentially leading to a £0.50 loss in Gross Value Added per hour per person in Britain (Confederation of British Industry, 2017). Efficient transportation infrastructure, crucial for economic productivity, requires up-to-date information on road assets, with road geometry playing a vital role in traffic safety and operational efficiency (Design Manual for Roads and Bridges, 2020). Road geometric digital twins (GDT) enable such monitoring of road conditions and infrastructure health, offering near real-time data integration and detailed replicas of road systems (Steyn and Broekman, 2022). The life-cycle of GDTs includes construction from point clouds and optionally images and ongoing maintenance to reflect changes in the road infrastructure (Drobnyi et al., 2023; Osadcha et al., 2023). This paper focuses on automating the stage of GDT maintenance, which is a substantial part of the GDT life-cycle.

The main challenge in integrating GDTs is their high cost of maintenance that outweighs potential benefits, introducing a significant barrier to their widespread adoption. This research aims to reduce these costs, eliminating the manual work involved. Specifically, maintaining kilometres of roads requires manual effort in identifying changes and modelling various changed road assets (Davletshina and Brilakis, 2023). Although current research has automated parts of this process, it falls short in scope, often failing to cover a wide range of objects and identifying specific changes that happened to the assets. Additionally, there is a need for a comprehensive and automatic end-to-end pipeline for change detection, followed by applying these changes to GDT.

This paper aims to facilitate efficient, proactive maintenance by providing a tool to update a geometric digital twin automatically, thus delivering a fresh version of GDT quickly and cheaply. We achieve this by first comparing point clouds for changes, identifying the types of changes and applying them in GDT in a streamlined, fully automatic pipeline. Exemplar input and outputs can be seen in Figure 1. The solution covers all the object classes that are already present in GDT. Thus, the solution is flexible and not hard-coded to the specific entities. Furthermore, we consider the wide range of change types in the scope of our work: a) the same object is there, but it underwent pose (rotation and translation) changes; b) the same object is there, but it underwent shape deformations; and c) the object is removed or added. As a result, we detect most of these changes and achieve 77% F1 score.

The organisation of this paper is as follows. The next section discusses the background literature on maintaining GDTs. Then, section Proposed Solution introduces the proposed automatic digital twin maintenance method. Research Methodology, followed by the Results and Discussion section, can be found further. The conclusions of the study are presented in the last section.

(a) GDT (old).

(b) GDT (new).

Figure 1: Examples of the old-version GDT and updated GDT achieved by our proposed solution. Pose transformations and removals are detected in the road furniture and applied in the new version.

Background

The evolution of Digital Twins (DTs) presents a complex challenge, particularly in their efficient updating, which remains unresolved. Schrotter and Hürzeler (2020) highlighted the difficulty in updating urban DTs, highlighting the vast volume of 3D data as a significant obstacle to process. Nevertheless, ongoing research within the built environment sector is actively investigating various methods to effectively maintain and manage digital twins and 3D models of infrastructure assets.

Updating GDTs of built assets starts with collecting appropriate data on the physical counterpart. Osadcha et al. (2023) found that photogrammetry and laser scanning are the primary means for data collection and capturing 3D information. While photogrammetric methods are cheaper, they produce less accurate depth data due to lighting and shadow artefacts (Stal et al., 2013). Mobile laser scanning could serve as a compromise between various methods from images to static scanning, allowing faster operation while driving but producing sparser point clouds than static counterparts (Soilán et al., 2019). Collected data then requires further processing.

A critical step in maintaining GDTs is effectively detecting changes in the data collected at different timestamps. Stilla and Xu (2023) explored this, identifying several types of changes in urban objects using 3D point clouds. They distinguished the following possible change types: appeared, disappeared, partially and fully moved, deformed and unchanged, highlighting that the most common types were the first two, while partial movement was used when detailed changes were required upon use case. Their work estimates point-based, segment- or object-based spatial differences after having point clouds registered. While point-based methods rely on accurate distance measurements, segment- and object-based methods depend on the accuracy of the classification provided. Recent works use deep learning solutions to provide semantic labels and discover further opportunities to match changes with machine learning.

In this realm, Zhu et al. (2023) introduced a method for multi-object relocalization and reconstruction in the segmented into instances 3D point clouds. They could match, register and reconstruct the objects of the input scenes using a deep learning approach, namely an encoder-decoder network. Although solving three tasks simultaneously, they envision the challenge of large-scale spatiotemporal changes yet to be addressed.

Developing dynamic Building Information Models (BIMs) from 3D point clouds is another key research direction in this field. Rausch and Haas (2021) developed a dynamic BIM by updating BIM elements' shape and pose from 3D point clouds, utilising optimisation of geometric object parameters. Despite achieving below-centimetre accuracy, their method's reliance on generic shapes limits its effectiveness with free-form objects. Alternatively, Ghahremani et al. (2018) used 3D point clouds for structural condition monitoring, comparing different-timestamp point clouds to detect deterioration. They identified changed areas by discrepancies between as-built and as-is point clouds, using meshing to update the model. Shellshear et al. (2015) adopted a simpler approach for 3D model maintenance, employing repeated scans from fixed viewpoints and markers to automatically update regions in the target model with point clouds. While effective for built assets, these methods highlight the need for focused approaches in maintaining GDTs for roads.

Knowledge Gaps, Research Objectives and Questions

Maintaining GDTs for roads involves two primary phases: detecting changes and applying these changes. However, a fully automated system for this task is yet to be realized. Also, current methods do not encompass most road assets, such as the frequent objects identified in the work of Davletshina and Brilakis (2023), highlighting a significant gap. While researchers have primarily focused on automating aspects such as identifying temporal changes, further development is still needed. Although advancements in built asset management provide insights, there remains a critical need for domain-specific adaptation and refinement of these solutions to cater specifically to road infrastructure.

This context sets the stage for our objective: to propose and develop a fully automated method for maintaining road GDTs. We outline the following key research questions in addressing our objective: 1) What approaches are most effective for detecting changes in road environments? 2) What strategies should be employed to efficiently apply detected changes to GDTs? 3) Is it possible to devise a versatile method adaptable to various road environments and infrastructures? Addressing these questions will be pivotal in advancing the field and ensuring the accuracy and relevance of road GDTs in a dynamic and evolving landscape.

Proposed Solution

Our approach to detecting and analysing temporal changes in geometric data employs a dual-layer methodology, as illustrated in Figure 2. This is necessitated by the requirement to initially discern the nature of any changes (Change Detection Layer), followed by appropriately addressing the identified changes (Change Application Layer). Initially, the method involves a comparative analysis of point cloud data, representing different states of a physical environment captured at distinct times. As the two point clouds will not contain the same point set, we take advantage of comparative distances to estimate high discrepancies. The framework commences with the older Geometric Digital Twin (GDT) version, incorporating the original point cloud and the newly acquired point cloud for subsequent change detection and processing.

The procedure begins with a thorough analysis of the input point clouds to identify changes by computing distances between points. The algorithm targets areas exhibiting significant alterations, i.e., large distances, indicative of substantial environmental changes. Thus, we filter these points using a distance threshold, which is a hyper-parameter to be tuned. Upon pinpointing these areas, we anticipate that most changes, unless due to extensive road reconstruction, will be isolated to individual objects, as road infrastructures typically undergo minimal alterations. As a result, we have these isolated change points, i.e., filtered with the threshold and as a result isolated in space point clusters featuring high distance values such as in Figures 3d and 3e. Then, we apply the DBSCAN clustering algorithm (Ester et al., 1996) on these points – known for its effectiveness with noisy data and not requiring pre-set cluster numbers – which facilitates the segregation and extraction of clusters (e.g., denoted by different colours in Figures 3d and 3e) and their associated objects as identified in the current GDT, enabling a more focused examination.

Subsequently, the nature of the detected changes guides the ensuing steps. For each cluster, the algorithm first finds the closest correspondences between the two versions and ascertains whether the change is an addition, a removal, or a transformation of an existing object. Additions and removals are identified similarly, as they can be seen as the same operation but in opposite sequences. We consider that addition or removal happened if the closest corresponding cluster was already linked to another cluster with a smaller distance. Otherwise, we consider that there was a pose change.

Objects identified as removed are excised from the dataset while new additions are integrated. In cases where existing objects have undergone modifications, the Iterative Closest Point (ICP) algorithm (Zhang, 1994) is deployed on the found cluster correspondences (i.e., correspondence pairs) to compute the transformation, encompassing both rotational and translational adjustments between them. This process begins with estimating the initial transformation for a global match, then minimising the distance between two shapes to locate the closest points. This phase is pivotal in deciphering how the object has shifted or evolved over time. Depending on the nature of the change, the algorithm either applies a transformation to align the old and new states of objects with pose changes if any pose changes were found or otherwise updates the GDT to mirror the new geometry of deformed objects by replacing the object. As a result, an automatic command is sent to perform these described changes to the GDT.

Consequently, our solution presents a systematic and flexible framework for temporal analysis of point cloud data concerning roads, resulting in an up-to-date GDT reflecting reality. This fully automated and scalable solution can be applied over extensive road lengths, significantly enhancing our comprehension and monitoring of environmental alterations across various applications, from asset maintenance to safety analysis.

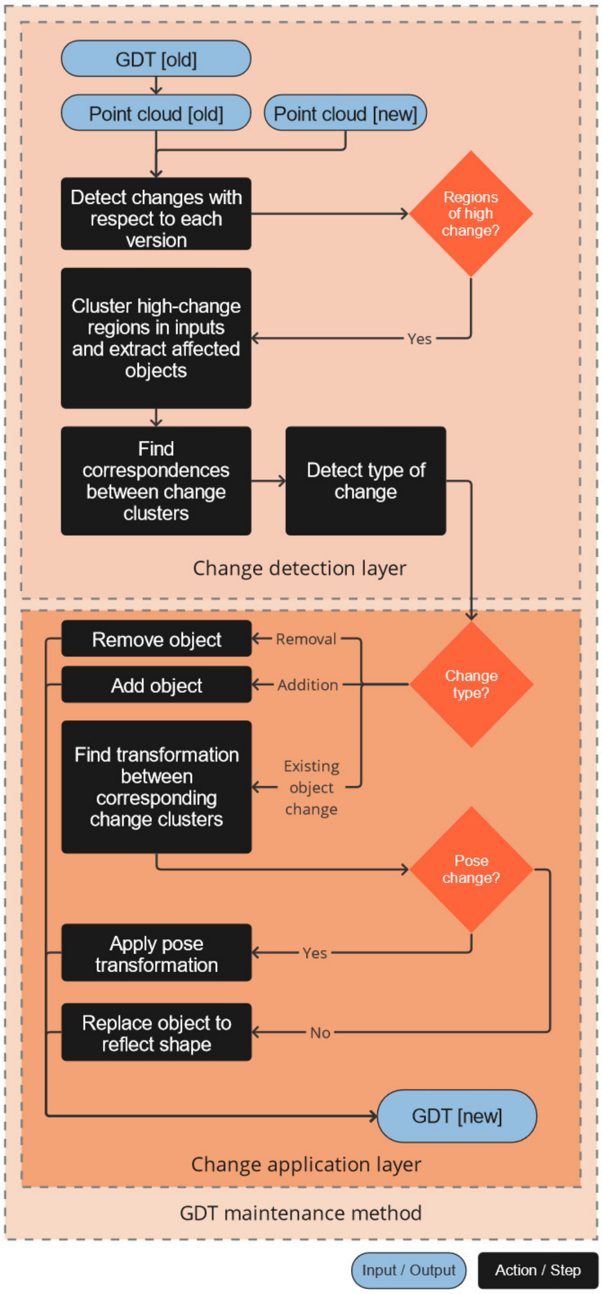


Figure 2: Solution overview.

Research Methodology

We use the Digital Roads (Digital Roads of the Future, 2023) point cloud dataset, which contains 11 labelled classes. We limit the data to the A11 road subset, which comprises 20 road segments, each 200 meters long. As the study conducted in this paper requires time-series data with changes, we synthetically generate the changed version of the point clouds (the *new* state) from the existing dataset (the *old* state) due to the unavailability of a real-life data. We introduce specific perturbations to the data for sufficient differences between data to simulate real-life scanning. For that, we add random noise to the coordinates of both versions of the data (the *old* and *new* states). The noise is normally distributed around mean 0 with a standard deviation of 0.01. Furthermore, we randomly downsample both versions such that 20% of the points are removed. We simulate the changes to the road infrastructure, namely road barriers, traffic signs and lamps, such as fallen assets, leaning assets, missing assets and moved assets in the *new* version of the point cloud. We consider removals of the existing objects and additions of the new objects in the case of the missing objects. These scenarios are seen to be equivalent as one object is missing in one of the versions of scanning.

Our assumptions are as follows. Firstly, we assume that the GDT is semantically labelled and has its original point cloud stored as meta information. If that is not the case, then a synthetic point cloud is expected to be created from the GDT's 3D model. Secondly, we assume that point clouds are registered with each other; data registration is out of the scope of this work. Lastly, we assume that new scanning is performed before new furniture is installed to avoid confusion between removed and newly added assets' pairs.

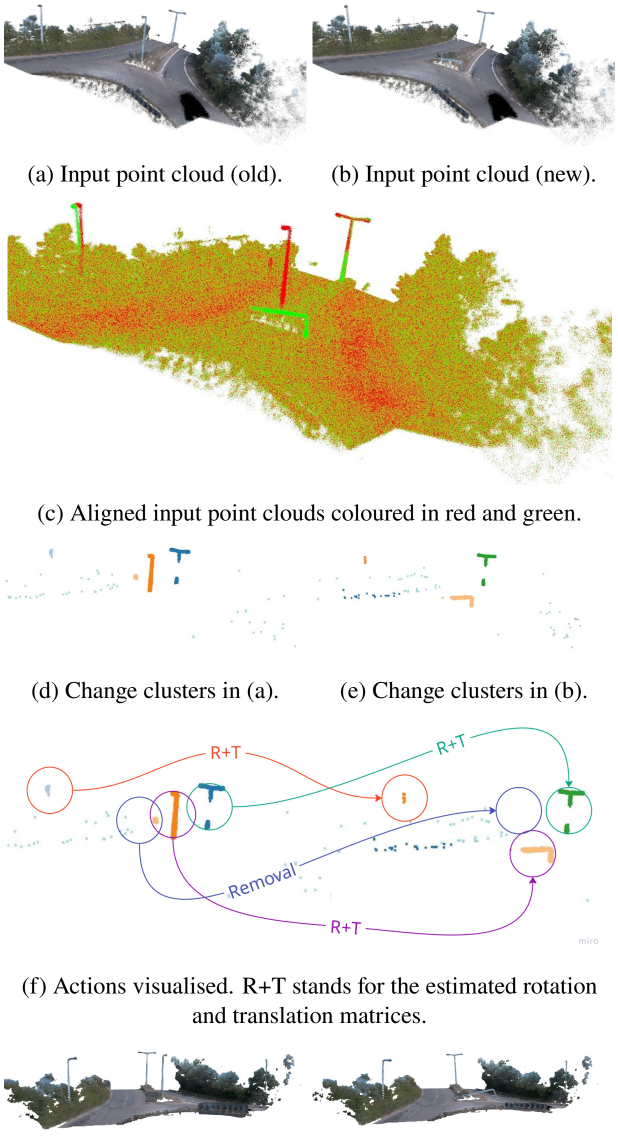
We use the threshold-based method to identify high-change regions, with the threshold hyperparameter equalling the 0.99 quantiles. Further, we use the DBSCAN clustering algorithm (Ester et al., 1996) to cluster these high-change regions with the following hyperparameters: epsilon equal to 3 and minimum points per cluster equal to 10.

We provide the metrics of precision (Formula 1) showing how many changes were correctly classified, recall (Formula 2) giving the number of correctly predicted items among all correct items and their harmonic mean as F1 score (Formula 3). In the formulas, *TP* refers to true positives, denoting the number of data points classified correctly as positives, *FP* – false positives, denoting the number of data points classified incorrectly as positives; *FN* – false negatives, denoting the number of positive points predicted as negatives. We calculate precision and recall values for each sample and report averaged final metrics.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

$$Recall = \frac{TP}{TP+FN} \tag{2}$$

$$F1 = \frac{2*Precision*Recall}{Precision+Recall} \tag{3}$$



(a) Input point cloud (old). (b) Input point cloud (new).

(c) Aligned input point clouds coloured in red and green.

(d) Change clusters in (a). (e) Change clusters in (b).

(f) Actions visualised. R+T stands for the estimated rotation and translation matrices.

(g) GDT (old). (h) GDT (new).

Figure 3: Examples the inputs, intermediate and final outputs of the method.

Table 1: Change detection evaluation, in %

Precision	Recall	F1 score
90.97	68.07	77.87

Results & Discussion

Figure 3 demonstrates the solution's inputs, intermediate and final outputs. Our method can find the changed regions, such as in Figure 3c, where the two input point clouds (Figures 3a and 3b) are aligned. These divergent regions are clustered, as in Figures 3d and 3e, and matched, as well as their transformations are estimated (Figure 3f). As a result, our solution finds pose changes with 90.97% precision and 68% recall (Table 1). The method misses out on a few samples where the changes are minimal, such as small translation and rotation angles. Further, vegetation represents a challenge, introducing additional high discrepancies. Thus, vegetation should be considered for filtering out beforehand using vegetation detection solutions. Additionally, the intermediate found change clusters (e.g., Figures 3d and 3e) come with a certain amount of noise either within clusters or outside of them, which might be considered as a potential scene alteration. As this confuses the method, noise removal at this stage must be included in future work.

Furthermore, the algorithm faces challenges in accurately matching changed clusters when they are significantly distant from each other, particularly in situations where multiple similar changes occur concurrently within the same scene. Handling such intricate change scenarios effectively requires more advanced algorithms for object matching. Incorporating prior knowledge could be crucial in these contexts, aiding in differentiating between scenarios where objects have been moved as opposed to those where objects were removed and replaced with similar ones.

Conclusions

In summary, this work presents a method for maintaining up-to-date geometric digital twins of roads in a fully automatic manner. This method is distinguished by its flexibility and adaptability, which is capable of handling a wide range of objects based on the semantic labels provided in the input GDTs. Our approach goes beyond mere change detection in point clouds; it integrates these changes directly into the GDTs, thereby forming a seamless, end-to-end pipeline that extends from initial scanning to the final GDT for road infrastructures. This development paves the way for the broader implementation of geometric digital twins in the transportation sector, offering significant benefits from the design phase to asset management. This is achieved with reduced costs, courtesy of the automation process. Enhancing the productivity of the sector will yield positive outcomes for road users, including improved connectivity, fewer road closures, and reduced emissions, ultimately leading to a more efficient and environmentally friendly transportation experience. Looking forward, we propose that future research should concentrate on collecting comprehensive datasets that capture real-life temporal changes, tackle the challenges presented by complex and large-scale transformation scenarios, and harness the potential of deep learning for enhanced accuracy in change detection.

Acknowledgments

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) Industrial CASE in partnership with National Highways and Costain [grant No EP/V056441/1] and the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie [grant agreement No 101034337].

References

Confederation of British Industry (2017). Shaping regional infrastructure. Priorities for growth.

Davletshina, D. and Brilakis, I. (2023). A review on constructing and maintaining geometric digital twins of highways.

Design Manual for Roads and Bridges (2020). CD 109 - Highway link design.

Digital Roads of the Future (2023). Digital Roads Dataset. `https://drf.eng.cam.ac.uk/research/digital-roads-dataset`.

Drobnyi, V., Hu, Z., Fathy, Y., and Brilakis, I. (2023). Construction and maintenance of building geometric digital twins: State of the art review. Sensors, 23(9):4382.

Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. In kdd, volume 96, pages 226–231.

Ghahremani, K., Khaloo, A., Mohamadi, S., and Lattanzi, D. (2018). Damage detection and finite-element model updating of structural components through point cloud analysis. Journal of Aerospace Engineering, 31(5):04018068.

National Highways (2023). Annual Benchmarking Report. `https://nationalhighways.co.uk/media/dnen52ja/abr-v5-171023.pdf`.

Osadcha, I., Jurelionis, A., and Fokaides, P. (2023). Geometric parameter updating in digital twin of built assets: A systematic literature review. Journal of Building Engineering, page 106704.

Rausch, C. and Haas, C. (2021). Automated shape and pose updating of building information model elements from 3d point clouds. Automation in Construction, 124:103561.

Schrotter, G. and Hürzeler, C. (2020). The digital twin of the city of zurich for urban planning. PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science, 88(1):99–112.

Shellshear, E., Berlin, R., and Carlson, J. S. (2015). Maximizing smart factory systems by incrementally updating point clouds. IEEE computer graphics and applications, 35(2):62–69.

Soilán, M., Sánchez-Rodríguez, A., del Río-Barral, P., Perez-Collazo, C., Arias, P., and Riveiro, B. (2019). Review of laser scanning technologies and their applications for road and railway infrastructure monitoring. Infrastructures, 4(4):58.

Stal, C., Tack, F., De Maeyer, P., De Wulf, A., and Goossens, R. (2013). Airborne photogrammetry and lidar for dsm extraction and 3d change detection over an urban area–a comparative study. International Journal of Remote Sensing, 34(4):1087–1110.

Steyn, W. J. V. d. M. and Broekman, A. (2022). Development of a digital twin of a local road network: a case study.

Stilla, U. and Xu, Y. (2023). Change detection of urban objects using 3d point clouds: A review. ISPRS Journal of Photogrammetry and Remote Sensing, 197:228–255.

Zhang, Z. (1994). Iterative point matching for registration of free-form curves and surfaces. International journal of computer vision, 13(2):119–152.

Zhu, L., Huang, S., Schindler, K., and Armeni, I. (2023). Living scenes: Multi-object relocalization and reconstruction in changing 3d environments. arXiv preprint arXiv:2312.09138.

DATA QUALITY IN THE BUILT ENVIRONMENT ACROSS EUROPE

José L. Hernández[1], Susana Martín[1], Vanja Lonac[1], Ignacio de Miguel[2] and Alessandro Rossi[3]
[1]CARTIF technology centre, Boecillo, Spain
[2]Universidad of Valladolid, Valladolid, Spain
[3]Engineering Ingegneria Informatica S.p.A., Palermo, Italy

Abstract

Energy efficient projects are usually based on data to take better-informed decisions. High-quality data is necessary to reduce the uncertainty, which is a challenge to solve data gaps and/or outliers, thus limiting the trustworthiness in the decision-making. High-quality data is essential for assessing energy usage, identifying improvement measures and implementing effective solutions. This paper presents a data quality methodology, based on seven dimensions and validated in ten pilots across Europe. The maturity levels in the monitoring stage are crucial. Completeness, accuracy and consistency values vary between 40%-99%. It indicates the need for creating correction procedures and increasing the data quality.

Introduction

The increasing adoption of leading-edge Information and Communication Technologies (ICTs), such as Internet of Things (IoT), Artificial Intelligence (AI), Distributed Ledger Technology (DLT), Blockchain or Big Data (BD), as well as the wider extension of Building Automation and Control Networks (BACN) existence (mainly in tertiary buildings), has motivated the generation of big amounts of data in the building domain.

Even though over the past two decades, there has been an increasing focus on the digitalization of building data, and more and more data are being generated by the building stock, data quality remains a challenge (Duvier et al., 2008). There are data gaps, errors and inconsistencies in registrations that still need to be addressed. In contemporary construction and building projects, data plays a pivotal role in decision-making processes, project management, and overall success. However, the quality of data used in these endeavours often falls short of desired standards, posing significant challenges throughout project lifecycles. Poor data quality in building projects presents multifaceted obstacles, ranging from increased costs and delays to compromised safety and functionality.

Additionally, nowadays data-driven approaches are being followed, but here the main challenge is related to the reliability of the results due to data issues (e.g. communication or infrastructure problems) (Yong et al., 2021). One of the major challenges is to be able to process and analyse the data traceability to detect the errors (Hossein et al., 2020). The big amounts of data being generated make this analysis more and more complex. In terms of data governance, there is no consensus about the meaning of this term (Ender, 2021). However, different types of data should be considered when determining the ownership, which is even more complex when non-human data is included. These data are produced and treated in silos, without correlation between pillars (Abraham et al., 2019), which will promote better exploitation potential.

Poor data quality can also hinder effective stakeholder communication and collaboration, exacerbating misunderstandings and conflicts among project participants. In the absence of reliable and up-to-date information, decision-makers may resort to subjective judgments or outdated assumptions, further compounding project risks and uncertainties.

All these challenges are contributing to move forward towards the creation of a high-quality data-driven Smart Buildings Landscape. Data concerns almost every aspect of the built environment: from how individuals and businesses use and interact with properties, to how the building's energy consumption and construction details are recorded and analysed to support informed decisions about construction and real estate processes. Data-informed decision-making and digital upgrading can help upyield operational efficiencies at low cost.

A proper validation and detection of gaps, wrong and/or inaccurate data in the whole value chain of the building monitoring is a crucial step towards transparency and trust of the involved actors during the decision-making process to achieve the objectives of the Energy Performance of Buildings Directive (EPBD, 2023) of ensuring low usage of energy, low rate of carbon footprint, maximizing thermal comfort or assessing air quality.

According to the BDVA (Big-Data Value Association), the data quality concept is too absolute and misleading (BDVA/DAIRO, 2021). Within this paper, the high-quality data concept is based on the following aspects, establishing performance metrics as BDVA also suggests:

Reliability & Credibility: a data quality methodology is defined, deployed and applied, based on ML algorithms to learn about historical data, for reducing and correcting data errors not only in the data gathering process, but also in the propagation. This methodology takes care of:

(1) Data gaps, in terms of missed data to detect non-completeness and creating interpolation methods when possible, to generate continuous data streams for dynamic timeseries. As well, for static and contextual data, completeness in terms of missing data to be completed with other sources (e.g. digital logbook missing data and cadastre information);
(2) Outliers that are produced by values that differ from the expected measurements, thus, reducing the uncertainties of data analytics.
(3) Consistency and accuracy by removing duplicates, aligning multi-source measurements (e.g. date/time in timeseries data from various sources) and cleaning ambiguous values.

(4) Model check, which is focused on static data (i.e. BIM-Building Information Modelling). Errors in the building modelling are usual and propagated to the tools for the building life-cycle management. These mistakes should be prior detected and solved whenever possible.

Interoperability: Better-informed decisions mean using combined data from heterogeneous sources (e.g. sensors, energy performance certificates databases, digital logbooks, BIM, CityGML, Level(s), etc.), but the lack of interoperability is an issue to merge these datasets. To solve this, dynamic and adaptative interoperability will be ensured by using standard Data Models (such as NSGI-LD, SAREF, BRICK or IFC, among others) to accommodate data into the requirements of the use cases or services to be deployed.

Privacy & Security: Last but not least, quality also involves privacy (personal data management) and security. DLT and blockchain framework can be applied here, where privacy and security are the main benefits.

Addressing the challenges stemming from poor data quality in building projects is paramount for the sustainable advancement of the construction industry. By assuring these aspects, buildings' stakeholders can rely on the data analytics and data-driven services thanks to high-quality data stocks. Additionally, the creation and use of data-driven business models will be built on top of high-quality data, providing more accurate and trustworthy value (e.g. ESCO-Energy Services COmpany models where energy prices are based on performance calculation and energy savings could be calculated using wrong data).

Real-world cases abound where the repercussions of inadequate data quality have been acutely felt, leading to suboptimal project outcomes. For instance, consider the scenario of a large-scale infrastructure development project where inaccurate site survey data resulted in improper foundation designs, ultimately leading to structural instability and substantial rework costs. Similarly, in the realm of building information modelling (BIM), incomplete or inconsistent data inputs have led to coordination errors, clashes between different trades, and inefficiencies in construction sequencing. Another example could be obtained from ESCO services, whose main challenge is the data collection from smart meters. Data gaps are usually appearing, leading to more complex billing procedures due to missing data, as well as the limitations in the energy savings calculation.

In summary, this paper presents how the previous challenges are overcome by deploying a data quality methodology that covers the data life-cycle of the building, delving into the complexities surrounding data quality issues. It will not be only focused on the data gathering at field level, but it will cover the different stages of the building and data lifecycle to reduce the error propagation. This methodology will be supported by a federated Data Lake (Hernández et al., 2023a) (deployed within a data platform) to collect cross-domain data, being able to benefit the data management and governance processes. Finally, the data platform will provide additional services, which will facilitate the management of data ownership, and together with blockchain, will allow traceability, privacy and security of data stocks.

The rest of the paper is structured as follows. The following section presents the literature review in terms of data quality (including real-case scenarios). Next, the data quality methodology is explained, to be later applied in 10 pilots composed by one or several buildings. Results of the data quality assessment are collected into the "Data quality analysis results" section, where statistics for the 7 data quality dimensions are presented per pilot. A "Discussion" sections follows, to highlight the dependency of the buildings' data quality values to the age of the building and monitoring maturity and stage. Finally, the main remarks of the paper are presented in the "Conclusions" section.

Literature review

The literature provides various definitions and dimensions of data quality in the context of smart buildings. While traditional data quality dimensions such as accuracy, completeness, consistency, and timeliness remain relevant, the dynamic nature of IoT-generated data introduces new challenges and considerations. The importance of additional dimensions such as interoperability, security, privacy, and relevance in assessing data quality in smart buildings is emphasized (Rao, 2023). Ensuring data compatibility and integration across heterogeneous systems while safeguarding privacy and security concerns emerges as critical areas of focus.

Despite the potential benefits, smart building deployments encounter numerous challenges related to data quality. These include issues such as data silos (Hernández et al., 2023b), interoperability gaps between devices and systems (Coujard, 2023), sensor inaccuracies, data integration complexities, and cybersecurity vulnerabilities. The literature highlights organizational barriers such as lack of data governance frameworks (Kaginalkar, 2023), limited expertise in data management, and resistance to change as significant impediments to achieving and maintaining high data quality standards in smart building initiatives.

Diagnosis of data quality is therefore a main challenge. Some authors have analysed the uncertainty in the calibration of the measurements (Morewood, 2023). It is reported that, in 43 out of 63 real-cases, overall accuracy is achieved, while precision is reduced up to four. Calibration is shown in 18, whereas measurement out-of-range is only reached in 23. Another real-case scenario is performed in the cites of Nantes, Hamburg and Helsinki (Hernández et al., 2022). The main conclusion is the real need for methods to increase data quality to foster better-informed decisions. The authors demonstrated the low quality of various data-sets in terms of completeness and out-of-range, highlighting the importance of proper commissioning mechanisms.

Complementary to these previous studies, the contribution of this work extends the data quality analysis to additional real-case sites across Europe to understand the real data-quality scenarios. Moreover, while previous researches limited the dimensions of the data quality, this manuscript increases the dimensions, covering accuracy, completeness, reliability, consistency, relevance, accessibility and timeliness. This work has been carried out within the EU HORIZON DigiBUILD project, whose main aim is to transform traditional silo approaches by making use of high-quality data and next generation digital building services for assuring trust and transparency, and better-informed decision-making processes. It counts on ten pilots across Europe where the technologies are tested.

Data quality methodology

Within DigiBUILD project, a multi-dimension strategy is approached in terms of data quality. Table 1 summarizes the list of dimensions that are part of the data quality analysis. From the 7 dimensions, accuracy, completeness and consistency are considered the pivotal ones.

Table 1: Data quality multi-dimension approach

Data quality dimension	DigiBUILD approach
Accuracy	Out of "expected" range detection
Completeness	Gap detection and interpolation
Reliability	Accuracy x completeness
Consistency	Detection of "normal" data patterns & Proof of Existence (PoE)
Relevance	Data filtering processes
Accessibility	Open APIs and protocols
Timeliness	Pre-analytics

Accuracy determines the percentage of data samples compliant with the "expected" ranges. It should be clarified the "expected" range differs from the measurement range. For instance, an indoor temperature sensor can obtain data within values from 0ºC to 50ºC (depending on the manufacturer). However, temperature are expected to be around 15-30ºC (depending on the building use, insulation level and other features). The accuracy dimension treats the detection of data samples, which, even correct, are out of this range to avoid its usage in the application of services (i.e., classified as outliers). Within DigiBUILD, the procedure calculates the data samples within the range specified by the building expert, manager and/or operator to extract the ratio with respect to the total number of samples.

Completeness goal is to determine the data gaps that are in a dataset. Normally, data samples are periodically collected according a sampling frequency. Therefore, the total samples to be gathered in a time span are known. Thus, the missing data can be calculated by the ratio between the available data samples and the theoretical ones to be saved.

Reliability focuses on the feasibility of data to carry out analysis, develop high-level services or provide high-quality data to third parties. This dimension correlates the completeness and accuracy to provide an overall mark.

Consistency looks for the data patterns within data. For that end, normal distribution is considered, where measurements are scattered between the mean and ± standard deviation.

Relevance dimension is helpful to determine the datasets that are really useful from the end-user's perspective (services). Data normally include data-points without any usability in the implementation of the services; therefore, the percentage of relevant data is the main purpose.

Accessibility provides the ability to access, gather and share data. Here, the use of open protocols and APIs (Application Programming Interfaces) fosters the accessibility to the datasets. Nevertheless, this is not always possible, limiting the access to some datasets. Thus, this dimension checks the ability to access the relevant data from the building.

Timeliness is related to up-to-date data or the delay to gather the information that is required for making informed decisions. It is expressed in time units in order to establish the maximum delay when data is made available.

Table 2: Data quality outcomes within DigiBUILD

Data quality dimension	DigiBUILD outcome
Accuracy	> 85%
Completeness	> 90%
Reliability	No duplicates (clean data)
Consistency	Increase 10%
Relevance	90% of useful datasets
Accessibility	90% of data accessible
Timeliness	Specific per pilot

With regard to the results to be obtained in DigiBUILD, Table 2 shows the outcome that should be achieved by the detection and the correction techniques to be applied. In this sense, the accuracy and completeness dimensions should reach 85% and 90% respectively to provide a clean dataset without duplicates (reliability). Moreover, the case of consistency aims to improve the quality of the datasets being used in a 10%, making relevant the 90% of the collected dataset with an accessibility of 90%. Finally, timeliness depends on the specifications of the pilots when data is polled / pushed.

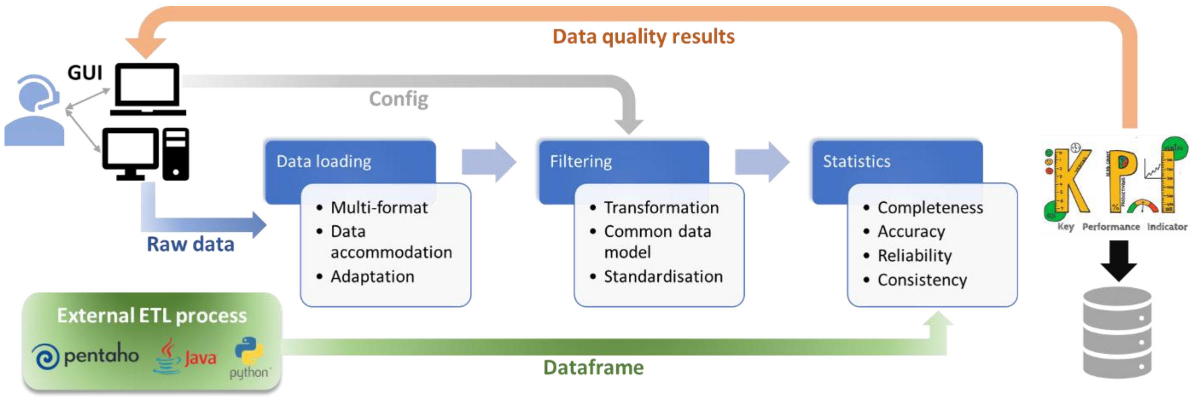


Figure 1: Data quality detection and check approach

Data quality check approach

Two procedures are approached for data quality assurance. On the one hand, the checker in charge of the analysis of the quality of datasets; on the other hand, the methodology for the correction of data (such as interpolation or other techniques for data cleansing). However, this paper is focused on the status of the data quality across Europe; therefore, the correction part is out of the scope. It will consist of a machine-learning approach to learn from historical data. Then, when a gap or an inconsistent data measurement is obtained, the corrector module will populate data according to the data-driven model.

Focusing on the checker, Figure 1 depicts the schema how it is approached within DigiBUILD. It works on two possibilities. The first one makes use of downloaded data from the pilots, for instance, csv files, Excel sheets, JSON documents, etc. This raw data is ingested by a data loading module that accommodates the multiple formats and adapts them in order to be transformed (filtering process) into a common data model. The goal is to provide a "standard" representation of data (following a python DataFrame). Complementary, a configuration file is used where the user provides the parameters and specific aspects related to data.

The second way is the use of data collection technologies, as for the case of Pentaho (Hitachi, 2023), python or Java scripts. In this case, the tools for data interfacing already provide the data loading and filtering processes. Hence, they are directly connected to the statistics module.

The statistics module is a common component working in the two paths. It basically calculates the indicators for the data quality in the multiple dimensions explained before. It accepts a DataFrame in order to make it interoperable, replicable and scalable; therefore, being able to calculate the data quality indicators independently of the source.

Data quality analysis results

As introduced, this paper presents the baseline results in terms of data quality for a set of buildings (pilots) across Europe. As extracted in Table 3, ten pilots from different European countries that are part of the EU DigiBUILD project, with different typologies and available datasets are analysed. Next sections describe the results for each one of these pilots in the various dimensions that are accounted in DigiBUILD.

Table 3: Pilot buildings within the EU DigiBUILD project for the data analysis

Pilot	Type	Datasets
UCL (UK)	University	Energy, photovoltaics, indoor air quality, occupancy, windows contact
EDF (France)	Offices	Indoor air quality, energy (heating, electricity and plugs)
IASI (Romania)	Various types	Indoor air quality, energy (heating and electricity)
VEOLIA (Spain)	Residential	Energy (district heating)
EMOT (Italy)	Offices	Energy, photovoltaics, charging station, electric vehicle, indoor air quality
FOCCHI (Italy)	Factory	Energy, photovoltaics, indoor air quality
HERON (Greece)	Various types	Energy, charging station, electrical vehicle
FVH (Finland)	Various types	Energy (district heating & heat pumps), photovoltaics
IEECP (The Netherlands)	Schools	Energy, indoor air quality, occupancy
NTUA (Greece)	University	Energy, indoor air quality, HVAC info

Accuracy

Starting with the accuracy dimension, The results show a dispersion of values, from 42.04% to 99%. In some cases, as for instance the UCL pilot, out- of-value ranges are due to changing needs of the heating/cooling systems. Other cases, malfunctioning of any sensor causes measurements out of range.

Table 4 summarizes the percentages obtained from a historical dataset sample of the aforementioned available datasets in the pilot building. For that end, minimum and maximum values for each data-point are obtained in order to reduce the range in 10%, being an acceptable value, such as stated by (Villada et al., 2008) and (Seyedzadeh et al., 2020). Then, values out of these range (known as expected) are considered anomalous.

The results show a dispersion of values, from 42.04% to 99%. In some cases, as for instance the UCL pilot, out-of-value ranges are due to changing needs of the heating/cooling systems. Other cases, malfunctioning of any sensor causes measurements out of range.

Table 4: Accuracy results of the data analysis

Pilot	Accuracy
UCL (UK)	70%
EDF (France)	44.43%
IASI (Romania)	n.a.
VEOLIA (Spain)	85%
EMOT (Italy)	56.18%
FOCCHI (Italy)	69.92%
HERON (Greece)	42.04%
FVH (Finland)	99%
IEECP (The Netherlands)	86.90%
NTUA (Greece)	55.69%

Completeness

Completeness aims the detection of data gaps in the timeseries. There are many cases where communication errors, malfunctioning of sensors or processing errors, among others, provoke data logging problems. Table 5 highlights the results of the analysis per pilot.

The results are diverse, remarking the case of EMOT, which is offering very low values (approximately 62% of data samples are missing). It should be considered the case of VEOLIA, with very high values, although this is not the reality. The reason behind lies in interpolation from the heat meter in a form of linear regression. The appearance is very complete data-set, but, with accurate populated values (e.g., boilers not working, but linear interpolation indicating energy consumption).

Figure 2: Monitoring stages

It is notable the maturity levels in the data collection. Pilots such as EDF or FVH with multiple years of data collection offers higher values with a very stable data gathering procedure. Less mature pilots are still in early stages where optimal operation is still not reached, as depicted in the monitoring stages in Figure 2 (Hernández et al., 2022).

Table 5: Completeness results of the data analysis

Pilot	Completeness
UCL (UK)	70%
EDF (France)	99.82%
IASI (Romania)	n.a.
VEOLIA (Spain)	99%
EMOT (Italy)	38.39%
FOCCHI (Italy)	~100%
HERON (Greece)	67.31%
FVH (Finland)	75.24%
IEECP (The Netherlands)	98.78%
NTUA (Greece)	99%

Reliability

The case of reliability focuses on the error-free data, without inconsistencies (accurate and complete). Then, its value is represented by the correlation between the completeness and accuracy, considering unique values from completeness (i.e., no data duplicates). Table 6 provides the results, where it is worth mentioning the case of EDF, in contrast to the completeness, lots of duplicates are removed, decreasing the reliability value.

Table 6: Reliability results of the data analysis

Pilot	Reliability
UCL (UK)	50%
EDF (France)	44.35
IASI (Romania)	n.a.
VEOLIA (Spain)	85%
EMOT (Italy)	21.57%
FOCCHI (Italy)	62.92%
HERON (Greece)	28.29%
FVH (Finland)	74.49%
IEECP (The Netherlands)	85.84%
NTUA (Greece)	55.63%

Consistency

Consistency relates to uniformity of data in terms of data behaviour. Considering data distribution follows a "normal" curve around the mean value, it is considered that data is consistent when a data-point complies with mean ± standard deviation. Table 7 illustrates the values for consistency. It should be taken the unexpected behaviour of the systems into account, such as the heating system. An example could be understood with COVID-19 situation, where energy demand increased, being a non-usual operation. Therefore, producing a deviation with respect to the “normal” distribution.

Table 7: Consistency results of the data analysis

Pilot	Consistency
UCL (UK)	70%
EDF (France)	34.11%
IASI (Romania)	n.a.
VEOLIA (Spain)	79%
EMOT (Italy)	32.97%
FOCCHI (Italy)	71.83%
HERON (Greece)	88.60%
FVH (Finland)	~99%
IEECP (The Netherlands)	82.14%
NTUA (Greece)	86.53%

Consistency is also addressed by a Smart Contract in charge of the notarization of data. A Proof of Existence (PoE) algorithm is developed, based on blockchain, to calculate the hash on the data provider side, as well as on the data storage side (see Figure 3). By comparing the hashes, the consistency in both sides can be obtained (i.e., detection of communication errors or data manipulation).

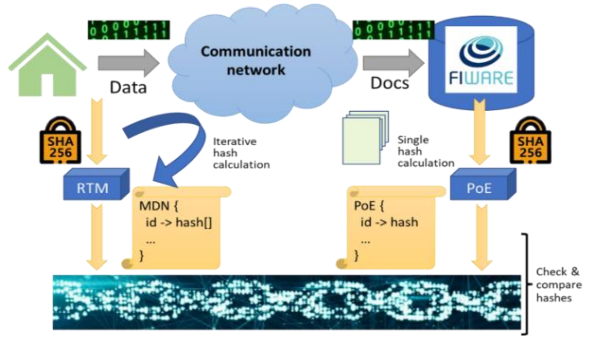
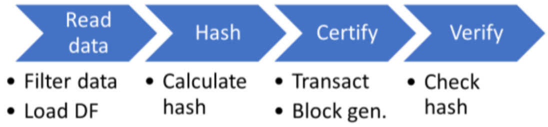


Figure 3: Proof of Existence concept

Figure 4 depicts the process, which has been applied within VEOLIA case. From the real data read, a data-set is filtered and loaded as a DataFrame, whose hash is calculated to be sent to the blockchain as a transaction and, thus, certify data. In the verification step, the received data is hashed and compared to the original hash (see Figure 4). The result of the transaction is shown below, where the output from the blockchain is generated, indicating the hash of the transaction and the block that has been generated.



AttributeDict({'args': AttributeDict({'': AttributeDict({'description': 'data from VEOLIA', 'provider': '0xb8f356ab125212CCd7017D8766B0463B531bA137', 'timestamp': 1700053800})}), 'event': 'Notarized', 'logIndex': 0, 'transactionIndex': 0, 'transactionHash': HexBytes('0x0310cf7829b8a11414ce1306acf6fe0dd8acd3908a13a9b2ce825bab40436c11'), 'address': '0x8a442905884c6a62938a77600aAeC7c9a073AAD7', 'blockHash': HexBytes('0x8548aabb4683617de8c49551b9f3a49ed36c94780a6baeea1f2d10cc9cd6717e'), 'blockNumber': 3641})

Figure 4: Consistency check from Smart Contract

Relevance

The relevance dimension focuses on the end-user and the usability of data. Nowadays, monitoring systems and digitalization have exponentially grown, obtaining bigger amounts of data without real usability in services or decision-making. That is why relevance plays an important role in DigiBUILD, determining the real usable datasets to provide high-level services. Table 8 specifies the percentage of datasets that are relevant for the DigiBUILD services and, thus, useful in this context.

Table 8: Relevance results of the data analysis

Pilot	Relevance
UCL (UK)	35.54%
EDF (France)	28%
IASI (Romania)	37.5%
VEOLIA (Spain)	55.73%
EMOT (Italy)	100%
FOCCHI (Italy)	55.73%
HERON (Greece)	100%
FVH (Finland)	30%
IEECP (The Netherlands)	n.a.
NTUA (Greece)	n.a.

Accessibility

Data, even with quality enough and relevant, sometimes is not accessible due to proprietary protocols. Accessibility determines the ability of remotely capture data and make it available for services and decision-making procedures. In this sense, Table 9 compiles the percentage of datasets that are accessible from pilot buildings. In some cases, although interfaces for access are being configured, these are not fully available; therefore, limiting the accessibility to data.

Table 9: *Accessibility results of the data analysis*

Pilot	Accessibility
UCL (UK)	50%
EDF (France)	50%
IASI (Romania)	n.a.
VEOLIA (Spain)	100%
EMOT (Italy)	80%
FOCCHI (Italy)	50%
HERON (Greece)	80%
FVH (Finland)	80%
IEECP (The Netherlands)	70%
NTUA (Greece)	80%

Timeliness

Last dimension in DigiBUILD is timeliness, responsible for setting the time when data is made available, which relates to the "real-time" operation. Some samples are obtained in 5 min periodicity, while others are in 15 minutes or even 1 hour, as specified in Table 10. Here, the event-driven data collection (in other words, change of value) is not considered, although it exists.

Discussion

Data quality is crucial to make decisions, such as improvement in the energy management of buildings, retrofitting investments or air quality enhancement. This is not always covered and decisions are made based on low quality data. Within DigiBUILD a methodology to check datasets quality, incorporating the improvement, is proposed. At this first stage, the initial data analysis has been performed to determine the baseline.

What is clearly remarkable is the maturity level and the age of the building. According to Figure 2, depending on the monitoring stage, optimal operation in the data collection can be achieved, as the cases of VEOLIA, IEECP or FVH pilots, with very high level of maturity; therefore, high-quality of data. Others, such as IASI pilot is still in the interventions stage, without providing data (0% in the indicators), as depicted in Figure 5. The age of the building is also important. New buildings, such as the UCL pilot with only 1 year old, contain smarter technology with capabilities of data storage and access.

To sum up, it is undoubtedly data quality is a pivotal aspect in the decision-making process and this aspect should be carefully treated. As observed in Figure 5, there exists a diversity of maturity levels with a clear room for improvement. The graph is only showing a subset of dimensions (the most relevant ones). Availability and accessibility of data must be ensured. Data processing for data filtering and cleansing is a very important task that would improve the current data quality.

Table 10: *Timeliness results of the data analysis*

Pilot	Timeliness
UCL (UK)	5/15 min
EDF (France)	5/10 min
IASI (Romania)	n.a.
VEOLIA (Spain)	15 min
EMOT (Italy)	3 min
FOCCHI (Italy)	1 hour
HERON (Greece)	30 sec
FVH (Finland)	10 min
IEECP (The Netherlands)	30 min
NTUA (Greece)	5 min

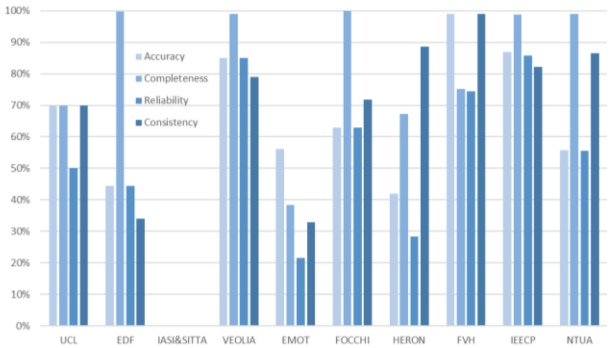


Figure 5: Data quality indicators per pilot building

Conclusions

This paper has described a data analysis carried out under the umbrella of the European DigiBUILD project to determine the baseline in terms of data quality for decision-making. From historical data samples in 9 different countries across Europe, 7 dimensions have been considered within the data analysis. The results provide an overview of the real maturity levels within the digitalization and data collection of the building stock.

Different levels of maturity could be observed, from pilots that already are in the optimal operation stage; then, obtaining high-values for the different data quality dimensions to other in very early stages, decreasing the trustworthiness on data. This deals with two issues. First of all, making decisions based on low-quality data; hence, increasing the uncertainty. Secondly, decision-makers are reluctant when addressing an energy efficiency project. In this sense, data quality checks are crucial.

A data analysis has been performed to check the data quality levels of multiple datasets. Due to the capability for improvement of the quality, the methodology is extended with correction techniques. At the current state of the project, they have not been fully implemented and it is the future work. The aim is to increase data quality and reach the outcomes that were previously explained.

Acknowledgments

The authors would like to thank DigiBUILD consortium and European Commission for funding the project under GA#101069658 of the Horizon Europe programme.

References

Abraham, R., Schneider, J., Vom Brocke, J. (2019). Data governance: A conceptual framework, structured review, and research agenda. International Journal of Information Management, 49, pp. 424-438.

BDVA/DAIRO position paper. (2021). Response to the European Commission's proposal for AI Regulation. https://www.bdva.eu/sites/default/files/BDVA_DAIRO%20response-feedback%20AI%20Regulation_Final.pdf, visited on 10th January 2024.

Coujard, C., Eloire, K. L., Zarli, A., & David, A. (2023). SmartBuilt4EU: Towards a strategic research and policy agenda for the European smart buildings community. In ECPPM 2022-eWork and eBusiness in Architecture, Engineering and Construction 2022 (pp. 785-794). CRC Press.

Duvier, C., Neagu, D., Oltean-Dumbrava, C., Dickens, D. Data quality challenges in the UK social housing sector, International Journal of Information Management, Volume 38, Issue 1, 2018, 196-200.

Ender, L. 'Data Governance in Digital Platforms: A case analysis in the building sector', Dissertation, 2021

EPBD (2023) https://energy.ec.europa.eu/topics/energy-efficiency/energy-efficient-buildings/energy-performance-buildings-directive_en#current-rules-to-improve-the-eus-building-stock

Kaginalkar, A., Kumar, S., Gargava, P., & Niyogi, D. (2023). Stakeholder analysis for designing an urban air quality data governance ecosystem in smart cities. Urban Climate, 48, 101403.

Hernández, J.; Quijano, A.; García, R.; Nouaille, P.; Risch, L.; Virtanen, M. and de Miguel, I. (2022). Analysis of Data Quality in Digital Smart Cities: The Cases of Nantes, Hamburg and Helsinki. In Proceedings of the 11th International Conference on Data Science, Technology and Applications - DATA; ISBN 978-989-758-583-8; ISSN 2184-285X, SciTePress, pages 353-360. DOI: 10.5220/0011271900003269

Hernández, J., Martín, S., Kapsalis, P., Katsigarakis, K., Sarmas, E., Marinakis, V. Building a Data Lake for Smart Building Data: Architecture for data quality and interoperability, 2023 14th International Conference on Information, Intelligence, Systems and Applications (2023a), University of Thessaly, Volos, Greece, 10-July 2023.

Hernández, J. L., Martín, S., Marinakis, V., & de Miguel, I. (2023b). From silos to open, federated and enriched Data Lakes for smart building data management. In 2023 IEEE International Workshop on Metrology for Living Environment (MetroLivEnv) (pp. 29-33). IEEE.

Hitachi, Pentaho Data Integration, online: https://help.hitachivantara.com/Documentation/Pentaho/Data_Integration_and_Analytics/9.1/Products/Pentaho_Data_Integration, visited on 27th December 2023.

Hossein Motlagh, N.; Mohammadrezaei, M.; Hunt, J.; Zakeri, B. Internet of Things (IoT) and the Energy Sector. Energies 2020, 13, 494. https://doi.org/10.3390/en13020494.

Morewood, J. (2023). Building energy performance monitoring through the lens of data quality: A review, Energy and Buildings, Volume 279, 2023, 112701, ISSN 0378-7788, https://doi.org/10.1016/j.enbuild.2022.112701.

Rao, P. M., & Deebak, B. D. (2023). Security and privacy issues in smart cities/industries: technologies, applications, and challenges. Journal of Ambient Intelligence and Humanized Computing, 14(8), 10517-10553.

Seyedzadeh, Saleh, Rahimian, Farzad Pour, Oliver, Stephen, Rodriguez, Sergio and Glesk, Ivan, Machine learning modelling for predicting non-domestic buildings energy performance: A model to support deep energy retrofit decision-making, Applied Energy, Volume 279, 2020, 115908, ISSN 0306-2619, https://doi.org/10.1016/j.apenergy.2020.115908.

Villada, Fernando; Cadavid, Diego Raúl and Molina, Juan David. Pronóstico del precio de la energía eléctrica usando redes neuronales artificiales. Rev.fac.ing.univ. Antioquia [online]. 2008, n.44 [cited 2023-12-27], pp.111-118. Available from: http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0120-62302008000200011&lng=en&nrm=iso. ISSN 0120-6230.

Yong Teng, S., Touš, M., Dong Leong, W., Shen How, B., Loong Lam, H., Máša, V. Recent advances on industrial data-driven energy savings: Digital twins and infrastructures, Renewable and Sustainable Energy Reviews, Volume 135, 2021, 110208, https://doi.org/10.1016/j.rser.2020.110208

TOWARDS DETECTING DAMAGE IN LIGHTWEIGHT BRIDGES WITH TRAVELING MASSES USING MACHINE LEARNING

Georgios Dadoulis[1], George D. Manolis[1], Konstantinos Katakalos[1], Thamer Al-Zuriqat[2], Kosmas Dragos[2], and Kay Smarsly[2]
[1]Aristotle University of Thessaloniki, Thessaloniki, Greece
[2]Hamburg University of Technology, Hamburg, Germany

Abstract

Lightweight bridges are subjected to moving loads (vehicular traffic), with vehicular masses typically being comparable to structural masses. Moving loads are thus regarded as "traveling masses", resulting in complex dynamic behavior, which is hardly covered by conventional damage detection strategies. This paper presents a concept towards damage detection in lightweight bridges with traveling masses using machine learning (ML). Specifically, a ML model for classifying structural damage is trained, using simulations, and applied using real-world structural response data. Preliminary tests of the proposed concept validate the power of the ML model in identifying structural damage, despite the non-stationarity of the problem.

Introduction

Research on advancing material technologies to develop high-strength materials has been conducted since the second half of the 20th century (Flaga, 2000). The construction industry has been increasingly leveraging high-strength materials in an attempt to meet structural design requirements while reducing material usage (Zhang et al., 2019). Furthermore, in bridge design and construction, practitioners frequently seek to reduce the size of structural elements, which, apart from the obvious financial benefit, also helps meet practical and aesthetical requirements.

High-strength materials, however, frequently result in lightweight flexible structures, which are prone to heavy oscillations. A notable example is the case of the Millennium Bridge in London, which exhibited large oscillations, due to the large crowds drawn during opening in 2000 (Dallard et al., 2001). Consequently, material reduction in bridges requires care to avoid unanticipated oscillations. Moreover, placing the discussion on bridge oscillations under operational conditions, bridges are expected to withstand oscillations induced by operational loads, such as vehicular traffic.

At a first glance, the problem of vehicles traversing bridge decks may be described as a "moving load" problem, the dynamics of which differ from conventional structural dynamics, in that the structural response is affected by the velocity of the vehicles (Firus, 2023). In lightweight bridges, however, vehicular masses may be comparable to structural masses (particularly in railway bridges) and are regarded as "traveling masses", resulting in a complex non-stationary problem. With respect to damage detection, which is the focal point of structural maintenance, conventional damage detection strategies hardly address the traveling-mass problem. Experimental techniques as well as structural health monitoring (SHM) strategies typically rely on data analysis methods suitable for stationary problems, e.g. identifying the frequency content of structural response data using the Fourier transform (Dragos et al., 2024). Studies on moving loads on beams also fail to capture the damage-detection traveling-mass problem in its totality.

Approaches describing the dynamic response of beam structures under traveling masses have been widespread in literature. Examples range from as far back as the late 1990s, with the work of Siddiqui et al. (1998) who investigated the motion of a cantilever Euler-Bernoulli beam under the effect of a moving mass-spring system, to recent works, such as the approach reported by Meher et al. (2019), who have used Green's function to describe the response of a beam traversed by a traveling mass. Furthermore, the vibration control of beams, considering vehicle-bridge interaction, has been investigated by Pi & Ouyang (2016), and the critical speed of a traveling mass, affecting the behavior of the coupled mass-beam system, has been the focus of work by Dehestani et al. (2009). A recent work, reported by Abu-Alshaikh et al. (2020), has focused on obtaining analytical responses of a functionally graded beam with a traveling mass using Caputo-Fabrizio fractional derivative models.

Nevertheless, damage detection on beams with traveling masses has received scarce attention. An early approach has been proposed by Billelo & Bergman (2004), basing damage detection on comparing the displacement histories obtained from an intact bridge model before and after inducing damage. In recent work, Cicirello (2019) has studied the response bounds of Euler-Bernoulli beams with structural damage from a theoretical standpoint. Moreover, Zhan et al. (2021), have used wavelet transform coefficients of experimentally derived modal parameters for indirect damage detection. Despite the solid background offered by the aforementioned approaches, damage detection relies on either displacement measurements, which, although frequently used in experimental techniques, are uncommon in state-of-the-art SHM systems employed for structural maintenance, or on the accuracy of experimentally derived modal parameters, which is hardly guaranteed in non-stationary problems. As a result, damage detection on beams with traveling masses stands to benefit from an approach compatible with measurements typically collected in modern SHM systems (e.g. acceleration response data) and with state-of-the-art computational tools.

The non-stationarity of the traveling-mass problem calls for advanced data analysis methods, drawing from the field of artificial intelligence and its subset machine learning (ML), which has been employed in civil engineering applications, such as in fault diagnosis of SHM systems (Al-Zuriqat et al., 2023). In this direction, this paper presents a concept for damage detection on lightweight bridges with traveling masses, using a ML model for classifying acceleration response data into structural damage scenarios. The reasoning behind using ML is that damage patterns manifested in acceleration response data may be either too subtle to discern or obscured in frequency-domain representations of the data, due to the non-stationarity of the problem. The ML classification model, used in this study, is a convolutional neural network (CNN), which is trained with "labeled" acceleration response data, i.e. data corresponding to predefined structural condition scenarios, obtained from simulations using a well-calibrated analytical model. Upon completing training, classification is performed using real-world acceleration response data. The proposed concept is validated in preliminary laboratory tests, showcasing the capability of the CNN in identifying structural damage, as represented by partial loss of fixity at one support of a steel beam. The remainder of the paper includes a presentation of the concept, followed by the validation tests. The paper ends with a summary and conclusions and an outlook on future research.

Damage detection in lightweight bridges with traveling masses

The cornerstone of the proposed concept is the coupling between analytical modeling and simulation with ML. It is therefore clear that creating a well-calibrated analytical model for producing the acceleration response data for training the CNN is particularly important. As such, the description of the concept in this section starts with a brief discussion on the analytical modeling and simulation method employed for producing acceleration response data, followed by a presentation of the CNN training, including the preprocessing of the acceleration response data and the functionality of the CNN.

Analytical modeling of the traveling-mass problem

In its simplest form, a lightweight bridge with a traveling mass can be modeled as a continuous Euler-Bernoulli beam, as shown in Figure 1. The equation of motion of the beam accounts for inertia forces both of the vibrating beam and of the travelling mass (Dadoulis & Manolis, 2023):

$$\rho A\ddot{w}(x,t)+c\dot{w}(x,t)+EIw''''(x,t)=f_c\delta\left(x-|v|t\right). \quad (1)$$

In Equation (1), ρ is the material density, A is the cross-section area of the beam, E is the modulus of elasticity, I is the moment of inertia of the beam in the vertical direction, c is the damping coefficient, w is the vertical displacement of the neutral axis of the beam, δ is the Dirac delta, x represents the spatial coordinate in the longitudinal direction of the beam, and t is time. The functional $f_c > 0$ represents the contact force between the traveling mass and the beam, and $|v|$ is the speed of the traveling mass. Finally, $\dot{w} = \partial w/\partial t$, $\ddot{w} = \partial^2 w/\partial t^2$, and $w'''' = \partial^4 w/\partial x^4$.

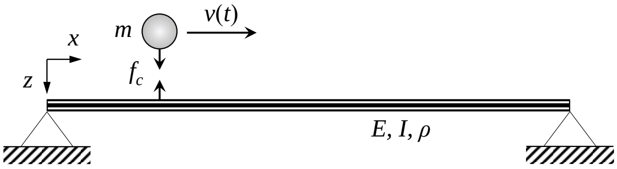


Figure 1: Bernoulli-Euler beam with traveling mass.

From Equation (1), it is evident that the factor that differentiates the traveling-mass problem from a regular Euler-Bernoulli-beam equation of motion is the contact force f_c. To estimate f_c, the equilibrium at the contact point of the traveling mass with the beam is considered. First, the static equilibrium is expressed as the total vertical displacement w_T, which is a summation of the vertical displacement of the beam and a component representing the surface roughness $r(x) = r(|v|t)$:

$$w_T\left(|v|t,t\right)=w\left(|v|t,t\right)+r\left(|v|t\right). \quad (2)$$

Furthermore, the dynamic equilibrium is provided by the following expression:

$$m\ddot{w}_T\left(|v|t,t\right)=mg-f_c, \quad (3)$$

where g is the gravitational acceleration. By solving Equations (2) and (3) with respect to f_c and by substituting f_c into Equation (1), the equation of motion is converted into:

$$\rho A\ddot{w}+c\dot{w}+EIw''''=m\left[g-\frac{d^2w}{dt^2}-\frac{d^2r}{dt^2}\right]\delta\left(x-|v|t\right). \quad (4)$$

In Equation (4), the time and spatial coordinate of w have been dropped for simplicity. It is noted that the right-hand side of Equation (4) includes total derivatives, which are expanded as follows:

$$\frac{d^2w}{dt^2}=\frac{\partial^2 w}{\partial t^2}+2v\frac{\partial^2 w}{\partial x\partial t}+v^2\frac{\partial^2 w}{\partial x^2}, \quad (5)$$

$$\frac{d^2r(x)}{dt^2}=\frac{\partial^2 r}{\partial x^2}\left(\frac{dx}{dt}\right)^2=v^2\frac{\partial^2 r}{\partial x^2}. \quad (6)$$

In Equation (5), the second term corresponds to the Coriolis acceleration, and the third term is the centrifugal acceleration, which is irrelevant to the problem being studied and is therefore neglected. Using Equations (5) and (6), Equation (4) becomes:

$$\rho A\ddot{w} + c\dot{w} + EIw'''' = m\left[g - \ddot{w} - 2v\dot{w}' - v^2 r''\right]\delta(u) \\ u = x - vt. \tag{7}$$

Equation (7) is solved via modal analysis, for which the vertical displacement is analyzed into a spatial component and a temporal component:

$$w(x,t) = \varphi_n(x)q_n(t), \qquad n = 1,2,...,p\,, \tag{8}$$

where $\varphi_n(x)$ is nth "eigenfunction", representing the mode of vibration in the form of a wave function, and $q_n(t)$ is the nth "generalized" coordinate function, which characterizes the temporal variation of the vibration. From Equation (8), it is clear that the complexity of the solution depends on the number of eigenfunctions (p) considered. For regular structures in structural dynamics, it is common to consider only a few eigenfunctions, which are capable of capturing the structural dynamic behavior. The nth eigenfunction is given by:

$$\varphi_n(x) = C_n \sin\left(\frac{n\pi x}{L}\right), \tag{9}$$

where L is the beam length, and C_n is a constant related to the boundary conditions of the beam. Eigenfunctions follow the orthogonality condition:

$$\rho A\int_0^L \varphi_i(x)\varphi_j(x)dx = \delta_{ij}\,, \tag{10}$$

where δ_{ij} is the Kronecker delta. As a result, by substituting Equation (8) into Equation (7) for p eigenfunctions arranged in vector format $w = \boldsymbol{\varphi}(x)\mathbf{q}(t)$, ($\boldsymbol{\varphi}(x) = [\varphi_1(x),\varphi_2(x),\ldots,\varphi_p(x)]$, $\mathbf{q}(t) = [q_1(t),q_2(t),\ldots,q_p(t)]^T$), pre-multiplying with $\boldsymbol{\varphi}(x)^T$, integrating over L, and exploiting the Dirac delta property:

$$\int_0^L \varphi(x)\delta(x - vt)dx = \varphi(vt), \tag{11}$$

proven in Dadoulis & Manolis (2022), the equation of motion is recast into a $p\times p$ system of differential equations of the following format:

$$\begin{aligned} &\mathbf{M}(t)\ddot{\mathbf{q}}(t) + \mathbf{C}(t)\dot{\mathbf{q}}(t) + \mathbf{K}(t)\mathbf{q}(t) = \mathbf{F}(t) \\ &\mathbf{M}(t) = \mathbf{I} + m\boldsymbol{\varphi}\boldsymbol{\varphi}^T \\ &\mathbf{C}(t) = diag(2\xi_n\omega_n) + 2mv\boldsymbol{\varphi}\boldsymbol{\varphi}'^T \\ &\mathbf{K}(t) = diag(\omega_n^2) + mv^2\boldsymbol{\varphi}\boldsymbol{\varphi}''^T \\ &\mathbf{F}(t) = mg\boldsymbol{\varphi} - mv^2 r''(x(t))\boldsymbol{\varphi}. \end{aligned} \tag{12}$$

In Equation (12), ξ_n is the critical damping ratio of the nth eigenfunction, and ω_n is the respective natural frequency, computed as follows:

$$\xi_n = \frac{c}{2\rho AL\omega_n} \qquad \omega_n = \frac{n^2\pi^2}{L^2}\sqrt{\frac{EI}{\rho A}}. \tag{13}$$

For producing the structural response to the traveling-mass problem, the generalized coordinate functions $\mathbf{q}(t)$ need to be computed. To this end, Equation (12), which represents a set of ordinary differential equations of second order, is converted into a set of first-order ordinary differential equations, using the state-space formulation:

$$\dot{\mathbf{y}}(t) = \left[-\mathbf{A}^{-1}(t)\mathbf{B}(t)\right]\mathbf{y}(t) + \mathbf{A}^{-1}(t)\mathbf{h}(t)\,. \tag{14}$$

The notations in Equation (14) are explained as follows:

$$\mathbf{A}(t) = \begin{bmatrix} \mathbf{0} & \mathbf{M}(t) \\ \mathbf{M}(t) & \mathbf{C}(t) \end{bmatrix}, \ \mathbf{B}(t) = \begin{bmatrix} -\mathbf{M}(t) & \mathbf{0} \\ \mathbf{0} & \mathbf{K}(t) \end{bmatrix} \\ \dot{\mathbf{y}}(t) = \begin{bmatrix} \dot{\mathbf{q}}(t) \\ \mathbf{q}(t) \end{bmatrix}, \quad \mathbf{h}(t) = \begin{bmatrix} \mathbf{0} \\ \mathbf{F}(t) \end{bmatrix}. \tag{15}$$

Assuming discrete time, Equation (14) is solved for every time instance k by performing eigenvalue analysis to matrix $(-\mathbf{A}_k)^{-1}\mathbf{B}_k$, which is non-symmetric and yields a matrix of complex eigenvalues $\boldsymbol{\Lambda}_k$ and eigenvectors matrix $\boldsymbol{\Psi}$. The eigenvalues matrix is given by:

$$\boldsymbol{\Lambda}_k = diag(\boldsymbol{\lambda}_{n,k}), \ \boldsymbol{\lambda}_{n,k} = \begin{bmatrix} e^{\lambda_{n,k}} & 0 \\ 0 & e^{\bar{\lambda}_{n,k}} \end{bmatrix}, \\ \lambda_{n,k} = \xi_{n,k}\omega_{n,k} \pm i\omega_{n,k}\sqrt{1 - \xi_{n,k}^2}\,, \quad \xi_{n,k} \neq \xi_n. \tag{16}$$

Using Equation (16), the solution of Equation (14) at time instance k is:

$$\mathbf{y}_k = \boldsymbol{\Theta}_k(\boldsymbol{\Theta}_{k-1})^{-1}\mathbf{y}_{k-1} + \boldsymbol{\Theta}_k\int_{k-1}^{k}(\boldsymbol{\Theta}_s)^{-1}\mathbf{A}_k^{-1}\mathbf{h}_k ds \\ \boldsymbol{\Theta}_k = \boldsymbol{\Psi}_k\boldsymbol{\Lambda}_k. \tag{17}$$

The generalized coordinate functions are used together with the eigenfunctions to reconstruct the vertical displacement functions $w(x,t)$ using Equation (8). In turn, the vertical displacement functions are used as to produce the acceleration response data using Equation (14) (hereinafter termed "simulation data"), which are processed to be used for training the CNN.

Training of a convolutional neural network for damage detection

In theory, the simulation data could be used "as is" (i.e. in time-history format) as input to the CNN, since simulation data is relatively "clean" from noise and random components. However, upon completing training, the

CNN is expected to be applicable with real-world acceleration response data, in which the presence of spurious factors, such as noise and random components, is likely. As argued in Fritz et al. (2022), to reduce the effect of spurious factors and expose features in the simulation data, data preprocessing is necessary. Considering the non-stationarity of the traveling-mass problem, common preprocessing methods, producing fixed-window frequency-domain representations are hardly informative. Instead, preprocessing that depicts the coupled time-frequency content of the simulation data is necessary. Furthermore, as evidenced in Equation (14), the eigenvalues of the coupled traveling-mass-beam system change at every time instance *k*. As a result, a preprocessing method that enables tracking the evolution of eigenvalues is used, namely the Gabor transform (Qian & Chen, 1993):

$$G_w\left(\kappa\right) = S\left[\sum_{j=1}^{N} \ddot{w}\left(x,t_j\right) z_j e^{-i2\pi\kappa j / N}\right]^2, \kappa = \frac{N \cdot f}{f_s}. \quad (18)$$

In Equation (18), G_w is the Gabor coefficient, κ is the index of the discrete frequency bin of frequency f, considering a discrete set of simulation data of length N, sampled at frequency f_s ($f_s = 1/\Delta t$). The "Gabor" window function is denoted as z_j and could be any window function, such as the Gaussian function or the Hann window. Finally, S is a scaling factor.

Gabor coefficients are typically illustrated as colormap images. An exemplary illustration of a Gabor-transform image of simulation data representing acceleration of an Euler-Bernoulli beam is given in Figure 2.

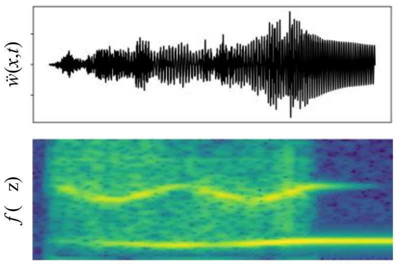


Figure 2: Gabor transform depiction of simulation data.

Upon producing the Gabor transform images, the damage detection problem is reduced to a classification problem, which is solved by the CNN. As shown Figure 3, a typical CNN consists of one *input* layer, a succession of *convolution* layers and *pooling* layers, (at least) one *fully-connected* layer, one *dropout* layer, and one *output* (classification) layer. Each Gabor transform image is passed in red-green-blue (RGB) format of size $N_I \times N_I \times 3$ from the input layer to the first convolution layer. Thereupon, the image is subjected to feature extraction, performed by progressively "sliding" square matrix "kernels" of size $n_I \times n_I \times 3$ (generally, $n_I << N_I$) over the image and by convolving RGB values of image areas (of dimensions matching the kernel size) with the kernel matrix elements. The outcome of each convolution layer is a so-called "feature map", consisting of convolutional products, and each product is fed to an activation function, which introduces non-linearity to the CNN, necessary for solving non-trivial problems. The outcomes of the activation function ("activations") are propagated to the pooling layer, which applies a sliding-window operation on the activations, typically computing either the average of each sliding window or the maximum value, to reduce the dimensionality of the activations. The last pooling layer is succeeded by a fully-connected layer, which comprises neurons connected to every outcome of the pooling layer. The fully-connected layer is followed by a dropout layer, in which a predefined percentage of features are "dropped" (i.e. ignored) to prevent overfitting. Finally, the dropout layer is connected to the output layer, which holds the classes.

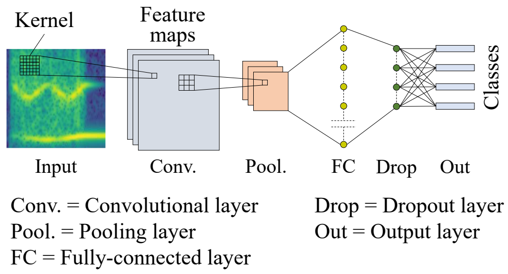


Figure 3: Layout of a typical CNN.

To perform classification, the CNN is trained, i.e. the kernels are fine-tuned so that the prediction accuracy of the output layer exceeds a predefined threshold. The training is performed using a "labeled" set of Gabor transform images, i.e. images that correspond to known structural conditions, including images from the intact structure as well as images from damage scenarios, simulated via changing values of structural parameters of the analytical model, resulting in stiffness reduction equivalent to reduction caused by damage. The labeled set is divided into a training set (70% of the images), a validation set (20% of the images), and a testing set (10% of the images). The training set is propagated sequentially (in "batches") through the CNN, and the prediction error gradient, obtained from the output layer is back-propagated to update the kernels using an optimization algorithm, such as gradient descent, and a learning rate, which controls the fine-tuning rate of the kernels. Each cycle of propagation and back-propagation of a batch is a training iteration, and once all the images in the training set are used, one epoch is completed. The validation set is passed through the CNN periodically during training, after a predefined number of iterations. The purpose of validation is to monitor the prediction accuracy improvement of the CNN with an independent set of images so as to avoid overfitting the CNN to the training set. The testing set is passed through the CNN at the end of training, i.e. upon achieving the predefined prediction accuracy, to check the CNN performance.

A proof-of-concept implementation and validation of the proposed concept, including the analytical modeling of a beam structure, the data preprocessing and the training of a CNN for damage classification are presented in the next section.

Implementation and validation

The prototype implementation and validation of the proposed concept involve defining software tools for simulations and CNN training, as well as devising a laboratory proof-of-concept test. In this section, first the implementation is discussed, followed by the presentation of the validation test and a discussion of the results.

Implementation

The first part of the implementation consists in developing an algorithm for analytical modeling and data pre-processing, based on Equations (1)-(18). The algorithm is written in the Python programming language and consists of functions *model_setup*, *eig_decomp* and *state_space_comp*, dedicated to (i) computing the matrices **A** and **B** for every time instance k, (ii) performing eigenvalue analysis to matrix $(-\mathbf{A}_k)^{-1}\mathbf{B}_k$, and (iii) to applying Equation (17) to compute the vector $\mathbf{y}_k$, respectively. A separate function, *data_gen*, is written to retrieve the displacement histories, using the eigenfunctions, and to compute the acceleration response data using Equation (14). Finally, the Gabor transform is realized by function *gaborTrsf*, and the corresponding images are produced.

The second part of the implementation involves developing and training the CNN. Since, only a proof-of-concept study is shown, an elaborate parametric analysis of the CNN performance and an optimal definition of the CNN architecture fall beyond the scope of this paper and will be investigated in future research. Instead, the MATLAB deep learning tool is leveraged, and a simple CNN architecture is devised, shown in Figure 4, adjusted to the input of Gabor transform images and to the number of classes being considered.

Validation tests

The laboratory validation test is devised on a simply supported HEB100 steel beam, representing a downscaled model of a lightweight bridge traversed by a traveling mass, shown in Figure 5. The cross section of the beam has dimensions 100×100×10×6 (mm) (width × height × flange thickness × web thickness). The total length of the neutral axis of the beam is 5830 mm. The beam is part of an experimental setup, which includes adjustable struts that ensure that the beam is level. It is noted that the beam is supported only at its ends, i.e. the struts are not used as supports to the beam. The experimental setup also consists of pulleys that enable masses to move along the longitudinal axis of the beam and a motor that controls the mass speed.

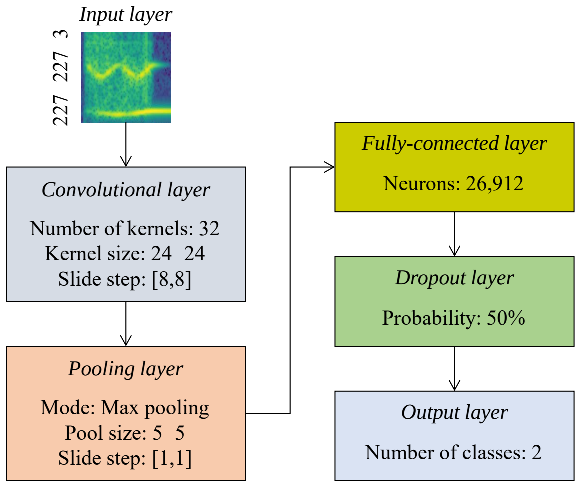


Figure 4: The CNN architecture defined in this study.

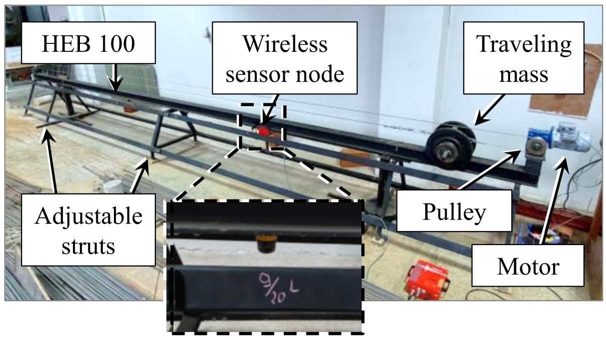


Figure 5: Experimental setup.

An analytical model of the beam is developed and used for simulating the motion of a traveling mass across the beam. The modulus of elasticity and the material density are computed via preliminary laboratory tests as $E = 198.5$ GPa and $\rho = 7.65 \cdot 10^3$ kg/m^3, respectively. To produce a labeled set covering a sufficient part of the available solutions space, the mass value and speed are randomly perturbed to devise several simulation scenarios. Specifically, the mass (m) assumes values between 10 kg and 40 kg, and the speed ($|v|$) assumes values between 0.2 m/s and 0.5 m/s. With respect to the structural condition, two structural “states” are considered, one for the intact beam and one assuming loss of fixity at one support, i.e. by substituting the vertical support with a translational spring, resembling erosion of foundation subsoil, e.g. as a result of scour. A total of 500 simulations are conducted for each state, computing acceleration response data at a distance equal to 9L/20 from the beam end and at a sampling frequency of $f_s = 128$ Hz. The reasoning behind avoiding the midspan of the beam is that the second eigenfunction, which contributes to the structural response, is characterized by an antisymmetric sinusoidal shape with a zero-crossing point at the midspan. Hence, collecting acceleration response data from the midspan would underestimate the contribution of the second eigenfunction. In turn, 500 Gabor transform images per class are produced, forming the labeled set used for training the CNN. Exemplary illustrations of Gabor transform images for both states are given in Figures 6 and 7. The variation of the first two eigenfrequencies is marked for both structural states considered, as a result of the variable structural dynamic properties of the beam due to the motion of the mass. As can be seen in the figures, the loss of fixity at the support reduces the eigenfrequencies; however, the variation trend is similar in both states.

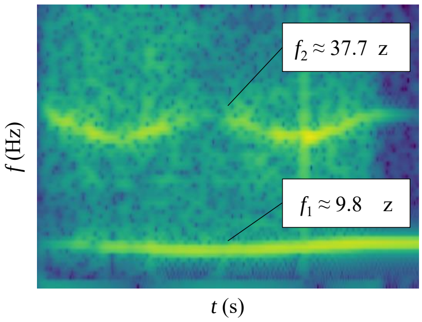


Figure 6: Gabor transform for the intact beam (m = 23.3 kg, $|v|$ = 0.351 m/s).

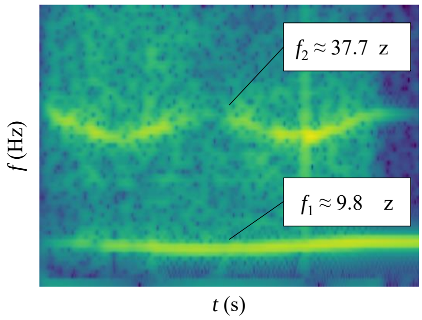


Figure 7: Gabor transform for the damaged beam (m = 15 kg, $|v|$ = 0.46 m/s).

The training of the CNN is performed by first splitting the labeled set into 70% training set (350 images), 20% validation set (100 images), and 10% testing set (50 images). During training, the "Adam" optimizer – a stochastic gradient-descent optimization algorithm – is employed for fine-tuning the kernels in each iteration (Kingma & Ba, 2015). The initial learning rate is set to 0.001, the batch size is 30 images, and the validation frequency is defined at every 30 iterations. The total number of epochs is set equal to 10, and the learning rate is reduced at every epoch by a factor of 0.5. A high training accuracy of 98% is achieved after 122 iterations, and the accuracy of the validation set reaches 100% after 30 iterations, as shown in Figure 8.

Prior to applying the CNN with acceleration response data from the experimental setup, the classification capability of the CNN is verified using the testing set. The classification test results are given in the form of a "confusion" matrix, shown in Table 1.

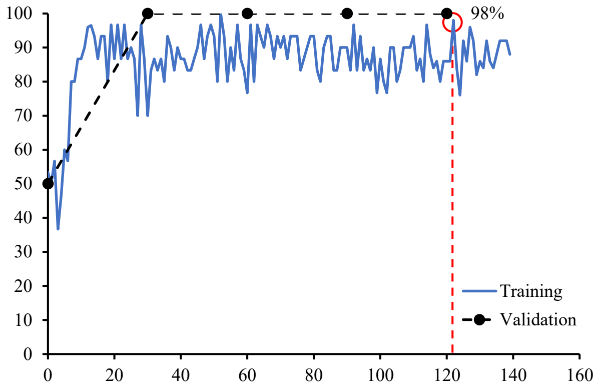


Figure 8: Training and validation accuracy history.

Table 1: Confusion matrix for the testing data set.

	Intact	Damaged
Intact	50	0
Damaged	0	50

The results for the testing dataset corroborate the training accuracy, since every image of the testing dataset has been correctly classified. The next step of the validation test involves using the CNN for classifying Gabor transform images produced from real-world acceleration response data obtained from the experimental setup. The data is collected using a Lord Microstrain G-Link-200 wireless sensor node capable of measuring acceleration in 3 axes at ranges up to ±8g and at sampling frequencies up to 4,096 Hz (Microstrain Sensing, 2020). Experiments are conducted for two structural states, matching the states used in the simulations, using 30 combinations of mass and speed values per state, as listed in Table 2.

Table 2: Combinations of mass and speed values for the laboratory experiments.

Nr.	Mass m (kg)	Speed $|v|$ (m/s)	Nr.	Mass m (kg)	Speed $|v|$ (m/s)
1	13	0.25	16	27	0.40
2	13	0.30	17	27	0.45
3	13	0.35	18	27	0.50
4	13	0.40	19	18	0.25
5	13	0.45	20	18	0.30
6	13	0.50	21	18	0.35
7	23	0.25	22	18	0.40
8	23	0.30	23	18	0.45
9	23	0.35	24	18	0.50
10	23	0.40	25	38	0.25
11	23	0.45	26	38	0.30
12	23	0.50	27	38	0.35
13	27	0.25	28	38	0.40
14	27	0.30	29	38	0.45
15	27	0.35	30	38	0.50

The loss of fixity is realized by placing a spring between the beam end and its support. Acceleration response data is collected for every experiment close to the midspan of the beam, at a point located at $9L/20$ from the beam end, as shown in Figure 5, at a sampling frequency of $f_s = 128$ Hz, as in the simulations. The acceleration response data from each experiment is used to produce a Gabor transform image, resulting in a set of 60 Gabor transform images to be fed to the CNN, an example of which is shown in Figure 9.

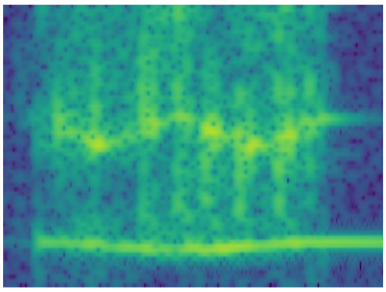


Figure 9: Gabor transform for experiment 10 (intact beam).

The similarity between Figure 9 and Figures 6 and 7 is indicative of the satisfactory proximity between the analytical model and the experimental setup. The Gabor transform images are fed to the CNN and classified into the two classes, corresponding to the states considered. The results are summarized in the following confusion matrix:

Table 3: Confusion matrix for the experiments.

	Intact	Damaged
Intact	30	0
Damaged	1	29

The results of the experiments clearly highlight the capability of the CNN to correctly classify real-world acceleration response data. From the Gabor transform images of the intact beam, only one experiment is misclassified as “damaged”, while the respective images of the damaged beam are all correctly classified. Furthermore, the results serve as proof of the validity of the proposed concept and of the “transferability” of the CNN training from the domain of simulations to the real world.

Summary and conclusions

Damage detection on lightweight bridges with traveling masses is a non-stationary problem that is hardly covered by conventional methods, developed for experimental testing or structural health monitoring. In this context, this paper has presented a concept towards detecting damage on lightweight bridges, represented by Euler-Bernoulli beams. The proposed concept is based on analyzing the behavior of the beam using analytical modeling for a set of predefined damage scenarios and on using simulation-derived acceleration response data to train a convolutional neural network to classify the data into the damage scenarios. Thereupon, the convolutional neural network is applied using real-world acceleration response data.

The proposed concept has been validated in a proof-of-concept laboratory test on a steel beam with a traveling mass, representing a downscaled model of a lightweight bridge. The results have shown that the convolutional neural network (i) can be trained with high accuracy using simulation-derived acceleration response data and (ii) can be transferred to real-world applications, assuming adequate calibration of the analytical model. Future work may focus on considering several damage scenarios, anticipated to occur over the lifetime of lightweight bridges, as well as on developing the proposed concept into a structural health monitoring approach.

Acknowledgments

This paper is the result of a collaborative effort of authors funded by different sources. The authors would like to gratefully acknowledge the support offered by the German Research Foundation (DFG) under grant SM 281/20-1, by the Hellenic Foundation for Research and Innovation (HFRI) under grant 6255, as well as by the German Federal Ministry for Digital and Transport (BMDV) within the mFUND program under grant 19FS2013B. Results presented in this work have been produced using the Aristotle University of Thessaloniki High-Performance Computing Infrastructure. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the DFG, the HFRI, or the BMDV.

References

Abu-Alshaikh, I.M. & Almbaidin, A.A. (2020) Analytical responses of functionally graded beam under moving mass using Caputo and Caputo–Fabrizio fractional derivative models. Journal of Vibration and Control, 26(19-20), pp. 1859-1867.

Al-Zuriqat, T., Chillón Geck, C., Dragos, K., & Smarsly, K. (2023) Adaptive fault diagnosis for simultaneous sensor faults in structural health monitoring systems. Infrastructures, 8(3), 39.

Bilello, C. & Bergman, L.A. (2004) Vibration of damaged beams under a moving mass: theory and experimental validation. Journal of Sound and Vibration, 274(3-5), pp. 567-582.

Cicirello, A. (2019) On the response bounds of damaged Euler–Bernoulli beams with switching cracks under moving masses. International Journal of Solids and Structures, 172-173(2019), pp. 70-83.

Dadoulis, G. & Manolis, G.D. (2022) On the detection of fracture within vibrating beams traversed by a moving force. Infrastructures, 7(7), 93.

Dadoulis, G. & Manolis, G.D. (2023) Dynamic response of a damaged bridge model traversed by a heavy point mass. Journal of Sound and Vibration, 551(2023), 117613.

Dallard, P., Fitzpatrick, T., Flint, A., Low, A., Ridsdill Smith, R., Willford, M., & Roche, M. (2001) London Millennium Bridge: Pedestrian-induced lateral vibration. ASCE Journal of Bridge Engineering, 6(6), pp. 412-417.

Dehestani, M., Mofid, M., & Vafai, A. (2009) Investigation of critical influential speed for moving mass problems on beams. Applied Mathematical Modeling, 33 (2009), 3885-3895.

Dragos, K., Magalhães, F., Manolis, G. D., & Smarsly, K. (2024). Frequency-domain synchronization of structural health monitoring data. Journal of Sound and Vibration, 571(2024), 118017.

Firus, A. (2023) A contribution to moving force identification in bridge dynamics. Wiesbaden, Germany, Springer Vieweg.

Flaga, K. (2000) Advances in materials applied in civil engineering. Journal of Materials Processing Technology, 106(1-3), pp. 173-183.

Fritz, H., Peralta, J., Legatiuk, D., Steiner, M., Dragos, K., & Smarsly, K. (2022) Fault diagnosis in structural health monitoring systems using signal processing and machine learning techniques. In: Cury, A., Ribeiro, D., Ubertini, F., Todd, M. D. (eds.). Structural health monitoring based on data science techniques. Pp. 143-164. Cham, Switzerland: Springer.

Kingma, D.P. & Ba, J. (2015) Adam: A method for stochastic optimization. In: 3rd International Conference on Learning Representations, San Diego, CA, USA, 05/07/2015.

Meher, S., Parida, S., & Behera, R.K. (2019) On the response of a beam structure to a moving mass using Green's function. In: 1st International Conference on Applied Mechanical Engineering Researh. Waragnal, India, 05/02/2019.

Microstrain Sensing (2020) Microstrain Sensing product datasheet: G-Link-200. Williston, VT, USA: Parker Hannifin Corp.

Pi, Y. & Ouyang, H. (2016) Vibration control of beams subjected to a moving mass using a successively combined control method. Applied Mathematical Modeling, 40(2016), pp. 4002-4015.

Qian, S. & Chen, D. (1993) Discrete Gabor transform. IEEE Transactions on Signal Processing, 41(7), pp. 2429-2438.

Siddiqui S.A., Golnaraghi, M.F., & Heppler, G.R. (1998) Dynamics of a flexible cantilever beam carrying a moving mass. Nonlinear Dynamics, 15(1998), pp. 137-154.

Zhan, Y., Au, F.T.K., & Zhang, J. (2021) Bridge identification and damage detection using contact point response difference of moving vehicle. Structural Control Health Monitoring, 28(12), e2837.

Zhang, Z., Mehmood, A., Shu, L., Huo, Z., Zhang, Y., & Mukherjee, M. (2018) A survey on fault diagnosis in wireless sensor networks. IEEE Access, 6(2018), pp. 11349-11364.

Zhang, Z., Yuvaraj, A., Di, J., & Qian, S. (2019) Matrix design of light weight, high strength, high ductility ECC. Construction and Building Materials, 210(2019), pp. 188-197.

AUTOMATIC EXTRACTION OF BUILDING FEATURES FOR BUILDING FAÇADES BASED ON LASER SCANNING TECHNOLOGY

Mudan Wang[1], Ya Wen[1], Yuandong Pan[1], Hamidreza Alavi[1]
Soheila Kookalani[1], Weiwei Chen[2], and Ioannis Brilakis[1]
[1]University of Cambridge, Cambridge, United Kingdom
[2]University College London, London, United Kingdom

Abstract

Feature extraction on façades from unstructured point clouds is a challenging work, especially in the presence of noise. Previous studies in reconstructing the building façade are limited to traditional urban buildings. However, facade with a big proportion of windows, like curtain wall, are less explored. This research aims to propose an automatic approach to extract building features of a facade with a big proportion of windows, such as curtain wall. This approach includes automated detection of the facades, openings detection and geometry information extraction. A real case is used for testing the framework and validating the effectiveness of the approach.

Introduction

Building façades are highly related to energy consumption. Energy-efficient retrofitting of existing buildings is a key aspect of reaching the energy consumption reduction targets fixed by public authorities in different countries (Previtali et al., 2014; Sun et al., 2019). Europe has numerous historic and old buildings that face challenges, such as energy loss and higher operational costs, and need to be more energy efficient (Massafra et al., 2022). The geometry information of facades is crucial for the energy simulation.

Recent developments in generating digital twins (DTs) of buildings and point cloud techniques have the potential to obtain the geometry of building façades and improve its efficiency and accuracy (Pan et al., 2023). Geometric building models can be created based on point cloud acquisition and processing techniques (Pan et al., 2022; Wang et al., 2020). However, obtaining 3D information of façades from point cloud of existing buildings is still a challenging and costly process (Wang et al., 2022).

Previous studies in reconstructing the building façade and their openings are limited to specific types of buildings, mainly urban traditional buildings. These building facades are typically composed of concrete or brick blocks with low reflectance materials. Building façades with high reflectance materials were not considered in previous studies.

This research aims to propose an automatic approach to extract building features of a facade with a big proportion of windows, such as curtain wall. The proposed framework includes automated detection of the facades and glass windows from the building's point cloud using emerging image processing technologies techniques. Specifically, the geometric information of the recognised windows could be automatically extracted from the point clouds. Compared with previous approaches, the approach proposed in this research could be more feasible for building façades with a big proportion of windows, such as curtain wall and the outcomes could be directly employed for the other applications requiring the geometry information of façade features, such as building retrofitting and energy simulation. A laser scanning point cloud dataset of the Civil Engineering Building in the University of Cambridge is used for validation, which is an appropriate case with curtain wall and a large number of windows on the façade and the results demonstrate the effectiveness of the proposed framework.

Literature review

The reconstruction of geometric models of building façades from point cloud data has been explored by many researchers. For example, Truong-Hong et al. (2012)introduced an algorithm designed to identify building boundaries and features, employing a method that transforms point cloud data into a robust solid model through the implementation of voxels within an octree representation. This approach was deployed on the façades of three masonry buildings, exhibiting proficiency in detecting all apertures and accurately reconstructing the façade boundaries. Previtali et al. (2014) developed an automated methodology to derive highly detailed 3D vector models of existing building facades from terrestrial laser scanning data. The methodology involves segmenting the point cloud of a building facade into its planar elements and then generating 3D vector models based on the facade break lines.

The openings, such as windows and doors are always used for façade segmentation. Pu and Vosselman (2009) proposed a knowledge-based approach for façade detection, utilising walls, roofs, protrusions, and doors as features for identifying façade elements. The geometric reconstruction was viewed as a process involving polygon fitting, along with the generation of knowledge-based assumptions to address occluded parts in the proposed approach. Zolanvari et al. (2018) introduced the improved slicing method for detecting opening boundaries, including those on roofs, such as chimneys, as well as determining a building's overall outer boundaries through a local density analysis technique.

The plane fitting and clustering methods, as well as the use of a projection method to simplify a 3D scenario into a series of 2D ones, are suitable for relatively simple structures. However, when dealing with a diverse range of existing building types, such as complex architectural buildings or multi-planar building façades, these approaches tend to lack robustness in façade segmentation. For example, Hamid-Lakzaeian (2020) developed a multi-planar algorithm to detect the principal façade across the entirety of multi-planar masonry building facades. The algorithm further identifies recessed and protruded sections within each portion of the principal façade.

To identify the actual geometries for non-rectilinear openings and extract complicated facade boundaries, Iman Zolanvari and Laefer (2016) introduced the slicing method for extracting overall façade and window boundary points, enabling the reconstruction of a façade into a geometry compatible for computational modelling. It detected free-form openings and overall boundaries in building façades, even when faced with complexities and out-of-plane protrusions. Hamid-Lakzaeian (2019) proposed the Gridded-RANSAC approach to segment point cloud façade, which is particularly applicable to highly ornamental masonry buildings with non-rectilinear openings. Xia and Wang (2019) proposed a facade separation method capable of dividing connected facades into distinct building instances. This method can effectively detect complex windows, including bay windows, and identify a subset of individual façades based on dividing lines.

In addition to making reconstructions using TLS data, some studies have proposed reconstructing building façades from images. Zhao et al. (2021) outlined a BIM-based image management framework for reconstructing as-is building 3D models from fragmented images. This is achieved by registering and aligning UAV images onto BIM, matching building façade components in the BIM model with those in the UAV images. This approach primarily focuses on flat building façades with windows and doors, rather than emphasizing non-flat facades or buildings without façade components. Murtiyoso et al. (2021) proposed an approach to semantically segment building façades using deep learning, deploying the DeepLabv3+ network and training it on a database of labelled building façade images. The proposed approach facilitates the transition between 2D orthoimages and 3D point clouds. Bacharidis et al. (2020) reconstructed building façades in urban environments by establishing and leveraging a relationship between stereoscopic images and tacheometry data. This approach involves combining image and georeferenced data to extract meaningful attributes for the structural elements of the façade, including positional, appearance, and depth-related features.

Some studies have explored combinations of various techniques or features extracted from point cloud data for façade detection. By leveraging the advantages of 3D point cloud and 2D optical images, Wang et al. (2018) described a building facade feature extraction method for extracting building facade features, mainly structural information, from 3D point cloud. The proposed method involves image feature extraction, exploring the mapping method between image features and 3D point cloud data, and optimising the initial 3D point cloud facade features while considering structural information. Macher et al. (2021) investigated the use of radiometric information, specifically colour and intensity, for segmenting windows from façade point clouds. Jarząbek-Rychard et al. (2020) proposed a supervised approach for extracting façade openings, specifically windows and doors, from photogrammetric 3D point clouds, utilising both RGB and thermal infrared information.

In summary, previous studies discussed various approaches for building façade reconstruction, including deep learning for semantic segmentation in point clouds, slicing and clustering methods, supervised extraction of façade openings utilising RGB and thermal infrared data, and extracting façade features by combining 3D point cloud data with 2D images. However, these studies are still limited to building façades with rectilinear and repetitive openings or façades with multi-planar surfaces and openings in complex shapes, which are mainly urban traditional buildings. These building facades are typically composed of concrete or brick blocks with low reflectance materials. There is a lack of studies focused on building façades with a big proportion of windows, such as curtain wall, and no studies have been conducted to further extract the geometry information of the recognised windows in an automatic manner.

Research methodology

To enrich the knowledge of building façade reconstruction from high-reflectance materials, this research proposed a digital twin-enabled framework for automatically generating the geometric information for building façades. The developed modelling methodology is applicable to unstructured point clouds containing tens of millions of points. Each point is parameterised by its spatial coordinates and may also include related attributes, such as intensity or colour, but it does not share any topological relationships with nearby points. In the proposed approach, the scan registration process is conducted in the CloudCompare software, and the rest steps are implemented in Python to process the point cloud data.

The input point cloud can be generated by a single or multiple laser scan. After scan registration, scans are merged into a specific data structure. The overall procedure is illustrated in Figure 1. Once all scans are collected and registered in the CloudCompare software, it is essential to determine the windows from the façade. As mentioned earlier, the glass window is high-reflective material, therefore it is difficult to be fully scanned by laser scanners. In this case, the points reflected by the windows could be sparse and random. To detect the windows from the point cloud, it is essential to remove the points of glass as these points will influence the object detection of windows and need to be regarded as noises. A filter process is developed to identify the points reflected by windows. Usually, the intensity feature can be applied for the detection of this glass window on the façade (Macher et al., 2021). The filtering

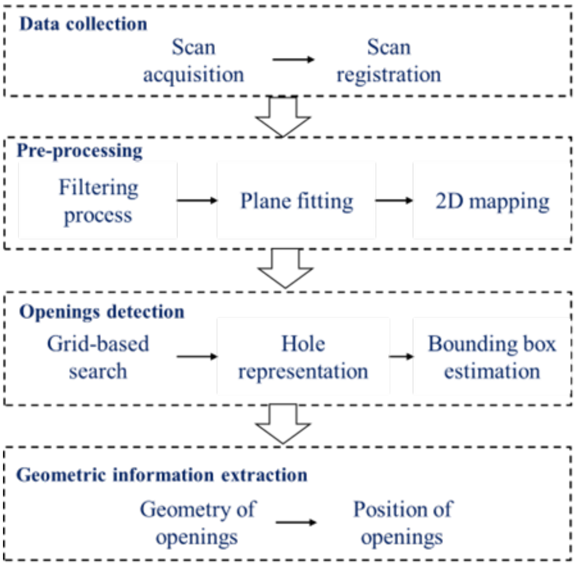


Figure 1: Pipeline of the facade reconstruction process

process is developed based on the intensity feature of the point cloud and implemented in Python.

After the removal of the glass window points, the next critical step is to detect the points belonging to the windows. The main elements constituting the façade are firstly identified by standard random sample consensus (RANSAC) algorithm. The RANSAC algorithm is an iterative method to estimate parameters of a mathematical model from point cloud data (Chum et al., 2003; Wang et al., 2021). The main plane of the façade can be computed based on RANSAC and this plane will represent the façade surface. At this stage, the point cloud data is still in 3D level. To efficiently detect the openings, including windows and doors, from the 3D point cloud data, this study adopts the mapping approach to project the 3D point cloud data onto the façade surface, which is a 2D point cloud data to identify the openings.

To detect the openings from 2D point cloud data, grid-based search for openings is performed and implemented in Python. The size of the grid needs to be determined. If the size is too big, details of the openings' boundary will be lost in the discretization and the accuracy of the dimensions could be decreased. If the grid size is too small, the grid loses its advantages over the point cloud. In general, the size will be determined by the resolution of the point cloud data and the application's requirements, which is set as 3 cm in this research.

A grid representation of point clouds for detecting hole boundaries is commonly used in previous research (Nguyen et al., 2012). The flowchart of the grid-based search approach for the openings detection is shown in Figure 2. A 3 cm $\times$ 3 cm grid is applied on the 2D point cloud to cover the whole data. Some areas may not be covered by points as they are located in the openings such as windows, therefore, the point cloud may include some holes which are regions with no points. Then, each grid is checked that

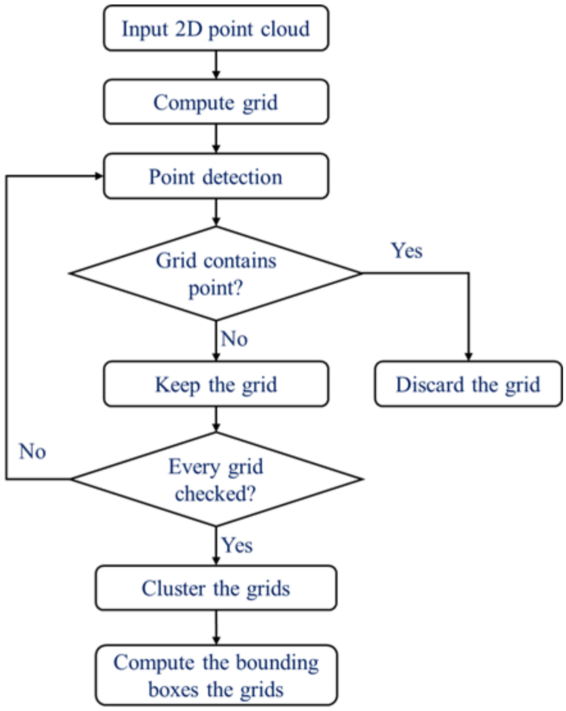


Figure 2: Flowchart of openings detection process

whether it contains point or not. If a grid has no point inside, the grid will be kept as the openings. After all grids have been checked, the openings from the façade will be detected.

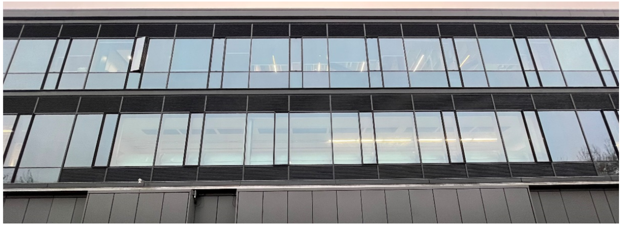
Finally, the remaining grids are clustered, and each cluster represents one opening on the façade. A convex oriented bounding box (OBB) is computed and considered as the boundary of each opening. In this case, a primitive geometry of rectangles is fitted to the opening. The height and length of the opening are be calculated by their minimum and maximum x and y coordinates. Therefore, the geometry information of openings can be extracted from point cloud data and this step is implemented in Python.

The proposed approach is capable of detecting the facades and the details, such as windows. Moreover, this research is also suitable for building façades with high reflectance materials. A real building with window façades is employed in this study to demonstrate the effectiveness of the proposed method.

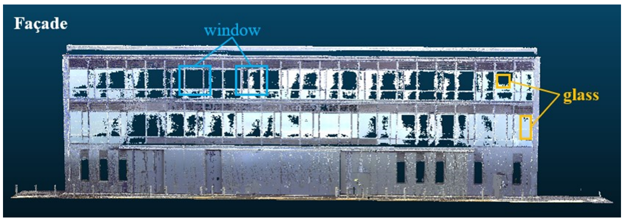
Results and discussion

Case background

The terrestrial laser scanner Faro was adopted for data acquisition in this study. After the scan registration, the building façade of university building was collected. To test the proposed framework, this research cropped one piece of façade for the experiments. The picture of the building in Figure 3 demonstrates the façade is composed of glass windows. From the point cloud, it indicates that the glass windows were scanned, but few points lie on the glass because of the high reflection.



(a) Picture of the building façade



(b) Point cloud of the building facade

Figure 3: Building façade of case study

Experiment results

The filter process is adopted to remove these noise points. As proposed in methodology section, the filtering of points of glasses are based on the intensity feature of the point cloud. The scalar field image of the point cloud intensity is shown in Figure 4. The points with intensity values out of the range will be displayed in grey. In Figure 4, the range of the intensity values is from 158 to 65,535. It is obvious that the intensity values of the glass in the windows are below 158, which can be used as the threshold for filtering the noises in the windows.

The point cloud after the filtering process is given in Figure 5. It demonstrates that the filtering process has effectively remove the points of glasses in the point cloud. Based on the comparison of the point clouds of the original one and the one after filtering process, most of the points, which are appeared grey in Figure 4 have been removed. The Figure 5 shows a cleaner point cloud without the influence of the high-reflection materials on the façade. The results proved that the intensity feature of the point cloud could be capable of removing the noises caused by high-reflection materials of building's façade, such as glasses.

As proposed in the Figure 1, the cleaned point cloud was fitted with an optimal plane based on RANSAC algorithm. The plane that represents the façade surface has been identified. Therefore, the points on the façade were projected on the plane. Therefore, a 2D point cloud was generated. This 2D plane was used to detect the edge of the window frames on the façade surface. As proposed in the Figure 2, the 2D point cloud will be processed with the openings detection process to find the windows from the point cloud.

After applying the openings detection process, the openings were detected from the 2D point cloud, as shown in the Figure 6. It is obvious that all windows can be successfully detected. As given in the Figure 6, the openings are represented as clusters in different colours. By comparison with the original point cloud data in Figure 3, most of the glass windows were detected despite the high-reflection issue caused by the façade. Among them, for the second floor, all the windows were successfully identified and separated. For the first floor, on the left side of the façade, all the windows were also segmented and separated. In regard of the right side of the windows, the lower smaller windows were not successfully separated, and this is because the original point cloud was not scanned completely, and the window frames of some smaller windows were not scanned. Thus, it is challenging to separate these small windows with the detection process. This situation is caused by the quality of the scanned data.

After the detection process, the final step is to extract the geometry information of these openings as given in Figure 2. The bounding boxing method was adopted to estimate the boundaries of these openings, so that each cluster can be represented as a rectangular box. The result of the bounding box estimation is given in Figure 7(a). After that, the geometry information of each bounding box can be used to estimate the geometry information of the windows, as seen in Figure 7(b). The coordinates of the four vertices of the bounding box were computed, thus, the location, width and length of each window were obtained. In summary, with the proposed framework, the geometry information of the openings in the façade can be extracted automatically.

Discussions

According to experiments using real case point cloud data, the effectiveness of the proposed framework was validated above. In contrast to prior studies centred on general facades, our research pioneers a novel method for detecting facades with high-reflectance materials. The filtering process successfully removed points related to glass by utilising the intensity feature, allowing for the identification of openings. The results demonstrated the successful detection of glass windows from the point cloud, with approximately 94% of openings detected correctly.

The grid size used in this case is 3 cm for the openings detection, as defined in the method section. However, based on the alignment between window clusters and the point cloud before the detection process, some variations were observed among window boundaries and the facade. This might reduce the accuracy of the extraction of the geometry information. Therefore, it is suggested that the grid size used might be too large for accurate detection. For more precise geometry information generation, the grid size needs to be set smaller, or further experiments need to be conducted to determine the optimal grid size.

According to the Figure 7(b), it indicated that some small windows were merged as one window. However, it is because that they were not completed separated. A filtering algorithm could be developed to filter these windows

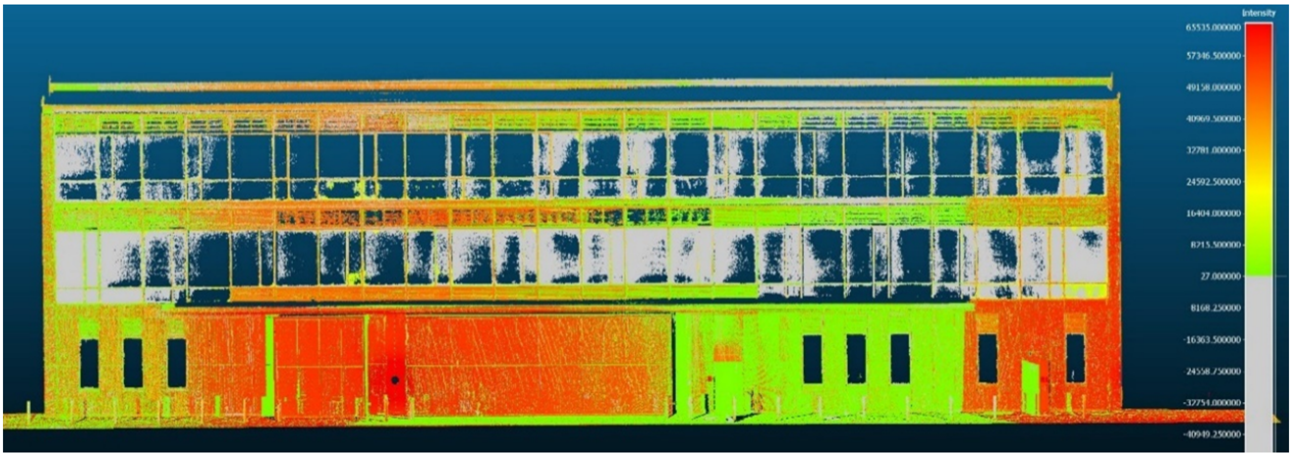


Figure 4: Scalar field of point cloud intensity

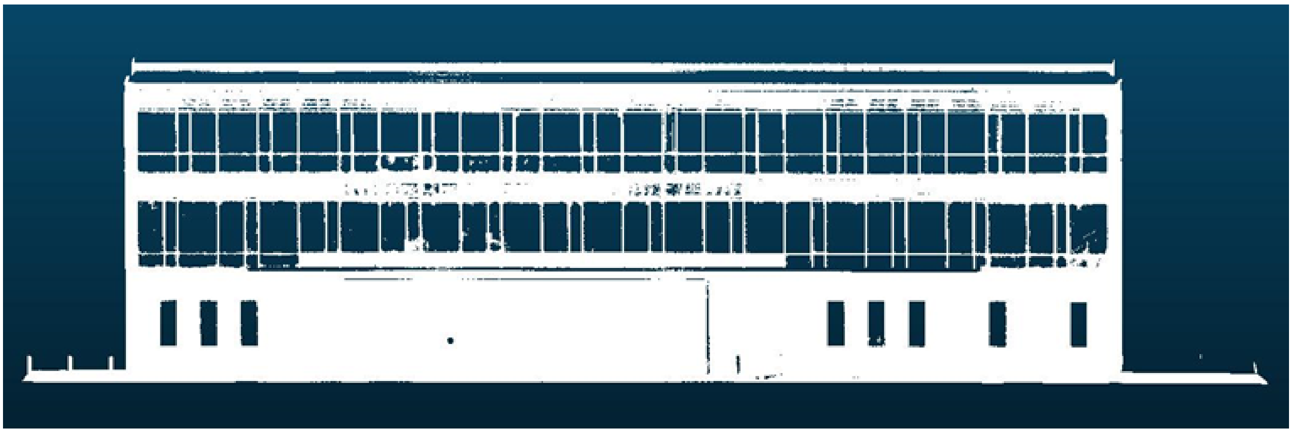


Figure 5: Point cloud after noise filtering process

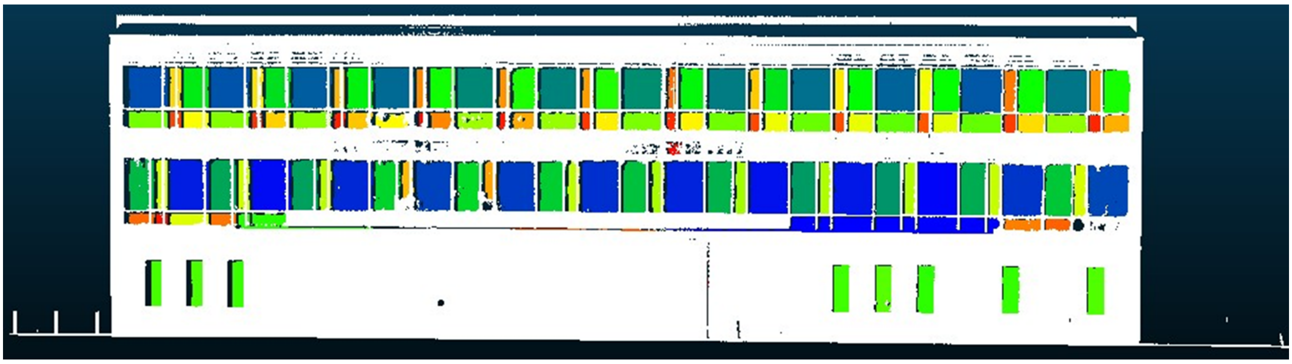


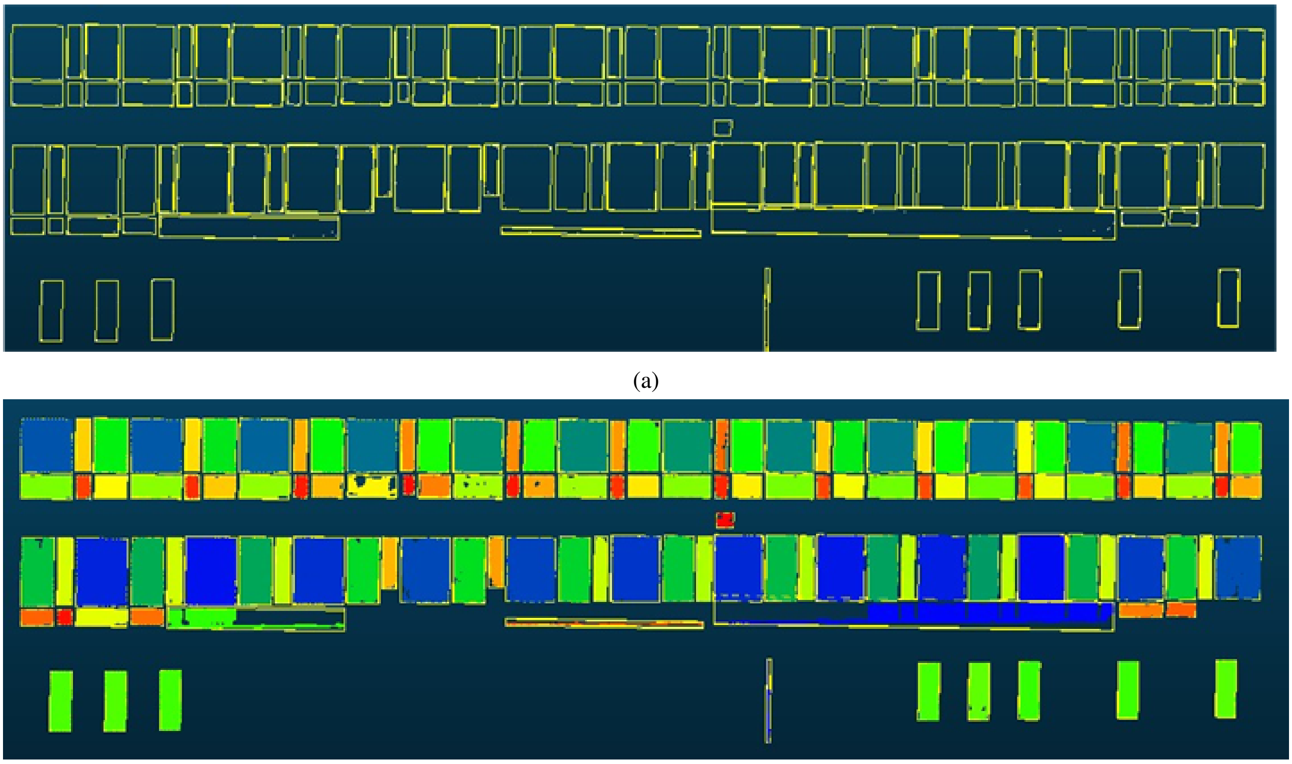
Figure 6: Point cloud after openings detection process

which need to be recognised as errors. Based on prior knowledge of building, some limits could be set for the width and length of the windows that can be used as criteria for filtering the errors, therefore improve the accuracy of the detection.

Regarding geometry information, this research assumes that windows are regular rectangular shapes, which aligns with actual practices in real cases. The extracted geometry information can be directly applied to future building energy simulations. This study opens the opportunity to automate the process from the point cloud to energy simulation, and further investigations could be conducted for this purpose.

Conclusions

The main contribution of this study is the development of the approach to extract building features of a facade with a big proportion of windows, such as curtain wall. The framework includes automated detection of the facades and glass windows from the building's point cloud. Specifically, the geometric information of the recognised windows could be automatically extracted from the point clouds. The proposed framework is validated in an existing building, and the results demonstrate its effectiveness. The



(a)

(b)

Figure 7: Bounding box estimation of the openings: (a) bounding boxes (b) bounding boxes and windows

outcomes could be directly employed for the other applications requiring the geometry information of façade features, such as building retrofitting and energy simulation.

Additionally, this research has the potential to enrich a geometric digital twin of the façade with details. This automatic approach enhances the accuracy and efficiency of the point cloud process. The limitation of this research is that the proposed approach was validated only on one case study. More types of façades could be also included to validate the effectiveness of the proposed approach. Future work in this area includes the extension of this technique to capture an entire building (exterior and interior objects) and to generate more complicated structures and building elements such as roofs, free-form surfaces, internal objects (e.g. ceiling, internal walls), shading elements, façades of non-terraced buildings, and skyscrapers.

Acknowledgments

This work was supported by funding from the Horizon Europe UKRI Underwrite Innovate, under the grants G115676, G115919, the Marie Sklodowska-Curie grant agreement No 101034337, and EPSRC via grant number EP/W018705/1. The authors gratefully acknowledge the collaboration of all academic and industrial project partners. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the institutes mentioned above.

References

Bacharidis, K., Sarri, F., and Ragia, L. (2020). 3d building façade reconstruction using deep learning. ISPRS International Journal of Geo-Information, 9(5):322.

Chum, O., Matas, J., and Kittler, J. (2003). Locally optimized ransac. In Pattern Recognition: 25th DAGM Symposium, Magdeburg, Germany, September 10-12, 2003. Proceedings 25, pages 236–243. Springer.

Hamid-Lakzaeian, F. (2019). Structural-based point cloud segmentation of highly ornate building façades for computational modelling. Automation in Construction, 108:102892.

Hamid-Lakzaeian, F. (2020). Point cloud segmentation and classification of structural elements in multi-planar masonry building facades. Automation in Construction, 118:103232.

Iman Zolanvari, S. M. and Laefer, D. F. (2016). Slicing method for curved façade and window extraction from point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 119:334–346.

Jarząbek-Rychard, M., Lin, D., and Maas, H.-G. (2020). Supervised detection of façade openings in 3d point clouds with thermal attributes. Remote Sensing, 12(3):543.

Macher, H., Roy, L., and Landes, T. (2021). Automation of windows detection from geometric and radiometric information of point clouds in a scan-to-bim process. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2021:193–200.

Massafra, A., Predari, G., and Gulli, R. (2022). Towards digital twin driven cultural heritage management: A hbim-based workflow for energy improvement of modern buildings. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 46:149–157.

Murtiyoso, A., Lhenry, C., Landes, T., Grussenmeyer, P., and Alby, E. (2021). Semantic segmentation for building faÇade 3d point cloud from 2d orthophoto images using transfer learning. Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLIII-B2-2021:201–206.

Nguyen, V.-S., Bac, A., and Daniel, M. (2012). Boundary extraction and simplification of a surface defined by a sparse 3d volume. In Proceedings of the 3rd Symposium on Information and Communication Technology, pages 115–124.

Pan, Y., Braun, A., Borrmann, A., and Brilakis, I. (2023). 3d deep-learning-enhanced void-growing approach in creating geometric digital twins of buildings. Proceedings of the Institution of Civil Engineers - Smart Infrastructure and Construction, 176(1):24–40.

Pan, Y., Braun, A., Brilakis, I., and Borrmann, A. (2022). Enriching geometric digital twins of buildings with small objects by fusing laser scanning and ai-based image recognition. Automation in Construction, 140:104375.

Previtali, M., Barazzetti, L., Brumana, R., Cuca, B., Oreni, D., Roncoroni, F., and Scaioni, M. (2014). Automatic façade modelling using point cloud data for energy-efficient retrofitting. Applied Geomatics, 6(2):95–113.

Pu, S. and Vosselman, G. (2009). Knowledge based reconstruction of building models from terrestrial laser scanning data. ISPRS Journal of Photogrammetry and Remote Sensing, 64(6):575–584.

Sun, J., Xiong, X., Wang, M., Du, H., Li, J., Zhou, D., and Zuo, J. (2019). Microalgae biodiesel production in china: A preliminary economic analysis. Renewable and Sustainable Energy Reviews, 104:296–306.

Truong-Hong, L., Laefer, D. F., Hinks, T., and Carr, H. (2012). Flying voxel method with delaunay triangulation criterion for façade/feature detection for computation. Journal of Computing in Civil Engineering, 26(6):691–707.

Wang, M., Wang, C. C., Sepasgozar, S., and Zlatanova, S. (2020). A systematic review of digital technology adoption in off-site construction: Current status and future direction towards industry 4.0. Buildings, 10(11).

Wang, M., Wang, C. C., Sepasgozar, S., and Zlatanova, S. (2022). A time efficient quality check method based on laser scanning for installation of prefabricated wall panels. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, X-4/W2-2022:273–280.

Wang, M., Wang, C. C., Zlatanova, S., Sepasgozar, S., and Aleksandrov, M. (2021). Onsite quality check for installation of prefabricated wall panels using laser scanning. Buildings, 11(9):412.

Wang, Y., Ma, Y., Zhu, A. x., Zhao, H., and Liao, L. (2018). Accurate facade feature extraction method for buildings from three-dimensional point cloud data considering structural information. ISPRS Journal of Photogrammetry and Remote Sensing, 139:146–153.

Xia, S. and Wang, R. (2019). Façade separation in ground-based lidar point clouds based on edges and windows. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(3):1041–1052.

Zhao, L., Zhang, H., Wang, Q., and Wang, H. (2021). Digital-twin-based evaluation of nearly zero-energy building for existing buildings based on scan-to-bim. Advances in Civil Engineering, 2021:6638897.

Zolanvari, S. M. I., Laefer, D. F., and Natanzi, A. S. (2018). Three-dimensional building façade segmentation and opening area detection from point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 143:134–149.

DIGITAL TWIN FOR SAFETY ON CONSTRUCTION SITE: A REAL TIME-TIME RISK MONITORING SYSTEM COMINING WEARABLE SENSORS AND 4D BIM

Wahib Saif [1,2], Tommy Williams[3], Charence Wong[4], Jozef Dobos[4], Pablo Martinez[1], and Mohamad Kassem[5] [1]Northumbria University, Newcastle, UK
[2]BIM Academy, Newcastle, UK
[3]Plinx, UK; [4]Asite 3D Repo, UK
[5]Newcastle University, Newcastle, UK

Abstract

This paper presents the outcomes of an innovative safety project integrating Health and Safety standards, Building Information Modelling (BIM) and wearable sensors to address safety on construction sites. The paper presents a framework and a prototype that integrates 4D BIM and a real-time tracking sensing system. The 4D BIM virtually identifies and collects the data of hazard zones which is synchronized with the tracking-based sensing system and is used to monitor the access to and interactions with risky zones on construction sites. Real-time alerts are automatically triggered in case of unauthorised trespassing to proactively prevent accidents. The collected monitoring data are then visualised within the 4D BIM model, enabling stakeholders to identify trends and evaluate the effectiveness of safety measures. The integrated system was validated in a construction site and showcased a shift to proactive safety measures, improving hazard identification and control.

Introduction

Construction sites are widely considered one of the most hazardous work environments due to their constantly evolving and unpredictable nature, encompassing high-risk activities, and the frequent interaction between workers and machinery (Jin et al., 2020; Kulinan et al., 2024) The latest report from the Health and Safety Executive (HSE, 2023) documented a total of 135 fatal injuries in work-related accidents in Great Britain in the year 2022/23, 45 of which were within the construction sector, making it by far the leading sector in fatal accidents. The report also highlighted a total of 53,000 non-fatal injuries within the construction sector for the same year. One of the major causes of injuries and fatal accidents in construction sites is "unauthorised incursion" (Jin et al., 2020; Shuang et al., 2019). Unauthorised incursion refers to the entry or presence of an onsite person in an identified hazard zone without a permit, reflecting his role or competence relative to that zone. According to HSE, many contractors in the UK address this issue by employing traditional engineering controls from the hierarchy of controls, Figure 1, by providing workers with Personal Protective Equipment (PPE) and installing physical barriers (e.g., safety bollards, tapes, barricades, toe boards, signs, cones, etc.) to protect and isolate workers from hazards. Although traditional barriers can mitigate some of the risks on site, they are associated with several limitations. These include labourious and time-consuming installation, vulnerability to trespassing or destruction, and a lack of proactive and early warning mechanisms (Ding et al., 2022; Kulinan et al., 2024). Hence, over the last two decades, many researchers and industry practitioners have sought different ways to address such limitations and improve safety measures by proposing various digital technologies which is referred to in this study as an innovative measure, i.e., digital control, to the Hierarchy of Controls, (see Figure 1). These digital technologies include Building Information Modelling (BIM), IoT devices, extended reality, wearable and embedded sensors, and robotics for identifying and controlling onsite hazards.

These emerging technologies are integrated into construction safety management owing to their capabilities for automated hazard identification, real-time monitoring, proactive warnings, and data-driven decision-making. Although technologies in general contributed significantly to enhancing safety measures in construction sites, it remains difficult to meet high safety standards especially when these technologies are solely employed or not incorporated with safety standards and protocols (Jin et al., 2020; Wang et al., 2022). Therefore, this paper introduces an integrated safety system designed to overcome previous limitations, aiming to identify and control hazard zones, fostering mutual interaction between the virtual design model and the physical site.

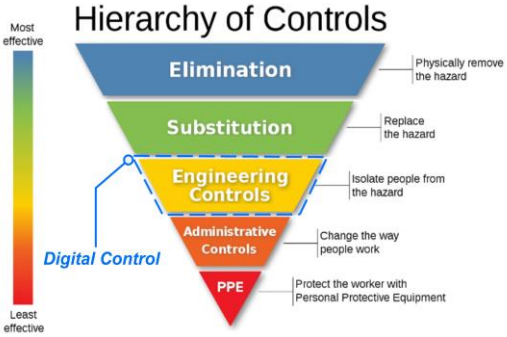


Figure 1. The Hierarchy of Controls Framework; with Digital Control integrated into the Engineering Control layer.

Related Studies

Various digitalization-based safety systems have been developed and introduced to enhance safety measures and prevent incidents in construction sites by identifying, monitoring, and controlling onsite hazards. Table 1 presents a summary of related studies and characterization of their safety systems to better highlight the innovation and contribution of our system. The characterization examines the studies based on several criteria that examine the level of automation, effectiveness, and responsiveness of the presented safety systems. The characterization examines the studies based on the below criteria that inspect the level of automation, effectiveness, and responsiveness of the presented safety systems.

- *Hazard identification mechanism* that investigates the type of technologies used to identify and highlight safety hazards whether during the design in virtual models or during construction on site. It also checks whether the project programme was considered as it is key to include the time dimension and the activities sequence when identifying hazards to be monitored. For instance, when identifying a hazard zone, its associated activities, and their start/end times should be available for the monitoring system to be activated/deactivated based on such data.

  In general hazard identification technologies include BIM, extended reality, and computer vision. BIM has proven to be an effective tool in safety management during the design phase by enabling designers and planners to highlight safety hazards in the design models and propose mitigations. However, BIM independently lacks dynamic data from the construction site that should be streamed in real-time to accurately depict what is currently happening on site (Kulinan et al., 2024). Therefore, BIM by itself is not sufficient for managing safety and requires the integration of IoT and sensing devices to provide timely data during construction.
  Extended Reality (XR) technologies have also been widely adopted in safety management to provide an immersive and interactive environment to identify and visualise onsite risks (Wu et al., 2022). Their adoption has mainly been for providing safety training to construction workers where safety instructions and guidelines are displayed for identified hazards, enabling the worker to manage onsite hazards effectively and interactively.

  Other technologies have also been used to identify onsite hazards including computer vision-based approaches. Many researchers proposed vision-based safety systems that collect and process visual data to recognise onsite risky behaviours of construction workers or equipment. Despite their automated approach to identifying potential hazards, computer vision techniques are limited to certain safety scenarios and types of objects that were trained to detect and recognise (Kulinan et al., 2024). Furthermore, these systems require a clear and uninterrupted visual data stream which in many construction sites is not easy to ensure due to their dynamic nature with various moving parts.
- *Hazard monitoring mechanism*, looking into the types of technologies used to collect and process data from the construction site and whether this mechanism is fully automated or not. There is a wide array of different sensing and IoT devices adopted in the literature depending on the type of data that needs to be collected. For example, in most safety applications, location data of onsite personnel or equipment is mainly needed, and it is generally collected using position sensors (e.g., GPS). In other applications where the motion data of equipment is required, motion sensors (e.g., IMU) are mainly used. In some studies, visual sensors (e.g., CCTV cameras) are favoured to acquire images or video streams that can be used to monitor hazard zones.
- *Hazard controlling mechanism*, ensuring the safety system leverages real-time and automated controlling mechanism that alerts people from potential hazards and proactively prevents injuries or accidents. For instance, when an unauthorized incursion occurs, it is expected for the system to trigger a fully automated warning in real time to alert the unauthorized person as well as HSE managers. This criterion also investigates whether a system employs visualisations that report and present monitoring data and analysed metrics (e.g., zone interactions metrics) in a timely manner. It also examines whether a learning feedback loop is integrated into the safety system that ensures a continuous learning loop based on the monitoring information from previous reports to enhance the safety measures of future onsite hazards.
- *Integration of Design and construction measures* which checks whether a given study integrated the safety measures during the design phase with those applied during the construction phase and if there is a feedback loop from the construction to the design that can enhance the safety planning of future sites.

As illustrated in Table 1, there is a gap in integrating all these elements into a unified system able to identify, monitor, and control on-site hazards in an automated and proactive manner while ensuring a learning/improving feedback loop between the design and construction.

Table 1. A summary and characterisation criteria of key related studies

Study	Hazard identification			Hazard Monitoring		Hazard Control			Integrating design and construction measures	Application description
	4D BIM	XR	Other	IoT sensing	Automated data stream	Real-time alert/warning	data/metrics visualisation	Feedback loop		
(Jin et al., 2020)				●	●	●	●	●		Employing IoT-based locating and alarming system for unauthorised incursions to risky workspaces in a construction site.
(Hossain et al., 2023)	●			●	●	●	●			GPS-based mobile application for tracking and alerting workers approaching predetermined hazard zones.
(Wu et al., 2022)	●	●		●		●	●	●		Integrating deep learning and mixed reality to generate real-time visual warnings.
(Chung et al., 2023)				●	●	●		●		RFID-based system for monitoring and alerting onsite personnel
(Liang & Liu, 2022)	●			●		●		●		BIM and IoT technologies are proposed to develop a safety risk warning system for underground projects.
(Kulinan et al., 2024)	●			●	●		●			Integrating BIM and computer vision for real-time tracking of construction workers for safety directions.
(Ding et al., 2022)	●			●	●	●	●			IoT-based BIM safety system for identifying and monitoring energy hazards in petrochemical construction.
(Zhang et al., 2023)				●	●	●				RFID-based system for preventing hazard zone incursion through real-time warning.
(Jiang et al., 2021)	●			●	●	●	●	●		A cyber-physical system for identifying hazards in virtual models and monitoring them using IoT sensing and networking devices
(Hong & Teizer, 2024)			●	●			●	●		Using spatial-temporal data of construction workers to derive behavioural patterns and predict potential hazard zones.
(Tran et al., 2021)	●		●						●	Integrating BIM-based safety planning and image stitching for hazard zone identification in construction sites.
(Kojima et al., 2020)				●	●	●	●	●		Bluetooth-based monitoring and warning system to control human-machine interaction in tunnelling sites.
This study	●			●	●	●	●	●	●	Integrating H&S standards, BIM, and wearable sensors for identifying and controlling onsite hazards.

Research Methodology

This research project was conducted as part of the Discovering Safety initiatives delivered by the Health and Safety Executive (HSE) in the UK. The main aim was to integrate current Health and Safety (H&S) standards, BIM technology, and advanced wearable sensing systems for enhancing safety measures during both the design and construction phases while maintaining a learning/improvement feedback loop between the two.

To achieve this aim, a conceptualized integrated safety framework was developed that incorporates safety measures during the design and construction following the "plan, do, check, and act (PDCA)" followed by the HSE and according to H&S standards. This system was developed with continuous feeding input from key industrial stakeholders including project owners, designers, contractors, HSE representatives, and technology suppliers. The developed framework also presents the integration workflow between the two technologies (i.e., BIM and wearable sensors) that will be employed to identify and manage safety hazards during both phases. In this project, the SafetiBase system provided by 3D Repo was used to identify and capture hazards within the 4D BIM model while the sensing system was used to monitor and control the identified hazards. The developed framework was later implemented in a real construction site using both SafetiBase and Plinx platforms to validate its level of effectiveness and responsiveness to prevent onsite incidents and enhance safety measures.

Construction Safety Integrated Framework

Figure 2 presents the construction safety framework, integrating the BIM technology and the sensing system for managing safety during the design and construction phases. In the design phase, 4D BIM models play a key role in hazard identification within the construction sequence following the PAS 1192-6 standard, as illustrated in Figure 3a. Pre-approved treatment strategies following H&S standards and contractors' best practices are then placed to address those hazards. Where possible, identified hazards should be eliminated or substituted by changes to the building sequence following the hierarchy of controls (recall Figure 1). There are, however, lots of instances where this is not practical and engineering controls must be employed to ensure the safety of the working area during the construction phase.

For each identified hazard, a zone is marked virtually within the BIM model encompassing that hazard (Figure 3b). For each identified hazard, detailed information, including XYZ coordinates, project reference point, and the start and end times of associated activities, is gathered by the Plinx broker through 3D Repo's API. These XYZ coordinates are then converted to WGS84 Lat/Lon by the broker for compatibility within the Plinx system. Plinx effectively activates and deactivates the designated zones at the specified times. The Plinx platform hosts metadata from the risk treatment ticket, involving zone type, purpose, and additional information available in the treatment notes.

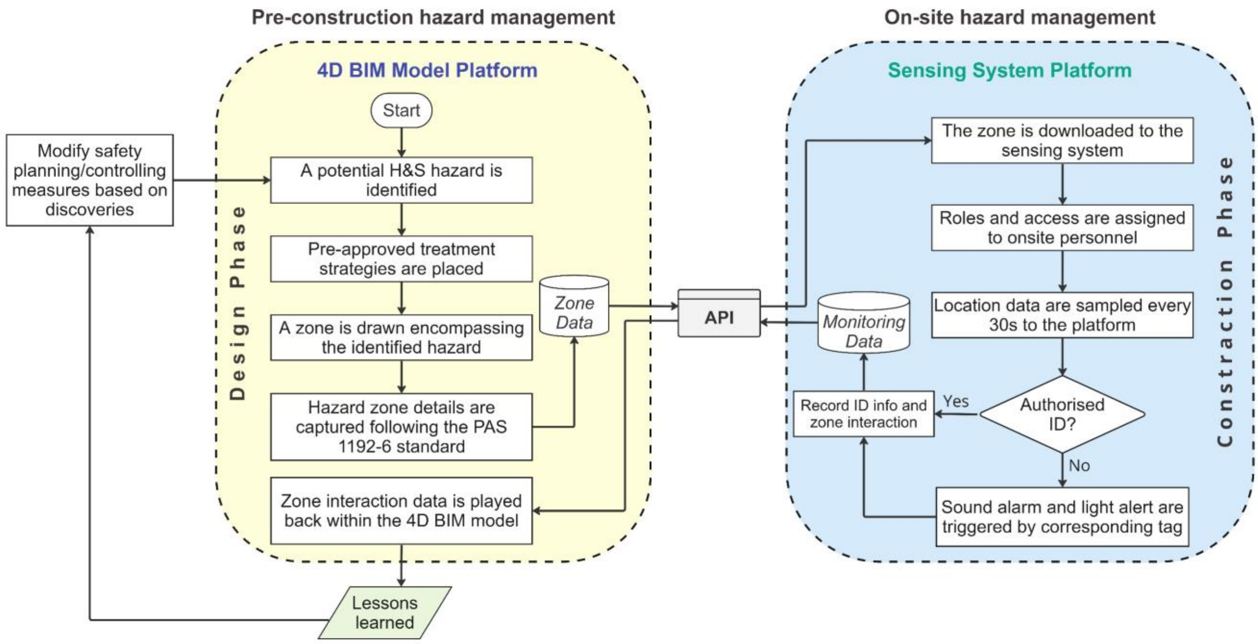
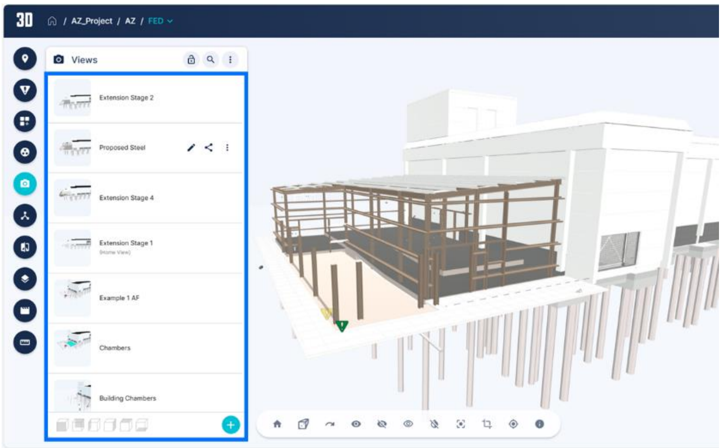
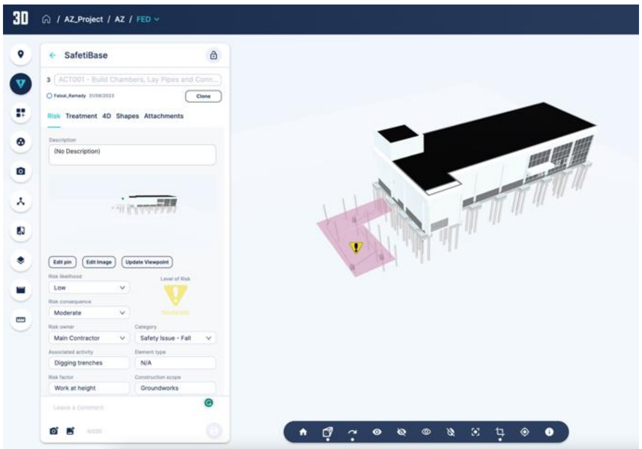


Figure 2. The proposed integrated safety system as a bi-platform management system



(a)



(b)

Figure 3. Identifying H&S risks virtually in a 4D BIM model using 3D Repo platform: (a) identified hazard activities within the construction sequence, (b) marking up a hazard zone associated with a drainage activity as an example.

Upon receiving zone data, Plinx establishes a connection with its corresponding physical wearable sensors. Before each construction shift, access permits are assigned to onsite personnel based on their roles and competencies, specifying permissions for entry and operation within the designated zone intended to be monitored. Unauthorised access triggers real-time alerts through the automatically activated sensors. The Plinx platform offers real-time visualisation of all monitoring data and zone interactions.

At the end of each shift, zone interaction data, both authorised and unauthorised entries, is promptly transmitted to 3D Repo via API. This includes location data recorded every 30 seconds, accessible through an API call typically conducted at the end of each working day. Plinx ensures the conversion of this data to XYZ coordinates before export. Within the 3D Repo environment, the movement and interactions within zones can be played back onto the BIM model, providing a contextual reference to the initially created zones. Doing so enables reflecting on the effectiveness of the current safety strategies and enhancing the safety planning and management of future hazard zones. This iterative process not only reinforces the practical implementation of integrated systems but also supports the overall efficacy of the H&S risk management framework.

Case Study

To validate and measure the effectiveness and responsiveness of the developed integrated safety framework, an on-site case study was implemented with the main intention of acquiring authentic, real-world data from a dynamic construction site. The on-site trial took place at AstraZeneca's quality assurance building project.

Hazard Zone Setup: AstraZeneca's 4D BIM model was imported into the SafetiBase platform for hazards to be identified and captured. Initially, four construction activities with potential risks were identified — footings excavation, drainage chamber installation, steel erection, and utility connection into the main road. However, given the compact site and the diverse expertise of the chosen subcontractors, establishing safety zones for all these activities based on roles and competencies posed a significant challenge. Consequently, only the steel erection task, which was executed by a specialised team, was selected to be monitored onsite (Figure 4). A safety zone was established around its locations with its hazard details captured in the BIM model following the PAS 1192-6 standard. The zone and its data were then extracted from SafetiBase by the API and sent to the Plinx platform to be monitored on-site during construction.

On the construction site, the WGS84 coordinates of the zone are precisely mapped using a survey rover, (Figure 5a). The survey rover outputs a GEOJSON file which is then uploaded directly to the Plinx platform to validate its alignment with the safety zone plotted in SafetiBase.

Figure 4. The safety zone selected around the steel erecting work to be monitored and controlled.

(a) **(b)** **(c)** **(d)**

Figure 5. The sensing system: (a) a GNSS Rover, (b) RFID card, (c) charging and allocation station, (d) wearable sensor connected to a safety helmet.

Sensing System Setup: The sensing system used to monitor the hazard zone was supplied by Plinx. The wearable sensors were connected to JSP construction helmets as illustrated in Figure 5d. Plinx RFID cards were used to assign operative roles and competencies of onsite personnel against their card IDs', see Figure 5b. Site operatives then scan these RFID cards to receive a sensor with personalised zone access based on their assigned roles relative to the zone.

Safety Zone Monitoring: Following the sensing system set-up, the safety zone around the steel erection work was monitored for one week. The wearable sensors and cards were assigned at the beginning of the working shifts. The wearable sensors transmitted the location data of onsite personnel every 30 seconds. These data points were visually represented on the platform as green dots, as illustrated in Figure 6. The monitoring system was intricately designed to generate alerts in the event of unauthorised individuals entering the designated zone. To assess the system's responsiveness and reaction time, deliberate unauthorised interactions were induced within specific areas of the monitored zone. Upon the entry of unauthorised individuals, their sensors emitted audible beeps and vibrations. The locations of these breaches were then visually depicted as dark blue dots (Figure 6).

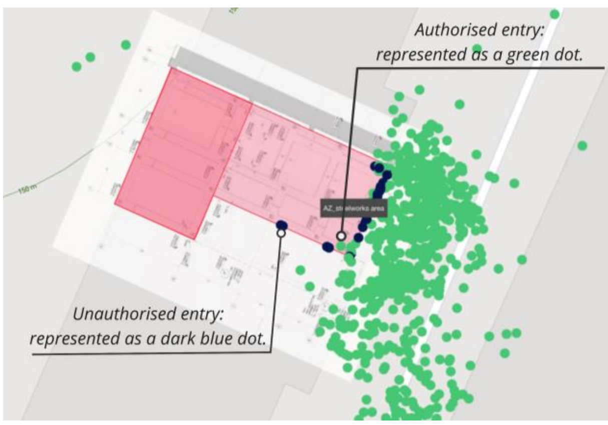
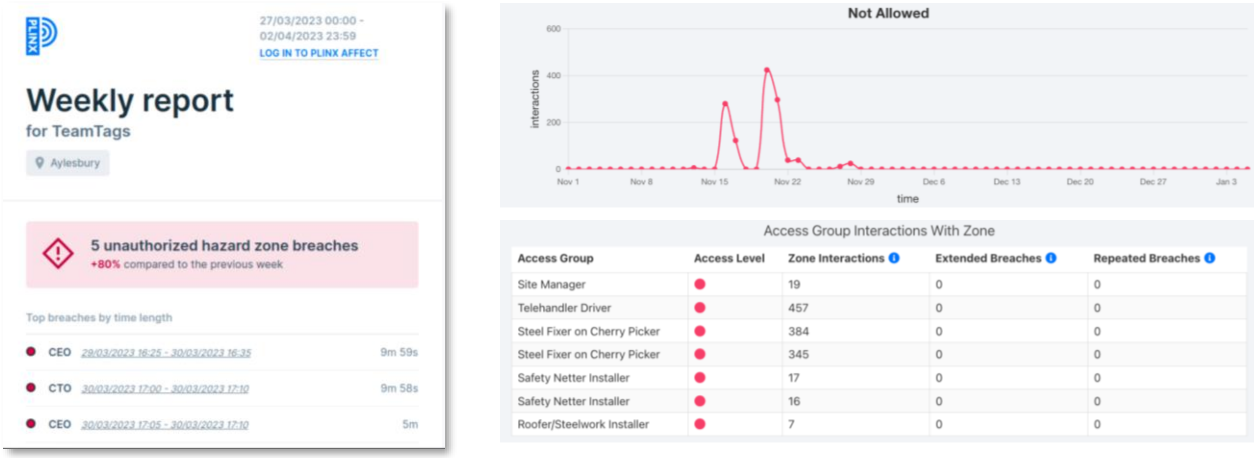


Figure 6. Zone interaction data for a week-period monitoring

All zone interaction data are recorded and presented in a comprehensive weekly report, see Figure 7. This report includes details such as the number of authorised/unauthorised entries and associated information, including personnel ID and role, date, and time of entry, as well as the duration of their presence within the zone. The gathered monitoring data are used to enhance week-on-week management strategies of safety zones by playing it within the 4D BIM model, as illustrated in Figure 8. More significantly, it plays a crucial role in providing insightful feedback to planners and designers, offering valuable insights into the efficiency of implemented preventive measures. This feedback loop ensures an ongoing enhancement of safety protocols, fostering a dynamic and responsive approach to maintaining and improving overall safety standards.

Results and Discussion

Trial Impact: This work has demonstrated the successful implementation of the developed safety framework integrating H&S standards with digital technologies for identifying and managing hazards on construction sites. The integration of digital technologies has significantly enhanced construction safety through proactive hazard identification and management. Unlike traditional reactive approaches, the digital control system, equipped with automated sensing and monitoring, predicts and prevents potential hazards in real-time. This ensures not only the effectiveness of safety measures but also responsiveness to dynamic site conditions. The system's proactive alerting further strengthens safety by triggering immediate responses to unauthorised individuals breaching safety zones, reducing the risk of accidents. The monitoring system has a detection accuracy within a 50 cm radius, significantly enhancing safety by triggering a warning when a worker approaches the zone at a proximity of just 50 cm at a walking speed of 4 mph. This level of precision surpasses the accuracy of systems documented in previous studies. For example, the system presented by (Jin et al., 2020) gets triggered only when a worker is within a 200 cm distance. Additionally, our system demonstrates real-time responsiveness, with an estimated average reaction time of just 0.3 seconds.



(a) **(b)**

Figure 7. Recorded unauthorised incursions to the zone: (a) an example of a weekly report summary, and (b) data visualisations of the steel work zone interactions over the monitoring period.

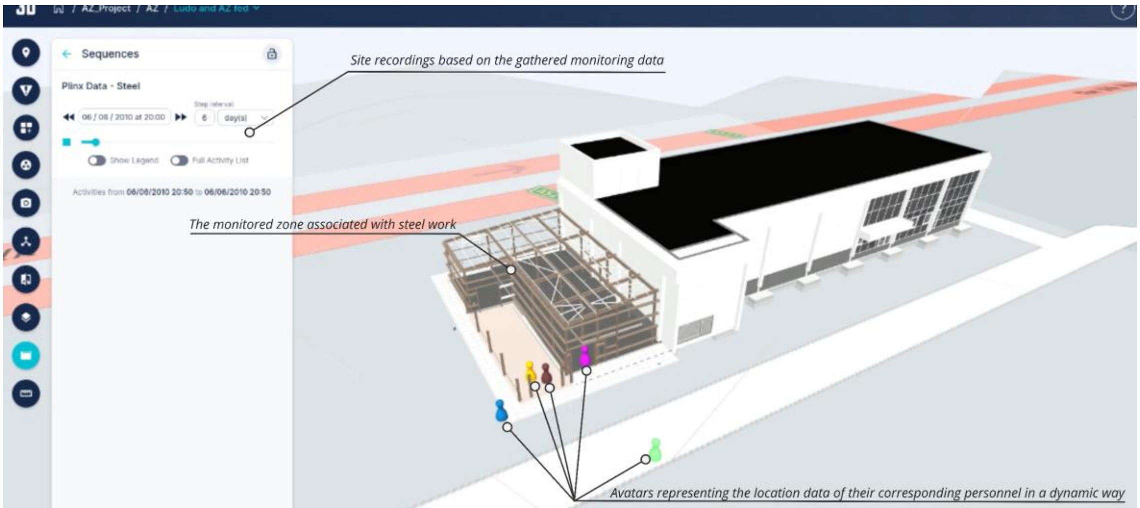


Figure 8. A screenshot from a video displaying the location data gathered during the zone monitoring within the 4D BIM model.

Moreover, the developed system extends data visualisation into the 4D BIM model, providing a comprehensive understanding of the construction site layout and potential hazards. This allows stakeholders to analyse safety data over time, identify trends, and make informed decisions for targeted improvements. The feedback loop in the system plays a crucial role in refining safety protocols by providing continuous, data-driven insights. Unlike traditional post-incident analysis, this system fosters a proactive safety culture, emphasizing ongoing refinement and optimization based on timely feedback.

Limitations and Future Recommendations: While this trial has shown promising advancements in safety management by integrating BIM and monitoring-based sensing, it is crucial to recognise the challenges and limitations faced during implementation. These insights are valuable for refining the system and optimizing its performance in larger-scale construction scenarios. One key challenge involved difficulties in obtaining and applying pre-approved treatment strategies for establishing and marking safety zones. This issue is widespread among contractors due to the absence of a standardized zonal control framework specifying attributes of a given zone based on the construction activity associated with it. The current process has a limitation where the shape drawn in SafetiBase is a freeform polygon. While this provides flexibility, it may lead to discrepancies between the digital zone and the physical segregation on site. A preferable solution would involve the system applying standardized rules to the highlighted hazard, assisting designers in comprehending the spatial impact and aiding the site team in deploying physical barriers more effectively.

It is vital to automatically align safety zones with planned activities to prevent sensor alerts from inaccurately triggering and misrepresenting actual locations. The plan involves developing a customised SafetyZone interface within 3D Repo to enable safety managers to authorise and restrict zone access to operatives, machines, or groups seamlessly as part of the system workflow. It is also worth noting that the deployment of sensing technologies, particularly monitoring personnel through wearable sensors and collecting location data, raises privacy and data security concerns. Achieving a careful balance between enhancing safety and respecting individual privacy rights is essential. Implementing robust data security measures is crucial for ethical and legal considerations, ensuring responsible use of the collected data and compliance with privacy regulations.

Conclusion

This work presented a construction safety framework that integrates H&S standards with digital technologies to enhance safety planning and management in construction sites. This innovative approach represents a shift from reactive to proactive safety measures. Utilizing BIM for virtual hazard identification and real-time monitoring through sensing technology, the integrated system predicts and prevents potential hazards, reducing the likelihood of accidents. The developed system also enables visualising monitoring data within the 4D BIM model, offering stakeholders a comprehensive understanding of the site and supporting continuous improvement through a dynamic feedback loop between the design and construction phases.

Acknowledgements

This safety project was built upon a five-way partnership involving: Discovering Safety—a five-year programme being delivered by HSE and funded by the Lloyd's Register Foundation, AstraZeneca's construction team, on-site support from Dalkia, and the two technology suppliers— Asite 3D Repo and Plinx.

References

Chung, W. W. S., Tariq, S., Mohandes, S. R., & Zayed, T. (2023). IoT-based application for construction site safety monitoring. *International Journal of Construction Management*. https://doi.org/10.1080/15623599.2020.1847405

Ding, L., Jiang, W., & Zhou, C. (2022). IoT sensor-based BIM system for smart safety barriers of hazardous energy in petrochemical construction. *Frontiers of Engineering Management*. https://doi.org/10.1007/s42524-021-0160-6

Hong, K., & Teizer, J. ; (2024). A Data-Driven Method for Hazard Zone Identification in Construction Sites with Wearable Sensors. *Citation*, 41–48. https://doi.org/10.24840/978-972-752-309-2

Hossain, M. M., Ahmed, S., Anam, S. M. A., Baxramovna, I. A., Meem, T. I., Sobuz, M. H. R., & Haq, I. (2023). BIM-based smart safety monitoring system using a mobile app: a case study in an ongoing construction site. *Construction Innovation*. https://doi.org/10.1108/CI-11-2022-0296

HSE. (2023). *Construction statistics in Great Britain, 2023*. https://www.hse.gov.uk/statistics/assets/docs/construction.pdf

Jiang, W., Ding, L., & Zhou, C. (2021). Cyber physical system for safety management in smart construction site. *Engineering, Construction and Architectural Management*. https://doi.org/10.1108/ECAM-10-2019-0578

Jin, R., Zhang, H., Liu, D., & Yan, X. (2020). IoT-based detecting, locating and alarming of unauthorized intrusion on construction sites. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2020.103278

Kojima, H., Fujii, T., Mihara, Y., & Ihara, H. (2020). Introduction of the new safety concept “safety2.0” to reduce the risk of machinery accidents. *Proceedings of the 37th International Symposium on Automation and Robotics in Construction, ISARC 2020: From Demonstration to Practical Use - To New Stage of Construction Robot*. https://doi.org/10.22260/isarc2020/0197

Kulinan, A. S., Park, M., Aung, P. P. W., Cha, G., & Park, S. (2024). Advancing construction site workforce safety monitoring through BIM and computer vision integration. *Automation in Construction*, *158*, 105227. https://doi.org/10.1016/J.AUTCON.2023.105227

Liang, Y., & Liu, Q. (2022). Early warning and real-time control of construction safety risk of underground engineering based on building information modeling and internet of things. *Neural Computing and Applications*. https://doi.org/10.1007/s00521-021-05755-8

Shuang, D., Heng, L., Skitmore, M., & Qin, Y. (2019). An experimental study of intrusion behaviors on construction sites: The role of age and gender. *Safety Science*. https://doi.org/10.1016/j.ssci.2019.02.035

Tran, S. V. T., Nguyen, T. L., & Park, C. (2021). A BIM Integrated Hazardous Zone Registration Using Image Stitching. *Proceedings of the International Symposium on Automation and Robotics in Construction*. https://doi.org/10.22260/isarc2021/0026

Wang, X., Liu, C., Song, X., & Cui, X. (2022). Development of an Internet-of-Things-Based Technology System for Construction Safety Hazard Prevention. *Journal of Management in Engineering*. https://doi.org/10.1061/(asce)me.1943-5479.0001035

Wu, S., Hou, L., Zhang, G. (Kevin), & Chen, H. (2022). Realtime mixed reality-based visual warning for construction workforce safety. *Automation in Construction*. https://doi.org/10.1016/j.autcon.2022.104252

Zhang, M., Ghodrati, N., Poshdar, M., Seet, B. C., & Yongchareon, S. (2023). A construction accident prevention system based on the Internet of Things (IoT). *Safety Science*. https://doi.org/10.1016/j.ssci.2022.106012

TOWARDS A STANDARDIZED PROCESS FOR VIRTUAL SENSORS DEVELOPMENT FOR INDOOR AIR QUALITY MONITORING IN DEMOCRATIZED MANUFACTURING

Jorge Merino[*1], Daiki Ikeuchi[2], Alena Puchkova[1], Luning Li[1],
Sebastian Pattinson[2], Stylianos Karatzas[1], Ajith Parlikad[1]
[1] Distributed Information and Automation Laboratory, Department of Engineering, University of Cambridge, Institute for Manufacturing, Alan Reece Building, 17 Babbage Road, CB3 0FS, Cambridge, United Kingdom
[2] Computer Aided Manufacturing Laboratory, Department of Engineering, University of Cambridge, Institute for Manufacturing, Alan Reece Building, 17 Babbage Road, CB3 0FS, Cambridge, United Kingdom
*jm2210@cam.ac.uk, di261@cam.ac.uk, ap823@cam.ac.uk, ll669@cam.ac.uk,
swp29@eng.cam.ac.uk, sk2095@cam.ac.uk, aknp2@cam.ac.uk

Abstract

Increased accessibility to additive manufacturing technology facilitates democratization of manufacturing, bringing it to habitable environments. The operation of additive manufacturing can be hazardous to human health mid-long term. Virtual sensing extends the capabilities of hardware sensors enabling affordable monitoring to ensure safe operation in democratized manufacturing environments. However, the development process has not yet been standardized for informally trained personnel to facilitate the adoption of virtual sensors. This paper presents a case study analysis to propose a standardized process for the data collection and development of virtual sensors for indoor air quality monitoring in democratized manufacturing environments.

Introduction

Democratization of manufacturing is driven by personal digital fabrication (Mota, 2011; Anderson, 2012; Gershenfeld, 2012), which enables informally-trained consumers to meet their production and customization needs (Browder et al., 2019). Increased access to fabrication technology like Additive Manufacturing (AM) or 3D printing (3DP) brings digital fabrication to ad-hoc manufacturing spaces at home, university, and start-ups that are not designed to manage the harmful emissions for human health (Ford & Despeisse, 2016). The widely adopted Material Extrusion (ME) 3DP technology typically utilizes thermoplastic or thermoset polymers as feedstock materials (Ligon et al., 2017), releasing Volatile Organic Compounds (VOC) and Particulate Matter (PM) during their operation. If inhaled, PM (PM2.5 or smaller is smaller than 2 µm) can reach the alveolar region of the lungs and be transported by the bloodstream to other organs (Pilou et al., 2015). (Afshar-Mohajer et al., 2015; Azimi et al., 2017) proved that VOC and PM in indoor environments for AM can greatly exceed the safety thresholds (World Health Organization, 2016).

Indoor environments where AM technology is deployed must be controlled to limit these health hazards since informally trained personnel are neither aware of the health hazards (Berger et al., 2023; Schieweck et al., 2018). Many authors have addressed the characterization of emissions in terms of VOCs and PM from 3D printing processes under controlled laboratory environments (Afshar-Mohajer et al., 2015; Bravi et al., 2019; Kim et al., 2015; Mendes et al., 2017; Pernetti et al., 2023). These works of characterization of pollutants mostly focus on single sampling of emissions using high-end equipment. Nevertheless, these solutions are not transferable to the built environment (e.g., in offices, classrooms, and makerspaces) due to the reduced number and cost of sampling equipment. In contrast, low-cost sensors are accessible and affordable solutions to monitor the emissions in additive manufacturing and larger environments. The caveat of using low-cost sensors is the lack of accuracy of the available technology to detect emissions. Plus, makers are not willing to add the extra cost of sensors due to their lack of awareness. Some authors try to make low-cost monitoring feasible by integrating it with signal processing and analysis (Stefaniak et al., 2021; Vogt et al., 2021; Martínez-Comesaña et al., 2022; Tagliabue et al., 2021). These techniques can be automated via virtual sensors to increase awareness on IAQ for 3D printing users.

Virtual sensors are online instruments of measurement that supplement other monitoring entities (Lin et al., 2007; Montague et al., 1992; Martin et al., 2021). Virtual sensors have been used to estimate NO, NO2, NOx, SO2, O3, CO, Cd, PM2.5, and PM10 concentration and virtually recreate the pollution in the built environment, in entire neighborhoods (Benabbas et al., 2019; Dogeanu et al., 2019) and undergrounds (Loy-Benitez et al., 2020; Loy-Benitez et al., 2022). Applications of virtual sensing to measure indoor air quality (IAQ) in the literature are still limited to controlled environments (Leidinger et al., 2014; Zaidan et al., 2020), but has proven to be a feasible option for monitoring comfort indoors (Kowli et al., 2023). Virtual sensors can also be used to monitor targets that are out of reach (i.e., it is not possible to install a hardware sensor), and to reduce the number of sensors required making the final solution more affordable. Given the limited literature in the context of democratized manufacturing, there is still no standard process to guide the development of virtual sensors for IAQ.

This paper aims to kick-off the design of a standard process for the development of virtual sensors for IAQ monitoring in ad-hoc democratized manufacturing spaces. The process learns from the ad-hoc solutions in the literature and the analysis of a case study that implemented a virtual sensor model in a 3D printing environment.

Context and case study analysis

The key elements of a case study design are (Priya, 2021):

(a) Purpose of study: To create a data collection protocol and a training and deployment process for virtual sensors to monitor IAQ in a democratized manufacturing environment.

(b) Type of case study research: exploratory case study.

(c) Research question: How to develop and utilize virtual sensors for IAQ monitoring?

The context of the case study is a democratized additive manufacturing environment in a commercial office building. Herein, the study focuses on a cabinet hosting four printers Creality CR20-PRO Fused Deposition Modelling (FDM), an ME technique. FDM 3D printing principle consists in building three-dimensional solid objects from their digital models by selectively accumulating liquified thermoplastic material (e.g., PLA, ABS) layer-by-layer (Anon, 2024; ISO/ASTM, 2013) into a bed where the material cools and cements, shaping a geometry. The cabinet is enclosed for the sake of safety in terms of limiting the emissions released into the rest of the shared working environment; yet this space represents a practical democratized manufacturing space in terms of size, purpose, and the level of available resources. The dimensions of the cabinet are 200x90x100 cm (see Figure 1). It is composed of a desk and eight methacrylate panels for two side panels, four access windows, two sliding windows in the back and the ceiling, plus a single extraction point in the center of the ceiling panel. Extraction works at the same speed but can be turned on and off. The goal of monitoring in this space is to raise awareness on health for the users in terms of utilization of the space, and the cabinet, and pollution escaping the cabinet during operation of the 3D printers.

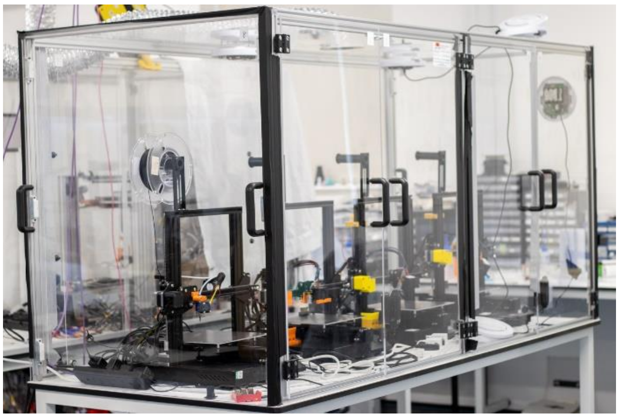


Figure 1. Cabinet hosting four 3D printers for relevant data collection under democratized manufacturing settings.

Virtual sensor data collection, training, and deployment process

This section explains the process for collecting IAQ data and training and deploying virtual sensors in the democratized additive manufacturing environment. This process is addressed to occupants operating democratized additive manufacturing environments and data practitioners interested in monitoring air quality indoors.

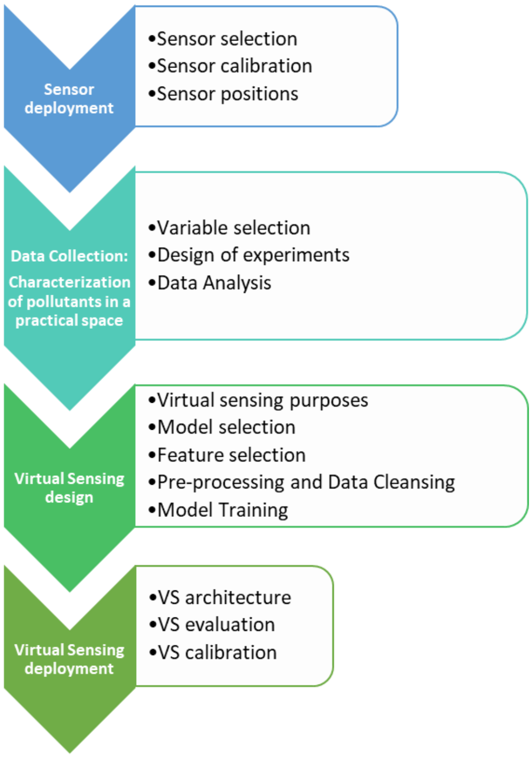


Figure 2. Process for data collection and development of Virtual sensors for IAQ monitoring.

Figure 2 summarizes these steps in the process, which are the result of the analysis of literature on IoT deployments (Bashir et al., 2022; Bibri, 2018), prior knowledge in other areas (Merino et al., 2022), and adaptations from the application of each step to the domain of IAQ.

Low-cost sensor deployment

An effective sensor deployment is launched by appropriate selection of sensors. Sensor selection must be informed by literature in the domain as well as a market review. For instance, in the context of 3DP, (Afshar-Mohajer et al., 2015; Bravi et al., 2019; Kim et al., 2015; Mendes et al., 2017; Pernetti et al., 2023) used chamber setups to characterized the emissions of popular 3DP filament materials like PLA, ABS, HIPS, and PC, resulting in particulate matter (PM) and diverse species of volatile organic compounds (VOCs).

Sensors shall be calibrated before deployment. Although manufacturers claim that sensors have been calibrated in laboratory environments at source, they need to be calibrated in the field since fabrication variance is not avoidable yet in low-cost sensing. The calibration methods for air quality monitoring vary (Chowdhury et al., 2023; Zaidan et al., 2020; Zarrar & Dyo, 2023; Tariq et al., 2020; Zimmerman et al., 2018), but share the notion of comparing against a source of truth. Often these sources of truth are high-quality sampling equipment.

This equipment comes at a high price tag both for purchasing and for single-time calibration from the manufacturer, which may be prohibitive for some makers. Furthermore, concept drift (De Vito et al., 2024) appears when operating conditions change from those encountered during training set recording, the derived calibration accuracy drops. Thus, calibration should be considered over longer-time of monitoring to account for seasonal differences and sensor drifts. Alternatively, some authors targeting scalable calibration have proposed median calibration, which is obtained through using datasets considering the median response of several sensor units to identical field recorded conditions (De Vito et al., 2024).

Finally, the position of the low-cost sensors in the space must be designated for deployment. Sensor positioning in air quality monitoring is informed by the identification of sources of pollution and the expected usage of data. The final positions and configuration of sensors must be logged into a repository for traceability and maintenance of the deployment, since hardware sensors can suffer network drops and reconfigurations, power problems, or even complete failure that prompts replacement. For deployments with limited number of sensors, can make them rotate around the intended positions, then later substituted for their virtual counterparts.

Data Collection: Characterization of pollutants in a practical space

The purpose of the data collection is to characterize how the pollutants are dispersed in a practical democratized additive manufacturing environment.

Variables are selected to enable virtual sensing training and operation. Generally, nearby locations measurements are used to estimate PM and VOC, but some authors have used humidity and temperature to improve the accuracy of CO2 estimations (Tagliabue et al., 2021; Zaidan et al., 2020), CO and CO2 and operational parameters of the 3D printers for estimating PM and VOC (Stefaniak et al., 2021).

Experiments must be contextualized in the process of the indoor environment that is causing the emissions. The goals of data collection and literature can support the design of experiments. In the case of semi-controlled or uncontrolled environments, experiments conducted shall cover data collection of a range of practical situations. Experiments should have the same starting conditions and they are repeated for validation of data collection under the same conditions. Experiments can also be repeated for identification of seasonality.

Data analyzed to test the validity and thoroughness of data collection and to identify potential missing conditions for new experiments. Experiments ought to include at least baseline identification (i.e., how is the signal for each variable when nothing is happening) and the range of practical situations devised during experiment design.

Virtual Sensing design

Virtual sensors are used as a backup for potential failures of low-cost sensors, to completely replace them, and to provide measurements in locations hardware sensors cannot reach (e.g., temperature inside a boiler). In the case study, the virtual sensor was designed as a backup for PM1.0 of the low-cost sensor in the same location and to replace it for use in a different space.

Diverse regression and machine learning models have been used to estimate PM and VOC: in terms of indoor air quality, (Stefaniak et al., 2021) used linear regression to estimate VOC and PM concentration in a large-format yet controlled AM lab. (Dogeanu et al., 2019) trained Neural Network virtual sensors on data from fixed air quality stations located in urban areas to estimate the concentration of PM2.5 and PM10 among other pollutants to virtually recreate the pollution in entire neighborhoods. (Leidinger et al., 2014) implemented a virtual sensor based on linear discriminant analysis (LDA) for selective detection of hazardous VOCs using a gas sensor array. Literature can influence model selection, but it is ultimately the decision of the data practitioner which model to select. Comparison between several models is recommended. It is important to understand that models trained in a given space may not be transferable to other spaces unless the geometry and operations are near to identical (e.g., another cabinet with the same setup). Models in the literature should be implemented within the context of the case study to test the validity and enable comparison under desired conditions.

Along with model selection, comes the features to be used to train the virtual sensors. Data analysis should undercover correlations between variables that can support estimations in the models. Literature can also inform known relationships between variables, but they should be confirmed through data analysis.

In most cases, raw data from hardware sensors needs some form of preparation before they are suitable for model training. Typical methods for data preparation are filtering, synchronization, and transformations (e.g., principal component analysis, logarithmic conversion, or derivatives of the original signal). Outliers due to sensor temporary malfunction and missing data due to power or network losses should be identified and removed, and only relevant controlled data from the experiments should be used for training.

With the advent of Machine Learning open-source frameworks, model training is becoming easier than ever. (Nguyen et al., 2019; Wu et al., 2022) compare popular frameworks in terms of capabilities and training and testing performance. The main goal of training is to obtain acceptable accuracy without incurring into overfitting models. This becomes an iterative process of trial and error, and continuous improvement in most cases. In some cases, this iterative process requires the practitioner to apply different pre-processing methods, selecting different features, or conducting further data analysis.

Virtual Sensing deployment

Having a trained model that can estimate the pollutants in a given position is not enough for using a virtual sensor in a practical case. Virtual sensors for IAQ monitoring need to be understood as standalone systems and behave similarly to low-cost sensors. A system architecture is necessary to deploy virtual sensors in a scalable manner (Merino et al., 2023; Martin et al., 2021; Leidinger et al., 2014; Zaidan et al., 2020).

Any type of virtual sensor needs evaluation (Goodwin, 2000), but more importantly when it comes to IAQ monitoring since they are aimed to monitor pollutants that can affect human health. Virtual sensors need long-term calibration (Koo & Yoon, 2022) since their accuracy can decrease because of various uncertainties in the system operation and virtual sensor model.

Case study application and Discussion

This section describes the application of the data collection, training, and deployment process for virtual sensors. Results of implementation are discussed including the identification of the challenges faced and adaptations made to the process.

Sensor deployment implementation

The literature on air quality in additive manufacturing environments guided the selection of sensors. The materials used in the 3D printers are PLA and ABS, which emit PM and diverse species of VOC (Steinle, 2016). The virtual sensor designed focuses on PM. Low-cost sensors available in the market to monitor these two variables are based on the light-scattering principle for PM. Sensors from two manufacturers are used: M1 can measure PM1.0, PM2.5, and PM10, and M2 can monitor PM1.0, PM2.5, PM10, tVOC, Temperature, Humidity, CO2, and Ozone.

Sensors were calibrated using the median approach. While other authors installed sensors in their final position to collect data for median calibration, sensors were collocated in stacks in the same center of the desk inside of the cabinet and shuffled once during the calibration data collection. This deviation from the literature moved the calibration step before the sensor positioning in the process. Median was calculated across all sensors and selected as a general calibration law. A data pre-processing module was added to the data pipeline to locally adjust the differences between each sensor signal and the median. Additionally, a sensor drift analyzer was set up to identify and correct potential sensor drifts due to aging. Calibration was performed per manufacturer.

The positions of the sensors are determined by the positions of the 3D printers in the cabinet. There are four 3D positions for the 3D printers evenly distributed along the Y axis (200 cm long) in the cabinet, centered at 25 cm, 75 cm, 125cm and 175cm. The purpose of devised virtual sensors is to estimate distribution of pollutants within the cabinet, therefore, the sensors (four M1, and fifteen M2) were evenly distributed in the cabinet.

Data collection implementation

Temperature, Humidity, and PM1.0 from all sensor locations are considered. Temperature and humidity have an impact and Particulate matter sensors, respectively and can help to identify seasonal bias.

The following experiment variables were selected: extraction of the cabinet (i.e., can be on/off, when it is on, it always extracts air at the same speed); windows opened (i.e., whether it is planned to open any window during the experiment, monitored by open closed sensors); printer positions in operation (i.e., takes the form of an array of Booleans, e.g., [1, 1, 0, 0] means that only printers in positions 1 and 2 are printing); material (i.e., type of material used throughout the printing, either PLA or ABS from the same manufacturer and model in all the experiments); and nozzle temperature. The experiments consisted of permutations of the variables. For example, experiment 1 included extraction off, windows closed, and 3D printer in position 1 printing a geometry using ABS for one hour. Before each experiment, the air inside the cabinet is exchanged for at least one hour. Experiments with the same conditions (i.e., variable setups) were repeated at least three times.

The baselines of PM1.0 in all sensor locations were identified over a period of 7 hours under non-printing conditions by repeating the experiment with windows open and closed, extraction on and off. For instance, Figure 3 depicts two baselines identified for PM1.0 in the target location (i.e., the sliding window used to access the printed geometries): Extraction Off and door closed (400-420 ppm), and extraction on/windows opened (380-400 ppm). This enabled a baseline correction reference for each sensor parameter in data.

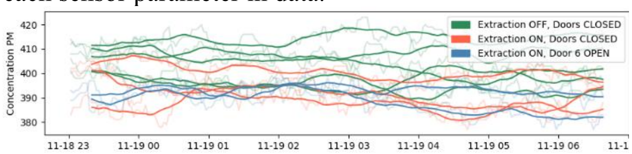


Figure 3. PM1.0 baseline identification in target location

Among the permutation of experiment variables under printing conditions, Figure 4 depicts the experiments for position 1 printer printing with PLA over a period of 1h:30'. Some conclusions of this plot are that extraction is confirmed to help with the clearance of the cabinet, both in concentration of PM and time for the concentration to decrease back to baseline after the end of the experiments.

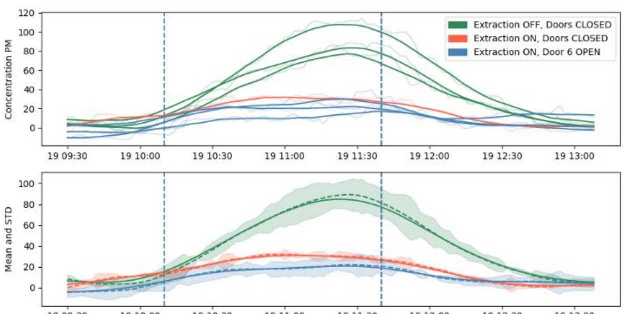


Figure 4. PM1.0 in target location under printing conditions

Having two different sensor manufacturers required repeating the analysis for repeated parameters. Sensors M1 were more reactive to changes in PM concentration than M2. Cross-manufacturer comparisons are required. Data collection is an iterative and cyclical process even with planning. Furthermore, some data collection impacted the selection of more sensors of Manufacturer M1 as well as the position (together with sensors M2).

Virtual sensing design implementation

Several models were selected both linear (e.g., lasso, elastic networks, Bayesian) and machine learning (e.g., boosted trees). Linear regression models have fair accuracy when supported by online measurements from hardware sensors (Stefaniak et al., 2021), whereas boosted trees have been used to estimate PM in (Koo & Yoon, 2022). A simple case is to remove the time variable of the series and estimate individual values based on the measurements of the reference low-cost hardware sensor in the same space, for which all linear models performed (Merino et al., 2023). This model lacks the capability of identifying trends since it is not based on time-series estimation.

Features selected for PM estimation are measurements of the same PM size from surrounding hardware sensors. For the simple case, the correlation between each of the PM variables (i.e., PM1.0, PM2.5, and PM10) is enough to train linear models when estimating single values. Thus, to estimate PM1.0 in sensor a target position, PM1.0 readings from all positions but the target are selected. This decision is in line with the purpose of the virtual sensor, to serve as a backup for and potentially replace the target low-cost sensor location (i.e., by the sliding door).

3D printer logs and low-cost hardware sensors required timestamp synchronization to compensate for the skew derived from internal clock drifts and different time server setups of the sensor and 3D printer platforms. Additionally, outliers identified (e.g., sudden jumps in raw signal or subsequent out-of-range measurements due to malfunction of PM and VOC sensors) are removed if the value is beyond two standard deviations. Uncontrolled data is discarded for the purposes of the virtual sensor. Finally, a Savitzky-Golay filter (Savitzky & Golay, 1964) was applied to smooth the curves. For the simplest case of the single-value estimations based on surrounding sensors, the time dimension is removed. Since the signals are previously synchronized that results into a set of arrays with values for each sensor position at a given time.

Models trained take the PM filtered and untimed arrays and having partitioned them into 80% training and 20% testing segmented sub-sets. Linear models R^2 scores reached 0.826 for Lasso, 0.826 for Elastic Networks, and 0.826 for Bayesian ridge for the simple approach. Non-linear approach using temperature and/or humidity as supporting estimators for both PM could not reach a significant R^2 score to be considered good estimators in a practical case, potentially due to the transformation into single-value estimation rather than time-series estimation, which enable trend identification and more accurate regression. Trained models can estimate the pollutant concentration in a specific point where the original low-cost sensor was positioned, therefore, they can be used to virtualize.

Virtual sensing deployment

The architecture published in (Merino et al., 2023) was designed to streamline virtual sensors training and operation. The architecture modularizes the process to facilitate data collection, preprocessing, training, model storage and loading, virtual sensor execution, and recalibration. In the simple approach for estimating PM based on surrounding sensors signals, the evaluation consisted in comparing the signal from the virtual sensor against the original sensor used to train the model after a period of dual operation (i.e., both the virtual sensor and the low-cost hardware sensor measuring PM values for a period of two weeks before replacing the hardware sensor).

The need for recalibration of virtual sensors can be identified by deploying a hardware sensor again in the original position. In a practical democratized additive manufacturing environment, or any other indoor building operation, hardware sensor may not always be available for recalibration, therefore, other approaches like (Koo & Yoon, 2022) may be more suitable.

Conclusions

Indoor air quality monitoring in democratized additive manufacturing environments is still not sufficiently explored. Virtual sensing approaches have demonstrated good performance for outdoors air quality and indoor comfort and pose a good opportunity to improve the awareness and control of hazardous pollutants. Beyond increasing awareness of emissions indoors at an affordable price, virtual sensors have the potential of informing ventilation systems operation to ensure operators safety while minimising building energy management. Virtual sensors can help understanding spatiotemporal distribution of emissions and emission flow into the rest of the built environment (e.g., adjacent offices, classrooms, or even bedrooms at home). Furthermore, IAQ monitoring can have the potential to inform on materials and design of indoor spaces as well as maintenance during the operation phase of the built environment (Abdalla & Peng, 2021).

This paper proposes a standard process for data collection and development of indoor air quality using virtual sensors in a practical democratized additive manufacturing environment, resulting of the analysis of the literature and the adaptations after the implementation in a practical democratized additive manufacturing environment. A simple example of a virtual sensor is explained throughout the implementation. The aim of the process is to facilitate the entry barrier to this technology to ad-hoc democratized manufacturing spaces, answering the research question. Additionally, the process should also address a second research question in the next iteration, on how to overcome lack of trust in low-cost monitoring and consequently in virtual sensing. Evaluation and calibration approaches should be investigated further to answer this question.

This process needs to be implemented in a set of case studies in similar democratized manufacturing spaces to test its validity and find adjustments. Additionally, a replication of this study in other democratized manufacturing environments would be beneficial for the confirmation of findings and to discover the casuistic across different setups and additive manufacturing technologies. Another limitation of this study is within the fact that indoor environments suffer from pollution from non-characterizable sources like outdoors pollution from construction sites or traffic. Case studies identified to complement this work in this direction and to ensure full transferability to the built environment context include domains like accommodation, offices, hospitals, universities, and public transport stations.

Acknowledgements

This work was supported by the UK's Department for Digital, Culture, Media and Sport through the project entitled “Digital Twins Enabled IAQ Management for Healthy Living-TwinAIR”. TwinAIR is funded by UK Research and Innovation (UKRI) through the Horizon Europe Guarantee Scheme for European Union’s Horizon Europe Research and Innovation Programme, grant number 101057779, HORIZON-HLTH-2021-ENVHLTH-02-02.

References

Abdalla, T. & Peng, C. (2021) Evaluation of housing stock indoor air quality models: A review of data requirements and model performance. Journal of Building Engineering, 43, p.102846.

Afshar-Mohajer, N., Wu, C.-Y., Ladun, T., Rajon, D.A. & Huang, Y. (2015) Characterization of particulate matters and total VOC emissions from a binder jetting 3D printer. Building and Environment, 93, pp.293–301.

Anderson, C. (2012) Makers: The New Industrial Revolution. Random House Business Books. Available from: <https://books.google.co.uk/books?id=zjAvjJ2 XKzEC>.

Anon (2024) Additive Manufacturing Glossary [Internet]. Available from: https://www.sme.org/additivemanu- facturing-glossary> [Accessed 9 August 2023].

Azimi, P., Fazli, T. & Stephens, B. (2017) Predicting Concentrations of Ultrafine Particles and Volatile Organic Compounds Resulting from Desktop 3D Printer Operation and the Impact of Po- tential Control Strategies. Journal of Industrial Ecology, 21 (S1), pp.S107–S119.

Bashir, M.R., Gill, A.Q. & Beydoun, G. (2022) A Reference Architecture for IoT-Enabled Smart Buildings. SN Computer Science, 3 (6), p.493.

Benabbas, A., Geißelbrecht, M., Nikol, G.M., Mahr, L., Nähr, D., Steuer, S., Wiesemann, G., Müller, T., Nicklas, D. & Wieland, T. (2019) Measure particulate matter by yourself: data-quality monitoring in a citizen science project. Journal of Sen- sors and Sensor Systems, 8 (2), pp.317– 328.

Berger, C., Mahdavi, A., Ampatzi, E., Bandurski, K., Hellwig, R.T., Schweiker, M., Topak, F. & Zgank, M. (2023) The role of user controls with respect to indoor environmental quality: From evidence to standards. Journal of Building Enginee-ring, 76, p.107196.

Bibri, S.E. (2018) The IoT for smart sustainable cities of the future: An analytical framework for sensor-based big data applications for environmental sustainability. Sustainable Cities and Society, 38, pp.230–253.

Bravi, L., Murmura, F. & Santos, G. (2019) Additive Manufacturing: Possible Problems with Indoor Air Quality. Procedia Manufacturing, 41, pp.952–959.

Browder, R.E., Aldrich, H.E. & Bradley, S.W. (2019) The emergence of the maker movement: Implications for entrepreneurship research. Journal of Business Venturing, 34 (3), pp.459–476.

Chowdhury, S.R., Parve, S., Chopda, B.N. & Gupta, R.K. (2023) Sensor calibration techniques for IoT networks. In: Chennai, India, p.090007. Availa <blhettp://aip.scitation.org/doi/abs/10.1f0r6o3m/5: .0136184> [Accessed 29 January 2024].

De Vito, S., D’Elia, G., Ferlito, S., Francia, G.D., Davidović, M.D., Kleut, D., Stojanović, D. & Jovaševic-Stojanović, M. (2024) A Global Multiunit Calibration as a Method for Large-Scale IoT Particulate Matter Monitoring Systems Deployments. IEEE Transactions on Instrumentation and Measurement, 73, pp.1–16.

Dogeanu, A., Nastase, I., Ursu, I., Enciu, D., Bode, F., Sandu, M., Croitoru, C., Zaharia, S. & Iana, G. (2019) Real time monitoring network demonstrator for air quality management. In: 2019 International Conference on ENERGY and ENVIRONMENT (CIEM). Timisoara, Romania, IEEE, pp.459–463. Available from: https://ieeexplore.ieee.org/document/8937625/ [Accessed 10 August 2023].

Ford, S. & Despeisse, M. (2016) Additive manufacturing and sustainability: an exploratory study of the advantages and challenges. Journal of Cleaner Production, 137, pp.1573–1587.

Gershenfeld, N. (2012) How to Make Almost Anything: The Digital Fabrication Revolution. Foreign Affairs, 91 (6), pp.43–57.

Goodwin, G.C. (2000) Predicting the performance of soft sensors as a route to low cost automation. Annual Reviews in Control, 24, pp.55–66.

ISO/ASTM (2013) ISO/ASTM 52921:2013 [Internet]. Available from: <https://www.iso.org/standard/62794.html> [Accessed 26 January 2024].

Kim, Y., Yoon, C., Ham, S., Park, J., Kim, S., Kwon, O. & Tsai, P.-J. (2015) Emissions of Nanoparticles and Gaseous Material from 3D Printer Operation. Environmental Science & Technology, 49 (20), pp.12044–12053.

Koo, J. & Yoon, S. (2022) In-situ sensor virtualization and calibration in building systems. Applied Energy, 325, p.119864.

Kowli, A., Rani, V. & Sanap, M. (2023) Data-driven virtual sensing for spatial distribution of temperature and humidity. Journal of Building Engineering, 67, p.105726.

Leidinger, M., Sauerwald, T., Reimringer, W., Ventura, G. & Schütze, A. (2014) Selective detection of hazardous VOCs for indoor air quality applications using a virtual gas sensor array. Journal of Sensors and Sensor Systems, 3 (2), pp.253–263.

Ligon, S.C., Liska, R., Stampfl, J., Gurr, M. & Mülhaupt,

R. (2017) Polymers for 3D Printing and Customized Additive Manufacturing. Chemical Reviews, 117 (15), pp.10212–10290.

Lin, B., Recke, B., Knudsen, J.K.H. & Jørgensen, S.B. (2007) A systematic approach for soft sensor development. Computers & Chemical Engineering, 31 (5), pp.419–425.

Loy-Benitez, J., Heo, S. & Yoo, C. (2020) Soft sensor validation for monitoring and resilient control of sequential subway indoor air quality through memory-gated recurrent neural networks-based autoencoders. Control Engineering Practice, 97, p.104330.

Loy-Benitez, J., Tariq, S., Nguyen, H.T., Safder, U., Nam, K. & Yoo, C. (2022) Neural circuit policies-based temporal flexible soft-sensor modeling of subway PM2.5 with applications on indoor air quality management. Building and Environment, 207, p.108537.

Martin, D., Kühl, N. & Satzger, G. (2021) Virtual Sensors. Business & Information Systems Engineering, 63 (3), pp.315–323.

Martínez-Comesaña, M., Eguía-Oller, P., Martínez-Torres, J., Febrero-Garrido, L. & Granada- Álvarez, E. (2022) Optimisation of thermal comfort and indoor air quality estimations applied to in-use buildings combining NSGA-III and XGBoost. Sustainable Cities and Society, 80, p.103723.

Mendes, L., Kangas, A., Kukko, K., Mølgaard, B., Säämänen, A., Kanerva, T., Flores Ituarte, I., Huhtiniemi, M., Stockmann-Juvala, H., Partanen, J., Hämeri, K., Eleftheriadis, K. & Viitanen, A.-K. (2017) Characterization of Emissions from a Desktop 3D Printer. Journal of Industrial Ecology, 21 (S1), pp.S94–S106.

Merino, J., Ikeuchi, D., Moretti, N., Mukherjee, A., Li, Luning, Karatzas, S., Pattinson, S. & Parlikad,

A.K. (2023) Virtual sensor architecture for indoor air quality monitoring. In: Proceedings of the 1st 1 1st Workshop on Low - Cost Digital Solutions for Industrial Automation. Cambridge, United Kingdom.

Merino, J., Sasidharan, M., Herrera, M., Zhou, H., Crespo del Castillo, A., Parlikad, A.K., Brooks, R. & Poulter, K. (2022) Lessons learned from an IoT deployment for condition monitoring at the Port of Felixstowe. IFAC-PapersOnLine, 55 (19), pp.217–222.

Montague, G.A., Morris, A.J. & Tham, M.T. (1992) Enhancing bioprocess operability with generic software sensors. Journal of Biotechnology, 25 (1), pp.183–201.

Mota, C. (2011) The rise of personal fabrication. In: Proceedings of the 8th ACM conference on Creativity and cognition. C&C ’11. New York, NY, USA, Association for Computing Machinery, pp.279–288. Available from: <https://doi.org/10.1145/2069618.2069665> [Accessed 25 January 2024].

Nguyen, G., Dlugolinsky, S., Bobák, M., Tran, V., López García, Á., Heredia, I., Malík, P. & Hluchý, L. (2019) Machine Learning and Deep Learning frameworks and libraries for large-scale data mining: a survey. Artificial Intelligence Review, 52 (1), pp.77–124.

Pernetti, R., Maffia, S., Previtali, B. & Oddone, E. (2023) Assessment of nanoparticle emission in additive manufacturing: Comparing wire and powder laser metal deposition processes. Journal of Occupational and Environmental Hygiene, pp.1–7.

Pilou, M., Mavrofrydi, O., Housiadas, C., Eleftheriadis,

K. & Papazafiri, P. (2015) Computational modeling as part of alternative testing strategies in the respiratory and cardiovascular systems: Inhaled nanoparticle dose modeling based on representative aerosol measurements and corresponding toxicological analysis. Nanotoxicology, 9 (sup1), pp.106–115

Priya, A. (2021) Case Study Methodology of Qualitative Research: Key Attributes and Navigating the Conundrums in Its Application. Sociological Bulletin, 70 (1), pp.94–110.

Savitzky, Abraham. & Golay, M.J.E. (1964) Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry, 36 (8), pp.1627–1639.

Schieweck, A., Uhde, E., Salthammer, T., Salthammer, L.C., Morawska, L., Mazaheri, M. & Kumar, P. (2018) Smart homes and the control of indoor air quality. Renewable and Sustainable Energy Reviews, 94, pp.705–718.

Stefaniak, A.B., Bowers, L.N., Martin, S.B., Hammond, D.R., Ham, J.E., Wells, J.R., Fortner, A.R., Knepp, A.K., Preez, S.D., Pretty, J.R., Roberts,

J.L., Du Plessis, J.L., Schmidt, A., Duling, M.G., Bader, A. & Virji, M.A. (2021) Large-Format Additive Manufacturing and Machining Using High-Melt-

Temperature Polymers. Part I: Real- Time Particulate and Gas-Phase Emissions. ACS Chemical Health & Safety, 28 (3), pp.190–200.

Steinle, P. (2016) Characterization of emissions from a desktop 3D printer and indoor air measurements in office settings. Journal of Occupational and Environmental Hygiene, 13 (2), pp.121–132.

Tagliabue, L.C., Re Cecconi, F., Rinaldi, S. & Ciribini,

A.L.C. (2021) Data driven indoor air quality prediction in educational facilities based on IoT network. Energy and Buildings, 236, p.110782.

Tariq, H., Abdaoui, A., Touati, F., E Al-Hitmi, M.A., Crescini, D. & Mnaouer, A.B. (2020) An Autonomous Multi-Variable Outdoor Air Quality Mapping Wireless Sensors IoT Node for Qatar. In: 2020 International Wireless Communications and Mobile Computing (IWCMC). Limassol, Cyprus, IEEE, pp.2164– 2169. Available from: <https://ieeexplore.ieee.org/document/9148392/ > [Accessed 29 January 2024].

Vogt, M., Schlichter, J., Aschersleben, F., Abraham, T., Wolf, L. & Herrmann, C. (2021) Integration of cyber-physical HVAC systems in Incremental Manufacturing to improve Energy Efficiency and Air Quality. Procedia CIRP, 104, pp.482– 487.

World Health Organization (2016) Ambient air pollution: a global assessment of exposure and burden of disease.

Wu, Y., Liu, L., Pu, C., Cao, W., Sahin, S., Wei, W. & Zhang, Q. (2022) A Comparative Measurement Study of Deep Learning as a Service Framework. IEEE Transactions on Services Computing, 15 (1), pp.551–566.

Zaidan, M.A., Hossein Motlagh, N., Fung, P.L., Lu, D., Timonen, H., Kuula, J., Niemi, J.V., Tarkoma, S., Petaja, T., Kulmala, M. & Hussein, T. (2020) Intelligent Calibration and Virtual Sensing for Integrated Low-Cost Air Quality Sensors. IEEE Sensors Journal, 20 (22), pp.13638–13652.

Zarrar, H. & Dyo, V. (2023) Drive-by Air Pollution Sensing Systems: Challenges and Future Directions. IEEE Sensors Journal, 23 (19), pp.23692–23703.

Zimmerman, N., Presto, A.A., Kumar, S.P.N., Gu, J., Hauryliuk, A., Robinson, E.S., Robinson, A.L., & R. Subramanian (2018) A machine learning calibration model using random forests to improve sensor performance for lower-cost air quality monitoring. Atmospheric Measurement Techniques, 11 (1), pp.291–313.

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

EQUIRECTANGULAR 360°IMAGE DATASET FOR DETECTING REUSABLE CONSTRUCTION COMPONENTS

Ana Bendiek Laranjo[1,2], Jens J. Hunhevicz[1,2], Karsten Menzel[3], Catherine De Wolf[2]
[1]EMPA, Dübendorf, Switzerland
[2]ETH Zürich, Zurich, Switzerland
[3]Technische Universität Dresden, Dresden, Germany

Abstract

Insufficient as-built data hinders the transition of the architecture, engineering and construction (AEC) sector to a circular system. Combining reality capture and machine learning (ML) could help better detect reusable components. However, a comprehensive image dataset of on-site inventory for circular economy strategies has yet to be developed. This study introduces and describes the generation of a purpose-built, 360°dataset. Initial validation using the YOLOv8 object detection model demonstrates a 63.4% mean average precision (mAP50), making it viable for computer vision. Further exploration of automating building stock inventory using 360-degree images and ML for urban mining is needed.

Introduction

The architecture, engineering and construction (AEC) industry must change from a linear to a circular system to minimize its harmful impact and meet climate targets. Reusing resources is a key circularity strategy for reducing waste (European Council and European Parliament, 2008). In the built environment, this strategy implies the non-destructive recovery and reuse of building components according to their original purpose without a loss in their value (Hillebrandt et al., 2018). The extensive reuse of building components is currently challenged by the lack of sufficient data about material and substance composition and the as-built condition of the existing building stock (Çetin et al., 2021; Iacovidou et al., 2018; Uotila et al., 2021).

Urban mining projects need detailed (digital) information on their composition and dimensions to reuse and recycle building resources efficiently (Çetin et al., 2021; Honic et al., 2019; Uotila et al., 2021). This as-built data needs to be generated in advance, typically using reality-capture techniques such as light detection and ranging (Lidar) (Gordon et al., 2023; Xiong et al., 2022), 360°cameras (Gordon et al., 2023), and public digital imaging mining using Google Streetview (Raghu et al., 2023). Digital data for information extraction is typically processed further using photogrammetry for extracting dimensions (Gordon et al., 2023; Xiong et al., 2022) and machine learning (ML) for determining the material composition of the building stock (Raghu et al., 2023).

Among reality-capturing technologies, 360°cameras proved to be the most viable technology for capturing reliable information for deconstruction purposes in a comparative case study, because of their high accuracy and low noise (Gordon et al., 2023). Panoramic images have advantages over planar photos because of their comprehensive view of spaces, device cost-effectiveness, and quick data capture (Barazzetti et al., 2018; Gordon et al., 2023; Chou et al., 2020). Furthermore, 360°images are increasingly used for quick image analyses in computer vision applications (Barazzetti et al., 2018; Gordon et al., 2023; Chou et al., 2020). However, computer vision and 360°images have yet to be combined for circularity strategies.

This research examines whether combining building inventory using 360 °cameras with ML can be used to detect reusable building components. Yet since ML applications are only as good as the training data on which they are trained (Géron, 2023), the quality of the dataset is paramount. The relevance of training data can be defined as the extent to which it aligns with the data that the model is likely to encounter in the production phase (Witt, 2023). This concerns what is depicted and how it is annotated. The quality of the dataset refers to aspects influencing the generalization capabilities of a model to unseen data. The generated dataset should be suitable for the application in terms of representativeness, image quality, dataset size, and variance (Géron, 2023). For the envisaged ML application, a real-world dataset that simulates a survey of end-of-life (EoL) scenarios is necessary.

However, in a preliminary field study we found that existing panoramic datasets of buildings are not suitable for the above application for the following reasons: they are either synthetically generated, they are a compilation of different scenarios without any repetition, they are limited to specific parts of buildings (e.g., the façade), they are too homogeneous in quality to be captured on-site, or they have only very broad labels (e.g., doors are not subdivided by function into entrance doors, emergency exits, or room doors)

Therefore, our study takes a first step towards using 360°panoramic images in reusability assessment by generating and validating a 360°image dataset in a case study on the Technical University of Dresden (TU Dresden) campus. The custom dataset ("TUDataset") is preliminarily validated using the state-of-the-art YOLOv8 object detection model to identify building components. The validation explores whether the data collected during an on-site audit is qualitatively sufficient for object detection applications. It also explores the differentiation of component types within a component category. In so doing, this

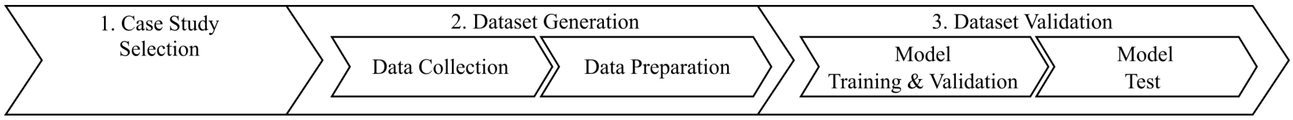


Figure 1: Research design of this study.

study gives valuable insights into the process of generating suitable 360°image datasets for building component reuse by discussing the chosen approach, potential for improvement, and next steps toward a full data-capturing ML pipeline.

Structure of the paper

The paper is structured according to the Figure 1. First, the following section describes the representative case study for EoL situations. It defines the data collection locations and the objects (i.e., the reusable building components) to be detected by the ML application. Second, the section "Dataset Generation" presents the approach and the image capture process to showcase how a 360°dataset can be created, and describes image processing for the specific use case of component reuse detection. Third, the "Dataset Validation" section describes how the approach is validated for computer vision. Experiments with the YOLOv8 object detection model verify the suitability of the dataset generation approach for the desired purpose of detecting reuse components. The methodology steps are presented in greater detail in the respective sections. Finally, the discussion reviews the contributions and further research steps toward a full pipeline of 360°image dataset creation for ML-based reuse detection.

Case Study Selection

As described in the introduction, the case study determines the relevance of the dataset for future input data from EoL buildings. The EoL refers to the final phase of a building's existence (after the production and use phase) in its current form (Hillebrandt et al., 2018). For this study, the case of "Technical University of Dresden (TU Dresden) campus" was chosen. It was designed to mirror urban mining projects by encompassing the entire building premises and emphasizing the frequent repetition of diverse building component designs within structures. It consists of five buildings on the TU Dresden campus that were selected according to the following considerations: around 32% of the residential and nonresidential buildings demolished in 2022 were built before 1948, 53% between 1949 and 1986, and only 14% between 1987 and 2010 (Statistisches Bundesamt, 2023). Accordingly, the case study was limited to the buildings on campus built between 1906 and 1990: two modern buildings, as well as three buildings of reform architecture from the early 20th century (see Table 1). Different EoL situations are presented: while the Beyer Building is partially gutted and currently unused, the Fritz-F¨ orster building is only partially in use due to renovation. The Nürnberger Ei and the Schumann building are contain operating offices and equipment (and, therefore, more data clutter).

Finally, only publicly accessible buildings were considered to comply with privacy compliance for publishing. Accordingly, the building was inventoried outside of business hours. People were asked to leave the field of view, and images with recognizable faces were removed during data cleaning.

Dataset Generation

While the relevance of the data was ensured in the design of the case study, the dataset quality (see Introduction) is addressed in the dataset generation process.

Data collection

A total of 1112 relevant images were captured for the case study (see 1). The data was generated from scratch using the specialized OpenExperience 360°camera helmet (Figure 2). The helmet has two installed 180°cameras, whose individual images are seamlessly merged into a spherical or panoramic view by a stitching algorithm. Capturing images in an onsite audit using a conventional 360°camera ensures that the data represents the expected input data. The final ML application is expected to generalize to input data collected under similar technical conditions.

The equirectangular projection (ERP) on a two-dimensional pane was used for the dataset. The ERP of the images has a resolution of 7000 × 3500 pixels, a horizontal and vertical resolution of 96 dpi, and a bit depth of 24 (see Figure 3). To address image variance deficiency observed in existing datasets, the data was captured at different times of the day and under varying weather conditions, resulting in different lighting and shading conditions. Furthermore, the images were taken without a fixed object distance, often capturing the same room from different positions and a person's viewpoint, generating a variance in the object's

Figure 2: DIGIBAU 360°helmet camera used in this study. Source: (OpenExperience, 2023).

Table 1: Data generation protocol on TU Dresden campus

Generation Date	Building	Location	Construction/ Renovation Year	#Images
30.03.2023	Fritz-Foerster-Bau	Mommsenstraße 6, 1069 Dresden	1926/2022	280
30.03.2023	Nürnberger Ei	Nürnberger Straße 31a, 01187 Dresden	1996	37
31.03.2023	Georg-Schumann-Bau	Münchner Platz 3, 01187 Dresden	1906	369
31.03.2023	Haus 116	August-Bebel-Straße 30, 01219 Dresden	1970/2013	159
14.04.2023	Beyer-Bau	George-Bähr-Straße 1, 01069 Dresden	1913	277
				1122

Figure 3: Unprocessed ERP image included in the TUDataset.

appearance, scale, and occlusion. Finally, the strongly varied resolution of the images resulted in different representations of the same object.

Dataset Preparation

The generated raw data was prepared using the Roboflow[1] online tool through data cleaning, annotation, set partition, and pre-processing and augmentation further described below.

Data Cleaning

In object detection, images or videos are used as inputs, and the features are extracted from the information in the pixels. The selected YOLO (You Only Look Once) model employs a feature extraction method that does not require the prior definition and cleaning of features. Therefore, the data cleaning was limited to eliminating duplicates and excluding low-quality and blurred images.

Data annotation

Object detection is a supervised machine learning application that requires a fully labeled dataset for the training. What is recognized and the granularity of differentiation is determined in the annotation process.

The EU policy fails to provide regulations and prerequisites for the reuse of components without prior (destructive) testing. Thus, to determine what could be reused within the case study, a field search including online reused material marketplaces (Concular, Restado. Bauteilnetz, Ebay Kleinanzeigen, SALZA, and Bauteilclick), practice reports, and guidelines was conducted to identify building components categories with assumed general or ”default” reuse potential.

Table 2: Selection criteria for component identification.

Criteria	Description
Relevance in practice	The building components should be found in large quantities in every building.
Post-processing	The component should be reusable per default, with little refurbishment and no testing needed.
Typology	The components should have standard features, but different configurations. Shapes and colors should differ
Indoor	The availability of data demands a restriction to indoor components.

Because this study is a proof-of-concept, the scope of the object detection model was set to recognize only a selection of reusable components. To fit the ML requirements on representativeness, quantity, and variance, the selection criteria in Table 2 were developed. Out of the pool of default reuse components, five categories adhere to the established selection criteria: doors, windows, radiators, sanitary objects, and lights. These component categories are considered reusable without further destructive testing. They present a large intra-class variation (e.g., many different types of windows) and were captured in large numbers in the case study.

The next step determines the granularity of differentiation within the selected component categories. Existing indoor datasets, such as the PanoContext Zhang et al. (2014), Indoor360 Chou et al. (2020) etc., differentiate only in rough component categories, such as doors and windows. However, function, material, and design are decisive for component reuse. The selected reusable component categories, ”windows, doors, sanitary objects, and electrical

[1]https://roboflow.com/

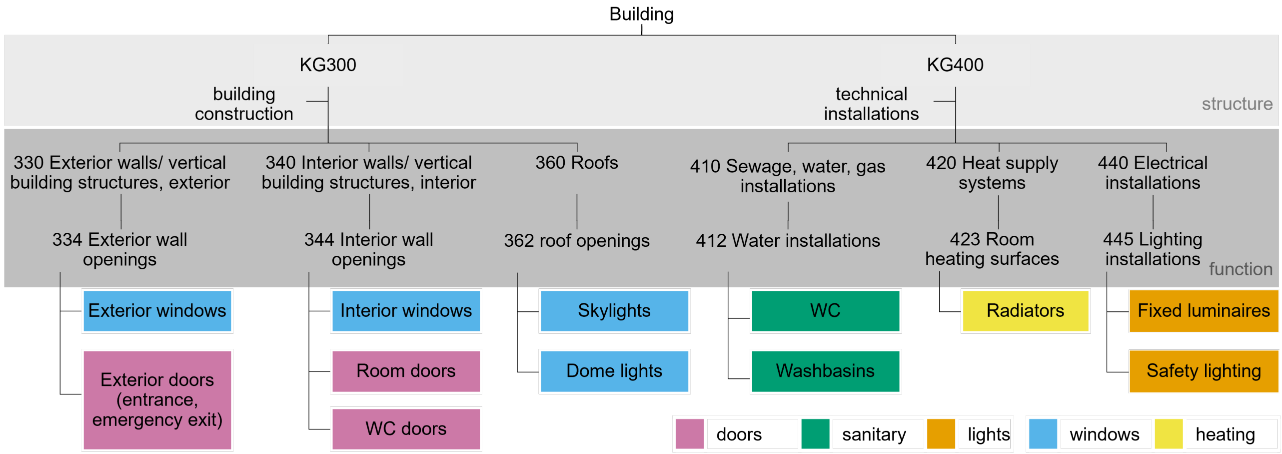


Figure 4: Categorization of reusable component selection according to DIN276.

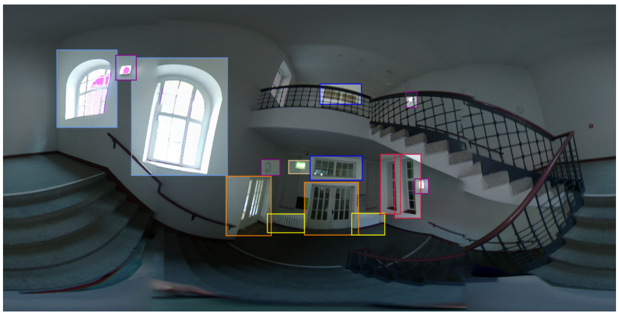


Figure 5: Each image label has a different bounding box indicating different component types.

installation", needed further differentiation to meet these requirements.

This study used the German AEC sector's building cost structure of the national standard DIN 276:2018-12 (Deutsches Institut für Normung, 2018) as a reference for the annotation. The component selection (windows, doors, plumbing, lights, and heating) was further differentiated into cost groups according to their associated building structure and function within the building or system, as seen in Figure 4. Categorization by cost group will not only standardize the work in different projects by reference to a public standard, but it could also allow for rapid cost determination for selective reconstruction.

In the annotation, the selected reusable component categories were set as super-categories, and the different cost groups were set as component classes. Furthermore, the component types were differentiated within these classes according to the material and design. For example, if the frame materials differed, two external glass doors with the same design would belong to different types. Hence, the labeling was structured according to component category, class (e.g., exterior door), and type (e.g., type 21). This resulted in 136 labels, ergo, different component types.

Dataset Partition

This study employed hold-out validation. It is an evaluation technique in which the dataset is split into three subsets: training, validation and test set. The training set is used to fit the models, the validation set is used to estimate the prediction error for model selection, and the test set is used to assess the generalization error of the final model configuration. The test data set is "held-out" and only used once, as reusing it can result in a substantial underestimation of the true test error. The TUDataset images were therefore divided into 70% training (792), 20% validation (226), and 10% testing (116) images.

Data pre-processing and augmentation

Applying pre-processing techniques to the training, validation, and testing sets ensures that the machine learning model learns and infers based on consistent image properties. Inference refers to the process of generating predictions.

First, the images were auto-orientated, removing the EXIF (Exchangeable Image File Format) data from images to ensure that they are displayed in the same manner as they are stored on the disk. Then, the image size was stretched to 640 × 640 pixels. Finally, the last step consists of data augmentation or training set expansion. This technique artificially increases the training set size and is a regularization method, reducing overfitting. To improve the model's tolerance to position, orientation, and size changes, the augmented instances should be as realistic as possible and ideally be indistinguishable from non-augmented instances by the human eye. Following Zhao et al. (2021), several augmentation techniques considering the particularity of the equirectangular projections were applied to the training and validation sets: the images were flipped horizontally and sheared $\pm$ 15°horizontally and $\pm$ 15°vertically both on the image level and bounding box level (see Figure 7). The outputs per training sample were set to 3, creating three altered images for every instance and resulting in a final dataset size of 2718 (3 × 792) images.

Dataset Validation

After generating the dataset, the suitability of the onsite 360°images for computer vision methods, and con-

(a) Scaling

(b) Horizontal flipping

(c) Image level shearing (+15°horizontally, +15°vertically)

(d) Image level shearing (15°horizontally, -15°vertically)

(e) Image level shearing (-15°horizontally, +15°vertically)

(f) Image level shearing (-15°horizontally, -15°vertically)

Figure 7: Different Augmentation Techniques

sequently for ML-aided reusability assessment, was validated using an object detection model.

The use of 360°images poses specific challenges for object detection. The projection of the spherical image to a plane distorts the objects depending on their distance and angle to the camera viewpoint. Integrating optimized distortion-aware convolution layers could handle these geometric deformations (Li et al., 2023). However, for the purpose of this proof-of-concept study, the conventional state-of-the-art object detection model YOLO (Redmon et al., 2016) was deemed appropriate. Different versions of this one-stage-detector have already been applied in the field of object detection in 360°images (Chou et al., 2020; Yang et al., 2018), as well as being frequently used as a benchmark model (e.g. in Zhao et al. (2019, 2020)). Due to constraints on Graphics Processing Unit (GPU) capacity, the Yolov8s model pre-trained on the COCO dataset (Lin et al., 2014) was selected.

Model Training

This study adopted an iterative training and validation process proposed by Géron (2023): the training configurations were tweaked based on the performance on an independent validation set to avoid overfitting, wherein a model becomes excessively optimized for the test set and fails to generalize to novel data. The model was evaluated on the test set only after selecting a final configuration. This iterative strategy is simplified in Figure 6. The training runs, and their validation and final testing were locally implemented in PyTorch using an NVIDIA RTX A4000 GPU.

The training process for the YOLOv8 model aims to minimize the training and generalization errors. In each run, the cost or loss function compares the predicted bounding box outputs with the actual outputs on the training set (training loss) and the validation set (validation loss). The generalization error is calculated as the difference between validation and training losses and is a key metric indicating the model's ability to perform well on new, unseen data.

The model's training performance, speed, and accuracy depend on using various hyperparameters and configurations that need manual configuration. This study adopted the validation approach proposed by Smith (2018) that examines the learning rate, batch size, momentum, weight decay, cyclical learning rates (cosine learning rate scheduler), and cyclical momentum. Accordingly, training and validation loss are analyzed during training to detect indications of underfitting and overfitting and determine the optimal combination of hyperparameters.

An initial configuration was used as a benchmark to adjust the training arguments, i.e., learning rate, batch size, momentum, weight decay, cyclical learning rates (cosine learning rate scheduler), and cyclical momentum. Nine configurations were trained and validated. The models were assessed by their mean average precision (mAP), precision, recall, and F1 score. A final configuration with the following hyperparameters was selected: a learning rate of 0.0007, weight decay of 0.001, the use of a cosine annealing learning rate scheduler, dropout regularization, and the Adam optimization algorithm.

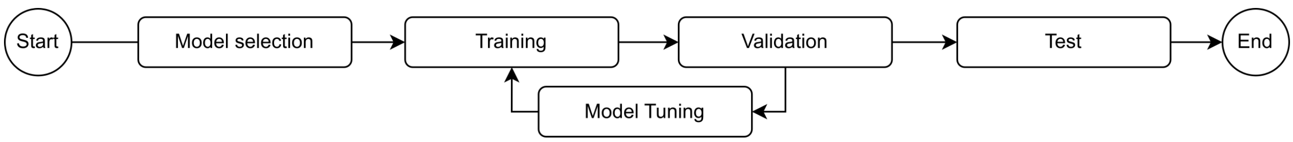


Figure 6: The applied iterative model training approach.

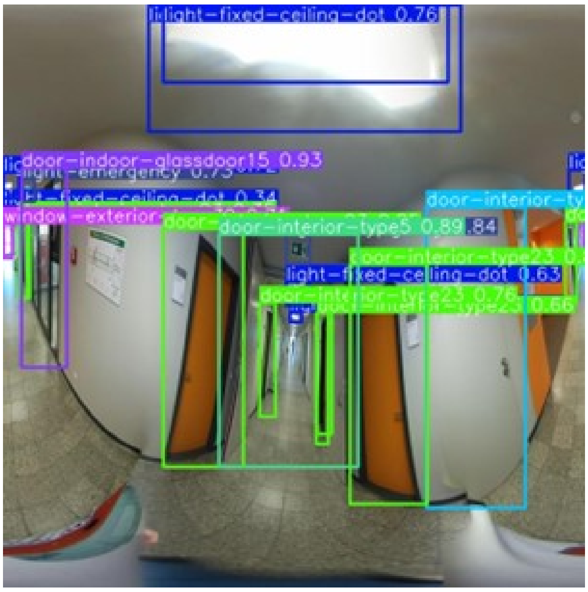


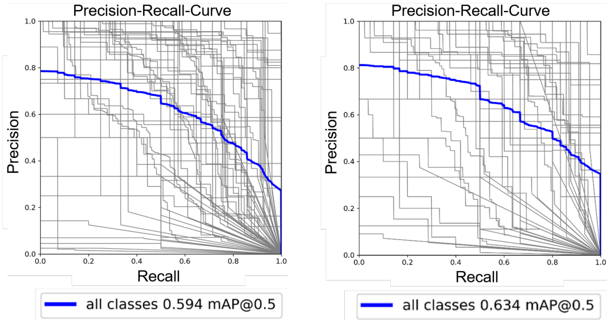
Figure 8: The model output with colored bounding boxes according to their label and superscribed with the confidence of the selected label.

Model Test

Finally, the chosen model configuration is tested on the hold-out test set, and its metrics are evaluated. The test run was performed on the test subset of 116 images and 975 instances. An example of the model's output in the test run can be seen in Figure 8. The model achieved an mAP at an intersection over union (IoU) threshold of 0.5 (mAP50) of 0.634, indicating a satisfactory overall performance in correctly detecting and localizing objects (see Table 3).
Furthermore, precision and recall have higher values than in the validation (see Figure 9). The PR curve reveals that a significant number of class labels exhibit a comparatively high PR score, indicated by the curves above the blue average PR score line, and a subset of classes with considerably inferior PR curves. This combination leads to an overall mAP@50 of 63.4% for all classes.
It indicates a model demonstrating relatively high precision at the beginning and overall good performance in precision and recall trade-off. In the class-wise evaluation of the model, in the worst-performing classes, zero instances were detected, resulting in low precision, recall, and mAP scores. This indicates that the model does not effectively detect and localize instances of these classes. These classes may have visual attributes that are difficult for the model to distinguish, leading to recognition errors. Furthermore, insufficient training data for these classes could also contribute to poor performance, as the model is not sufficiently exposed to different examples for effective learning.

Table 3: Test set performance metrics.

Class	Images	Instances	Box(P	R	mAP50	mAP50-95)
all	116	975	0.721	0.596	0.634	0.371



(a) Precision-Recall-Curve in training
(b) Precision-Recall-Curve in test-run

Figure 9: Performance metrics of the test run.

The diagonal in the confusion matrix (Figure 10) suggests that most classes are confidently predicted. However, the outliers indicate that some classes are incorrectly assigned to another specific class with high confidence. This can be recognized by a single point in a row and the coloring of the point in the graph. For these classes, more images or higher quality images are needed so that the model extracts features that lead to differentiation from the other component types more efficiently. Certain classes, such as class 135 (window-interior-type8) and class 136 (window-interior-type9), lack fundamental distinguishing features, which is visible in the erroneous assignment to very different component types (doors, lights, windows).

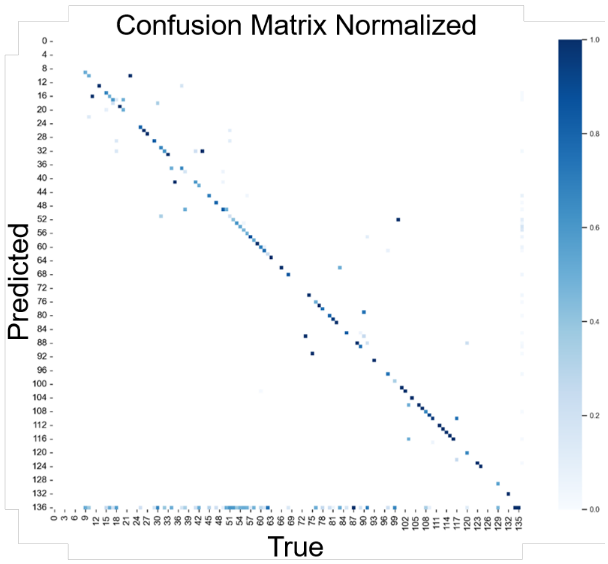


Figure 10: Normalized confusion matrix in test-run. Class IDs 1–52 = door types, 53–59 = lights, 60–64 = sanitary objects, 65–136 = window types.

Discussion

The implemented YOLOv8 object detection model shows that satisfactory accuracy can be achieved using conventional object detection models on equirectangular projections for differentiating component types within a class. The test metrics indicate effective assignment to building component categories (see Fig. 4), with most errors occurring in differentiating component types within the same component class. For instance, misclassified windows are mostly labeled as another window type rather than a different class, like doors or lamps. The model's overall performance indicates that the differentiation into different component types needed for the reusability assessment is possible. However, refining misclassifications, minimizing false positives and negatives, and augmenting training data for complex classes would improve its accuracy in recognizing component types.

A limitation of this study is the component-level annotation, which is time-consuming and, thus far, only includes component types from the case study. Using the object detection model trained on the TUDataset for other buildings will face challenges in detecting new building component types. Therefore, for subsequent research, the dataset annotation will be edited to be broader by eliminating further typification according to design and material. Only the categories of the DIN276 (see Figure 4 should be considered: e.g., exterior windows, skylights, dome lights, exterior doors (entrance, emergency), interior doors (room doors, WC doors), toilets, sinks, radiators (vertical, horizontal), lights (fixed luminaires), and safety lights. Combining existing labels into broader ones will introduce further intra-class variation, promising a better generalizing model and ensuring applicability to new buildings and EoL situations. While a simplified annotation process is considered to facilitate the up-scaling of the proposed approach to further building inventories, future studies to assess and validate scalability to other building contexts and larger datasets are needed.

This study aimed to detect the component category and type. However, more as-built information, such as the components' dimensions, material, and condition, should be extracted for reuse assessment. Therefore, photogrammetry and further feature extraction will be explored in upcoming research.

Furthermore, this study only used the object detection model to validate the approach. Further research will focus on the functionality and performance of the model. Different established object detection models should be compared in terms of robustness in dealing with large datasets and detection performance for the final model selection. Furthermore, various approaches integrating spherical layers into convolutional neural networks have been proposed in research (Cohen et al., 2018; Li et al., 2023; Zhao et al., 2019). Considering these additional models is promising for effectively addressing the distortions in ERP images and potentially improving the detection performance.

Finally, some limitations stem from the case study selection. First, resource constraints required a limitation on the included architectural styles, potentially compromising the generalizability of the results to buildings from different eras. Therefore, the selection criteria should be reviewed to integrate diversity of architectural styles and building functionalities, and the TUDataset should be continuously adapted to represent the variety of existing building stock. Second, the selection of component categories examined in this research represents only a small subset, and a broader range of non-destructive, non-toxic component categories could potentially be reintegrated into a circular economy. Third, it is based on the assumption that all areas and spaces within a building are accessible. Consequently, only the components captured by the camera can be identified. Therefore, the data introduces uncertainties in the number of components similar to those in traditional on-site inspections and component identification processes.

Conclusion

The main contribution of this research is the introduction of an approach to generate well-suited 360°image datasets for circular component reuse. The selected case study, centered around five buildings on the TU Dresden campus, served as a representative scenario for end-of-life sites. The dataset generation process involved using a specialized OpenExperience 360°camera helmet, capturing 1112 relevant images at different times of the day and under varying weather conditions. Data preparation, including cleaning, annotation, set partition, and pre-processing, was conducted to ensure the technical and relevance requirements of datasets.

The resulting fully-labeled TUDataset consists of approximately 2,400 panoramic images in equirectangular projection and 136 object classes, facilitating research and practical applications in object detection. Data quality and relevance requirements were addressed in the case study's design and during the dataset generation process. The YOLOv8 object detection model validated the real-world dataset, achieving a mAP50 of 0.634 on the test set.

The workflow developed in this research can be extended to other reusable building components, thereby contributing to developing new cost-effective and scalable approaches for building component recovery. Exploiting machine learning and 360°images for the inventory and reuse assessment could save time and costs, making circular strategies more competitive.

Overall, this study demonstrates that 360°images generated on site are suitable for effectively detecting different building component classes (windows, doors, sanitary, lamps, radiators), and provides the foundation for the next steps toward automated reusable component detection.

Acknowledgements

We thank the Technical University of Dresden for facilitating this study's data generation and implementation.

References

Barazzetti, L., Previtali, M., and Roncoroni, F. (2018). Can we use low-cost 360 degree cameras to create accurate 3D models? pages 69–75. Copernicus GmbH.

Chou, S.-H., Sun, C., Chang, W.-Y., Hsu, W.-T., Sun, M., and Fu, J. (2020). 360-Indoor: Towards learning real-world objects in 360° indoor equirectangular images. Proceedings - 2020 IEEE Winter Conference on Applications of Computer Vision, WACV 2020. Publisher: Institute of Electrical and Electronics Engineers Inc.

Cohen, T. S., Geiger, M., Koehler, J., and Welling, M. (2018). Spherical CNNs. Proceedings of the International Conference on Learning Representations.

Deutsches Institut für Normung (2018). DIN 276 Kosten im Bauwesen.

European Council and European Parliament (2008). Directive 2008/98/EC of the European Parliament and of the Council of 19 November 2008. Official Journal of the European Union, (L312).

Gordon, M., Batallé, A., Wolf, C., Sollazzo, A., Dubor, A., and Wang, T. (2023). Automating building element detection for deconstruction planning and material reuse: A case study. Automation in Construction, 146:104697.

Géron, A. (2023). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow. O'Reilly, Beijing; Boston; Farnham, third edition, oktober 2022 edition.

Hillebrandt, A., Riegler-Floors, P., Rosen, A., and Seggewies, J.-K. (2018). Atlas Recycling. Edition Detail. Detail Business Information GmbH, München, er ste auflage edition.

Honic, M., Kovacic, I., and Rechberger, H. (2019). Improving the recycling potential of buildings through Material Passports (MP): An Austrian case study. Journal of Cleaner Production, 217:787–797. Publisher: Elsevier BV.

Iacovidou, E., Purnell, P., and Lim, M. K. (2018). The use of smart technologies in enabling construction components reuse: A viable method or a problem creating solution? Journal of Environmental Management, 216:214–223.

Li, M., Meng, M., and Zhou, Z. (2023). RepF-Net: Distortion-Aware Re-projection Fusion Network for Object Detection in Panorama Image. In Wang, L., Gall, J., Chin, T.-J., Sato, I., and Chellappa, R., editors, Computer Vision – ACCV 2022, pages 508–523, Cham. Springer Nature Switzerland.

Lin, T.-Y., Maire, M., Belongie, S., Bourdev, L., Girshick, R., Hays, J., Perona, P., Ramanan, D., Zitnick, C. L., and Dollár, P. (2014). Microsoft COCO: Common Objects in Context.

OpenExperience (2023). DIGIBAU 360° Helmkamera.

Raghu, D., Bucher, M. J. J., and Wolf, C. (2023). Towards a 'resource cadastre' for a circular economy – Urban-scale building material detection using street view imagery and computer vision. Resources, Conservation and Recycling, 198:107140.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December. Publisher: IEEE Computer Society.

Smith, L. N. (2018). A disciplined approach to neural network hyper-parameters: Part 1 – learning rate, batch size, momentum, and weight decay. US Naval Research Laboratory Technical Report.

Statistisches Bundesamt (2023). Weniger Abriss: 2022 fielen so wenige Wohnungen aus dem Bestand wie noch nie seit 1992. Pressemitteilung N050, Statistisches Bundesamt, Deutschland.

Uotila, U., Saari, A., and Junnonen, J.-M. (2021). Investigating the barriers to laser scanning implementation in building refurbishment. Journal of Information Technology in Construction, 26:249–262.

Witt, J. (2023). How to Improve Your Computer Vision Model.

Xiong, Z., Gordon, M., Byers, B., and Wolf, C. (2022). Reality capture and site-scanning techniques for material reuse planning. page 88.

Yang, W., Qian, Y., Cricri, F., Fan, L., and Kamarainen, J.-K. (2018). Object Detection in Equirectangular Panorama.

Zhang, Y., Song, S., Tan, P., and Xiao, J. (2014). PanoContext: A Whole-Room 3D Context Model for Panoramic Scene Understanding. pages 668–686. Springer, Cham.

Zhao, P., You, A., Zhang, Y., Liu, J., Bian, K., and Tong, Y. (2020). Spherical Criteria for Fast and Accurate 360° Object Detection. Proceedings of the AAAI Conference on Artificial Intelligence, 34(07):12959–12966.

Zhao, Q., Chen, B., Xu, H., Ma, Y., Li, X., Feng, B., Yan, C., and Dai, F. (2021). Unbiased IoU for Spherical Image Object Detection.

Zhao, Z. Q., Zheng, P., Xu, S., and Wu, X. (2019). Object Detection with Deep Learning: A Review. IEEE Transactions on Neural Networks and Learning Systems, 30(11):3212–3232. Publisher: Institute of Electrical and Electronics Engineers Inc.

Çetin, S., Wolf, C., and Bocken, N. (2021). Circular Digital Built Environment: An Emerging Framework. Sustainability, 13(11):6348.

BENCHMARK FOR TOPOLOGICAL AND SPATIAL ASSESSMENT OF INDOOR RESIDENTIAL BUILDINGS

Abdullah Elsafty[1], Timo Hartmann[1]
[1]Civil System Engineering, Technische Universität Berlin, Germany
a.elsafty@tu-berlin.de

Abstract

Our paper introduces a new benchmark for evaluating indoor residential buildings, addressing the lack of standardization in building assessment methods. Utilizing laser-scanned point clouds, we capture diverse architectural styles, enhancing accuracy in maintenance, navigation, and safety planning. Our approach, underpinned by precise reference models and metrics, allows for comprehensive method comparisons. The accompanying datasets ensure reproducibility and wide applicability. This benchmark, available publicly, fosters collaborative research and advances building assessment techniques, ultimately improving the resilience and sustainability of residential structures.

Introduction

In the domain of building assessment, understanding, and analyzing indoor residential environments is crucial. These environments, central to daily human experiences, are characterized by a myriad of spatial and topological complexities (Zhang *et al.*, 2022).

Accurate spatial analysis is vital for effective space management, enhancing living comfort and functionality, while topological assessment is crucial for efficient navigation, particularly in emergencies, to ensure safety and accessibility. The point cloud technology marks a significant leap in accurately representing these environments. With its ability to provide precise 3D representations, point cloud offers an unprecedented level of detail and accuracy(Cheng *et al.*, 2018). This advancement has facilitated more realistic modeling of indoor environments, impacting fields ranging from architectural design to emergency planning (Gankhuyag and Han, 2020; Guo *et al.*, 2021).

Despite its importance, the field of building assessment faces significant challenges, primarily due to a lack of comprehensive and precise benchmarks specifically tailored for residential buildings. Existing methods and datasets often fall short in capturing the unique complexities and diversity inherent in residential environments (Chen *et al.*, 2023). This gap in accurate and scalable representation and assessment tools hinders advancements in architectural design, maintenance, and safety planning, ultimately impacting the efficacy of residential space utilization.

Our research addresses this critical gap by introducing a novel benchmark for the topological and spatial assessment of indoor residential buildings. Utilizing advanced laser-scanned point cloud datasets, this benchmark is meticulously crafted to capture diverse architectural styles and layouts of residential environments. The primary objective is to provide a robust and versatile tool that enhances the accuracy and comprehensiveness of building assessments, thereby facilitating effective maintenance, navigation, and safety planning in residential spaces.

In developing this benchmark, we employed a multifaceted approach that integrates various scanning technologies and environmental conditions. This approach ensures a broad and realistic representation of residential spaces, reflecting the complexities encountered in real-world scenarios. The benchmark's methodology emphasizes the precision, accuracy, and chromatic fidelity of the point cloud data, making it a highly reliable resource for evaluating building assessment methodologies.

Our primary goal is to bridge the existing research gap by offering a new database and robust evaluation metrics. This benchmark, featuring diverse point cloud data, is designed to advance algorithms tailored for residential settings, contributing to more efficient and accurate building maintenance, safety planning, and architectural design, ultimately enriching residential living spaces. This paper is structured as follows: Section 2 provides a detailed review of related work addressing the study foundation. Section 3 discusses the benchmark setup, from data acquisition to evaluation. Finally, Section 4 concludes the paper, summarizing our contributions and suggesting directions for future research.

Point of Departure

In this section, we discuss existing datasets and benchmarks in the field of residential indoor building assessment, specially from laser scanning. This frames our study within the broader landscape of architectural technology and highlights the value of our research focus. In the exploration of indoor datasets, various collections have emerged, each with unique attributes and limitations. ISPRS dataset (Khoshelham *et al.*, 2017) emerges as a notable collection, focusing on indoor modeling through point clouds, while rich in indoor point cloud data from varied environments like universities and a fire brigade office, presents certain limitations for our research focused on residential environments. The dataset's emphasis on institutional settings differs markedly from the complexities inherent in residential spaces, making it less applicable for studies centered on residential layouts. Similarly, the ScanNet dataset (Dai *et al.*, 2017) and the SceneNN dataset (Cheng, Meng and Sun, 2020) have made strides in capturing a range of indoor scenes. However, most of their scans are constrained to one or two rooms. This limitation significantly hinders their applicability in the domain of scalability, where a broader understanding of entire residential units is essential.

The ETH3D dataset [13], for instance, includes a modest collection of 16 indoor scans. While valuable for multi-view stereo applications, its limited scope and primary focus on stereo rather than 3D point-cloud processing render it less suitable for our scope. Contrasting with real-world data, the SUNCG dataset (Song *et al.*, 2017) offers a synthetic alternative, presenting a wide array of indoor scenes with CAD-quality geometry and annotations. However, the synthetic nature of SUNCG means it cannot fully capture the complexity and unpredictability of real-world residential settings, a crucial element for authentic floorplan generation.

Diverging from these, Floor-SP (Chen *et al.*, 2019) and Matterport3D (Yadav *et al.*, 2022) bring RGBD image sets. Matterport3D dataset brings forth high-quality panorama RGBD image sets encompassing 90 luxurious houses. While offering rich visual data, its format and focus do not align with the specific needs of laser scanning. In a similar vein, the Stanford 2D-3D-S dataset (Armeni *et al.*, 2016) provides large-scale scans of office spaces using the same Matterport camera technology. Despite its scale, the dataset's emphasis on 2D and 3D semantic annotations does not address the nuanced requirements of residential environments.

Among these, the Floor-NET dataset (Liu, Wu and Furukawa, 2018) stands out with its provision of full floorplan annotations and corresponding RGBD videos from smartphones for 155 residential units. Despite its closer alignment with residential applications, its format and the mode of data collection may pose challenges in scalability and accuracy when applied to broader residential scenarios. In addition to these datasets, numerous ad-hoc datasets have been developed (Fang *et al.*, 2021; Tang *et al.*, 2022; Kim and Lee, 2023), each tailored to specific tasks or applications. While these datasets contribute valuable insights to their respective areas, their specialized nature often limits their utility in a broader context, particularly in studies focused on scalability and general applicability in residential environments.

This examination of various indoor datasets underscores a prevalent issue: while numerous datasets provide valuable insights into indoor environments, they notably lack in addressing the comprehensive needs of residential building representation. A key challenge is the absence of datasets offering extensive scans of real-world residential spaces, including furnishings, a gap often attributed to privacy concerns. Moreover, the tendency for many datasets to be ad-hoc or overly specialized limits their utility in research focused on scalability and broad applicability in diverse residential settings. Considering these limitations, our research introduces a novel benchmark utilizing laser scanning (point cloud) technology. This benchmark is specifically designed to capture the intricate details and complexities of residential buildings, addressing the critical need for a more accurate and scalable approach in the topological and spatial assessment of indoor residential environments. This advancement not only fills a significant void in the current research landscape but also sets a new standard for precision and comprehensiveness in the field of building assessment.

Benchmark Setup

Benchmark dataset

In constructing this benchmark, we implemented a wide selection process for point cloud data, adhering to meticulously defined criteria to ensure a comprehensive and accurate depiction of a variety of residential buildings. This evaluation focused on key factors such as precision, accuracy, chromatic fidelity, and the granularity of details within individual data points.

We also prioritized diversity in data acquisition methodologies, incorporating a range of scanning technologies and environmental conditions to authentically mirror complex, real-world data collection scenarios. Moreover, the curation of point cloud content was strategically executed to encompass a wide array of residential spaces, each with their unique architectural styles and spatial layouts. Table 1 delineates these point cloud specifications and criteria. First, we classify the criteria into three main categories: point criteria (PCr.), scan condition (SCd.), and scan content (SCt.). then, accuracy, colors, number of points in millions (No.), and sensor type were added to the point criteria. Acquisition complexity (AC) and moving objects (MO) were added to scan condition category. Finally, building type (BT) and furniture (F) were added to the scan content. This systematic approach in dataset compilation not only addresses the shortfall in publicly available datasets but also establishes our benchmark as a model of versatility and practical applicability in residential building assessment.

Table 1:Point cloud data criteria

		D1	D2	D3	D4	D5
PCr.	Acc.	6-8 mm				1 cm
	Color	✓	✓	x	✓	✓
	No.	316	321	319	6.6	252
	Sensor	TLS / MLS		TLS	MLS	TLS
SCd.	AC	✓	✓	x	✓	✓
	MO	x	x	x	x	✓
SCt.	BT	H	H	H	D	A
	F	✓	✓	x	✓	✓

Our dataset includes a curated selection of five residential environments, each representing different living settings and contexts. It features datasets from a multi-floor house (D1, D2), offering an urban residential perspective with challenges like clutter, occlusions, and elaborate interior designs.

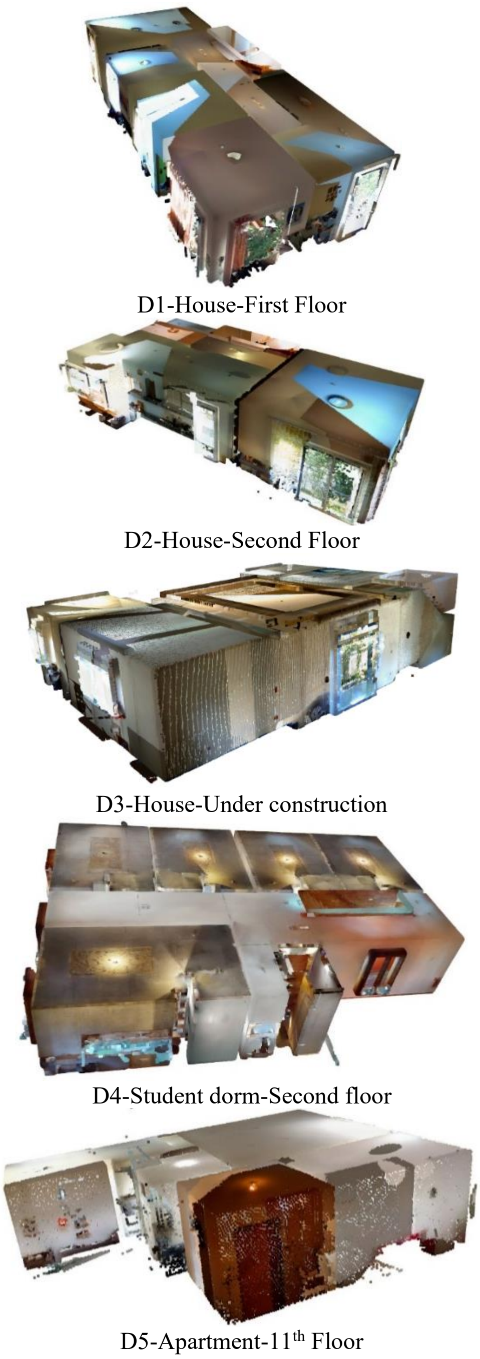


Figure 1:Point cloud data screenshots.

An additional dataset from a fully furnished apartment (D5), situated on the 11th floor, showcases the intricacies of high-rise living environments. We also included a ground floor dataset from a house under construction (D3), highlighting the complexities of unfinished architecture. Complementing these, a dataset from the second floor of a fully furnished student dormitory (D4) adds to the architectural and contextual diversity. Figure 1 provides a visual representation of these diverse point cloud data.

For data acquisition, we employed both Terrestrial Laser Scanning (TLS) and Mobile Laser Scanning (MLS) methodologies, tailored to the specific characteristics of each residential setting in our dataset. The Trimble X9 TLS system, renowned for its extensive range and high precision, was used, achieving an exceptional accuracy of 8 mm, particularly in interior scans. In contrast, for the dormitory dataset (D4), we utilized the MLS capabilities of the iPhone’s LiDAR sensor, which offers suitable navigation and accuracy (±1 cm) within apartment spaces (Luetzenburg, Kroon, and Bjørk, 2021). Notably, datasets D1 and D2 were captured using a combination of TLS and MLS technologies, leveraging the comprehensive range of the Trimble X9 and the agility of the iPhone's LiDAR sensor. This dual-technology approach for D1 and D2 ensures a richly detailed representation of the intricacies inherent in residential spaces.

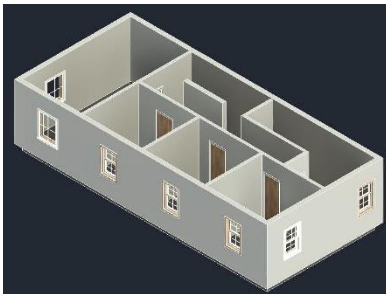
Table 2 provides detailed technical characteristics of these two scanning devices. By integrating both TLS and MLS methods, our benchmark accurately reflects the varied challenges and scenarios present in residential buildings, demonstrating the efficacy and adaptability of diverse laser scanning technologies in different residential contexts. Contrasting with many existing benchmarks that typically engage in preprocessing to remove noise and reflections, our benchmark adopts a novel approach. We emphasize the importance of preserving and incorporating indoor noise elements within each residential space. This decision stems from our understanding that noise, particularly in indoor environments, is an inherent characteristic of real-world data. It offers valuable insights into the everyday use and dynamics of residential spaces.

Table 2: Laser scanners technical characteristics

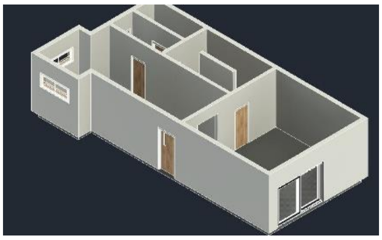
Specs.	TLS	MLS
Range	120 m	3 m
Accuracy	6 – 8 mm	± 1 cm
Acquisition time (one floor)	60 min	3 min

Reference models

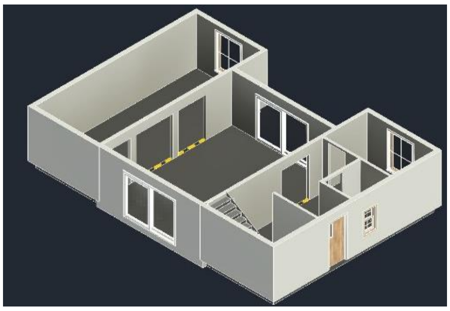
In the development of our benchmark, a fundamental component was the construction of models using Autodesk Revit™. This step involved the detailed creation of 3D models that are essential in residential buildings. While our primary focus was on the 3D models themselves, these models inherently facilitate the generation of crucial elements like adjacency graphs, navigation plans, and evacuation paths. As depicted in Figure 2, these models provide a comprehensive foundation for understanding and analyzing the spatial and topological aspects of residential environments. By focusing on 3D modeling, we ensure that the essential characteristics of residential buildings are accurately captured, laying the groundwork for subsequent analytical processes that are vital in assessing and understanding the dynamics of these spaces.



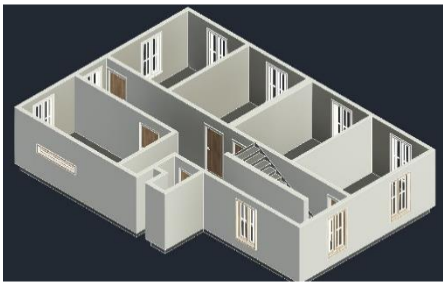
D1



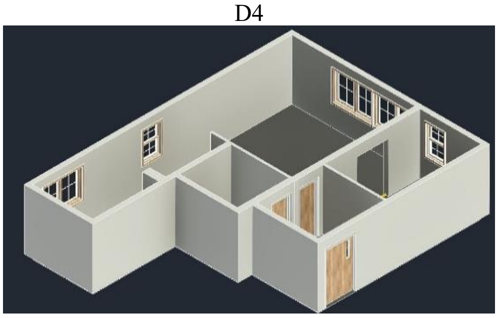
D2



D3



D4



D5

Figure 2: 3D reference models.

In the geometric reconstruction phase, we undertook a series of systematic procedures. These included adjusting levels, constructing walls, and other architectural elements, and integrating doors, windows, and circulation components. Notably, this reconstruction process did not rely on manual on-site measurements but solely on the precision inherent in the point cloud data. This strategy highlights the effectiveness of modern scanning technologies in accurately capturing indoor residential environments. By depending exclusively on point cloud data, our reference models not only demonstrate the capabilities of advanced scanning technologies in building assessment but also challenge the traditional dependency on labor-intensive manual measurement methods.

Additionally, it is crucial to address the concept of 'ground truth' within the context of point cloud processing. Typically, 'ground truth' refers to a model manually constructed in software, representing an 'as-built' or 'as-is' scenario. While these models are highly accurate, they are not without limitations. In our benchmark, these models are utilized more as a comparative standard rather than an absolute ground truth. They provide a robust and reliable foundation for evaluation but are acknowledged as not being infallible. This nuanced understanding of 'ground truth' is fundamental in our approach, ensuring that our benchmark provides a realistic and practical standard for evaluating the performance of various methodologies in the realm of residential building assessment.

Evaluation metrics

In our benchmark, specifically designed for topological and spatial assessments in residential environments, we incorporate a range of tasks including area measurement, floorplan generation, adjacency graph construction, navigation, and evacuation planning. Recognizing the integral role of geometric evaluation in understanding topological and spatial data (Pintore *et al.*, 2020), our metrics also encompass the assessment of geometric accuracy in reconstructed models. This approach ensures a thorough evaluation of algorithm performance beyond just computational efficiency and automation level, with a focus on the quality of algorithms in relation to indoor residential environments.

For geometric accuracy assessment, we utilize the Root Mean Square Error (RMSE), a well-established statistical measure in spatial data analysis (Li *et al.*, 2017). RMSE quantifies the variance between predicted and observed spatial dimensions, thereby evaluating the precision of area measurements in a reliable and quantifiable manner.

In evaluating floorplan generation, our benchmark employs accuracy metrics: Recall, Precision, and F1-score. These metrics, prevalent in pattern recognition and information retrieval, offer a multidimensional perspective on the accuracy and completeness of generated floorplans (Kim and Lee, 2023).

To address the evaluation of topological relations and spatial assessments, such as adjacency graph construction, navigation, and evacuation planning, we adopt a dual approach. **Qualitatively**, usually experts conduct a visualization evaluation, subjectively assessing the clarity, accuracy, and usability of the representations (Knauff, Rauh and Renz, 1997). **Quantitatively**, we implement graph similarity measures, like graph edit distance, to evaluate topological relations. This measure calculates the minimum number of edits needed to transform one graph into another, thus gauging the fidelity of the algorithm-generated graphs compared to our reference models. Additionally, spatial accuracy metrics should be employed to assess deviations in spatial configurations, and consistency checks are conducted to ensure the logical coherence of spatial arrangements. This combination of subjective and objective evaluations forms a comprehensive assessment framework, capturing both the perceptual quality and the measurable accuracy of spatial models in residential environments.

Accessibility and Participation

In line with our commitment to collaborative scientific progress, our benchmark will be made publicly accessible. We invite scholars, practitioners, and members of the scientific community to engage with this dataset, which will be available for download from the Deposit once Repository. This open-access initiative is designed to stimulate comprehensive utilization and analysis of the dataset, thereby contributing significantly to advancements in the field.

We strongly encourage scholarly engagement from researchers interested in applying and investigating various methodologies and algorithms with our database. This invitation extends to a diverse spectrum of researchers from both academic and industrial spheres. We advocate for the exploration of a wide range of application scenarios, including emergency response planning and architectural design. The insights gained from these explorations are expected to be invaluable in enhancing our understanding of current algorithmic capabilities and limitations and could lead to innovative developments in the field.

We are introducing a set of guidelines for its effective use and implementation. These are designed to bridge the gap between theoretical research and practical application, providing clear, actionable steps for accessing and utilizing the benchmark. Users can navigate the point cloud raw data and reference Revit files via the repository, enabling straightforward integration into their projects. For each specific application, practitioners are encouraged to refer back to this paper for the relevant proposed evaluation metrics to ensure accurate and effective application. This commitment to support aims to foster a collaborative environment, encouraging the widespread adoption and application of the benchmark across various professional domains.

Conclusion

This research has established a new benchmark for the topological and spatial assessment of indoor residential buildings, an area that has been inadequately represented in existing datasets. By incorporating a variety of point cloud data acquired through laser scanning, this benchmark serves as a crucial tool for the precise evaluation of algorithms related to this domain. Our benchmark, anchored in constructed reference models, provides a solid criterion for both geometric and topo-spatial assessments. This addresses significant gaps in existing methodologies and sets a new standard in the field. Looking forward, this benchmark opens the door to a multitude of research avenues and practical applications, marking a significant advancement in the field of spatial data analysis. Its potential extends across various domains, from enhancing emergency response strategies to innovating architectural design and urban planning. A key direction for our future work will involve conducting real-world case studies to present and validate the practical applicability and effectiveness of our benchmark across a diverse range of architectural styles and environments. we will illustrate how the application of our benchmark can lead to tangible improvements in building performance, such as increased energy efficiency, improved occupant well-being, and reduced environmental impact, thereby fulfilling our vision of fostering more sustainable and resilient living environments. The benchmark is poised to be a cornerstone in the development of more accurate and efficient algorithms, fostering interdisciplinary collaborations that blend expertise from architecture, urban planning, data science, and beyond. Additionally, it offers a platform for integrating emerging technologies like augmented reality for immersive visualizations and AI to develop advanced evaluation metrics, we aim to set new standards in spatial analysis, ensuring our benchmark remains at the forefront of architectural assessment and urban planning. While our dataset encompasses a variety of architectural styles, we aim to expand its diversity by incorporating more settings, featuring unique design elements and circulation patterns, to enhance its applicability. Also, by integrating socio-cultural aspects, recognizing the impact of social norms and cultural values on residential architecture. This inclusion will allow our benchmark to evaluate and guide the design of spaces that are not only physically sound but also resonate with the cultural and social fabric of their communities more accurately. In addressing the limitations of our benchmark, we acknowledge challenges in the acquisition of point cloud data for complex residential indoor environments, like furniture, which may not fully capture the intricacies of such spaces. However, this journey also invites us to confront and navigate the challenges and limitations inherent in such pioneering work, ensuring continuous improvement and relevance in an ever-evolving landscape.

Acknowledgments

We acknowledge the support of the ReconTOP project, and the funding provided by the Deutsche Forschungsgemeinschaft (DFG), which were crucial in the realization of this research.

References

Armeni, I. et al. (2016) '3D semantic parsing of large-scale indoor spaces', Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-Decem, pp. 1534–1543.

Chen, J. et al. (2019) 'Floor-SP: Inverse CAD for floorplans by sequential room-wise shortest path', in Proceedings of the IEEE International Conference on Computer Vision, pp. 2661–2670.

Chen, X. et al. (2023) 'ReCo: A Dataset for Residential Community Layout Planning', in Proceedings of the 31st ACM International Conference on Multimedia. New York, NY, USA: Association for Computing Machinery (MM '23), pp. 397–405.

Cheng, J., Meng, M.Q.-H. and Sun, Y. (2020) 'Robust Semantic Mapping in Challenging Environments', Robotica. 2019/05/21, 38(2), pp. 256–270.

Cheng, L. et al. (2018) 'Registration of laser scanning point clouds: A review', Sensors (Switzerland). MDPI AG.

Dai, A. et al. (2017) 'ScanNet: Richly-annotated 3D reconstructions of indoor scenes', Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua, pp. 2432–2443.

Fang, H. et al. (2021) 'Floorplan generation from 3D point clouds: A space partitioning approach', ISPRS Journal of Photogrammetry and Remote Sensing, 175, pp. 44–55.

Gankhuyag, U. and Han, J.H. (2020) 'Automatic 2D floorplan CAD generation from 3D point clouds', Applied Sciences (Switzerland), 10(8).

Guo, Y. et al. (2021) 'Deep Learning for 3D Point Clouds: A Survey', IEEE Transactions on Pattern Analysis and Machine Intelligence. IEEE Computer Society, pp. 4338–4364.

Khoshelham, K. et al. (2017) 'The ISPRS benchmark on indoor modelling', in International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives. International Society for Photogrammetry and Remote Sensing, pp. 367–372.

Kim, M. and Lee, D. (2023) 'Automated two-dimensional geometric model reconstruction from point cloud data for construction quality inspection and maintenance', Automation in Construction, 154(March), p. 105024.

Knauff, M., Rauh, R. and Renz, J. (1997) 'A cognitive assessment of topological spatial relations: Results from an empirical investigation BT - Spatial Information Theory A Theoretical Basis for GIS', in S.C. Hirtle and A.U. Frank (eds). Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 193–206.

Li, G. et al. (2017) 'Quality evaluation of point cloud model for interior structure of a common building', International Geoscience and Remote Sensing Symposium (IGARSS), 2017-July, pp. 6040–6043.

Liu, C., Wu, J. and Furukawa, Y. (2018) 'FloorNet: A unified framework for floorplan reconstruction from 3D scans', in Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 203–219.

Pintore, G. et al. (2020) 'State-of-the-art in Automatic 3D Reconstruction of Structured Indoor Environments', Computer Graphics Forum, 39(2), pp. 667–699.

Song, S. et al. (2017) 'Semantic scene completion from a single depth image', Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017-Janua, pp. 190–198.

EXPLAINABLE ARTIFICIAL INTELLIGENCE IN GENERATIVE DESIGN FOR CONSTRUCTION

Sanaz Zarghami[1], Hanieh Kouchaki[2], Longzhi Yang[1] ,and Pablo Martinez Rodriguez[1]
[1]Northumbria University, Newcastle, United Kingdom
[2]Tabriz Islamic Art University, Tabriz, Iran

Abstract

As artificial intelligence rapidly advance, their growing complexity enables more sophisticated applications across sectors, including construction. However, the opaque nature of algorithms, such as generative AI, reduces human interpretability and trust. While providing benefits like enhanced efficiency, generative design's black-box processes hamper adoption. Explainable AI can elucidate how AI algorithms generate outputs, thereby improving understanding and confidence. Despite explainable AI's potential, construction has given it limited focus. This research systematically reviews the application of explainable AI in generative design in construction, with an aim to allay risks and enable wider utilization of these emerging technologies for improved engineering design.

Introduction

The expanding integration of artificial intelligence (AI) in the field of architecture and construction, driven by the exponential growth in data, is reshaping traditional practices. The manual analysis of vast datasets and reliance on rule-based computing methods pose challenges, prompting the adoption of AI for systematic data analysis through predictive modeling. This transformation influences various facets of the industry, including architectural and structural design, construction safety, sustainability, affordability, speed, return on investment, and operational performance. Generative design, a departure from traditional methods, empowers computers to semi-autonomously explore design spaces, presenting designers with diverse options for analysis and consideration (Baduge at al., 2022; Junk and Burkart, 2021; Krish, 2011).

While the adoption of AI in construction is gaining recognition, challenges arise in understanding and interpreting AI model outputs, often considered "black boxes". Concerns about bias, fairness, trust, and reliability, particularly in critical domains such as recruitment, real-time progress monitoring, cyber-security, risk management, and safety, warrant attention. Human decision-making in these domains is also susceptible to bias, and the reluctance to embrace AI is often rooted in a lack of understanding. Establishing trust in AI models, crucial for widespread acceptance, is explored through explainable artificial intelligence (XAI). This involves methodologies and processes to enhance comprehension and confidence in the outcomes and outputs of AI algorithms, addressing the industry's need for transparency and reliability (Matthews et al., 2022; Gunning et al., 2019; Sokol et al., 2022; Love et al., 2023). While XAI has gained traction in fields like law and medicine , its exploration in construction remains limited despite the rise of generative AI.

This study aims to fill this gap by providing a thorough exploration of XAI for generative design in construction. Through a systematic review, this paper advocates for generative AI in construction, raises awareness of XAI limitations and gaps, and harnesses its advantages. To achieve this, the study provides comprehensive answers on how XAI tackles transparency and bias, differentiates generative AI, elucidates the mechanisms of XAI for generative design in construction, and examines how XAI influences the generative design framework in this industry.

Methodology

The methodology employed in this study aims to provide a comprehensive analysis of the research objectives. In essence, this paper begins by elucidating the fundamentals of AI followed by a systematic review of 31 academic papers on generative design and the emerging field of generative AI design. In essence, drawing from the mentioned comprehensive review, this section has been concluded by presentation of a categorization of generative design techniques, methods, and models. Subsequently, the concept of XAI is introduced and to enhance the clarity, terminologies are provided, outlining its significance and methodologies. This section derives from an analysis of 30 academic papers to establish a taxonomy for the preliminary stages of applying XAI. Moving forward, the paper conducts a critical review of the construction industry, focusing on the potential applications of generative AI within this domain. Illustrative examples of generative processes in architectural design further enrich the discussion. Finally, the paper proposes a comprehensive taxonomy titled "Explainable Generative AI Design," defining various approaches to render generative design algorithms interpretable and transparent. The mentioned taxonomy was developed by exploring major types of GENAI models based on their mechanisms. They were then categorized according to AI models. Finally, post-hoc expandability techniques were provided for these GENAI models. Each step of the methodology is meticulously created to provide a comprehensive understanding of the intersection between generative AI and construction, elucidating the theoretical underpinnings and practical implications for future research and applications.

Artificial Intelligence

Over the past decade, the construction sector has lagged in adopting transformative AI advancements that have revolutionized other industries. Notably driven by machine learning and its specialized subset, deep learning, these breakthroughs harness historical data to predict patterns, enabling automated generative design and analysis processes. Deep learning techniques, exemplified in Baduge et al. (2022), facilitate automated structural evaluation and design, emphasizing prestressed components. AI in construction promises generative and optimized layouts, safety risk prediction and mitigation, energy cost reductions, accelerated project timelines, enhanced financial returns, and sustainable building practices. The replicative capabilities of AI address feasibility and scalability challenges posed by big data in the construction sector. Despite being in an experimental stage, the rapid evolution of AI applications in construction, supported by expanding research and consistently impressive results, anticipates widespread integration and commercial viability in the near future (Love et al., 2023; Angelov et al., 2021).

Generative Design and Generative AI

The architectural design process is conventionally divided into three integral phases: the initial conceptual or scheme design, the detailed preliminary design incorporating optimization, and the creation of construction drawings. The conceptual phase holds paramount significance as it molds the ultimate design and heavily relies on the designer's expertise. Addressing prevalent issues in structural design, such as inefficiencies, data underutilization, and knowledge transfer challenges, has prompted a growing emphasis on smart design strategies. Leveraging historical design data, architectural concepts, and various forms of knowledge, the evolving approach employs generative AI to efficiently produce new designs (Kanyilmaz et al., 2022). The GD process comprises three components: a design schema, creating variations, and selecting desirable outcomes (Krish, 2011).

Generative design, rooted in facilitating innovation, utilizes algorithmic software driven by parameters set by designers, including materials, geometric shapes, and loads. The incorporation of artificial intelligence into generative design processes utilizes metaheuristic search algorithms, such as genetic algorithms, to explore and identify efficient solutions within predefined design frameworks. This transformative approach comprises three core elements: a generative geometry model, criteria defining design objectives, and a metaheuristic search algorithm. Generative AI has emerged as a pivotal technological development, prominently featured in intelligent design to overcome inefficiencies in the structural design industry (Avital, 2007; Junk and Burkart, 2021; Liao et al., 2024). This article delves into data feature representation, intelligent structural design creation, and robust design outcome assessment within the realm of generative AI for architectural design.

Singh's explorations in 2012 of generative design techniques, detailed in Table 1, encompasses SG, CA, GA, LS, and SI, considering their technical, design, and system development characteristics. Integrating these computational systems poses challenges due to the necessity for a common language across diverse generative design (GD) techniques. Singh envisions an interactive expert system, placing the designer at the core of the GD process. Meanwhile, Krish (2011) proposes generative design methods (see Figure 1 and outlined at the top of Table 1). Contemporary research emphasizes adapting generative AI models, employing deep learning methods like CNNs, GNNs, and RNN/LSTM for domain-specific problem-solving, as elucidated by advanced algorithms in the last section of the table (Liao et.al, 2024). Generative AI holds the potential to revolutionize data management in architecture, engineering, and construction (Ghimire et al., 2023). However, its transformative impact in the construction industry is hindered by issues, including a gap between research and industry, hesitancy due to interpretability, feasibility, reliability concerns, and ethical considerations (Arrieta et al., 2020).

Explainable Artificial Intelligence

Explainable artificial intelligence (XAI) is a collection of methodologies designed to enable human users to understand and trust the outcomes of artificial intelligence systems. The objective of any XAI method is to make the decision-making of AI models transparent and interpretable. Anand et al. (2020) have persuasively argued that AI algorithms often resemble inscrutable "black boxes" due to their intricate, nonlinear structures that are hard to elucidate, even for their creators. The inability to fully understand machine learning and deep learning models entails risks when relying on their unexplained outputs. XAI has thus been developed to help users grasp the rationale behind the specific outputs of AI systems. The diminished explainability of AI models has led Gerling et al. (2021) to assert that the inability to elucidate algorithms means we cannot contest, validate, enhance, or derive knowledge from them. Therefore, generating explanations that clarify the functioning of DL and ML models, as well as the basis for their decisions or predictions, is crucial for fostering trust in these technologies. Arrieta et al. (2020) have characterized XAI to provide insights or justifications that render the operations of a model transparent and easily comprehensible. While a consensus on the definition remains elusive, it is widely accepted that XAI systems ought to articulate their capabilities, recount their past actions, describe their current activities, predict future outcomes, and disclose the pertinent data upon which they operate. To achieve these ends, XAI addresses concerns related to interpretability and transparency.

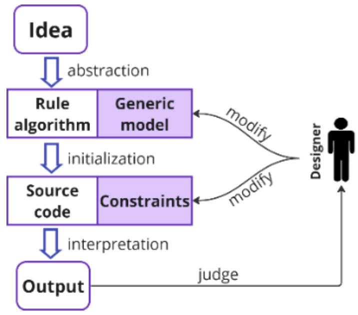


Figure 1: Generative design process (Krish, 2011)

Table 1: Generative design methods and techniques and generative AI models.

Ref	Category						
Krish. 2011	**Generative Design Method (GDM)**						
	Genotype	**Phenotype**	**Exploration envelope**	**Design Table**	**Design Generation Software**	**CAD system**	**Performance filters**
	Is composed of a generic parametric CAD model, list of design parameters and their initial value and initial exploration envelope.	Generated CAD files (that may include build history, built-in relationships and built-in equations).	A list of minimum and maximum values of the driving parameters, specifying the limits of the design space to be explored.	A data table that stores the driving design parameters, their initial values and limits.	It generates random variations of the driving design parameters within limits set by the exploration envelope.	Is a parametric CAD engine with a transparent and editable build history.	A pass/fail software filter that is able to evaluate the performance of generated designs based on preset performance criteria.
Singh. 2012	**Generative design techniques across their technical, design and system development characteristics**						
	Cellular automata (CA)	**Genetic algorithms (GA)**	**Shape grammars (SG)**	**L-systems (LS)**	**Swarm intelligence (SI) and multi-agent societies**		
	A grid of cells that change over time based on predefined rules influenced by the states of adjacent cells.	Simulate natural selection by evolving a population of design solutions using a fitness function.	Rules for manipulating shapes to produce a variety of designs.	Create structures resembling natural growth patterns, resulting in self-similar forms.	Collective behavior of simple agents that interact with their local surroundings. Leading to the formation of organized, global patterns.		
Ghimire et.al. 2023	**Major types of GenAI models based on their generative mechanism**						
	Generative Adversarial Networks (GAN)		**Variational Auto Encoders (VAE)**		**Autoregressive models**	**Diffusion Models**	**Flow-based Models**
	Two neural networks, a generator, and a discriminator, compete with each other to generate realistic data.		Encodes data into a latent space and then decodes it back into the original space.		Generate data one step at a time, using the previously generated data as input.	Start with a noisy image and gradually refine it to a realistic image.	Transform data from one distribution to another using a series of invertible functions.

XAI is a complex concept encompassing intrinsic or post hoc nature, local or global scope, and model-specific or model-agnostic applicability (see Figure 2).

Defining XAI definitively demonstrates challenging due to context-dependent interpretability, varying across domains, as illustrated in Figures 2 and 3. Although these terms are frequently used interchangeably, they may convey distinct nuances in various settings. To enhance clarity, the distinctions and commonalities among terminologies frequently employed within the realms of ethical AI and XAI are elucidated in this section:

- Understandability refers to a model's inherent quality that enables a human to grasp its functionality—how the model operates—without necessitating an exposition of its internal architecture or the algorithmic processes it utilizes to manage data internally.

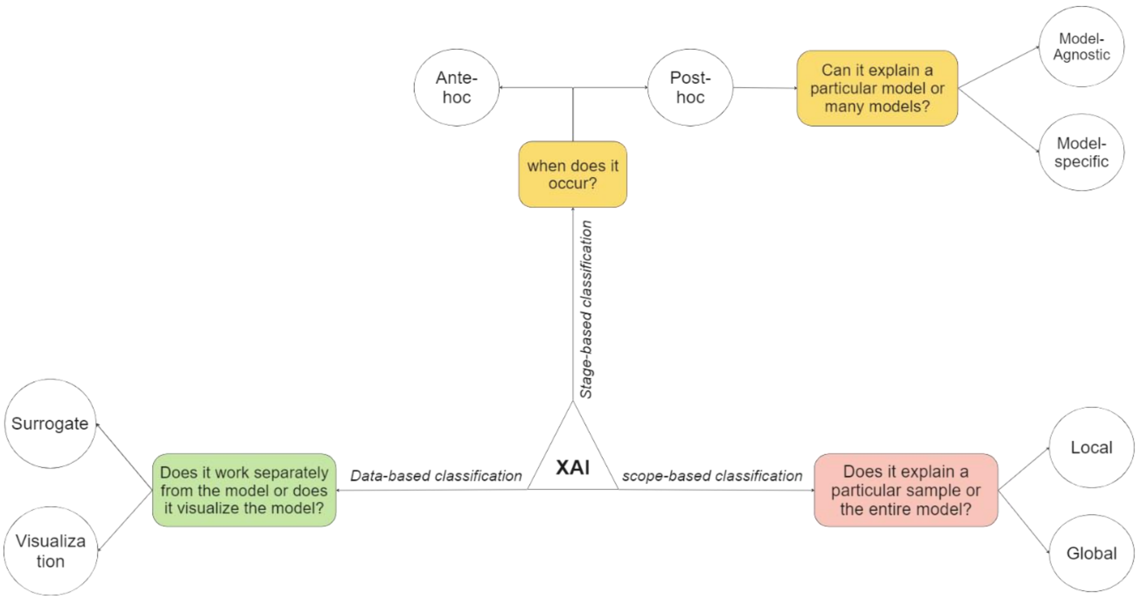


Figure 2. Preliminary stages in application of XAI (Source: Authors).

- Comprehensibility, when applied to machine learning models, pertains to the capacity of a learning algorithm to present its acquired knowledge in a manner that is intelligible to humans.
- Interpretability is described as the facility to convey significance in terms accessible and clear to a human observer.
- Explainability is crucial, acting as a bridge between humans and decision-making entities. It involves post-hoc methods to make opaque models intelligible. In this document, we prioritize explainability as the primary design goal, recognizing its comprehensive importance.
- Model transparency hinges on inherent comprehensibility, categorized into simulation models (step-by-step emulation), decomposable models (examining components and interconnections), and algorithmically transparent models (visible internal logic and decision-making processes).

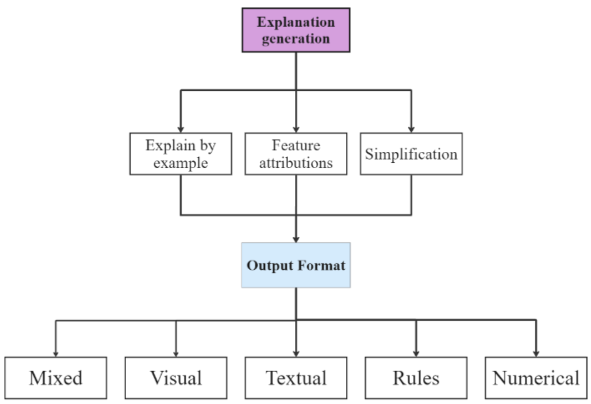


Figure 3: Various formats of XAI results (Source: Authors).

Explainable generative AI design in construction

As AI becomes increasingly integrated into applications that directly impact humans and as the ramifications of algorithmic decisions grow, the construction industry must broaden its focus beyond merely the predictive accuracy of ML and DL models to include their explainability (Love et al., 2023). Generative AI has the potential to transform data management in architecture, engineering, and construction, offering efficient processing of unstructured data through large language models (LLMs). A few examples can be found in Figure 4. Generative AI fosters seamless information exchange between tangible and digital realms in construction projects, enhancing comprehension and decision-making in project management by accessing previously inaccessible unstructured data (Ghimire et al., 2023). As the authors mention in their article, the potential of Gen AI in construction is:

1. To extract project information from construction documents such as dimensions, materials used, responsible person, point of contact, etc.
2. To generate new documents. Examples- proposals, reports, etc.
3. To classify & cluster documents based on project types, internal departments, sub-contractors, project phases, document types, materials, supply chain, etc.
4. Generating code to automate tasks.
5. Optimizing cost estimation workflow.
6. To help quality control by comparing completed tasks to project specifications to identify defects and deviations.
7. To generate an optimal schedule path.

Construction organizations have been reluctant to adopt artificial intelligence (AI) systems, not because of unwillingness, but due to the lack of explainability of autonomous decisions and actions made by AI, which undermines confidence and trust (Love et al., 2023). XAI offers significant benefits for construction, including reduced model bias, enhanced trust, actionable insights, and risk reduction. XAI elucidates the rationale behind AI decisions and capabilities (Love et al., 2023). Explanations can facilitate collaborative co-design among stakeholders, empowering them to make informed decisions about system use and appropriate timing for performance assessments (Khosravi et al., 2022). Consequently, evaluating XAI requires comprehending diverse socio-political environments of stakeholders using an interpretive approach that enables deeper understanding and generates motivation and commitment (Love et al., 2022).

As it is depicted in Figure 5, there are various approaches that can be selected for implementing XAI solutions tailored to generative design algorithms, contingent upon the differing characteristics of the models and their associated computational processes. In the following section, different types of XAI approaches that can be used to make generative design algorithms explainable are presented.

- **Local Explanations:** It focusses on explaining individual predictions or decisions made by the AI model. In the context of generative design, this could involve explaining why certain design choices were made for a specific instance. Techniques such as Local Interpretable Model-agnostic Explanations (LIME) or SHapley Additive exPlanations (SHAP) can be used to provide these local insights.
- **Global Explanations:** Unlike local explanations, global explanations aim to provide an understanding of the model's overall behavior. For generative design algorithms, this could involve explaining the general principles or rules the model uses to generate designs across different scenarios.
- **Feature Attribution Methods**: These methods explain the output of a model by attributing the prediction to the input features. For generative design, this could mean identifying which features of the input data are most influential in the design generated by the AI.
- **Visual Explanations:** Visualization techniques can be used to illustrate how the generative model is transforming input data into a design. This could include visualizing the activation of different layers in a neural network or showing intermediate steps in the design process.
- **Example-Based Explanations:** Providing examples of similar designs or design decisions made by the AI can help users understand the rationale behind the generated designs. This could include showing nearest neighbors or variations of the design that were considered by the algorithm.
- **Feature Attribution Methods**: These methods explain the output of a model by attributing the prediction to the input features. For generative design, this could mean identifying which features of the input data are most influential in the design generated by the AI.
- **Visual Explanations:** Visualization techniques can be used to illustrate how the generative model is transforming input data into a design. This could include visualizing the activation of different layers in a neural network or showing intermediate steps in the design process.
- **Example-Based Explanations:** Providing examples of similar designs or design decisions made by the AI can help users understand the rationale behind the generated designs. This could include showing nearest neighbors or variations of the design that were considered by the algorithm.
- **Counterfactual Explanations:** Such explanations provide insights into how the input could be changed to achieve a different outcome. In generative design, counterfactuals can help users understand how altering certain design parameters could lead to different design results.

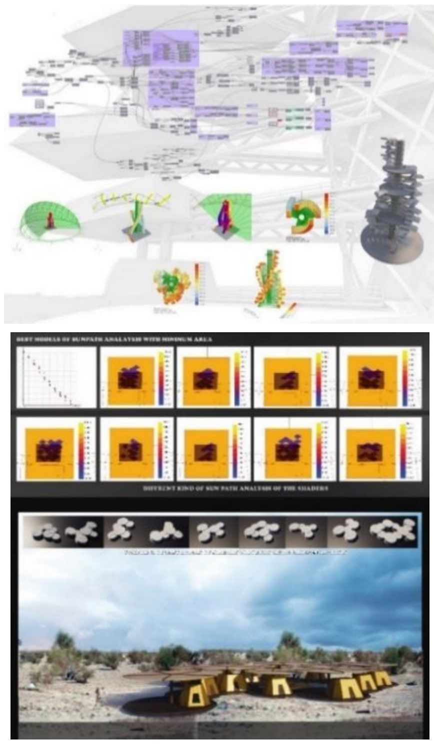


Figure 4: Examples of generative design process in the initial conceptual or scheme design (Source: Authors).

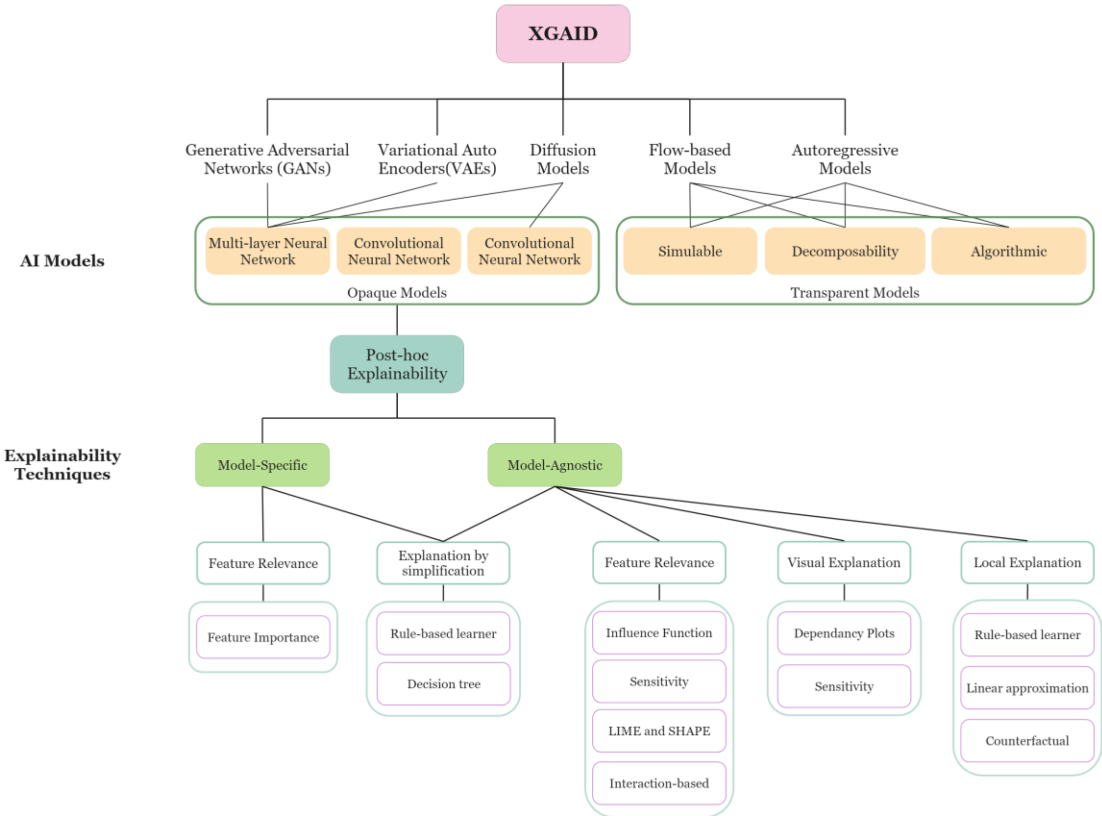


Figure 2: Proposed explainable generative AI design taxonomy (Source: Authors).

- **Interactive Explanations:** These tools allow users to explore and manipulate the AI model's inputs and observe the changes in the outputs. This hands-on approach can be particularly effective in helping users understand the generative process and the impact of their design choices.
- **Model Simplification:** Simplifying complex models into more interpretable forms can also serve as an XAI approach. For instance, a complex generative design model could be approximated by a simpler model that captures the main trends and patterns, making it easier to understand and explain.
- **Causal Inference:** Causal inference methods aim to understand the cause-and-effect relationships within the data that the AI model uses. In generative design, this could involve identifying which elements of the design are causing certain features or outcomes in the generated product. Techniques like causal diagrams or intervention analysis can help in providing these insights.
- **Process Tracing:** This approach involves tracing the decision-making process of the AI model step by step. For generative design, this could mean detailing the sequence of operations the model performs to arrive at a particular design, which can help users understand the generative process in a more granular way.
- **Natural Language Explanations:** Some XAI systems can generate textual descriptions that explain the AI's decisions in human-readable form. For generative design, the AI could provide annotations or descriptions of why it made certain design choices, which can be more intuitive for users without technical expertise.
- **Sensitivity Analysis:** This involves studying how changes in the input affect the output of the AI model. In generative design, sensitivity analysis can help determine which design parameters are most influential and how variations in these parameters can alter the final design.
- **Uncertainty Quantification:** Providing information about the uncertainty and confidence of the AI's decisions can be a form of explanation. In generative design, this could mean indicating the level of certainty the AI has about each aspect of the design it generates, which can help users make informed decisions about whether to accept, modify, or reject the AI's suggestions.
- **Rule Extraction:** Some XAI approaches aim to extract human-understandable rules from complex models. For generative design, this could involve distilling the AI's decision-making

process into a set of rules or guidelines that explain how different design outcomes are achieved.

Also, there is the option to make hybrid approaches by including more than one XAI algorithm at once. Hybrid approaches aim to leverage the strengths of both ante-hoc and post-hoc explainability. For example, a model may be trained with regularization techniques that promote sparsity, making it easier to identify which features are most influential, while still allowing for post-hoc analysis to explain individual predictions. Another hybrid strategy might involve training a complex model to achieve high performance and then using a simpler, inherently interpretable model to approximate the complex model's behavior in a way that is understandable to humans (Sun et al. 2022, Amershi et al.2014, Ross et al. 2021).

Discussion

While AI has been implemented across various industries, those systems cannot yet fully replace construction managers, engineers, and clients in building projects due to concerns around replacing human judgment with computational processes. However, this paper proposes a framework integrating explainable AI techniques into generative design workflows, which could improve process clarity for users. By increasing trust in AI through explainable methods, these systems may gain wider acceptance in construction. However, within current literature, evaluation frameworks for explainable AI remain underdeveloped, as the notion of explainability lacks consensus.

The proposed taxonomy aims to provide researchers and practitioners with a simple yet effective way of evaluating the XAI approaches available for the GD AI algorithm used in a systematic manner. The selection of the XAI algorithm from an effective manner has not yet been evaluated and remains a work in progress.

Nonetheless, basic tenets like interpretability and transparency are agreed upon. Although explainability has received limited focus in AI research on construction generative design, future work could employ existing post-hoc explanation methods to elucidate the functioning of machine and deep learning models. Qualitative approaches may also evaluate explanations as a precursor to developing quantifiable metrics. Moreover, while this paper presented a broad taxonomy for most generative AI models in design, more specific taxonomies for all model types require investigation. The relationship between individual mechanisms and their combination can be separately considered and presented. Additionally, balancing model performance and explainability poses challenges. More complex models like deep neural networks can generate higher quality outputs yet tend to be less interpretable, whereas simpler models may be more explainable but have limitations in producing novel optimized designs. Hence, hybrid approaches leveraging both complex and simple models warrant exploration.

Conclusion

This paper proposes a taxonomy for the integration of explainable AI (XAI) into generative design, with the aim of increasing user trust and paving the way for future AI applications in the construction industry. XAI enables transparency and accountability, building trust in AI-enabled construction solutions. The adoption of XAI is necessary to maintain a competitive edge and comply with evolving regulatory mandates. By merging XAI with generative algorithms, design quality, efficiency, and stakeholder satisfaction can be enhanced. However, realizing XAI's full potential requires investment in professional training and research refining XAI for construction-specific deployment. In conclusion, explainable AI has the potential to revolutionize construction design through human-centered generative practices and a unified commitment to its progression.

References

Amershi, S., Cakmak, M., Knox, W.B. and Kulesza, T., (2014). Power to the people: The role of humans in interactive machine learning. *Ai Magazine*, *35*(4), pp.105-120.

Anand, K., Wang, Z., Loog, M. and van Gemert, J., (2020). Black magic in deep learning: How human skill impacts network training. *arXiv preprint arXiv:2008.05981*.

Angelov, P.P., Soares, E.A., Jiang, R.M., Arnold, N.I., and Atkinson, P.M. (2021). Explainable Artificial Intelligence: An analytical review. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 11, 5, Article e1424 (2021). https://doi.org/10.1002/widm.1424.

Arrieta, A.B., Díaz-Rodríguez, N., Ser, J.D., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities, and challenges toward responsible AI. Information Fusion 58, 82–115. https://doi.org/10.1016/j.inffus.2019.12.012.

Avital, M., (2007). Innovation through generative systems design. In Cleveland, OH, USA, NSF Science of Design Workshop–technical report.

Baduge, S.K., Thilakarathna, S., Perera, J.S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A. and Mendis, P., (2022). Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*, *141*, p.104440.

Gerlings, J., Shollo, A., and Constantiou, I. (2021). Reviewing the need for explainable artificial intelligence (xAI). Proceedings of the 54th Hawaii International Conference on System Science, 5th-8th January, On-line. https://doi.org/ 10.24251/HICSS.2021.156.

Ghimire, P., Kim, K. and Acharya, M., (2023). Generative AI in the Construction Industry: Opportunities & Challenges. arXiv preprint arXiv:2310.04427.

Gunning, D., Stefik, M., Choi, J., Miller, T., Stumpf, S. and Yang, G.Z., (2019). XAI—Explainable artificial intelligence. *Science robotics*, *4*(37), p. eaay7120. https://doi.org/10.1126/scirobotics.aay7120.

Junk, S. and Burkart, L., (2021). Comparison of CAD systems for generative design for use with additive manufacturing. Procedia CIRP, 100, pp.577-582.

Kanyilmaz, A., Tichell, P.R.N. and Loiacono, D., (2022). A genetic algorithm tool for conceptual structural design with cost and embodied carbon optimization. *Engineering Applications of Artificial Intelligence*, *112*, p.104711. doi: https://doi.org/10.1016/j.engappai.2022.104711.

Khosravi, H., Shum, S.B., Chen, G., Conati, C., Tsai, Y.S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S. and Gašević, D., (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, *3*, p.100074. https://doi.org.10.1016/j.caeai.2022.100074.

Krish, S., (2011). A practical generative design method. Computer-Aided Design, 43(1), pp.88-100, Doi: 10.1016/j.cad.2010.09.009.

Liao, W., Lu, X., Fei, Y., Gu, Y. and Huang, Y., (2024). Generative AI design for building structures. Automation in Construction, 157, p.105187. https://doi.org/10.1016/j.autcon.2023.105187.

Love, P.E.D., Fang, W., Matthews, J., Porter, S., Luo, H., Ding, L., (2023). Explainable artificial intelligence (XAI): Precepts, models, and opportunities for research in construction. Advanced Engineering Informatics 57, 102024. https://doi.org/10.1016/j.aei.2023.102024.

Matthews, J., Love, P.E.D. Porter, S. and Fang, W. (2022). Smart data and business analytics: A theoretical framework for managing rework risks in mega projects. International Journal of Information Management, 65,102495, doi.org/10.1016/j.ijinfomgt.2022.102495.

Ross, A., Chen, N., Hang, E.Z., Glassman, E.L. and Doshi-Velez, F., (2021), May. Evaluating the interpretability of generative models by interactive reconstruction. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).

Singh, V. and Gu, N., (2012). Towards an integrated generative design framework. Design studies, 33(2), pp.185-207. DOI: http://dx.doi.org/10.1016/j.destud.2011.06.001.

Sokol, K., and Flach, P. (2022). Explainability is in the beholder's mind: Establishing the foundations of explainable artificial intelligence. ArXiv: 2112.14466v2.

Sun, J., Liao, Q.V., Muller, M., Agarwal, M., Houde, S., Talamadupula, K. and Weisz, J.D., (2022), March. Investigating explainability of generative AI for code through scenario-based design. In *27th International Conference on Intelligent User Interfaces* (pp. 212-228).

TOWARDS A SOLUTION FOR THE PRE-FAILURE CONCRETE ELEMENT CRACK DETECTION PROBLEM

Karolina Tomaszkiewicz[1], Tomasz Owerko[1]
[1]AGH University of Krakow, Poland

Abstract

Crack detection in concrete bridge elements is critical to the bridge's durability and safety. The ability to link cracks with the type of damaged element, location, and the moment of occurrence is critical for understanding the structure's behaviour. While the detection of cracks in concrete representing the failure condition is currently a relatively straightforward task, the identification of narrow cracks representing the pre-failure state has not yet received a satisfactory solution. This paper discusses a solution for segmenting structural elements on images and segmenting cracks using a deep learning network trained on a prepared dataset of pre-failure concrete cracks.

Introduction

The occurrence of cracks in reinforced concrete structures is a natural phenomenon, which in many cases cannot be avoided. In order to be able to assess whether the occurrence of cracks may affect the durability of the structure and pose a risk to its safety, it is necessary to determine the probable cause of the damage, based on factors such as the nature of the element's work, the place where the damage occurred or the time at which the damage occurred (e.g. after performing load tests). At the same time, the formation of cracks can be caused by the effects of both mechanical and environmental factors, as well as being the result of design or construction mistakes. It should also be noted that according to Eurocode 2 (CEN, 2004) the crack limit value (i.e. the value above which the serviceability limit state of the structure is exceeded) is a very small value and is 0.3mm.

For this reason, during the engineer's site inspection, these cracks may be difficult to detect by human eye, particularly in areas that are difficult to access such as pylons, spans over rivers and the inside of box girders. At the same time, the engineer inspecting the technical condition of the structure by reviewing individual elements, assessing the size of the damage, and documenting with photographs is the person who is responsible for assessing the impact of the damage on the structure's safety, and who decides on the necessary repair works or taking the structure out of service. The inspection process is labour-intensive, time-consuming and, in particular, is subject to the engineer's subjective opinion and experience. It has been estimated that approx. 50% of condition assessments are incorrect or vary depending on the engineer who performs the inspection (Hüthwohl et al., 2019).

This paper presents a solution based on machine learning algorithms that can support the engineer's work and condition assessment of structures. This solution allows for the detection of bridge structure elements on images and then the detection of cracks for a particular structure element in relation to the location of the damage. Particularly, by training a neural network based on a dataset of cracks representing the pre-failure state of the structure, the presented solution can be used to detect damage at an early stage of its occurrence. Detecting the cracks of reinforced concrete or asphalt structures representing the failure state is a relatively straightforward task in the current state of the art (Ali et al., 2022; Li et al., 2022; Nguyen et al., 2023), but identifying the pre-failure state indications in the form of very narrow cracks has not yet received a satisfactory solution. By diagnosing the occurrence of a crack at the pre-failure stage and linking it to a possible cause based on where the damage occurred, it is possible to take the necessary repair steps (e.g. protecting the damage), to monitor the development of the damage and, consequently, to extend the service life of the structure.

The first part of this paper discusses the causes of cracks in reinforced concrete elements of bridges and indicates for which elements a given cracking caused by a given factor is characteristic. The second part discusses the authors' dataset "NCCD-PF - A pre-failure narrow concrete cracks dataset for engineering structures damage classification and semantic segmentation" (Tomaszkiewicz & Owerko, 2023b). The segmentation of concrete elements of a bridge structure for which cracks are detected is presented in the third part. The fourth part, in turn, presents the solution of a deep learning algorithm trained based on the discussed dataset, for the identified structural elements.

The problem of cracks in reinforced concrete elements

The bridge condition inspection procedure varies from one country to another. The inspection procedure varies in terms of the frequency of inspections, the types of elements that need to be checked in a particular type of inspection, the value of the condition ratings of the elements. Regardless, the procedure for conducting the inspection itself is because an engineer performs an on-site inspection. In the case of damage, it is the engineer's responsibility to note this fact in the inspection protocol and to make documentation of the damage. At this stage, the engineer is required to decide whether the damage represents a risk to the bridge, whether additional expert work should be carried out and the urgency to do these works. This stage when the engineer decides whether the damage represents a risk to the structure is supported by the approach presented in this paper, in relation to concrete cracking.

Cracking in reinforced concrete structures is a natural phenomenon, which in many cases cannot be avoided. In order to be able to assess whether the occurrence of cracking may affect the durability of the structure and pose a risk to its safety, it is necessary to determine the probable cause of the damage, based on factors such as the nature of the given structural element's work, the place where the damage occurred or the time at which the damage appeared (e.g. after carrying out test loads on bridges).

Among the most characteristic cracks that occur in reinforced concrete elements, we can point out (Tomaszkiewicz & Owerko, 2023a):

- Cracks caused by the compression of a reinforced concrete element - in particular abutments, pillars
- Cracks caused by the excessive bending moment stress - especially in girders, in zones of extreme bending moment
- Cracks caused by nonuniform settlement of supports in particular abutments, pillars, foundations
- Cracks due to plastic shrinkage - especially observed for bridge slabs, pedestrian paths, concrete pavements
- Cracks in prestressed elements caused by corrosion of tendons, corrosion of tendon anchorages - characteristic for girders or anchorages of cable elements suspending the superstructure.

A particular type of cracking occurring in bridge structural elements is that of massive elements such as foundation slabs, abutments, pillars and pylons. The massiveness of a structural element can be defined as the ratio of the surface area of the element to its volume. In the process of cement hydration during concrete setting, the ratio of the surface area (i.e. the area through which the heat of hydration is removed) to the volume in which it is emitted is small. These cracks already appear during the construction phase, i.e. when the concrete strength is significantly lower than the design strength and when the bridge structure is not yet operating under full-service load. In addition, under further loading, cracks may propagate through the entire thickness of the concrete, resulting in a loss of structure monolithicity and a change in its static scheme.

The link between the type of structural element, the location of the damage and the causes of the damage will be discussed using the example of an abutment of a bridge structure. As a result of the cement hydration process, the temperature inside a concreted abutment rises to a maximum and then decreases to equalise with the ambient temperature. This results in tensile stresses that exceed the strength of the 'young' concrete. At the same time, the development of shrinkage phenomena occurs, resulting in cracking of the element. During the heating phase, the central part of the element heats up to a higher temperature than its edge, which causes the edge zones to be stretched. In the cooling phase, the central part of the element shrinks to a higher temperature than the edge, which causes stretching of the central zone of the element. Abutment walls are the elements in which the deformation of the lower edge is limited by external constraints. Abutment walls are concreted after the foundation has hardened and cooled, hence the curing of the wall takes place when the bottom of the wall is restrained in the foundation and there is no possibility of deformation. Therefore, thermal-shrinkage cracks of abutments are vertical cracks that begin above the abutment wall-foundation interface and disappear at the top of the abutment wall.

Description of the dataset for pre-failure cracks in concrete elements

The dataset developed by the authors "NCCD-PF - A pre-failure narrow concrete cracks dataset for engineering structures damage classification and semantic segmentation" (Tomaszkiewicz & Owerko, 2023b) has been made available as an open dataset and described in detail in the publication (Tomaszkiewicz & Owerko, 2023a). This dataset is dedicated to the problems of classification and segmentation of reinforced concrete element cracks. One of the dataset's characteristics is that it only represents those cracks which do not exceed a width of 0.3mm. This value is defined in EC 1992-1-1 (CEN, 2004) as the limit value of a crack with respect to typical construction methods for engineering elements and the conditions under which they operate. The crack detection using the algorithm trained on the presented dataset therefore allows the detection of cracks at the pre-failure stage. In addition, the dataset is characterised by the presence of a complex concrete surface finish (Figure 1), which can significantly influence to the results of the machine learning algorithms. The dataset's characteristic concrete surface complexity is shown in Figure 2.

The images that compose the dataset have been differentiated by:

- The type of element on which the cracks occurred (e.g. abutment, pillar, concrete barrier, tunnel wall)
- The cause of the cracks (e.g. thermal and shrinkage stresses in young concrete)
- Stage in the life cycle of the structure at which the image was obtained (from the construction stage to the operational stage)
- Quality of the reinforced concrete element construction, its maintenance and conservation.

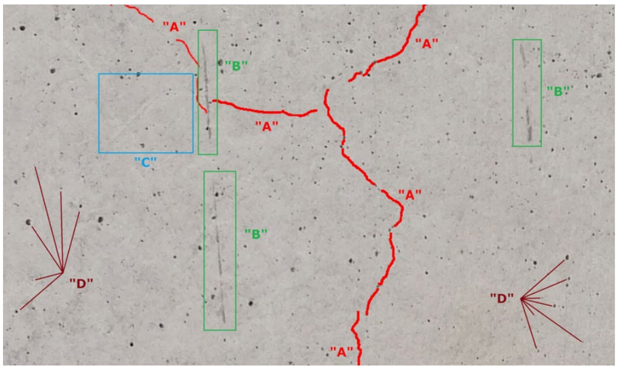


Figure 1: Example of concrete surface texture complexity. Marked as: "A" - cracks, "B" - dirts, "C" - mechanical damage, "D" – potholes (*Tomaszkiewicz & Owerko, 2023a*)

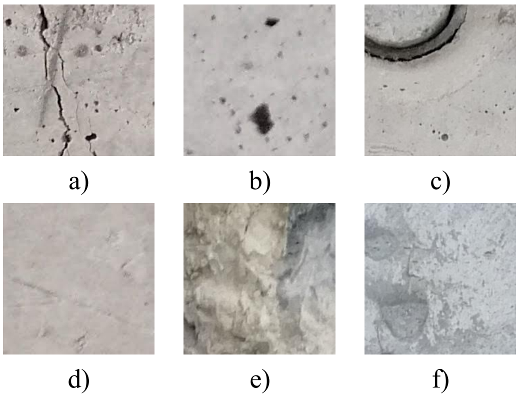


Figure 2: Example of concrete surface texture complexity, where a) dirt, b) bugholes, c) formwork marks, d) mechanical damage e) background obstructions f) troweling marks (Tomaszkiewicz & Owerko, 2023b)

The method of data acquisition for the dataset construction was chosen in such a way that the data acquired corresponds to the conditions under which bridge inspections are conducted. The images were obtained from cameras, without prior conditioning.

The dataset development started with the preparation of segmentation masks. An example of an image and its corresponding segmentation mask is shown in Figure 3 and 4. The images prepared in this way were segmented into sub-images of 224x224 pixels. On their basis, a classification dataset was created, examples of which are shown in Figure 5 (images of cracked concrete) and Figure 6 (images of uncracked concrete), respectively. The size of the dataset is shown in Table 1.

Figure 3: Example of image and corresponding segmentation mask (white pixels - crack, black pixels – background)

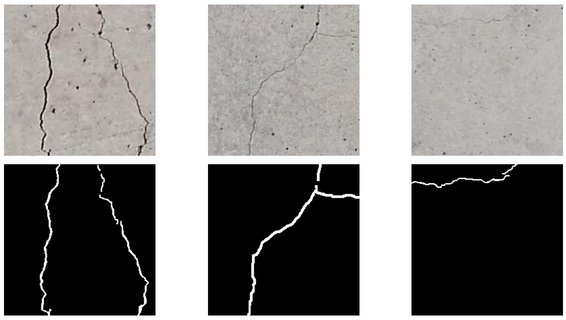


Figure 4: Example of subimages and corresponding segmentation masks (Tomaszkiewicz & Owerko, 2023b)

Figure 5: Example of cracked concrete images (Tomaszkiewicz & Owerko, 2023b)

Figure 6: Example of uncracked concrete images (Tomaszkiewicz & Owerko, 2023b)

Table 1: Number of images in the „NCCD-PF - A pre-failure narrow concrete cracks dataset for engineering structures damage classification and semantic segmentation" dataset (Tomaszkiewicz & Owerko, 2023a)

	Dataset for image classification	Dataset for image segmentation
Number of images	5388	5388
	Including: Cracked – 668 Uncracked – 4720	

Crack location based on segmentation of the structural element on images

To preserve the link between the bridge structure element and the damage, it is possible by segmenting an image of the single structural elements of the bridge structure and then providing data about a specific element as input information for a solution based on deep machine learning networks. For this purpose, the authors adapted the SegmentAnything algorithm (Kirillov et al., 2023) for iterative segmentation of bridge structure elements. Images identified as a structural element of a specified class are then segmented into sub-images and entered into a neural network. As a result, once the concrete surface cracks have been segmented using a deep learning algorithm, it is possible to link this damage to the element on which it occurred. Using this approach, a crack image of a given structural element is obtained, with the identification of those cracks that may pose a risk to the structural element and should be monitored and those that can be considered as neutral phenomena. An example of segmentation of reinforced concrete bridge elements is shown in Figure 7.

Figure 7: Results of bridge element segmentation

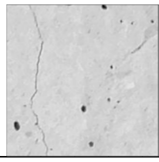
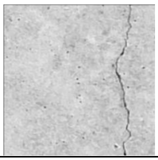
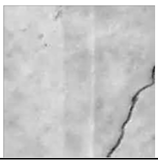
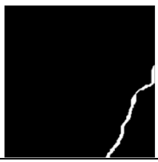
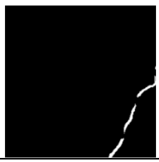
Crack detection solution using deep machine learning

Previous authors' experience (Tomaszkiewicz & Owerko, 2022) has shown that using networks trained on datasets for cracks with larger widths and fine-tuning them is not a proper solution for obtaining correct segmentation of cracks with widths below 0.3mm. The authors have developed a dedicated solution for pre-failure crack segmentation. The computer implementation was done using the PyTorch library.This solution used a modified and adapted UNet architecture originally proposed by (Weng & Zhu, 2015). The individual components of the network such as activation layers, convolution blocks etc. were implemented manually, the networks available in the PyTorch library were not used.

The network training process was based on a subset of 791 images from the presented dataset. The image size adopted for analysis was 224x224 pixels, which is also the original size of the images in the developed dedicated dataset. The images were randomly divided into a training and validation dataset in a ratio of 80:20. The dataset was divided into a batch size of 16 and trained for 125 epochs.

The results obtained, presented in Table 2, showed a high match between the network's prediction and the segmentation masks derived from the dataset. The information extracted in this way is sufficient information to be included on the image of the whole structural element and be used to assess its technical condition.

Table 2: Segmentation result using a trained deep learning network

Image	Label	Prediction
		
		
		

Limitations and future work

The presented deep learning solution covers one architecture that is effective for this problem but is not a most recent type of architecture (such as transformers). Furthermore, a problem that exists worldwide is that the amount of publicly available datasets related to the detection of different types of objects is severely limited. Data is not collected in a structured way, regarding FAIR Principles (Wilkinson et al., 2016).

Further research should be conducted in such a way that good quality datasets for deep learning algorithms can be built partly automatically. The procedures for performing bridge inspections should be developed or adapted in such a way that they meet the requirements for conducting inspections, but also allow the collecting of structured data that can be used for training deep learning algorithms. An important direction is also to conduct tests on different types of concrete mixes, keeping in mind the different potential causes of cracking. It may be helpful here to link the computational model to the location of damage in the structure, e.g. as a result of excessive loads). Such data can provide support for research related to the application of recent deep learning architectures based on multi-dimensional data as a source of knowledge (including consideration of IFC class, observation history from IoT sensors).

Acknowledgments

K.T. would like to acknowledge the support from the project Europe's Rail Flagship Project 1 - MObility management multImodal envirOnment aNd digitAl enabLers that has received cofounding from the Horizon Europe Framework Programme (HORIZON) under grant agreement number 101101973 and Polish Ministry of Science and Higher Education under grant agreement number MEiN 5339/HE/2023/2 within the programme International Co-Financed Projects. The project is being implemented at the AGH University of Krakow in 2022-2026.

T.O. would like to acknowledge the support from the project Europe's Rail Flagship Project 4 - Sustainable and green rail systems that has received cofounding from the Horizon Europe Framework Programme (HORIZON) under grant agreement number 101101917 and Polish Ministry of Science and Higher Education under grant agreement number MEiN 5337/HE/2023/2 within the programme International Co-Financed Projects. The project is being implemented at the AGH University of Krakow in 2022-2026.

References

Ali, R., Chuah, J. H., Talip, M. S. A., Mokhtar, N., & Shoaib, M. A. (2022). Structural crack detection using deep convolutional neural networks. *Automation in Construction*, *133*, 103989.

CEN. (2004). *Eurocode 2: Design of concrete structures - Part 1-1: General rules and rules for buildings*. CEN.

Hüthwohl, P., Lu, R., & Brilakis, I. (2019). Multi-classifier for reinforced concrete bridge defects. *Automation in Construction*, *105*, 102824.

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A. C., Lo, W.-Y., Dollár, P., & Girshick, R. (2023). *Segment Anything*.

Li, H., Wang, W., Wang, M., Li, L., & Vimlund, V. (2022). A review of deep learning methods for pixel-level crack detection. *Journal of Traffic and Transportation Engineering (English Edition)*, *9*(6), 945–968.

Nguyen, S. D., Tran, T. S., Tran, V. P., Lee, H. J., Piran, M. J., & Le, V. P. (2023). Deep Learning-Based Crack Detection: A Survey. *International Journal of Pavement Research and Technology*, *16*(4), 943–967.

Tomaszkiewicz, K., & Owerko, T. (2022, June 20). Deep machine learning in bridge structures durability analysis. *Proceedings of the 5th Joint International Symposium on Deformation Monitoring - JISDM 2022*.

Tomaszkiewicz, K., & Owerko, T. (2023a). A pre-failure narrow concrete cracks dataset for engineering structures damage classification and segmentation. *Scientific Data 2023 10:1*, *10*(1), 1–12.

Tomaszkiewicz, K., & Owerko, T. (2023b). *NCCD-PF - A pre-failure narrow concrete cracks dataset for engineering structures damage classification and semantic segmentation*. Zenodo

Weng, W., & Zhu, X. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. IEEE Access, 9, 16591–16603.

Wilkinson, M., Dumontier, M., Aalbersberg, I. et al. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, 3, 160018.

DIAGNOSING FACILITY MANAGEMENT WORK SYSTEMS FOR FACTORS IMPACTING SITUATION AWARENESS

Uttam Jodawat[1], Prajwal Prabhu[1], Vishal Singh[1] and Siva Kumar T[2]
[1]Indian Institute of Science, Bangalore, India
[2]Bosch Limited, India

Abstract

This paper introduces a method for diagnosing facility management (FM) work systems to improve situation awareness (SA). Identifying SA as a critical factor of success in FM, we propose a method to diagnose FM work systems, integrating multiple methods and offering a comprehensive sociotechnical systems perspective. The findings show the internal validity of the method, its application, and the identification of SA-impacting factors in FM work systems. Implications include continuous improvement and a proactive approach to addressing SA challenges in the FM work systems.

Introduction

The Facilities Management (FM) industry faces mounting pressure to align with sustainable development goals, exacerbated by the escalating complexity of digital transformation in the built environment (Nielsen et al., 2016; Okoro, 2023). This complexity necessitates a shift from traditional social systems to sociotechnical systems within FM work systems, wherein technical and social components are integral to decision-making processes (Yalcinkaya and Singh, 2014). However, this transformation has introduced challenges, particularly in maintaining up-to-date and comprehensive situation awareness (SA), which is crucial for informed decision-making.

SA, a critical factor in various domains such as nuclear power plants, healthcare, and aviation, remains relatively underexplored in the FM domain despite its evident importance (Gheisari and Irizarry, 2011; Gheisari, 2013; Akinci, 2014). Existing literature suggests two dominant and distinctive views on situation awareness; one defines SA as a cognitive product of individuals, and the other views it as an emergent property from interactions in a system. Acknowledging the multifaceted nature of SA, there is a need to adopt a sociotechnical systems perspective, where both views are accounted for, considering FM actors, FM technologies, and their interdependencies in achieving and maintaining SA. Evidence suggests that many individual and system-level factors interdependently influence SA in sociotechnical systems (Lau and Boring, 2016; Kurapati, 2017).

Improving FM work systems necessitates diagnosing the work systems and addressing the problems to support situation awareness. In this direction, the research community has put significant efforts into assessing the situation awareness of individuals and groups, taking a descriptive approach where only the social aspect of the work system is considered (Gawron, 2019; Alhaider, 2022). We argue that a normative approach with a sociotechnical systems perspective is required to effectively guide the interventions to address the SA-related issues in complex work systems. Moreover, existing methods expect a certain level of expertise in the SA domain (Salmon et al., 2006), presenting a barrier for the practitioners in the FM industry to utilize these methods and take an interest in them.

This research aims to bridge these gaps by establishing a problem-solving approach to address situation awareness-related issues in the FM sociotechnical (work) systems. In this regard, we propose a method to diagnose FM work systems to identify underlying patterns of SA-impacting factors.

By integrating insights from individual and system-level SA processes, this approach seeks to enhance SA within FM work systems, ultimately improving facility performance and supporting informed decision-making.

This work presents empirical findings from applying our proposed method and discusses implications for research and practice in the FM industry.

Literature review

This literature review delves into the evolution of contemporary Facilities Management (FM) work systems, understanding the shift from social systems to complex sociotechnical systems. Then, we inspect the challenges this transformation poses for achieving and maintaining high situation awareness (SA) levels among FM actors. Additionally, we explore strategies for diagnosing FM work systems to effectively address SA challenges, thereby augmenting informed decision-making and overall system performance.

Contemporary FM Work Systems and Rising Complexity

As the FM domain approaches contributing to sustainability, it is expected to play a strategic role beyond achieving efficiencies in operations and maintenance of the buildings and infrastructure (Collins and Junghans, 2015). In this broader context, FM can be considered "an integrated approach to operating, maintaining, improving and adapting the buildings and infrastructure of an organization to create an environment that strongly supports the primary objectives of that organization" (Barrett and Baldry, 2003). Towards addressing the sustainability challenges, i.e., achieving operational efficiencies, gaining competitive advantages, maintaining stakeholders' satisfaction, adhering to regulatory compliances and standards, and achieving resilience and sustainability of the built environment, the FM domain is going through a technological transition where advanced technologies like BIM, IoT, AI-ML, AR-VR, Building Automated Systems, ICT-based tools, Digital Twin-CPS, Blockchain/ DLT and emerging technologies, are becoming integral parts of FM work systems (Araszkiewicz, 2017; Yalcinkaya and Singh, 2019; Lee et al., 2021; Elyasi et al., 2023). Figure 1 explains FM as a sociotechnical (work) system adopting the framework proposed by (Brandt and Cernetic, 1998) where complex networks of social and technical systems (i.e. large teams of technicians, managers at different organizational levels, third-party companies and various technical systems associated with FM process and smart infrastructure) need to work collaboratively for effective management of the built environment. The attempts to bring sophistication to the built infrastructure and the intricacy of the FM work systems have posed challenges to FM actors in achieving and maintaining situation awareness, leading to poor decision-making and negative outcomes for facilities while increasing complexity.

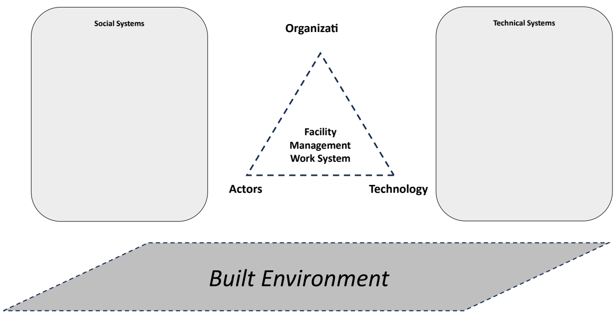


Figure 1:Facility Management as a sociotechnical system.

To further emphasize our claim, we analyze the dimensions of complexity suggested by (Vicente, 1999) in Table 1 to characterize FM work systems as a complex sociotechnical system.

Table 1: Characterizing FM work systems as a complex sociotechnical system with dimensions of complexity by (Vicente, 1999)

Dimensions of complexity	Definitions	Application to FM
Large problem space	Many different elements and forces	Many multidisciplinary functions shall be met by integrating people, processes, technology and places by effectively understanding the needs of the organizations and their peoples.
Social System	Composed of many actors who must work together	Inhouse social system + Outsourced social system + Customers + Owners
Heterogeneous perspectives	Actors with different backgrounds and disciplines	Operational and strategic perspectives, different disciplines, and conflicts of interests
Distributed system	Sub-systems located in different places	Different building systems at different locations, remote operations, teams at different locations, Cloud Computing
Dynamic system	Effects of actions, changes with times	Complex and dynamic built environment
Hazards	The high degree of potential hazards	Failures of critical infrastructures
Coupling	Highly coupled interactive subsystems	Tight coupling with organization and loose coupling between technical systems
Automation	Highly automated systems	Building Automated Systems, AI in FM
Uncertainty	Uncertainty in data available to actors	Rely on first responders, manual collection of data, and fragmented subsystems.
Mediated interactions	Systems not observable directly by actors	CAFM, Observable by Digital Interfaces
Disturbances	Actors dealing with unanticipated events	Unpredictable behaviours and activities of occupants, facility failures

Situation Awareness in FM Work Systems

The primary function of FM work system is to ensure the functionality, safety, comfort, efficiency and sustainability of the built environment (IFMA, 2024). To achieve this, the FM work system, i.e. the actors and FM technologies, continuously interacts with the facilities, its occupants and the core organization to get required situation awareness, which supports decision-making and performance of actions to deliver the agreed services successfully (see Figure 2). SA enables teams and individuals to keep an accurate mental picture of their environment and helps them anticipate changes and make well-informed decisions in complex and dynamic settings (Sorensen and Stanton, 2016). In other words, an adequate level of SA for FM actors will mean that they are aware of what has happened (Perception), what is happening (Comprehension), and what could happen (Projection) at any point, which is fundamental to making well-informed decisions. For instance, in a large office building with a centralized HVAC system, FM actors continuously monitor temperature, airflow, and energy usage to be aware of what has happened and what is happening. If they detect a sudden temperature increase in a zone, they anticipate a potential malfunction or high occupancy. This awareness of the current situation enables them to make informed decisions and take proactive actions to maintain comfort and optimize resource usage.

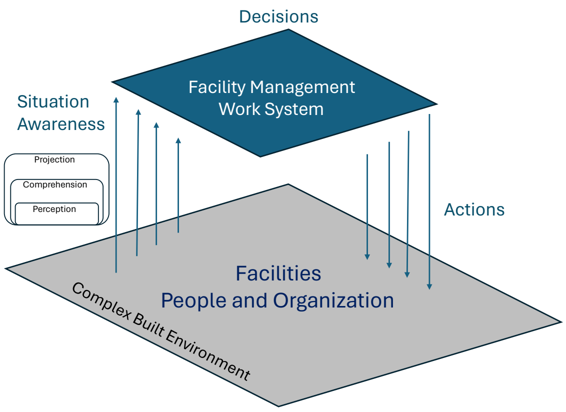


Figure 2 : Situation Awareness in Facility Management Work Systems

However, contemporary FM work systems can be characterized as complex sociotechnical systems where achieving and maintaining SA is challenged by an intricate interplay of many SA-impacting factors (Salmon, 2008; Boy, 2013; Lau and Boring, 2016). This necessitates a case for a sociotechnical systems approach to study SA in FM work systems.

The most influential work of situation awareness defines it as the perception and understanding of elements within an environment, coupled with the comprehension of their meaning and the projection of their status in the near future (Endsley, 1995). Apart from this, (Sarter and Woods, 1991) describe situation awareness as built on working memory and mental models: "Situation awareness is based on the integration of knowledge resulting from recurrent situation assessments". Additionally, (Tenney et al., 1992; Smith and Hancock, 1995)'s models are some of the most cited situation awareness models. Most of these theoretical constructs of situation awareness take an individualist view where situation awareness is considered a product of the active situation assessment process. This view highlights task-related, individual, and environmental factors impacting SA. Stanton presents another dominant view for situation awareness with a systems perspective, i.e. Distributed Situation Awareness (DSA). DSA expands on the idea of SA to include human and technological agents that work collaboratively, and SA emerges from SA transactions amongst them (Stanton et al., 2006; Salmon, 2008). This viewpoint strongly emphasizes team members' shared mental models, communication, and coordination to preserve a thorough grasp of the operating environment. Apart from the Stanton model on situation awareness, Endsley's and Jones's model on situation awareness and Artman and Garbis's distributed cognition model are some of the models that account for situation awareness from a team perspective (Artman and Garbis, 1998; Endsley and Jones, 2001). The reviewed literature suggests five dimensions for factors impacting SA in a sociotechnical system, i.e. individual, team, task/domain, technological, and environmental. Table 2 presents factors identified in literature impacting situation awareness in the sociotechnical system.

Table 2: List of factors impacting SA from the literature.

Categories	Factors Impacting SA
Individuals	Attention Tunneling, Errant mental models, Requisite Memory Trap, Out-of-the-loop syndrome, Goals and Expectations, Experience and training, Stress and fatigue, Training, Cognitive Load
Technology	Complexity Creep, Misplaced salience, Interface Design, System Capability, System Design, Procedures, Reliability and Robustness, Data Integrity
Team	Communication, Coordination, Roles and responsibilities, Team Cohesion, Attitude, shared mental models, Trust and Collaboration.
Tasks/ Domain	Information Overload, Lack of required information, Quality of information, Fragmented Information, Complexity, Automation, Workload, Workflow
Environment	Environmental stressors (Loud Noises, Lighting Conditions, Thermal Discomfort, Air Quality, Large Crowds, etc.), Complexity of the environment

Diagnosing FM work system for SA-related issues

To be innovative and efficient in the fast-changing demands and needs of client organizations, FM work systems must adopt strategies for continuous adaptation and improvement of the work systems (Duffy, 2000; Roberts, 2001; Nazali and Pitt, 2009). This necessitates adopting a problem-solving approach to identify poor performance outcomes/challenges within the work system, define problems, and design and implement solutions throughout the work system's life cycle. Poor performance outcomes in the work system can directly be attributed to poor SA, provided the actors can and are willing to decide and act with their SA (Endsley, 2020). This indicates that addressing SA-related issues in the work system can improve its performance. In order to achieve that, one should conduct a comprehensive diagnosis, which can guide interventions aimed at enhancing SA. The existing literature predominantly focuses on descriptive assessment of SA, which can describe whether a work system holds required and accurate SA or not. There are also few supports available in terms of guidelines for system design to support situation awareness (Endsley et al., 2003; Alhaider, 2022). However, we do not find any evidence of support available for diagnosing the work system for SA-related issues. Diagnosis in complex work systems demands systems thinking (Senge, 1994; Testa and Sipe, 2006; Wilson, 2014). The literature review discovered the iceberg model of systems thinking and its potential and applicability in diagnosing various work systems (Al-Homery et al., 2019; Ttr and Sivakumar, 2019). The researchers modified and adapted the iceberg model to guide the diagnosis of SA-related issues in the FM work system. (Refer to Figure 3 for details). For the scope of this research, we limit the diagnosis to one layer down to the iceberg model, where it uncovers patterns of individual, environmental, team, task/domain, and technological factors impacting situation awareness in FM work systems. Future work should focus on identifying mental models and sociotechnical interactions responsible for the emergence of SA-impacting factors in the FM work systems.

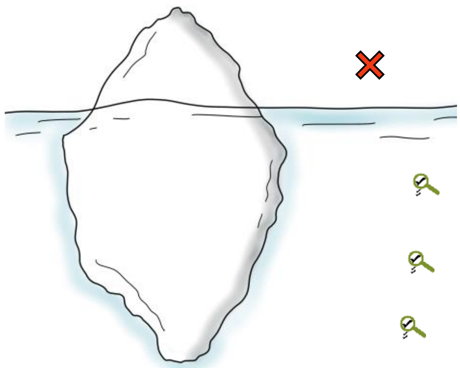


Figure 3:Iceberg model of systems thinking for diagnosing FM work system for SA-related issues.

Methodology

The ergonomic and human factors domain suggests integration of methods with a systems approach to address complex problems of sociotechnical systems (Stanton, 2018; Salmon and Read, 2019). With this inspiration, based on the literature review of situation awareness, sociotechnical systems, and facility management domains, we propose a qualitative method to support diagnosing FM work systems for situation awareness-related problematic situations. Figure 4 shows overall research methodology to be adopted for this research.

With a focus on improving situational awareness in FM work systems, the literature review on situational awareness comprehensively lists the factors that impact SA. The literature review on situation awareness touches on the different aspects of the sociotechnical systems, i.e., individuals, teams, technologies, tasks/domains, and the environment. Based on the different aspects of the sociotechnical systems, the corresponding factors that impact SA have been reviewed and listed in the Table 2. Based on the factors listed from the literature, the FM work system is investigated using the proposed method, which integrates multiple methods (see Table 3). Furthermore, while evaluating the FM work system with the proposed method, there could be a possibility of new factors that result from the diagnosis that are also included while diagnosing the work system. Overall, at meta level the method itself can be evaluated and improved for its effectiveness for FM work system but this is beyond the scope of this research work.

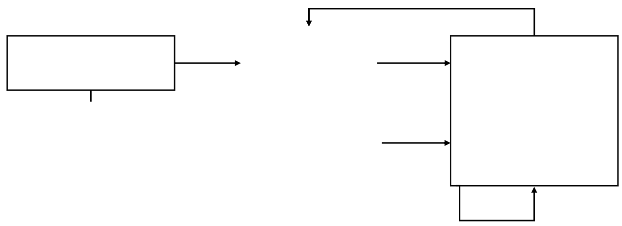


Figure 4:Adopted Research Methodology

Table 3: Proposed research method for diagnosis of the FM work system for SA related issues

Phases	Objectives	Activity	Outputs	Involvement of Participants
Data Collection	To understand the work system thoroughly	In-depth interviews, direct observations, system and operators' logs	Recordings, Notes and documents	Provide relevant information
Data Processing	To prepare data for analysis	Transcription and documentations	Documents	Not Required
	To identify and define potential problematic situations in the work system	Qualitative Analysis Who? What? When?	List of Problems	Verification of Problems for their occurrence
Data Analysis	To identify problems relevant to Situation Awareness and related possible causal factors	Reasoning How? Why?	Problems relevant to SA and related causal factors	Active Involvement in Reasoning
Interpretation of results	To interpret the results of the analysis	Visualization	Frequency Distribution of Responsible SA impacting factors	Not required

Case Study

The applicability of the proposed method was tested in an industrial FM setting. The industrial setting had multiple work systems like water management, helpdesk, energy, access control, security, etc. Out of the different work systems, the helpdesk work system is central to the efficient functioning of the different work systems together. Recent trends in FM emphasize the adoption of a helpdesk for the large and complex built environment with many occupants to process work orders, complaints, and service requests of the clients effectively. The principle of helpdesk systems is similar to the IT helpdesk – "to respond to a customer's inquiry as quickly as possible and follow it through until it has been satisfactorily resolved". However, the functioning of the helpdesk work system in the FM domain can be complex and challenging, as it deals with a large problem space with multiple interrelated support team networks and a complex-built environment. In this regard, the helpdesk work system was chosen as the place of interest for the study. The unit of analysis for this study was the whole helpdesk work system with two helpdesk executives and a facility manager responsible for its functioning, as shown in the Figure 5

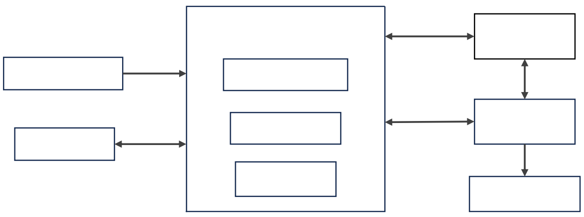


Figure 5:Basic structure of the FM helpdesk work system.

The study began with an informal discussion among the researchers and the help desk executives with two major objectives: for the researchers to familiarize themselves with the operations of the help desk work system and for executives to familiarize themselves with the study plan. After the informal discussion, the researchers conducted observational studies at different times to understand and record the basic functioning of the work system. Two separate questionnaires (See Table 4 and Table 5) were prepared to conduct semi-structured interviews with helpdesk executives and the facility manager to thoroughly understand the work system and its problems. The three and half-hour overall interviews were recorded and transcribed. The transcripts of the interview and the notes of the observational study were used together to do a qualitative analysis to understand the detailed functioning of the help desk work system and identify potential problematic situations. A more precise description of the problems was prepared with an understanding of who is involved, what the problem is, and when it occurs in the work system. Precise description helps to improve understanding of the problems, which supports the relevant stakeholders' diagnosis of the work system. These problems were revisited to check their uniqueness, relevance to SA and duplications were removed. The helpdesk executives further verified the problems identified for their occurrence. Diagnosing each problem was performed with reasoning to identify all possible causal factors that can impact SA in the work system by the researcher. This was done carefully with the factors identified in the literature and newly realized factors from the reasoning activity.

Questionnaires:

These questionnaires were prepared for semi-structured interviews to understand the helpdesk work system thoroughly as shown in Table 4 and Table 5.

Table 4:Questionnaire for Semi-structured Interview of Helpdesk Executives

Questions
What is your association with the Facility Management Work System?
Can you briefly explain the helpdesk work system?
Can you describe your role and responsibilities related to the helpdesk work system?
What is your work experience related to your current role? How do you gain experience?
What are your day-to-day operations at work? Are there any specific tools or software you use?
Do you work as a team? What is your specific role in this team?
Have you or your team identified any training or skill gaps required to improve your performance?
What reporting and documentation processes do you follow to keep track of system performance, issues, and resolutions?
How do you communicate with other stakeholders to address issues or to seek support?
How do digital tools and equipment impact your daily operations?
What issues do you face in the daily operations related to digital tools and equipment?
What are the anticipated risks if the issues are not resolved?
What improvements or changes could be made to enhance the efficiency and effectiveness of your role as an operator of this solution/system?

Table 5:Questionnaire for Semi-structured Interview of Helpdesk Work System Manager

Questions
Can you provide an overview of the facility management helpdesk work system's structure and its role within the organization?
What are the primary objectives and goals of the facility management helpdesk work system in ensuring smooth facility operations?
Could you describe the key processes and workflows the helpdesk team follows to manage facility requests and issues?
What are the key performance indicators (KPIs) or metrics used to measure the success and efficiency of the facility management helpdesk?

What are the most common challenges or bottlenecks the helpdesk team encounters in addressing facility-related issues?

How are facility requests and issues assigned and tracked within the helpdesk system, and what tools or software are used for this purpose?

How does the helpdesk team coordinate and communicate with facility management staff and other relevant departments to resolve issues promptly?

How is the helpdesk equipped to handle unexpected facility-related emergencies or disruptions effectively?

What technologies or tools support the facility management helpdesk's work processes and decision-making?

How do you ensure your helpdesk team has the resources and training to manage facility requests and issues effectively?

What strategies and practices are in place to continuously improve the facility management helpdesk's operations and service delivery?

How are conflicts or issues typically managed and resolved within the helpdesk team or with other facility-related stakeholders?

Can you share recent examples of successful facility management initiatives or improvements facilitated by the helpdesk team?

What future developments or changes do you foresee for the facility management helpdesk, and how are you preparing for them?

How does the helpdesk gather and utilize feedback from facility users and staff to enhance its service quality and efficiency?

What leadership and management style do you employ to ensure the success and satisfaction of the helpdesk team?

Are there any specific training or development needs you believe would benefit the helpdesk team's performance in managing facility requests and issues?

How does the facility management helpdesk contribute to fostering a positive facility management culture and ensuring high levels of user satisfaction?

Can you share experiences where the helpdesk effectively responded to critical facility-related situations or challenges?

What are the most important lessons you have learned as a leader of the facility management helpdesk, and how have they influenced your approach?

Results:

A total of seventy-one potential problematic situations related to poor performance outcomes were identified, and diagnosis was performed to identify related SA-impacting factors obtained from the literature and through analysis. Figure 6 shows the frequency distribution of the SA impacting factor relevant to the identified problematic situations in the FM helpdesk work system in decreasing order where ineffective communication and lack of knowledge and experience are the highest occurring factors, this could be the case as the industrial setting is in under digital transformation and very frequently new facilities and related teams are being introduced. Ineffective feedback mechanisms, the complex network of actors, lack of accountability, lack of multidisciplinary knowledge, work constraints, lack of standardization, and poor data handling are the new factors realized with reasoning activity in the data analysis phase which can impact SA.

The environmental stressor, team attitude, goal and expectations and out-of-the-loop syndrome did not show up in the analysis. The environmental conditions were well maintained in terms of lighting and air quality. Also, a dedicated space is allocated to the helpdesk team which lacks large crowd and loud noises. Team attitude and goal and expectations has complex relationships with SA which is difficult to put into the reasoning. The helpdesk executives were proactively engaging with the other relevant actors to get and provide updates which avoids out-of-the-loop syndrome.

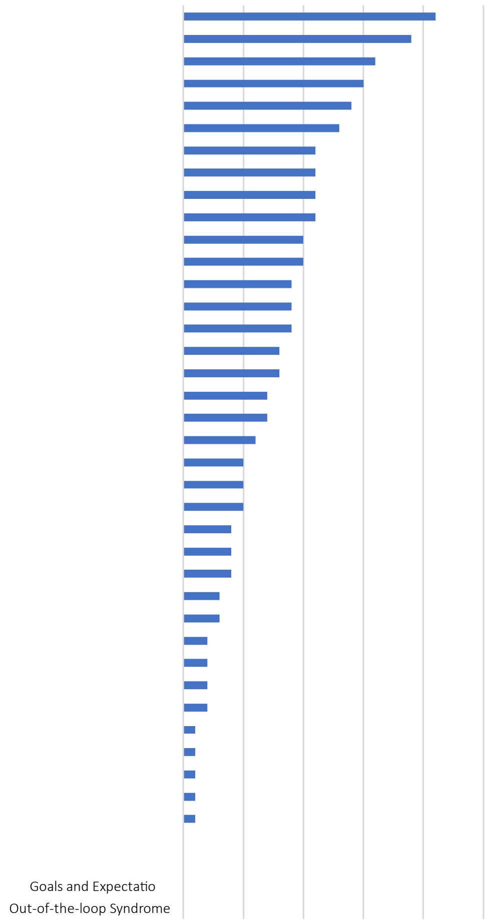


Figure 6: Frequency Distribution of the SA impacting factors from analysis.

Discussion and Conclusions

The method discussed in the study stemmed from the initial observations and discussion with relevant FM stakeholders of a help desk system in an industrial FM setting. These observations underscored the significant impact of poor situational awareness (SA) on the efficacy and functionality of the help desk system, prompting a comprehensive diagnosis of the underlying work system to enhance SA. The work system managers spend substantial time and effort in diagnosing the work system through routine meetings and discussions, but they were not found to have a situation awareness-oriented approach. From Figure 3, they mostly react to the poor performance outcomes for surface-level solutions, but the literature suggests it requires systems thinking to understand the underlying cause and treat them. This highlighted the critical need for thorough analysis and diagnosis of FM work systems to address underlying challenges for SA effectively.

An adequate understanding of all the factors impacting situational awareness is essential for the stakeholders to diagnose the FM work system. Some of the factors that impact the FM work system may be difficult to reason out, requiring comprehensive reasoning ability of the stakeholders, for example, how goals and expectations and team attitude can affect SA. The result of the analysis represents how responsible each factor is in impacting SA; it should not be confused with how much impact a particular factor causes to SA in the work system. This analysis of the FM work system acts as a reference point to further investigate the underlying interaction through which the SA-impacting factors emerge in the system. While the involvement of stakeholders in the process is ideal, practical considerations such as time constraints may necessitate analysis by the researchers alone, where they are expected to have a detailed understanding of the work system. Furthermore, it is necessary to ensure that stakeholders are well-supported and informed about the different steps and factors in understanding the method in advance to facilitate their effective participation.

Although the suggested approach has shown itself to be useful in the setting of a help desk work system, more research should be done to see whether it can be scaled up to include a greater number of participants and other work systems. Expanding the methodology's scope to include a wider range of stakeholders and organizational contexts would offer important insights into its effectiveness for various FM work systems while it may also demand large efforts. We anticipate the potential for developing digital tools to support the effective use of the method by system stakeholders where the tool's knowledge base can be enhanced with each study, it can support reasoning, and the involvement of the researcher can be avoided. Furthermore, the analysis presented in the study is from the researcher's perspective owing to the constraints of time and availability of the systems stakeholders. As the study is a work in progress, a comprehensive analysis of the problematic situations of the help desk work systems from the active participation of stakeholders is something that needs to be looked at in future works.

Acknowledgments

The project is supported by the SMART research project [SRG/2021/002285] funded by SERB, DST, India, and Bosch Limited.

References

Akinci, B. 2014. Situational Awareness in Construction and Facility Management. *Frontiers of Engineering Management*. **1**(3), p.283.

Alhaider, A. 2022. *Distributed Situation Awareness Framework to Assess and Design Complex Systems*. Blacksburg, VA: Virginia Polytechnic Institute and State University.

Al-Homery, H.A., Ashari, H. and Ahmad, A. 2019. The Application of System Thinking for Firm Supply Chain Sustainability: The Conceptual Study of the Development of the Iceberg Problem Solving Tool (IPST). . **8**(6).

Araszkiewicz, K. 2017. Digital Technologies in Facility Management – The state of Practice and Research Challenges. *Procedia Engineering*. **196**, pp.1034–1042.

Artman, H. and Garbis, C. 1998. Situation Awareness as Distributed Cognition *In*: *ECCE΄98, Limerick*.

Barrett, P. and Baldry, D. 2003. *Facilities management: towards best practice* 2nd ed. Osney Mead, Oxford, OX ; Malden, MA: Blackwell Science.

Boy, G.A. 2013. Orchestrating Situation Awareness and Authority in Complex Socio-technical Systems *In*: M. Aiguier, Y. Caseau, D. Krob and A. Rauzy, eds. *Complex Systems Design & Management* [Online]. Berlin, Heidelberg: Springer Berlin Heidelberg, pp.285–296. [Accessed 20 August 2023]. Available from: http://link.springer.com/10.1007/978-3-642-34404-6_19.

Brandt, D. and Cernetic, J. 1998. Human-centred appraoches to control and information technology: European experiences. *AI & SOCIETY*. **12**(1), pp.2–20.

Collins, D. and Junghans, A. 2015. Sustainable Facilities Management and Green Leasing: The Company Strategic Approach. *Procedia Economics and Finance*. **21**, pp.128–136.

Duffy, F. 2000. Design and facilities management in a time of change. *Facilities*. **18**(10/11/12), pp.371–375.

Elyasi, N., Bellini, A. and Klungseth, N.J. 2023. Digital transformation in facility management: An analysis of the challenges and benefits of implementing digital twins in the use phase of a building. *IOP Conference Series: Earth and Environmental Science*. **1176**(1), p.012001.

Endsley, M.R. 2020. The Divergence of Objective and Subjective Situation Awareness: A Meta-Analysis. *Journal of Cognitive Engineering and Decision Making*. **14**(1), pp.34–53.

Endsley, M.R. 1995. Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. **37**(1), pp.32–64.

Endsley, M.R., Bolté, B. and Jones, D.G. 2003. *Designing for situation awareness: an approach to user-centered design*. London ; New York: Taylor & Francis.

Endsley, M.R. and Jones, W.M. 2001. A Model of Inter and Intra-Team Situation Awareness: Implications for Design, Training and Measurement. , SA Technologies, Inc.

Gawron, V.J. 2019. *Human performance and situation awareness measures* Third edition. Boca Raton, FL: CRC Press/Taylor & Francis Group.

Gheisari, M. 2013. *An ambient intelligent environment for accessing building information in facility management operations; a healthcare facility scenario*. Dissertation, Georgia Institute of Technology.

Gheisari, M. and Irizarry, J. 2011. Investigating facility managers' decision making process through a situation awareness approach.

IFMA 2024. What is Facility Management - Internationa Facility Management Association. *What is Facility Management*. [Online]. Available from: https://www.ifma.org/about/what-is-fm/.

Kurapati, S. 2017. *Situation Awareness for Socio Technical Systems: A simulation gaming study in intermodal transport operations*.[Online] TRAIL Research School. [Accessed 26 August 2023]. Available from: https://repository.tudelft.nl/islandora/object/uuid%3A0f9fe428-baa0-4e8c-948f-e30a1c289727.

Lau, N. and Boring, R. 2016. Situation Awareness in Sociotechnical Systems *In*: *Human Factors in Practice: Concepts and Applications* [Online]. London: CRC Press. Available from: https://doi.org/10.4324/9781315587370.

Lee, J.Y., Irisboev, I.O. and Ryu, Y.-S. 2021. Literature Review on Digitalization in Facilities Management and Facilities Management Performance Measurement: Contribution of Industry 4.0 in the Global Era. *Sustainability*. **13**(23), p.13432.

Nazali, M.N.M. and Pitt, M. 2009. A critical review on innovation in facilities management service delivery. *Facilities*. **27**(5/6), pp.211–228.

Nielsen, S.B., Sarasoja, A.-L. and Galamba, K.R. 2016. Sustainability in facilities management: an overview of current research. *Facilities*. **34**(9/10), pp.535–563.

Okoro, C.S. 2023. Sustainable Facilities Management in the Built Environment: A Mixed-Method Review. *Sustainability*. **15**(4), p.3174.

Roberts, P. 2001. Corporate competence in FM: current problems and issues. *Facilities*. **19**(7/8), pp.269–275.

Salmon, P., Stanton, N., Walker, G. and Green, D. 2006. Situation awareness measurement: A review of applicability for C4i environments. *Applied Ergonomics*. **37**(2), pp.225–238.

Salmon, P.M. 2008. *Distributed situation awareness: advances in theory, measurement and application to team work*. Brunel University.

Salmon, P.M. and Read, G.J.M. 2019. Many model thinking in systems ergonomics: a case study in road safety. *Ergonomics*. **62**(5), pp.612–628.

Sarter, N.B. and Woods, D.D. 1991. Situation Awareness: A Critical But Ill-Defined Phenomenon. *The International Journal of Aviation Psychology*. **1**(1), pp.45–57.

Senge, P.M. 1994. *The fifth discipline: the art and practice of the learning organization* 1. Currency paperback ed. New York, NY: Currency Doubleday.

Smith, K. and Hancock, P.A. 1995. Situation Awareness Is Adaptive, Externally Directed Consciousness. *Human Factors: The Journal of the Human Factors and Ergonomics Society*. **37**(1), pp.137–148.

Sorensen, L.J. and Stanton, N.A. 2016. Keeping it together: The role of transactional situation awareness in team performance. *International Journal of Industrial Ergonomics*. **53**, pp.267–273.

Stanton, N.A. 2018. *Human factors methods: a practical guide for engineering and design*. Abingdon, Oxon: Routledge, Taylor & Francis Group.

Stanton, N.A., Stewart, R., Harris, D., Houghton, R.J., Baber, C., McMaster, R., Salmon, P., Hoyle, G., Walker, G., Young, M.S., Linsell, M., Dymott, R. and Green, D. 2006. Distributed situation awareness in dynamic systems: theoretical development and application of an ergonomics methodology. *Ergonomics*. **49**(12–13), pp.1288–1311.

Tenney, Y.J., Adams, M.J., Pew, R.W., Huggins, A.W.F. and Rogers, W.H. 1992. *A Principled Approach to the Measurement of Situation Awareness in Commercial Aviation*. BBN Systems and Technologies Cambridge, Massachusetts.

Testa, M.R. and Sipe, L.J. 2006. A Systems Approach to Service Quality: Tools for Hospitality Leaders. *Cornell Hotel and Restaurant Administration Quarterly*. **47**(1), pp.36–48.

Ttr, J. and Sivakumar, M. 2019. ICEBERG Metaphor – A Tool for Healthcare Quality Management Systemic Structure. *International Journal of Management Studies*. **VI**, p.118.

Vicente, K.J. 1999. *Cognitive work analysis: toward safe, productive, and healthy computer-based work*. Mahwah, N.J: Lawrence Erlbaum Associates.

Wilson, J.R. 2014. Fundamentals of systems ergonomics/human factors. *Applied Ergonomics*. **45**(1), pp.5–13.

Yalcinkaya, M. and Singh, V. 2014. Building Information Modeling (BIM) for Facilities Management – Literature Review and Future Needs *In*: S. Fukuda, A. Bernard, B. Gurumoorthy and A. Bouras, eds. *Product Lifecycle Management for a Global Market*. Berlin, Heidelberg: Springer, pp.1–10.

Yalcinkaya, M. and Singh, V. 2019. VisualCOBie for facilities management: A BIM integrated, visual search and information management platform for COBie extension. *Facilities*. **37**(7/8), pp.502–524.

ENHANCING CONSTRUCTION SITE SAFETY USING AI: THE DEVELOPMENT OF A CUSTOM YOLOV8 MODEL FOR PPE COMPLIANCE DETECTION

Mohamad Iyad Al-Khiami[1,2], Mohamed M ElHadad[3],
[1]Australian University - Kuwait, Mishref, Kuwait
[2]Aalborg University, Aalborg, Denmark
[3]Gulf University for Science & Technology, Mishref, Kuwait

Abstract

This study addresses construction safety by deploying computer vision techniques, specifically a YOLOv8 model by Ultralytics, to monitor PPE compliance. Targeting helmets, vests, and safety shoes, it aims to mitigate accident risks. The model was trained with 2934 images and validated with 816, achieved a 95% mAP. Emphasizing AI's potential in safety management and occupational health in the construction industry. This research lays groundwork for future AI-based safety enhancements in construction sector, highlighting the industry's pressing need for innovative approaches to reduce occupational hazards and improve compliance standards.

Introduction

The construction industry, known for being the primary driver to infrastructure development been consistently ranked among the most dangerous industries worldwide, with a high incidence of accidents and fatality rates making it one of the most hazardous sectors (Lingard, 2013; Pinto et al., 2011; Waehrer et al., 2007). High rates of accidents and fatalities have been consistently reported, many attributable to non-compliance with safety measures, particularly the use of Personal Protective Equipment (PPE) (Memon et al., 2023; Sehsah et al., 2020). In many incidents, the lack of PPE or improper use of safety gear such as helmets, vests, and boots has been a critical factor. (Kang, 2018) reported that more than 70% of all fatal accidents had some form of incompliance with PPE. This persistent challenge highlights a gap in safety protocols and enforcement on construction sites, highlighting the need for more stringent and effective monitoring tools to ensure worker safety and reduce the risk of accidents.

Given the high importance of maintaining safety standards and reducing injuries from accidents in construction sites along with the prevalent safety challenges, it is necessary to rethink traditional methods and employ innovative technologies to enhance safety compliance rates on sites (Zhang, 2021). Among these technologies, artificial intelligence (AI) takes the lead when it comes to the development of object detection systems specifically for PPE in the site (Abioye et al., 2021). On a site, monitoring systems using AI can assist safety engineers in achieving higher compliance of safety due to the fact that traditional human supervision can sometimes be expensive, prone to error and insufficient in maintaining safety standards (Yi and Wu, 2020). Such systems can aid in the detection of workers who are not complying with safety standards mainly, wearing proper PPE while working on sites. The necessity for improved safety compliance on construction sites, coupled with the inadequacies of traditional safety monitoring methods dictates the need to start integrating AI-driven object detection systems in construction sites. The integration of AI-driven systems in construction sites represents an important opportunity and a significant leap forward in terms of technology adoption within the construction industry.

This paper is part of a larger project that aims to utilize AI in the construction sector. The project is divided into several phases where the objective of this phase is to answer the following research question (RQ).

RQ – Are fine-tuned object detection models, specifically YOLOV8 efficient and effective in identifying safety helmets, safety shoes and vests in construction sites?

Literature Review

Safety risks in the construction sector

The construction engineering sector is a key driver of economic growth in both developed and developing nations (Sánchez et al., 2017). Despite advances in workplace safety within the construction sector, it still faces a greater risk of injuries and deaths than many other industries (Johansson et al., 2019). According to the US bureau of labor statistics (BLS), more than 1 in 5 deaths occurred in the workplace was within the construction industry in the year 2020 with a reported number of 1,008 construction workers that were killed on the job (A Look at Workplace Deaths, Injuries, and Illnesses on Workers' Memorial Day, 2022). Each year, more than 100,000 individuals suffer from fatal injuries each year within the construction industry as per the International Labor Organization (ILO) which alone, represents about 30% of all occupational fatal injuries ("Construction," 2015).

A recent study (Memon et al., 2023) highlights that substandard quality of PPE is a leading cause of accidents in the construction industry. This study also found that the use of PPE can reduce accidents related to falls by 30%. Another study indicated that many accidents on construction sites occur due to the lack of PPE or failure to wear it properly (Ammad et al., 2021). Despite approximately 62% of construction workers being at risk of falls, only about half use PPE, as reported by the Bureau of Labor Statistics (BLS) (A Look at Workplace Deaths, Injuries, and Illnesses on Workers' Memorial Day, 2022). Furthermore, it was noted that over 70% of fatal fall accidents involved workers not wearing PPE (Kang, 2018). Additionally, according to the Health and

Safety Executive (HSE), there are more than 9,000 PPE-related accidents annually on construction sites in the United Kingdom. Understanding the frequency of these incidents underscores the need to educate employees on the importance of proper PPE usage (Martin et al., 2021). Severe brain injuries on construction sites, primarily caused by falls and falling objects, are a significant concern (Kamardeen and Hasan, 2022). Furthermore, the Centers for Disease Control and Prevention (CDC, 2011) estimates that almost half (49%) of all fatal injuries in this sector are due to head injuries (Occupational Ladder Fall Injuries — United States, 2011).

The concerning statistics and studies highlighted in this section emphasize the urgent need for more stringent and effective enforcement of PPE safety compliance in the construction industry (Ebekozien, 2021; Gattuso, 2021). It is imperative to develop and implement reliable strategies to ensure that workers are adequately protected, thereby reducing the high incidence of injuries and fatalities that currently plague this sector.

Technology adoption and integration in construction industry

The potential of Artificial Intelligence (AI) is increasingly being recognized across various sectors. However, its adoption and application in the construction industry are scarce compared to other industries. As a matter of fact, the construction industry ranks among the least digitized sectors globally, and a common misconception among stakeholder exists regarding the industry's longstanding culture of resistance to change (Young et al., 2021). Additionally, the lack of technology integration in the construction industry is often associated with health and safety concerns (Nikas et al., 2007). In an effort to address this slow growth in adoption, many companies are now turning to Artificial Intelligence (AI) as a means to streamline their processes and boost productivity within the working environment (Yigitcanlar, 2021; Yigitcanlar and Cugurullo, 2020). The adoption of AI technology grants a competitive edge in terms of automation when compared to conventional approaches (Chien et al., 2020). Within the wide variety of AI-Based technologies, the application of computer vision through deep learning has shown promising potential in construction safety management. The object detection capability of AI provides flexibility in terms of classifying and recognizing objects, which is something to be capitalized upon to improve safety compliance. This technology, serves a foundation to effectively substitute human vision for many tasks across the construction safety workflow (Abioye et al., 2021). This sets the stage for exploring advancements in AI for PPE compliance monitoring in the next section.

Advancement in AI for PPE compliance monitoring

In recent years, the construction industry has seen significant advancements in the application of Artificial Intelligence (AI) for safety management, particularly in monitoring Personal Protective Equipment (PPE) compliance. The effectiveness of AI, specifically deep learning, and computer vision, in real-time monitoring of safety helmets and PPE compliance, showing promise for enhanced on-site safety have been demonstrated in the literature (Delhi et al., 2020; Kisaezehra et al., 2023).

Recent advancements in the construction industry's approach to safety management have been significantly influenced by the application of Artificial Intelligence (AI). A focus on enhancing Personal Protective Equipment (PPE) compliance has been evident, with AI-driven systems, particularly those incorporating YOLO models for object detection, demonstrating notable accuracy and real-time capabilities. This shift towards AI-based methodologies for safety gear recognition, especially through the use of advanced YOLO v5 and v8 models, underscores a growing trend in leveraging technology to improve on-site safety measures (Chen et al., 2021; Kim et al., 2023; Wang et al., 2023).

The advancements in AI for construction safety have seen significant strides in the development of systems for detecting safety helmets and protective clothing. A notable approach involves the enhancement of YOLOv3 methods, specifically tailored to improve the detection of smaller-sized safety gear. This innovation, focusing on the addition of a large-size input layer for multi-scale prediction, represents a crucial step in fine-tuning AI models to meet the unique demands of construction site applications, underscoring the critical role of AI optimization in specific industrial contexts (Wang et al., 2020).

The exploration of AI in the construction industry has further expanded with the introduction of rapid PPE detection systems for actual construction sites, utilizing deep learning techniques. This advancement, as presented in the literature, signifies the practicality and effectiveness of AI in enhancing real-time safety management on construction sites. It addresses the critical requirement for advanced and efficient safety monitoring tools within the industry, showcasing the potential of AI to significantly improve construction safety practices (Wang et al., 2021).

Together, these studies underscore the potential of AI and machine learning, particularly YOLO models, in revolutionizing safety compliance in the construction industry. They highlight the technical feasibility and practical implications of deploying AI systems for real-time, accurate PPE monitoring, marking a significant step forward in occupational safety management.

Closing remarks

As indicated by the literature, to the best of the authors' knowledge, there has been limited empirical research examining the adoption of AI technologies in the construction industry. As such, this study aims to contribute to the growing body of knowledge surrounding the integration of artificial intelligence, namely, PPE compliance detection systems in the construction industry.

Methodology

Development of the PPE Compliance AI model – YOLOV8

You Only Look Once V8 (YOLOV8), developed by Ultralytics in January 2023, served as the foundation for our AI model. YOLOV8 is a convolutional neural network (CNN), that is a category of deep learning neural networks, commonly used in analyzing visual imagery. YOLO was trained and validated using a dataset, namely Common Objects in Context (COCO). The COCO dataset contains more than 330 thousand images of 80 different common objects, including but not limited to, humans, bicycles, cars and animals. A total of 118 thousand images were used for the training, 5000 for validation and 20 thousand for testing. The model was then benchmarked against the validation dataset using the mean average precision (mAP) which is basically a percentage precision of the number of detected objects correctly identified across multiple objects (Ultralytics, 2023a). YOLOV8 can be used for different purposes, including object detection, object tracking, object classification and segmentation. The project utilizes the object detection capabilities of YOLOV8.

Currently, there exists five YOLO models with varying sizes, (1) Nano, (2) Medium, (3) Large, (4) Extra Large. Simply put, smaller models compromise accuracy for speed, and are useful where computational power is limited and speed is a necessity. On the contrary, the larger models are the most accurate, but also the most resource intensive. According to Ultraltytics documentations, mAP is 37.3 and 53.9 for the Nano and Extra-large model respectively (Ultralytics, 2023b).

Based on the limited computational power available, and the fact that the model is aimed to run in real-time, the YOLOV8 (m) model was used, with a mAP of 50.2. The (m) model offers a middle ground between speed and accuracy with a good balance between performance and efficiency.

YOLOV8 (m) fine tuning

While COCO is a well-established dataset, it fails to serve the purpose of the project, therefore, a custom dataset was required. Following Ultralytics recommendations, two main folders were created, a folder dedicated specifically to the training dataset and a folder dedicated to the model validation. Within each directory, two subfolders were created, namely “Images” and “Labels”. All the training images that were collected were inserted in the “Images” subfolder under “Train” main folder. On the other hand, the images used for validations were inserted in “Images” subfolder under the “Validation” main folder. A total of 2934 images and 816 images were for training and validation respectively. Figure 1 shows the breakdown of the dataset organization.

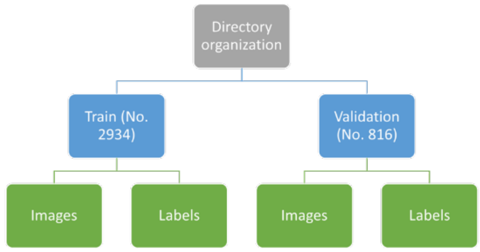


Figure 1: Directory organization for the customized training

The 2934 images collected were to fine tune the model to detect if a construction worker, within a construction site setting is adhering to the PPE requirements. The object detection model aims to detect if workers are wearing their safety helmets, vests, and safety shoes. The open-source images were collected in addition to taking photos using a phone camera in construction sites after taking consent from safety officers. To avoid bias in the image collection, the data collected was made sure to be as diverse as possible, encompassing different colors and shapes of helmets, vests, and safety shoes with varying backgrounds ensuring transferability of the model across different countries.

After the collection of the photos, the fine-tuning process begins. First, Conda, an open-source package management and environment management system was downloaded and installed in which all the machine learning is managed. A dedicated environment was created using Conda where all the packages required for YOLOV8 were installed. Using the command “pip install ultralytics” downloads all the packages and dependencies required to run YOLOV8. Prior to installing the packages concerning YOLOV8, it is necessary to annotate the images in the “Train” and “Validation” folders and save the output into the “Labels” subfolder of both “Train” and “Validation”. For that, “LabelImg”, an open-source graphical image annotation tool, was downloaded and installed. Labelling the images using “LabelImg” outputs a .txt file for each image with the location of the label within an image. Figure 2 shows the user interface and the labelling using “LabelImg".

Figure 2: LabelIMG annotation user interface

The annotation process involved six classes:

- Class 0 – “Helmets”

- Class 1 – "Vests"
- Class 2 – "Safety shoes"
- Class 3 - "No vests"
- Class 4 - "No helmets"
- Class 5 - "No safety shoes"

Upon annotating all the images collected, a ".yaml" file was prepared where the train and validation directory were set, the number of classes and the names of each class in order. This file is necessary as it contains all the necessary information required to override the existing trained YOLOV8.

An important parameter to consider before initiating the training, is the number of epochs required, that is basically one complete pass of the entire training dataset through the algorithm. Zhang et al., (2019) underscores the importance of setting the number of epochs to an acceptable and reasonable number. For example, a very small number of epochs can result in an underfitted model, meaning that the model has not been trained enough on the trained data, thus resulting in a poor performance against validation or testing data. Conversely, overfitting phenomena can occur in the cases of an exaggerated number of epochs. In such cases, the model memorizes the training set rather than generalizing. The model would ultimately perform well on the trained data but poorly on unseen data.

Selecting the appropriate number of epochs is an iterative process requiring several trials. The number of epochs for the project was set to 100 and the performance was constantly checked against the validation dataset setting an early stopping parameter in case there is no improvement in the performance as the number of epochs continues to increase. The command used in the Conda environment to conduct the training was as follows "yolo task=detect mode=train epochs=100 data=data_custom.yaml model=yolov8m.pt imgsz=640".

Figure 3 summarizes the whole processes followed to create the custom model.

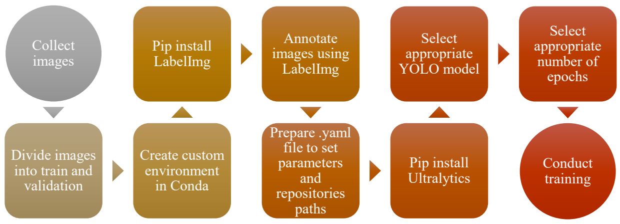


Figure 3: Summary of the fine-tuning on custom dataset.

Results and Discussion

The results and discussion section summarizes the findings of the methodological approach conducted.

YOLOV8 (m) was selected as the foundation for the fine-tuning process. A total of 2934 and 816 photos were used for the training and validation respectively. A processor of Intel® Core (TM) i9-9980HK CPU @ 2.40GHZ (16CPUs), with a dedicated graphics card of NVIDIA RTX 2060 and 32 GB RAM served as the training hardware.

By default, a patience value = 50 is set, where the numerical value represents the number of epochs. The patience parameter simply means that while the training is in process, the model shall check its performance against the validation dataset, in any case where no improvement is perceived in the last 50 completed epochs, an early stop is employed. As a result, the YOLOV8 fine-tuning process took 23.33 hours and had an early stop at 96 epochs, as there was no improvement seen beyond 46 epochs.

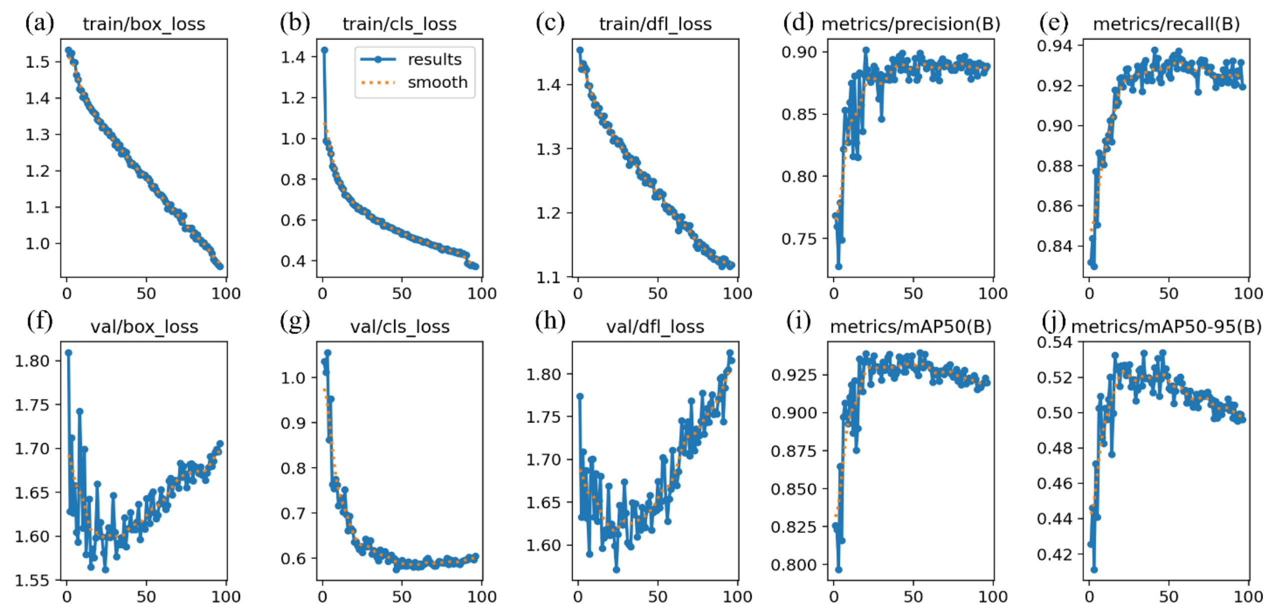


Figure 4: Training dashboard against the validation dataset

Figure 4. illustrates the results of the custom trained model from the validation dataset. The x-axis represents the number of epochs, while the y-axis varies depending on the graph it represents.

A total of 10 graphs (a) – (j) are shown in Figure 4. Figure 4 (a) – (c) and Figure 4 (f) – (h) shows the “box loss”, “cls_loss” and “dfl_loss” that correspond to box loss, classification loss, and directional focal loss on the training and validation dataset respectively. Box loss measures how well the model is predicting the bounding box coordinates for each detected object. It can be seen as the number of epochs increases, Figure 4 (a) – (c) decreases illustrating a downwards trend. This means that the model gets better at defining the coordinates of the binding boxes, improving its ability to detect the orientation of the objects along with their presence. Though, when considering Figure 4 (f) – (h) it can be seen that all three figures see improvement in the box_loss, cls_loss, and dfl_loss up to the 46th epochs. This confirms the early stop and patience parameter were beyond the 46th epochs, there was no improvement seen. The change in the figures trend indicates a sign of overfitting. When the validation loss starts to increase while the training loss continues to decrease, it means that the improvements in the model are specific to the training data and are not improving the model’s predictive ability for new, unseen data.

Figure 4 (d) – (e) shows the precision and recall changes through 96 epochs for the validation dataset. Both Figures (d) and (e) show a positive and linear trend against the epochs. Once the training hits the 46th epochs mark, the precision reaches its highest value of 0.89/1 and 0.92/1 for the recall. The training continues all the way to the 96th epochs where the value of the precision and recall falls to 0.88/1 and 0.91/1 respectively. The precision level shows that the model, at the 46th epochs is precisely detection the correct object 89% of the time. On the other hand, a recall of 92% indicates that the model is able to recall 92% of the objects.

Figure 4 (i) – (j) shows the mean average precision and the mean average precision at 95% at Intersection over Union (IoU). Similar to Figure 3 (d) – (e), the figures here look at the average precision per class. In addition, IoU of 95% is considered a very stringent threshold, it means that for a detection to be considered to be a true positive, the predicting bounding box must overlap with the ground truth bounding by at least 95%. Only detections that satisfy this threshold is considered true positive. On the 46th epochs, the mean average precision at 95 % IoU is 0.53/1.

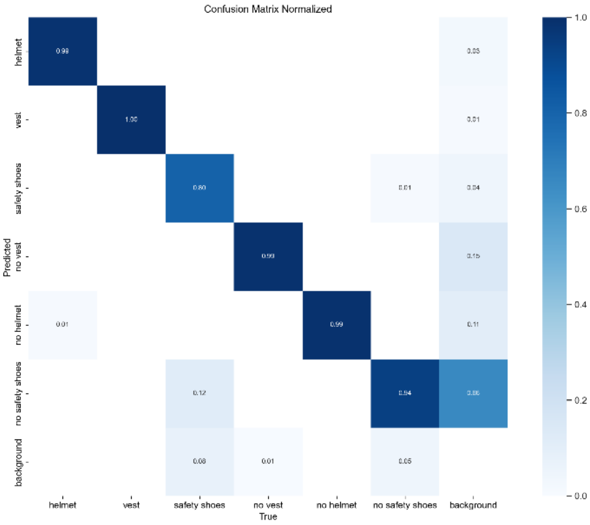


Figure 5: Normalized confusion matrix

To further understand the model reliability, and to visualize the performance of the algorithm, a normalized confusion matrix is shown in Figure 5. The matrix shows that the model predicted 'helmet' with 99% accuracy indicating true positive. Moreover, the 'vest' class has been predicted by the model with high accuracy. As for 'safety shoes', the true positives were 80%, but there were some instances where it predicted 'safety shoes' when there were none indicating false positives, and some instances where it failed to predict 'safety shoes' when there were some (false negatives). This can be due to the fact that normal shoes may emulate the look of safety shoes designs which can lead to false positives. It can also be mentioned that there are very few cases where the model indicated a 1% of false positives in helmets. While this result indicates high precision, it is limited to the used dataset.

The confusion matrix suggests that the model is quite effective at predicting 'helmet' and 'vest' classes, is fairly good at predicting 'safety shoes', and generally does not confuse items with the background. However, there are some areas where the model can generate false predictions, particularly with the 'safety shoes'. This information can be used to refine the model further, potentially by providing it with more training data for the classes where it is less accurate or adjusting the model's parameters.

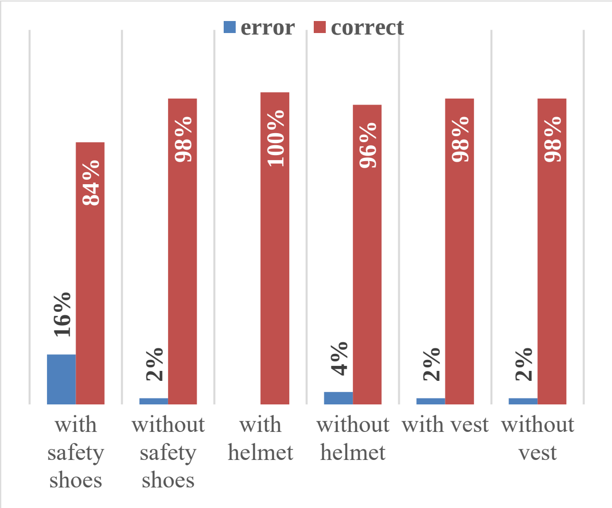


Figure 6: Results of YOLOV8 (m) "Best.pt" model against testing dataset

Figure 6 presents the performance results of the YOLOV8 (m) "Best.pt" model when evaluated against a testing dataset to detect various types of PPE. The figure illustrates and compares the percentage of correct identifications against the percentage of error across the six classes discussed before across 300 photos of workers complying or not in construction sites. All the 300 photos used for testing were exclusive to the testing dataset and were not used in the training nor the validation dataset.

The model clearly demonstrates high levels of accuracy in detecting the presence of a helmet, with a score of 100% and no perceived errors. Similarly, detecting vests achieved a 98% accuracy and success rate. The 'without helmet' category shows a slight decrease in accuracy of 4% error rate only which could be improved by further training the model. However, the model's performance exhibits a notable decline in the 'with safety shoes' category, with a correct identification rate of 84% and a corresponding error rate of 16%. This suggests that while the model is highly effective at identifying vests and helmets, it finds safety shoes more challenging, which may indicate a need for further model training or data augmentation in this category specially since safety shoes can exhibit a diverse number of models, colours and shapes.

The mAP for the testing dataset can be calculated by finding the average of all the precision from the 6 classes. A percentage of 95.6 was obtained, indicating a similar mAP to the validation dataset.

The error rates presented in the graph are essential for understanding the model's limitations and guide future improvements to enhance its predictive capabilities for PPE compliance on construction sites. It is important to note that since the testing data was only from 300 photos, the results cannot be considered reliable. Furthermore, construction sites are dynamic with variations in settings, lighting conditions and working environment. This illustrates the need to diversify the collected data to cover wider landscapes of conditions. A possible solution is using data augmentation techniques to transform images and simulate different lighting conditions.

Figure 7: PPE detection of sample in-test photo

Figure 7 demonstrates the PPE detection model's output when using the YOLOV8 (m) fine-tuned model. Each class is bounded by a box which states the confidence level of the PPE detection. The confidence level represents how accurately is the model detecting and determining the class of the PPE in use within the detection frame. While the testing was only conducted on images, the model can be utilized with a high-resolution camera to be tested and implemented in real-time scenarios.

Conclusion and Limitations

This research endeavor has illustrated the core hazards within the construction industry, examining the critical concerns concerning Personal Protective Equipment (PPE) safety standards. It highlights the necessity of upholding stringent safety compliance on construction sites to mitigate the risk of accidents and enhance worker protection.

Additionally, this study has detailed the capabilities of object detection technologies, namely, YOLO technique's robust framework. The analysis revealed that the YOLOV8 is balanced between precision and computational efficiency, particularly when utilizing a dataset of medium size to fine-tune the trade-off between speed and accuracy. The model's effectiveness at identifying compliance with helmet, vest, and safety shoes requirements in PPE protocols was notable, although it did exhibit a potential for enhancement in detecting 'safety shoes' class type.

The model is limited to the collected dataset which illustrates a need to increase the size of the training data, specifically, safety shoes. In addition, it is worth noting that the testing data was limited to 300 images which does not necessarily cover all real-world scenarios. Data augmentation techniques can further enhance the collected dataset to cover wider working conditions in construction sites. These insights not only validate YOLOV8's utility in practical applications but also identify specific possibilities for refining the model to achieve even higher levels of accuracy in PPE detection in future phases.

This stage of the conducted research was limited to the development of the AI-based PPE detection system. In the forthcoming stage of our research, a case study approach can be implemented to further investigate the model reliability. The focus will be on evaluating the impact of deploying the AI-based PPE compliance monitoring system within construction environments in real time. This assessment will illustrate the system's efficacy in reinforcing adherence to PPE usage standards and protocols. Moreover, we intend to conduct a thorough investigation into the sector's behavior in response to the system's implementation and acceptance.

While AI-based PPE compliance monitoring comprises privacy, it is essential to consider an ethical framework. Construction companies willing to implement this technology must obtain consent of workers being surveilled and implement anonymity approaches such as face blur techniques to preserve the privacy of workers. In addition, the reports generated by the AI model should be inspected against bias. Decision makers within the construction industry must be aware that the intention of such systems is to augment manual inspection and not replace it.

Acknowledgment

The authors would like to express their sincere gratitude to the Australian University – Kuwait for their support of the ongoing research. Special thanks are extended to Abdullah Jalal Jamal, Zainab Hani Jamal, Sherifa Meshal Borhamah, and Noryah Saeed Alqabadni for their exceptional efforts, without which this project would not have been possible.

References

A look at workplace deaths, injuries, and illnesses on Workers' Memorial Day : The Economics Daily: U.S. Bureau of Labor Statistics [WWW Document], 2022. URL https://www.bls.gov/opub/ted/2022/a-look-at-workplace-deaths-injuries-and-illnesses-on-workers-memorial-day.htm (accessed 1.6.24).

Abioye, S.O., Oyedele, L.O., Akanbi, L., Ajayi, A., Davila Delgado, J.M., Bilal, M., Akinade, O.O., Ahmed, A., 2021. Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. J. Build. Eng. 44, 103299. https://doi.org/10.1016/j.jobe.2021.103299

Ammad, S., Alaloul, W.S., Saad, S., Qureshi, A.H., 2021. Personal protective equipment (PPE) usage in construction projects: A scientometric approach. J. Build. Eng. 35, 102086. https://doi.org/10.1016/j.jobe.2020.102086

Chen, B., Wang, X., Huang, G., Li, G., 2021. Detection of violations in construction site based on YOLO algorithm, in: 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE). Presented at the 2021 2nd International Conference on Artificial Intelligence and Computer Engineering (ICAICE), pp. 251–255. https://doi.org/10.1109/ICAICE54393.2021.00057

Chien, C.-F., Dauzère-Pérès, S., Huh, W.T., Jang, Y.J., Morrison, J.R., 2020. Artificial intelligence in manufacturing and logistics systems: algorithms, applications, and case studies. Int. J. Prod. Res. 58, 2730–2731. https://doi.org/10.1080/00207543.2020.1752488

Construction: a hazardous work [WWW Document], 2015. URL http://www.ilo.org/safework/areasofwork/hazardous-work/WCMS_356576/lang--en/index.htm (accessed 1.6.24).

Delhi, V.S.K., Sankarlal, R., Thomas, A., 2020. Detection of Personal Protective Equipment (PPE) Compliance on Construction Site Using Computer Vision Based Deep Learning Techniques. Front. Built Environ. 6.

Ebekozien, A., 2021. Construction companies' compliance to personal protective equipment on junior staff in Nigeria: issues and solutions. Int. J. Build. Pathol. Adapt. 40, 481–498. https://doi.org/10.1108/IJBPA-08-2020-0067

Gattuso, A., 2021. Common Issues of Compliance with Personal Protective Equipment for Construction Workers. Constr. Manag.

Johansson, J., Berglund, L., Johansson, M., Nygren, M., Rask, K., Samuelson, B., Stenberg, M., 2019. Occupational safety in the construction industry. Work 64, 21–32. https://doi.org/10.3233/WOR-192976

Kamardeen, I., Hasan, A., 2022. Occupational Health and Safety Challenges for Sustaining Construction Apprentice Programs. J. Manag. Eng. 38, 04022042. https://doi.org/10.1061/(ASCE)ME.1943-5479.0001059

Kang, Y., 2018. Use of Fall Protection in the US Construction Industry. J. Manag. Eng. 34, 04018045. https://doi.org/10.1061/(ASCE)ME.1943-5479.0000655

Kim, Kyunghwan, Kim, Kangeun, Jeong, S., 2023. Application of YOLO v5 and v8 for Recognition of Safety Risk Factors at Construction Sites. Sustainability 15, 15179. https://doi.org/10.3390/su152015179

Kisaezehra, Farooq, M.U., Bhutto, M.A., Kazi, A.K., 2023. Real-Time Safety Helmet Detection Using Yolov5 at Construction Sites. Intell. Autom. Soft Comput. 36, 911–927. https://doi.org/10.32604/iasc.2023.031359

Lingard, H., 2013. Occupational health and safety in the construction industry. Constr. Manag. Econ. 31, 505–514. https://doi.org/10.1080/01446193.2013.816435

Martin, H., Mohan, N., Ellis, L., Dunne, S., 2021. Exploring the Role of PPE Knowledge, Attitude, and Correct Practices in Safety Outcomes on Construction Sites. J. Archit. Eng. 27, 05021011. https://doi.org/10.1061/(ASCE)AE.1943-5568.0000501

Memon, M.H., Abas, N.H., Sohu, S., 2023. Ranking the critical causes of accidents in construction projects of Sindh: Perspective of safety professionals. IOP Conf. Ser. Earth Environ. Sci. 1205, 012035. https://doi.org/10.1088/1755-1315/1205/1/012035

Nikas, A., Poulymenakou, A., Kriaris, P., 2007. Investigating antecedents and drivers affecting the adoption of collaboration technologies in the construction industry. Autom. Constr. 16, 632–641.

Occupational Ladder Fall Injuries — United States, [WWW Document], 2011. URL https://www.cdc.gov/mmwr/preview/mmwrhtml/mm6316a2.htm (accessed 1.6.24).

Pinto, A., Nunes, I.L., Ribeiro, R.A., 2011. Occupational risk assessment in construction industry – Overview and reflection. Saf. Sci. 49, 616–624. https://doi.org/10.1016/j.ssci.2011.01.003

Sánchez, F.A.S., Peláez, G.I.C., Alís, J.C., 2017. Occupational safety and health in construction: a review of applications and trends. Ind. Health 55, 210–218. https://doi.org/10.2486/indhealth.2016-0108

Sehsah, R., El-Gilany, A.-H., Ibrahim, A.M., 2020. Personal protective equipment (PPE) use and its relation to accidents among construction workers. Med. Lav. 111, 285–295. https://doi.org/10.23749/mdl.v111i4.9398

Ultralytics, 2023a. COCO [WWW Document]. URL https://docs.ultralytics.com/datasets/detect/coco (accessed 12.18.23).

Ultralytics, 2023b. Home [WWW Document]. URL https://docs.ultralytics.com/ (accessed 12.18.23).

Waehrer, G.M., Dong, X.S., Miller, T., Haile, E., Men, Y., 2007. Costs of occupational injuries in construction in the United States. Accid. Anal. Prev. 39, 1258–1266. https://doi.org/10.1016/j.aap.2007.03.012

Wang, L., Zhang, X., Yang, H., 2023. Safety Helmet Wearing Detection Model Based on Improved YOLO-M. IEEE Access 11, 26247–26257. https://doi.org/10.1109/ACCESS.2023.3257183

Wang, X., Niu, D., Luo, P., Zhu, C., Ding, L., Huang, K., 2020. A Safety Helmet and Protective Clothing Detection Method based on Improved-Yolo V 3, in: 2020 Chinese Automation Congress (CAC). Presented at the 2020 Chinese Automation Congress (CAC), pp. 5437–5441. https://doi.org/10.1109/CAC51589.2020.9327187

Wang, Z., Wu, Y., Yang, L., Thirunavukarasu, A., Evison, C., Zhao, Y., 2021. Fast Personal Protective Equipment Detection for Real Construction Sites Using Deep Learning Approaches. Sensors 21, 3478. https://doi.org/10.3390/s21103478

Yi, X., Wu, J., 2020. Research on Safety Management of Construction Engineering Personnel under "Big Data + Artificial Intelligence." Open J. Bus. Manag. 8, 1059–1075. https://doi.org/10.4236/ojbm.2020.83067

Yigitcanlar, T., 2021. Greening the Artificial Intelligence for a Sustainable Planet: An Editorial Commentary. Sustainability 13, 13508. https://doi.org/10.3390/su132413508

Yigitcanlar, T., Cugurullo, F., 2020. The Sustainability of Artificial Intelligence: An Urbanistic Viewpoint from the Lens of Smart and Sustainable Cities. Sustainability 12, 8548. https://doi.org/10.3390/su12208548

Young, D., Panthi, K., Noor, O., 2021. Challenges involved in adopting BIM on the construction jobsite. EPiC Ser. Built Environ. 2, 302–310.

Zhang, H., Zhang, L., Jiang, Y., 2019. Overfitting and Underfitting Analysis for Deep Learning Based End-to-end Communication Systems, in: 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP). Presented at the 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6. https://doi.org/10.1109/WCSP.2019.8927876

Zhang, Y., 2021. Safety Management of Civil Engineering Construction Based on Artificial Intelligence and Machine Vision Technology. Adv. Civ. Eng. 2021, e3769634. https://doi.org/10.1155/2021/3769634

Education, Policy and Standardisation

TECHNOLOGICAL SOLUTIONS TO LABOR SHORTAGES IN CONSTRUCTION: ASSESSING PRODUCTIVITY AND INNOVATION ADOPTION

Hooman Sadeh[1], Sadegh Haghighat[2], Aram Pirayesh[3], Dimitar Todorov[1], Min Jae Lee[2], and Farzad Shahbodaghlou[4]
[1]Utica University, Utica, NY, USA
[2]Virginia Tech, Blacksburg, VA, USA
[3]Politecnico di Milano, Milan, Italy
[4]California State University, East Bay, Hayward, CA, USA

Abstract

This study assesses the integration of technology in the construction sector to address labor shortages. Traditional methods show deficiencies in coping with workforce deficits. A survey design was employed, utilizing exploratory and inferential statistical techniques, including a two-step cluster analysis, Spearman's rho, and Mann-Whitney U tests, to examine the relationship between firm size, technological implementation, and productivity in the construction industry. Results indicate that larger firms leveraging technology report higher productivity. This study's contribution lies in highlighting technology's role in mitigating labor shortages and enhancing productivity, with implications for strategic industry decision-making and policy development.

Introduction

The construction industry is currently experiencing a severe global labor shortage (AGC, 2022), resulting in increased costs and delays in projects. This scarcity of skilled workers poses widespread challenges, escalating expenses, and prolonging timelines for construction projects globally. Addressing this issue is pivotal for mitigating its impact on finances and project completion timelines (ABC 2023).

Since the 1980s, the North American construction industry has consistently faced a shortage of skilled labor, exhibiting cyclical patterns over the past three decades (Karimi et al., 2018). This shortage, influencing the national economy and citizens' well-being, is tied to the overall performance of the U.S. economy, with construction demand fluctuating accordingly (Al-Bayati et al., 2020). The increased demand for construction services has led to a critical labor issue, affecting both skilled and unskilled workers, resulting in project delays, cost escalations, and compromised project quality. In this context, the research conducted by Sadeh et al. (2021) becomes particularly relevant. It highlights the industry's division into conventional and avant-garde firms in response to technological and market changes. Sadeh (2023) further underscores the industry's productivity lag due to insufficient digitalization, a challenge particularly acute for small and medium-sized firms. Additionally, Sadeh et al. (2022) reveal the underutilization of Building Information Modeling (BIM) in large U.S. construction firms, indicating a gap in digital proficiency. These studies collectively underscore the need for increased technology integration and digital proficiency in the construction industry. At the same time, the construction labor market's complexity requires a comprehensive approach to understand and address these challenges (Cheung et al., 2011). Additionally, the Department of Labor's concern over the aging workforce and the need for approximately 240,000 new workers underscores the importance of attracting and training a new generation to sustain industry growth and adapt to evolving demands (Kim et al., 2020). Despite its substantial contribution to global economic development—accounting for 9% of GDP and employing 7% to 8.5% of the workforce worldwide—the construction industry struggles with low productivity and minimal technological advancements (Hossain et al., 2020). Global construction spending, which reached USD 11.4 trillion in 2018 and is expected to rise to USD 14 trillion by 2025, contrasts starkly with the industry's mere 1.5% investment in technology in the U.S., significantly lower than that of manufacturing (3.3%) and the overall economy (3.6%) (Changali et al., 2015). This discrepancy is particularly notable given that productivity in construction has largely stagnated, in contrast to the manufacturing sector, which has nearly doubled its productivity. Contributing to these performance challenges, as identified by Gupta (2019), are factors such as regular payment delays, advancements in technology, labor efficiency, and the availability of skilled professionals, all of which are pivotal in shaping the outcomes of construction projects.

In response to the prolonged labor shortage and improving production quality, the construction industry is increasingly embracing technology. Major firms are not only establishing innovation departments and investing in new technologies but also conducting on-site product testing (Jackson, N.D., 2020). There is a growing interest and application in various forms of automation such as virtual reality, augmented reality, drones, robotic arms, reduced scanning/photogrammetry, and 3D printing, all actively researched and utilized, albeit often confined to specific projects (Hossain et al., 2020). In addition to these technologies, the Internet of Things (IoT), drones, and prefabrication are emerging as game-changers due to their cost-effectiveness and minimal disruption to traditional construction practices, thereby shaping the future of construction job sites (Jackson, N.D., 2020). Complementing these technological advancements Cai et al.'s research (Cai et al., 2018; Cai et al., 2019) delves into the application of automation and robotics in high-rise construction, addressing labor shortages and safety risks. Moreover, Halder et al. explored the integration of inspector assistant quadruped robots (Halder et al., 2023) and developed a computational framework for remote navigation of these robots, integrating live-streaming and AR with BIM models (Halder et al., 2022). Goh et al. (2019) conducted a detailed simulation study of modular construction operations, known as Prefabricated Prefinished Volumetric Construction (PPVC), in Singapore. Their work provides valuable insights for future applications in offsite construction research. These advancements highlight the industry's increasing reliance on technology to tackle labor shortages and improve productivity. However, the path to technological integration is not without its challenges. These include the critical task of evaluating the investment value of new technologies and overcoming resistance from tradesmen who might view certain technologies as inefficient or unnecessary (Jackson, N.D., 2020). In addition to these practical challenges, the industry faces broader issues such as low productivity and limited technological progress, especially evident in Digital Twin (DT) applications, as highlighted by Opoku et al. (2021). Naderi & Shojaei (2022) further underscore the complexities of infrastructural projects and the nascent nature of digital twins, pinpointing the lack of consensus among stakeholders as a major barrier to the adoption of infrastructure digital twins (IDTs). To address production issues in construction, Antunes et al. (2018) proposed a framework integrating automatic supervisory control and data acquisition, while also noting the often-isolated implementation of information technology and automation within the industry. Melenbrink et al. (2020) explored the challenges in achieving fully autonomous construction in unstructured environments, stressing the need for development across all construction task groups and coordination between task-specific robots. On the sustainability front, Adaloudis et al. (2021) utilized grounded theory methods to assess the benefits of 3D concrete printing (3DCP), with a focus on balancing environmental, economic, and social sustainability aspects. Their research indicates that firms are increasingly motivated to invest in technologies like 3DCP to enhance automation and address skilled labor shortages. Furthermore, Liu-Lastres et al. (2022) analyzed the causes and effects of the Great Resignation, concentrating specifically on labor shortages in the construction industry and proposing various strategies to manage these challenges effectively.

The primary aim of this study is to investigate how construction firms are utilizing technology to navigate the challenges of labor shortages and how this strategy influences their operational dynamics. It is driven by two core research questions: First, it seeks to understand the ways in which technology is leveraged to address labor shortages within the construction industry and how this approach interacts with variables such as firm size, project type, and productivity to shape the firms' operational profiles. Second, it examines the impact of technology usage on the productivity of construction firms, exploring the potential for a notable relationship between the size of these firms and their levels of productivity. These research questions will provide insights into strategic decision-making and policy development in the face of labor shortages.

Research Questions

RQ1: How are construction firms leveraging technology to address labor shortages, and how does this strategy interact with firm size, project type, and productivity to define their operational profiles?

RQ2: How does the use of technology to mitigate labor shortages affect the productivity of construction firms, and is there a relationship between the size of these firms and their productivity levels?

Hypothesis 2A (Firm Size and Productivity)

- **H_1:** There is a monotonic relationship between firm size and productivity in construction firms.

Hypothesis 2B (Technology Usage and Productivity)

- **H_1:** There is a statistically significant difference in the median productivity scores between construction firms that use technology to mitigate labor shortages and those that do not.

Methods and Materials

The methodology is comprised of the design and implementation of a survey instrument as a component of a comprehensive electronic questionnaire. The questionnaire was developed based on an extensive literature review to assess workforce development within the construction sector and was formed as part of a broader research initiative. In the comprehensive survey conducted, a total of thirty-seven questions were presented to participants, encompassing demographic information, organizational insights, and practices, as well as technology adoption strategies. However, for the scope of this article, focused analysis was conducted on a subset of these questions that directly contribute to our examination of workforce development challenges and the role of technology in mitigating labor shortages within the construction sector. These questions were selected based on their direct relevance to our research objectives, which underscored their importance in addressing the core themes of our investigation. The rationale for focusing on these questions is further supported by their potential to illuminate the key dynamics of workforce development and technology integration in the construction industry. The first section contained multiple-choice questions related to the demographics of respondents and company profiles. The second section comprised closed-ended yes/no questions and Likert scale items ranging from 1 to 5. The respondents were asked to rate the productivity of both their craftsmen and office personnel. Additionally, they were also questioned about their adoption of technology as a strategy to counter labor shortages. For those affirming the use of technological solutions, the survey further delved into identifying the specific types of technologies being employed. Prior to its distribution, a pilot study was conducted to evaluate the validity and reliability of the survey questions, ensuring the instrument's effectiveness in accurately capturing the relevant data. Based on the findings, certain questions were revised or removed. Data collection occurred over a three-week period at CM Expos and job fairs organized for construction management students on three university campuses across the United States, located in Virginia, New York, and California. These events attracted 270 companies, including both national and regional contractors, and provided a valuable opportunity for data gathering. The process was designed to secure a representative sample of the industry. The questionnaire, however, was only distributed to around 125 construction firms, of which 92 responses were received, and 86 were complete and deemed suitable for analysis, resulting in a response rate of 73.6%. Based on the research goals and questions, appropriate exploratory and inferential statistical techniques were employed and analyzed using SPSS 29. A two-step cluster analysis was conducted for Research Question 1, while Spearman's rho and Mann-Whitney U were used for Research Question 2. The demographics of the respondents and the companies' profiles are shown in Tables 1, 2, 3, 4, and Figure 1 below. Most of the firms were commercial contractors with revenues encompassing $500 million per year.

Table 1:Roles of Respondents

Role	Frequency	Percent
Project Manager	20	23.3
Field Engineer	6	7.0
Project Engineer	17	19.8
HR	10	11.6
Superintendent	6	7.0
Executive Leader	25	29.1
BIM Specialist	2	2.3
Total	86	100.0

Table 2: Experience of Respondents

Duration	Frequency	Percent
Under 5	36	41.9
6 - 10	14	16.3
11 - 15	10	11.6
16 - 20	6	7.0
21 +	20	23.3
Total	86	100.0

Table 3: Firm Size

Employees	Frequency	Percent
1 - 49	7	8.1
50 - 249	26	30.2
250 - 499	20	23.3
500 - 999	15	17.4
1000+	18	20.9
Total	86	100.0

Table 4: Project Type

Type	Frequency	Percent
Commercial	50	58.1
Residential	8	9.3
Heavy Civil	15	17.4
Mix-Use	8	9.3
Institutional	5	5.8
Total	86	100.0

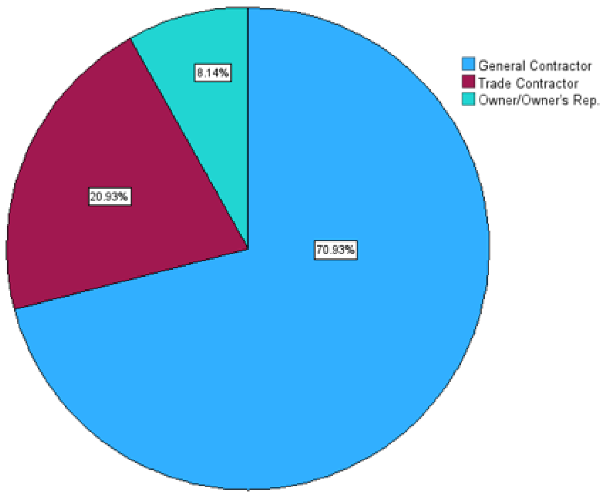


Figure 1:Firm Type

Results and Discussion

Usage of Technologies in the Industry (RQ1)

Figure 2 below shows the types of technologies that construction contractors are utilizing to mitigate the impact of workforce labor shortages.

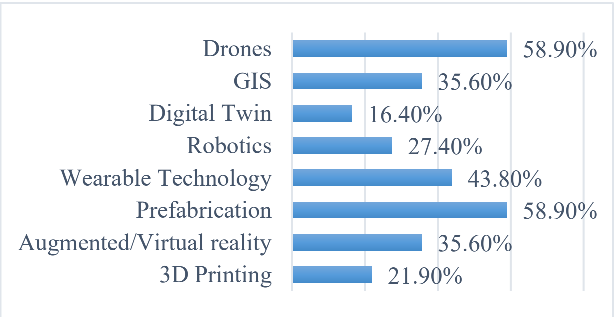


Figure 2:Technology Type

Drones and Prefabrication are the most widely adopted technologies, both at 58.9%. This suggests a significant trend in the construction industry towards automating and streamlining processes. Drones can be used for surveying, monitoring job sites, and ensuring that projects are progressing as planned while offsetting tasks usually performed with manual labor. Prefabrication indicates a shift towards constructing parts of buildings in a controlled environment before being transported to the site, improving efficiency, and requiring less on-site labor. Wearable technology emerges as another significant contributor, with 43.8% of industry professionals integrating it into their operations. This category includes advanced gear such as exoskeletons, which alleviate physical strain on workers, and smart helmets that enhance operational efficiency through real-time data delivery. The integration of such technology not only compensates for labor deficits but also promotes worker safety and productivity. Both Geographic Information Systems (GIS) and Augmented/Virtual Reality (AR/VR) are being utilized by 35.6% of respondents. GIS is used for planning and managing construction projects, while AR/VR can be used for training purposes and to visualize projects before they are built, which can reduce the need for rework and thus the demand for additional labor. Further, Robotics are being used by 27.4% of the respondents. Robots can perform tasks such as bricklaying, welding, and even more complex construction activities, which helps to alleviate the need for skilled labor. At 21.9%, 3D printing technology is adopted by a fifth of the respondents. This technology can be used to create building components or even entire structures with less manual labor required. Lastly, Digital Twin technology, while currently the least adopted at 16.4%, holds a significant promise for the future of construction and represents a significant undeveloped potential. This technology creates a virtual replica of physical construction sites, allowing for meticulous planning, real-time monitoring, and maintenance which, in turn, can lead to a more efficient allocation and utilization of labor.

Firms' Segmentation and Operational Profile (RQ1)

The two-step cluster analysis was utilized to categorize firms into five distinct groups, leveraging the inherent groupings that arise from variations in key variables within the dataset. This analysis considered four variables: Firm Size; Technology Usage, which indicates whether firms employed technology to mitigate labor shortages; Project Type; and Productivity, a continuous variable that measures the output of craftsmen and office personnel, including project managers, estimators, schedulers, and project engineers. The analysis resulted in five unique clusters as shown in Figure 3, with Figure 4 providing a graphical representation of the relative distribution of these clusters based on the four variables. Cluster 1, accounting for 17.4% of the sample, included medium-sized firms with 250-499 employees. These firms uniformly utilized technology and primarily engaged in heavy civil projects, recording a productivity total of 6.93. Cluster 2, which accounted for 15.1% of the firms, consisted mainly of small firms with 1-49 employees. Most of these firms did not use technology and focused on commercial projects, achieving a productivity total of 6.00. Cluster 3, representing 24.4% of the sample, included medium-scale commercial technological firms with 500-999 employees. These firms had a complete adoption of technology and closely matched Cluster 1 in productivity total with 6.95. Cluster 4 comprised 25.6% of the firms, including small-to-medium-sized firms (50-249 employees) that fully adopted technology and specialized in commercial projects, with a productivity total like Cluster 3 of 6.95. Finally, Cluster 5, which included very large firms with over 1000 employees, formed 17.4% of the sample. Most of these firms were technology users focusing primarily on commercial projects and showed the highest productivity total of 7.47. The segmentation indicated by the results is marked, with firm size, technology adoption,

Cluster	4	3	1	5	2
Size	25.6% (22)	24.4% (21)	17.4% (15)	17.4% (15)	15.1% (13)
Inputs	Firm Size 50 - 249 (100.0%)	Firm Size 500 - 999 (57.1%)	Firm Size 250 - 499 (46.7%)	Firm Size 1000+ (93.3%)	Firm Size 1 - 49 (46.2%)
	Technology Usage Yes (100.0%)	Technology Usage Yes (100.0%)	Technology Usage Yes (100.0%)	Technology Usage Yes (86.7%)	Technology Usage No (92.3%)
	Project Type Commercial (68.2%)	Project Type Commercial (81.0%)	Project Type Heavy Civil (60.0%)	Project Type Commercial (73.3%)	Project Type Commercial (53.8%)
	Productivity 6.95	Productivity 6.95	Productivity 6.93	Productivity 7.47	Productivity 6.00

Figure 3: Grouping of Clusters

project type, and productivity levels acting as distinguishing characteristics. Clusters 3, 4, and 5 demonstrated a strong technological integration, suggesting a trend towards digital transformation, especially among medium to very large firms. Clusters 1 and 5, which include the larger firms, reported higher productivity totals, implying a possible link between firm size and productivity, potentially attributable to economies of scale or greater resource investment in adaptation and process optimization. Notably, Cluster 2 stands out as the sole group where most firms did not embrace technology, coinciding with the smallest firm size and the lowest productivity, highlighting the critical influence of technology adoption on productivity.

The analysis of predictor importance, presented in Figure 5, provides valuable insights into the relative impact of each variable on the cluster assignments. Firm Size emerged as the most significant predictor, suggesting that the number of employees is a fundamental factor in differentiating the clusters. This emphasizes the role of firm size in operational practices, as well as its potential influence on the adoption of technology and the types of projects undertaken. Following in significance was Technology Usage, which stands as the second most critical predictor. This underscores the influence of digital adoption on how firms are classified, reflecting the centrality of technology in shaping business operations and productivity within the industry. Project Type also played a moderate role in the clustering, indicating that the nature of projects—whether commercial or heavy civil—has a discernible, though lesser, effect on the grouping of firms compared to size and technology usage. Productivity, while relevant, was the least influential predictor in the clustering process. This suggests that productivity, as a performance metric, does not contribute to cluster differentiation as strongly as the other variables, potentially because it is impacted by firm size and technology usage to an extent that limits its independent variation across clusters.

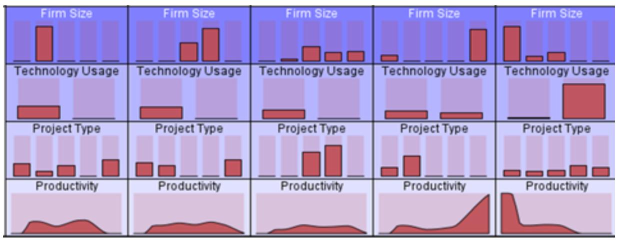


Figure 4:Relative Distribution of Clusters

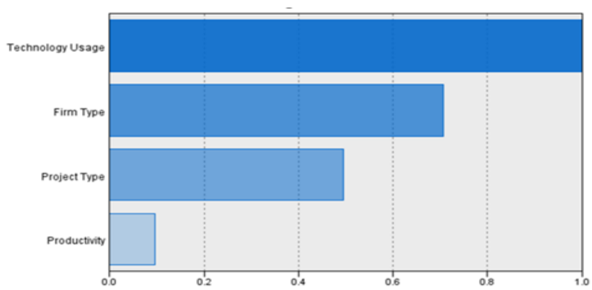


Figure 5: Predictor Importance

Impact of Technology on Productivity (RQ2)

In the context of examining the impact of technology usage on the productivity, the construction firms, particularly as a strategy to combat labor shortages, the Mann-Whitney U test provides a non-parametric means for assessing whether the median productivity scores differ between technology-using firms and non-technology-using firms. This test is particularly apt for situations where the data does not satisfy the assumptions necessary for a traditional t-test, such as non-normal distribution or ordinal measurements, as is the case with the sample for this study. The analysis incorporated 72 firms, shown in Table 5, that reported using technology to mitigate labor shortages, compared to 14 firms that did not employ technology for this purpose. The Mann-Whitney U test results revealed a mean rank of productivity of 46.19 for the technology-using firms, with their cumulative ranks summing up to 3326.00. Conversely, the non-technology-using firms exhibited a lower mean rank of 29.64, with a summed rank of 415.00. The marked disparity in the mean ranks suggests that firms engaging in technology tend to report higher productivity. The Z-score associated with the test statistic was -2.341, which indicates that the observed ranks are 2.341 standard deviations below the mean rank sum that would be expected under the null hypothesis of no difference between the groups. Crucially, the asymptotic significance (two-tailed) value was 0.019. This p-value is indicative of the probability of obtaining the observed results, or more extreme, under the null hypothesis that there is no difference between the two groups' median productivity. Since the p-value is less than the conventional alpha level of 0.05, the null hypothesis can be rejected at the reference confidence level. This implies that the difference in productivity ranks between firms using technology and those that do not is statistically significant. Therefore, the statistical analysis suggests a clear association between the implementation of technology and increased productivity among construction firms. The implication of this finding is that technology usage appears to be an effective measure in addressing labor shortages within the industry, as evidenced by the higher productivity rankings among firms that adopt technological solutions. This conclusion aligns with the ongoing narrative within the sector that technology not only serves as a stopgap for labor deficit but also enhances operational efficiencies and outputs. These results can guide industry stakeholders in decision-making processes regarding investments in technology to combat the pervasive challenge of labor shortages.

Impact of Firm Size on Productivity (RQ2)

This analysis sought to determine the relationship between firm size and productivity within the construction industry. Utilizing Spearman's rho, shown in Table 6, to measure the strength and direction of association between these two ordinal variables, we observed a correlation coefficient (rho) of 0.215.

Table 5:Mann Whitney Test

Technology Usage		N	Mean Rank	Sum of Ranks	Mann-Whitney U	Wilcoxon W	Z	Asymp. Sig. (2-tailed)
Productivity	Yes	72	46.19	3326.00	310.000	415.000	-2.341	0.019
	No	14	29.64	415.00				
	Total	86						

a. Grouping Variable: Technology Usage

This positive correlation, while weak, indicates that larger firms tend to have slightly higher productivity levels. Notably, this association was statistically significant (p = 0.047), suggesting that the observed relationship is unlikely to have occurred by chance. Therefore, we reject the null hypothesis in favor of the alternative hypothesis, concluding that there is evidence of a monotonic relationship between firm size and productivity in the population from which the sample of construction firms was drawn. Given the sample size of 86 firms, these findings provide some evidence that firm size may be a factor in productivity outcomes in the construction sector. However, given the modest strength of the correlation, further research should explore additional variables that might influence productivity more strongly.

Table 6:Spearman's rho Test

			Firm Size	Productivity
Spearman's rho	Firm Size	Correlation Coefficient	1.000	.215*
		Sig. (2-tailed)	.	.047
		N	86	86
	Productivity	Correlation Coefficient	.215*	1.000
		Sig. (2-tailed)	.047	.
		N	86	86

*. Correlation is significant at the 0.05 level (2-tailed).

Conclusion

The comprehensive analysis presented in this study offers important insights into how construction firms are leveraging technology to mitigate labor shortages and the subsequent impact on productivity. Particularly, the study's findings underscore the prevalence of technological adoption across the construction sector, with a particular emphasis on drones and prefabrication as leading strategies. The data suggests a strong relationship between firm size, technology usage, project type, and productivity, with larger firms demonstrating an inclination towards higher productivity levels, potentially reflecting the benefits of economies of scale and enhanced resource allocation.

The segmentation of firms into five distinct clusters highlights the nuanced ways in which these variables interplay to shape operational profiles within the industry. Firms that embrace technology, particularly those of medium to large scale, show a clear trend towards higher productivity totals, reinforcing the narrative that digital transformation is a key driver of efficiency and output in the face of labor challenges. In contrast, the standout characteristic of the smallest firms—many of which have not adopted technology—correlates with the lowest productivity scores, emphasizing the pivotal role of technology in enhancing productivity. Empirical findings from the Mann-Whitney U test and Spearman's rho analysis further validate the hypotheses set forth at the outset of the study, i.e., confirming a statistically significant difference in productivity between firms that utilize technology and those that do not, as well as a positive correlation between firm size and productivity. However, the study's conclusions are offered with the acknowledgement of several limitations. The method of data collection, primarily through university job fairs, could potentially introduce a selection bias. This is because firms participating in these fairs might be inherently more inclined towards innovation and technological advancements, while leveraging employment shortages, possibly skewing the sample towards more technologically progressive companies. Another key limitation lies in the reliance on self-reported data, which carries an inherent risk of reporting biases, as firms might tend to overestimate their productivity or the impact of their technological implementations. Lastly, the cross-sectional design provides a snapshot that cannot capture the dynamic nature of technology adoption over time. Despite these limitations, the findings suggest that the construction industry is actively adapting to labor challenges through technology, and the high response rate reinforces the reliability of this conclusion. This research has significant implications for industry stakeholders, suggesting that investments in technology are not only a viable response to labor shortages but also a critical component of enhancing overall firm productivity. The study’s insights are valuable for informed strategic decision-making and policy formulation aimed at improving growth and sustainability within the construction sector. Future research should aim to build upon these findings, exploring the broader implications of technology adoption and firm size on productivity, with an eye towards identifying other contributory factors that could further refine our understanding of these complex dynamics.

References

ABC. (2023, February). News Releases | Construction Workforce Shortage Tops Half a Million in 2023, Says

ABC. https://www.abc.org/News-Media/News-Releases/construction-workforce-shortage-tops-half-a-million-in-2023-says-abc. Accessed May 2023.

Adaloudis, M., & Bonnin Roca, J. (2021). Sustainability tradeoffs in the adoption of 3D Concrete Printing in the construction industry. Journal of Cleaner Production, 307. https://doi.org/10.1016/j.jclepro.2021.127201.

Al-Bayati, A. J., Tafazzoli, M., York, D. D., & Umar, T. (2020). Cyclical Construction Workforce Shortage: An Evaluation of the Current Shortage in Western North Carolina. Construction Research Congress 2020. https://doi.org/10.1061/9780784482872.070.

Antunes, R., & Poshdar, M. (2018). Envision of an integrated information system for project-driven production in construction. IGLC 2018 - Proceedings of the 26th Annual Conference of the International Group for Lean Construction: Evolving Lean Construction Towards Mature Production Management Across Cultures and Frontiers, 1, 134–143. https://doi.org/10.24928/2018/0511.

Associated General Contractors of America. (2022). 2022 Workforce Survey Analysis. https://www.agc.org/sites/default/files/users/user22633/2022_AGC_Workforce_Survey_Analysis.pdf. Accessed May 2023.

Cai, S., Ma, Z., Skibniewski, M. J., & Bao, S. (2019). Construction automation and robotics for high-rise buildings over the past decades: A comprehensive review. Advanced Engineering Informatics, 42, 100989. https://doi.org/10.1016/j.aei.2019.100989.

Cai, S., Ma, Z., Skibniewski, M., Guo, J., & Yun, L. (2018). Application of Automation and Robotics Technology in High-Rise Building Construction: An Overview. https://doi.org/10.22260/ISARC2018/0044

Changali, S., Mohammad, A., & Van Nieuwland, M. (2015). The construction productivity imperative.

Cheung, S., Yazdani, S., & Ghafoori, N. (2011). CONSTRUCTION LABOR SHORTAGE, CHALLENGES, AND SOLUTIONS: A SURVEY-BASED APPROACH. Proceedings of International Structural Engineering and Construction, 6. https://www.isec-society.org/ISEC_PRESS/ISEC_12/pdf/CON-16.pdf.

Goh, M., & Goh, Y. M. (2019). Lean production theory-based simulation of modular construction processes. Automation in Construction, 101, 227–244. https://doi.org/10.1016/j.autcon.2018.12.017.

Gupta, S. (2019). Identification of KPIs in Construction Industry: A Review. International Journal of Construction Engineering and Planning, 5(1), 28–38. https://civil.journalspub.info/index.php?journal=IJCE&page=article&op=view&path%5B%5D=500.

Halder, S., Afsari, K., Chiou, E., Patrick, R., & Hamed, K. A. (2023). Construction Inspection & Monitoring with quadruped robots in future human-robot teaming: A preliminary study. Journal of Building Engineering, https://doi.org/10.1016/j.jobe.2022.105814.

Halder, S., Afsari, K., Serdakowski, J., DeVito, S., Ensafi, M., & Thabet, W. (2022). Real-time and remote construction progress monitoring with a quadruped robot using augmented reality. Buildings, https://doi.org/10.3390/buildings12112027.

Hossain, M. A., Zhumabekova, A., Paul, S. C., & Kim, J. R. (2020). A review of 3D printing in construction and its impact on the labor market. In Sustainability (Switzerland) (Vol. 12, Issue 20, pp. 1–21). MDPI. https://doi.org/10.3390/su12208492.

Jackson, N. (2020). How New, Innovative Technology Can Be Implemented to Combat the Construction Industry Labor Shortage. Associated General Contractors of America. https://www.agc.org/sites/default/files/Files/Foundation/CLEAN%20Noah%20Jackson%20%281st%29.pdf.

Karimi, H., Taylor, T. R. B., Dadi, G. B., Goodrum, P. M., & Srinivasan, C. (2018). Impact of Skilled Labor Availability on Construction Project Cost Performance. Journal of Construction Engineering and Management, 144(7). https://doi.org/10.1061/(asce)co.1943-7862.0001512.

Kim, S., Chang, S., & Castro-Lacouture, D. (2020). Dynamic Modeling for Analyzing Impacts of Skilled Labor Shortage on Construction Project Management. Journal of Management in Engineering, 36(1). https://doi.org/10.1061/(asce)me.1943-5479.0000720.

Liu-Lastres, B., Wen, H., & Huang, W. J. (2023). A reflection on the Great Resignation in the hospitality and tourism industry. International Journal of Contemporary Hospitality Management, 35(1), 235–249. https://doi.org/10.1108/IJCHM-05-2022-0551.

Melenbrink, N., Werfel, J., & Menges, A. (2020). On-site autonomous construction robots: Towards unsupervised building. In Automation in Construction (Vol. 119). Elsevier B.V. https://doi.org/10.1016/j.autcon.2020.103312.

Naderi, H., & Shojaei, A. (2022). Civil Infrastructure Digital Twins: Multi-Level Knowledge Map, research gaps, and future directions. IEEE Access, 10, 122022–122037. doi:10.1109/access.2022.3223557.

Opoku, D. G. J., Perera, S., Osei-Kyei, R., & Rashidi, M. (2021). Digital twin application in the construction industry: A literature review. In Journal of Building Engineering (Vol. 40). Elsevier Ltd. https://doi.org/10.1016/j.jobe.2021.102726.

Sadeh, H. (2023). Digitalization of the construction sector (AEC) : The case of small and medium sized contractors. Politecnico di Milano. https://www.politesi.polimi.it/handle/10589/207413.

Sadeh, H., Mirarchi, C., & Pavan, A. (2021). Technological transformation of the construction sector: A conceptual approach. International Journal of Construction Management, 23(10), 1704–1714. doi:10.1080/15623599.2021.2006400.

Sadeh, H., Mirarchi, C., & Pavan, A. (2022). CLASSIFICATION OF CONSTRUCTION FIRMS BASED ON BIM ROLES AND BIM LEVELS USING MACHINE LEARNING TECHNIQUES. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVI-5/W1-2022, 205–210. https://doi.org/10.5194/isprs-archives-xlvi-5-w1-2022-205-2022.

DIGITAL BUILDING LOGBOOKS AND STAKEHOLDER MAPPING: A CROSS-SECTORAL APPROACH TO DATA COLLECTION AND SHARING

Sun-Ah Hwang[1]
[1]Department of Management in the Built Environment, Delft University of Technology, the Netherlands

Abstract

Digital building logbooks (DBLs) are designed to provide a platform to ease data sharing among stakeholders between and beyond the architecture, engineering, construction and operation (AECO) sectors. However, there is a gap in understanding among stakeholders regarding their roles and the overlaps and differences in the data they need. With the findings from qualitative research, the paper provides a new framework to DBL stakeholder mapping that facilitates the identification of stakeholders necessary for effective data exchange and efficient decision-making at different phases across the building lifecycle. This approach defines the roles of the stakeholders as data providers, data users or both, providing an initial impression to how building data can be sourced into and out of the DBL. The paper ultimately demonstrates the relevance of building data for all identified stakeholders, emphasizing the importance of a cross-sectoral involvement in DBLs.

Introduction

The European goal to increase the current renovation rate of 0.4-1.2% to 3% has been challenged by several barriers (Gómez-Gil et al., 2022). One such barrier often overlooked in literature is the lack of building data.

The limitations in the availability, reliability and storage thereof have been viewed as recurring impediments in the decarbonization of the built environment; namely when evaluating the real state of buildings, measuring decarbonization progress, advocating the urgency of retrofits, and during decision-making for planning renovation and maintenance (Monzón & López-Mesa, 2018; Blum, 2009; Talamo & Bonanomi, 2015).

The lack of building data sees great urgency, though it has rarely been addressed in literature (Gómez-Gil et al., 2022). While there is a continuous inflow of new building data, there still seems to be a lack of data particularly on the existing building stock. Gomez-Gil et al. (2022) describe the problem of data loss and unavailability as a "multidimensional problem which affects all scales: local, national, and European". Locally at municipal level, it is still a problem in European member states that the issuing of construction and renovation licenses is paper based, increasing risks of information loss and limiting possibilities of information processing and sharing. Where old but digital data is available, they are "available in spreadsheets or flat tables [and] linkages between them [are] not possible" (Formosa & Pace, 2022).

To provide an indication on the potential extent of the shortage in building data, the Housing Europe Observatory (2023) reported that dwellings built before the oil crisis in the 1970s accounted for 25-67% of the total stock in the 27 European countries studied. It can be deduced that if data and information on these dwellings are available at all, the likelihoods of them being dated, out of context, one-off, not in digital or easily changeable formats, (very) partially covered and/or has adopted a questionable research methodology are significantly high.

What should be stressed here is that the existing housing stock exceeds that of the new in most developed countries (Housing Europe Observatory, 2023). For instance, in the Netherlands, new constructions account for less than 1% of the existing housing stock each year (Centraal Bureau voor de Statistiek, 2023a, 2023b). This means that today's existing housing stock will account for at least 70% of the total building stock in 2050. Therefore, decarbonizing the available housing stock should be prioritized when concerning the 2050 climate goals.

To overcome building data loss and to increase the data accessibility and interoperability, Digital Building Logbooks (DBLs) were first introduced at the European level to address the fundamental issue of data loss and unavailability that is grounded on the asymmetric and obsolete way of data collection by all relevant stakeholders involved (Gómez-Gil et al., 2022). The DBL is designed to record and trace every change made to a building across its lifecycle, including but not limited to changes in ownership, tenure or use, maintenance and refurbishments. It is intended to be a dynamic tool that enables improved logging, management and accessibility of all data, information and documents on buildings. For that purpose, it holds "administrative documents, plans, description of the land, the building and its surrounding, technical systems, traceability and characteristics of construction materials, performance data such as operational energy use, indoor environmental quality, smart building potential and lifecycle emissions, as well as links to building ratings and certificates" (European Commission et al., 2020, p. 12).

The proposal to use the DBLs to promote and design building renovation processes, on the other hand, is recent (Gómez-Gil et al., 2022). It was recognized in the Renovation Wave Strategy as an autonomous tool to contribute to "create better conditions for staged renovation[s]" (European Commission, 2020). From what started off as a digital repository, DBLs have since evolved to bear the capacity to convert building data into actionable information for all relevant stakeholders. Essentially, the DBL is founded on the grounds that all information and data stored in DBLs are easily shared between and beyond the stakeholders of the architecture, engineering, construction and operation (AECO) sectors.

However, the built environment and construction sector are traditionally fragmented with the immense numbers of stakeholders involved. These stakeholders often have conflicting interests; the difference in information needs, use of data and purposes is inevitable. Data that is relevant and of value to every stakeholder involved is often unavailable in one place and a systematic approach to organizing and managing it is largely missing. The challenge thus lies in the gap in the understanding among stakeholders of the overlaps in the data they need in common, as well as the significance of their roles in DBLs and the larger ecosystem.

This paper therefore aims to address the gap in stakeholder understanding in the context of building data in DBLs, emphasizing the importance of cross-sectoral collaboration in data collection and sharing. This paper results in a new framework mapping DBL stakeholders that clearly highlights the importance of engaging stakeholders from both within and beyond the AECO sectors to ensure that: (1) the DBL is used at its full capacity, (2) reliable data is available and can be accessed in DBLs and (3) stakeholders recognize the urgency for cross-sectoral collaboration in data collection and sharing.

Research Methods

The research involved a literature review and qualitative research, in addition to the synthesis of the results from the two methods through stakeholder mapping.

To understand the scope of all internal and external stakeholders involved, this paper first builds on the findings from several EU studies on DBLs. As a starting point, the early report on the development of a European Union framework for DBLs by European Commission et al. (2020) defined and identified the key stages of a building lifecycle wherein the role of DBLs is significant, as well as 17 stakeholder groups from across the entire construction and built environment value chain involved in the process.

This list of stakeholders was refined and categorized in the subsequent DG-GROW project led by Ecorys et al. [1], which aims to develop a European model for DBLs to promote tools and protocols that enable data sharing and usage throughout the construction ecosystem. In their final report, Ecorys et al. (2023) grouped the stakeholders into five main user categories and outlined their potential DBL use based on their needs.

This paper then expands on the initial desk findings through two-part qualitative research: a survey questionnaire and a series of focus group sessions. The online survey questionnaire (65 respondents) was conducted to explore the (desired) roles of the stakeholders in the various DBL systems.

The second part of the qualitative research is conducted in collaboration with Demo-Blog [2]. Short for *Development and Demonstration of Digital Building Logbooks*, Demo-BLog is a Horizon Europe funded project which sees to cultivate a decision support tool that enables existing initiatives to identify optimal sustainability transition pathways that are driven by clustered DBL data. It aims to further exhibit how DBL data can serve to advance the current evaluation strategies for climate and energy transition implemented at various levels in Europe.

Demo-BLog brings together five currently operational DBLs in Europe to develop and demonstrate four diverse functionalities addressing key societal challenges, ranging from ‘quick wins’ to complex industrial transaction objectives including circularity. Table 1 provides and overview of the five DBLs and its active regions.

Table 1: The five DBLs studied in Demo-BLog (Grant agreement ID: 101091749).

Name DBL	Owning bodies	Active regions
Chimni	Chimni	United Kingdom
CLÉA	Qualitel	France
CAPSA	Chillservices	Germany, Scotland, the Netherlands, and Italy
Woningpas	VEKA, OVAM, Wonen-Vlaanderen and Departement Omgeving	Flanders (Belgium)
Cirdax	Re-Use Materials	The Netherlands and Belgium

With Demo-BLog, a series of focus group sessions were conducted to further identify individual stakeholder contributions and significance. The sessions were conducted in four regions where the DBLs studied had the most influence in: United Kingdom, Flanders (Belgium), France and Germany. Table 2 provides an overview on how the DBL is positioned in the four study regions.

Table 2: Overview of the positioning of the DBL in the four study regions (Findings from Demo-BLog).

Study regions	Positioning of the DBL
United Kingdom	The use of a DBL is not mandatory. There are six commercial providers* that together, at the UK government’s instigation, form the Residential Logbook Association (RLBA) to bring the DBLs to a common standard and promote data interoperability.
Flanders (Belgium)	The use of a centralized, government owned DBL, *Woningpas*, is mandatory for all homes in Flanders. It is automatically available for all building owners, both natural persons as well as businesses.
France	The use of DBLs for homes has been mandatory for all new buildings since 1 January 2023. Since then, a DBL must be established at moments of construction or renovation, equipped with an analysis on the environmental impact.
Germany	The use of a DBL is not mandatory. In comparison to the other regions studied, the stakeholders indicated that the lack of building data is most severe in Germany.

**Note that there are other property logbook providers that exist outside of RLBA.*

The synthesis of the review results and the empirical analysis resulted in the subsequent mapping of the stakeholders according to the phase(s) of the building lifecycle in which they are (potentially) engaged to the DBL. The analysis showed that stakeholders can be largely categorized into two groups in the context of DBL data: data users and/or providers. This approach was essential when identifying the type of data exchange that is necessary at various points in the building lifecycle, as well as how the gaps in data sourcing can be resolved with the larger involvement of stakeholders.

Literature Studies

DBLs: EU's solution to data loss and unavailability

The European Commission et al. (2020) describe that "the capturing and maintenance of data and information is the backbone of the DBL". They claim that a systematic, well-organized and standardized scheme for data collection and storage has the capacity to alleviate the following barriers that is observable in the current scene:

"Firstly, due to the absence of a systematic approach to capturing, storing, analysing and organising it, valuable data and information are lost. Secondly, the storage of data is fragmented and scattered across several organisations (and even departments within the same organisation). Thirdly, data that is collected and stored by one individual actor is not necessarily accessible and available to other actors in the value chain." (European Commission et al., 2020, p. 15).

A DBL is thus designed to improve transparency and trust between the various stakeholders involved.

DBL stakeholders as defined by early EU studies

Early European studies on DBLs conducted identified 17 stakeholder groups that are (potential) benefiters and/or contributors of DBLs (European Commission et al., 2020). Table 3 provides an overview of the list.

Table 3: 17 stakeholder groups as identified by the European Commission et al. (2020).

Stakeholder groups
Landlords and owner-occupiers (incl. prospective buyers and sellers)
Tenants
Designers
Developers
Construction contractors
Investors
Banks and insurers
Material suppliers (incl. urban miners)
Facility and building managers
Demolition contractors
Public authorities and policymakers (ie. urban planners)
Real estate agents
Researchers
Utilities
Certifiers
Lawyers, solicitors and conveyancers
Valuers

Lifecycle approach and users of a DBL

Opportunities for data collection are present at every stage of the building lifecycle, and they each exhibit different needs of data usage (European Commission et al., 2020). Naturally, every stakeholder requires access to accurate data that can seamlessly address their needs at different points of the building lifecycle. For the AECO sectors — and any other relevant— to leverage the full potential of structured data, the processed information should be transferred and available integrally throughout the entire building lifecycle. Figure 1 presents the preliminary stakeholder mapping in relation to the DBL conducted by the European Commission et al. (2021).

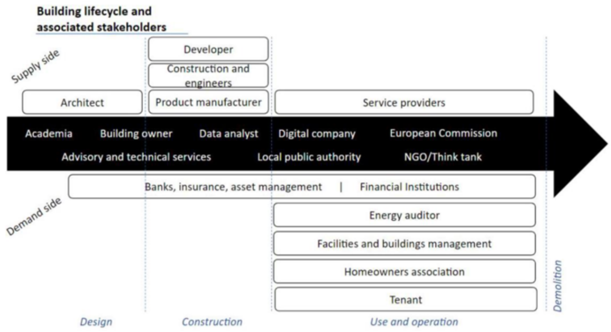


Figure 1: Stakeholder mapping in relation to the DBL (European Commission et al., 2021).

Figure 2 illustrates the simplified building lifecycle, with the phases in which the various stakeholders interact with the building defined.

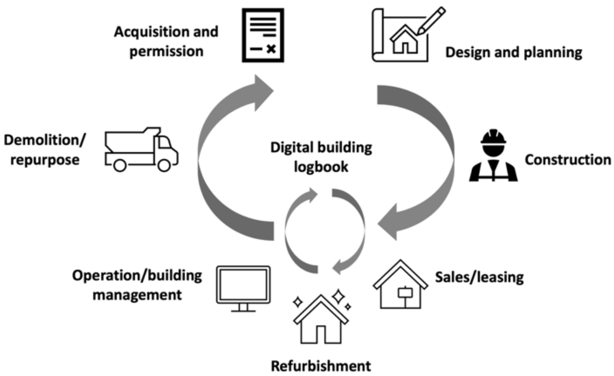


Figure 2: Simplified building lifecycle and the role of the DBL (European Commission et al., 2020).

In the context of DBLs and data, the various phases can be grouped to three larger phases (European Commission et al., 2020). Firstly, the design, planning and construction phases present the prime moment for the collection of data on the physical characteristics of the building, including information on the materials used and where they are located in the building. The collected information can then be used to prove compliance with building regulations and certification schemes (European Commission et al., 2020). Functionalities such as the construction project management can be used to ease information sharing between the various stakeholders involved, and the primary stakeholders in these phases can be said to be designers, architects, developers, contractors and material suppliers.

In the sales and/or leasing, operation and property management phases, data collection largely revolves around the operation and performance of the building. Such data may touch on maintenance, ownership transfer and change of use. The processed information can then facilitate, for instance, the identification of maintenance and renovation needs, the monitoring of user behaviour and the tracking of administrative requirements (European Commission et al., 2020). The information is likewise important for financing, as well as transaction underwriting and execution. Key stakeholders include building owners, tenants, facility managers, utility companies, real estate service providers, energy auditors, contractors and the financial sector.

The repurpose or demolition phases make use of all the data collected on the building, its composition and materials to support decision making processes when strategizing the next best step for the building; be it to refurbish, repurpose or demolish, or to optimize or extract the most value from recycling the building materials (European Commission et al., 2020). Key stakeholders include building owners, demolition companies, product maintenance service companies and recycling companies.

Considering the complexity and fragmented nature of the AECO sectors, it is evident that the data architecture of DBLs should place extra emphasis on the diversity in the data source including both legacy systems and state-of-the-art, the interoperability of the tool, as well as potential third party data processing tools and sources.

Recent findings from Ecorys et al. (2023)

In the subsequent DG-GROW project, the stakeholders that were identified by the European Commission et al. (2020; see Table 3) were grouped into five main user categories as can be seen in Table 4 (Ecorys et al., 2023).

Table 4: Five user categories as identified and defined by DG-GROW (Ecorys et al., 2023).

User categories	Primary use of a DBL
Governmental agencies (GA)	To obtain insights for policymaking, the issuing of licenses and enforcement of regulation and disaster management
Construction sector (CS)	To obtain building-related information on the design and build.
Building owners (BO)	To collect information about their buildings for exploitation and maintenance purposes.
Financial institutions (FI)	To perform analysis based on formulas to gain insight into the assets market and its developments and gain an understanding for building transactions.
Utility companies (UC)	To obtain information on the connections of utilities to the building and the analysis of its uses, and provide information related to building use or performance.

The study later developed to creating user journeys of the five user categories that highlights the (potential) use of the DBLs of each category. The exemplified user journeys of the four user groups can be seen in Figure 3.

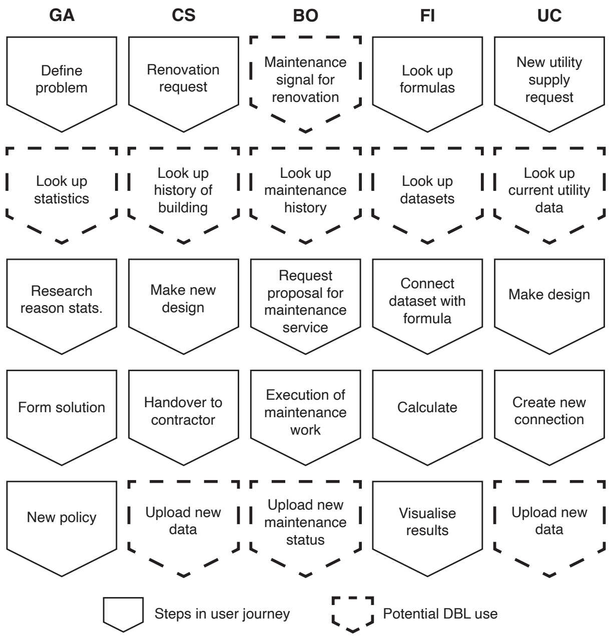


Figure 3: The exemplified user journeys of the four user categories and their potential use of the DBL (Ecorys et al., 2023).

A key finding that is significant in this study is that some stakeholders, in this case the CS, BO and UC user categories, are seen as both data users and providers. This multiplicity in the roles of the stakeholders at different phases of the building lifecycle in the DBL system provides a new light to the research on DBLs, that the two-way interaction between stakeholders and the tool, and the need for stakeholders to proactively engage with the DBL to ensure that the data stored therin is updated, accurate and remains valuable.

Survey Findings

We conducted a survey among those stakeholders identified from available literature —including but not limited to building professionals, building owners, data owners, government agencies— to identify the requirements, functionalities and benefits that are relevant to the stakeholders; for instance, data, equipment and technology, standards, and benefits such as decarbonisation, cost efficiency, resource allocation and access to trusted and traceable information. Table 5 provides a breakdown of the survey participants.

Table 5: Breakdown of survey participants in stakeholder groups.

Stakeholder group	%
Landlords and/or leaser	5
Owner-occupiers	18
Tenants (Lessees)	5
Architects and designers	5
Developers (Real-estate)	2
Banks and/or insurers	1
Building material suppliers	2
Facility and/or building managers	2
Valuers	1
Certifiers	3
Researcher	20
Public authorities	11
Policy makers	4
Other*:	21

**Other stakeholder groups as indicated by the respondents include DBL providers, service designers, data(base) companies and providers, as well as IT developers.*

Findings from the survey that is relevant to this paper can be largely categorized into two themes: data collection and data accessibility. On the former, the survey helped prove a valuable hypothesis; while stakeholders can benefit as users of the tool and the data therein, they can simultaneously play the role of data providers. Though the level of responsibilities across the data fields may differ, building owners, public authorities, building experts, financial institutions, and both public and private registers are perceived as equally valuable (potential) data providers for the efficient collection of building data.

For instance, homeowners, construction contractors, architects and facility managers are not only identified as crucial data providers for their role in updating every change made to the homes into the DBL, but are equally active data users when assessing and planning the necessary retrofit works. Government authorities can provide access to public databases that would supplement the data in the individual DBLs, while policymakers need the overview of the housing stock to efficiently distribute resources. Financial institutions, in contrast, would be active data users for efficient valuations of homes and encourage a wider roll out of DBLs in the market.

The survey also emphasized on the importance to note that the role of the stakeholder as data providers is not one-off. To ensure that the building data provided by the DBL is up to date, the regular revisions of building data is fundamental, imposing that the continuous engagement of the relevant stakeholder to the DBL throughout the entire building lifecycle is crucial.

When concerning the issue of data accessibility, the survey identified a list of third parties that could and/or should be granted access to (part of) the data a DBL holds. Several stakeholders raised here include buyers, architects, energy advisors, notaries, material and utility suppliers, estate agents, property lawyers, enterprises, financers, assurers, housing advisors and service providers. It can be deduced from this list that stakeholders across the entire construction and built environment value chain should orbit around the DBL, for the benefit of greater overall sectoral transparency, value chain integration, innovation and circularity.

Focus Group Sessions

Findings from both literature and the survey were thereafter employed in the composition of the focus group sessions with key stakeholders from four DBL-active regions: United Kingdom, Flanders (Belgium), France and Germany. Participants for the sessions were gathered through consultations with the DBL providers. 19 key stakeholders were identified as crucial contacts in the implementation and development of the DBLs across the four regions. Table 6 identifies the location and domains of the focus group participants.

Table 6: Breakdown of focus group participants.

Location	Participant	Stakeholder domain(s)
United Kingdom	1	Quality scheme
	2	Quality scheme
	3	DBL provider and DBL trade body
	4	Energy accreditation scheme I
	5	Energy accreditation scheme II and energy professional
	6	Independent climate organization
Flanders (Belgium)	1	Government climate agency I
	2	Government climate agency II
	3	Homeowner association
France	1	DBL provider
	2	DBL provider
	3	Certification scheme
	4	Consumer association
Germany	1	DBL provider
	2	Building professional (academia)
	3	Independent consultant and surveyor
	4	Climate professional
	5	Independent organization
	6	Energy innovation community

The aim of the focus groups was to study the optimal DBL performance in the respective regions. Part of the session was reserved to further distinguish individual contributions and significance of the stakeholders relevant to the operational DBLs and construction systems of the four regions. This was conducted in the form of a collective stakeholder mapping exercise, where participants of various domains in the study region collaborated in creating the matrix together.

From the focus group sessions, it was deduced that the differences in the language, operations and organizations of the construction sector, as well as the legal framework of a specific geographical region influence the defining and mapping of DBL stakeholders to a large extent. The variance in the regional positioning of the DBL and the developments of the tool in relation to the implementation and technological levels, public awareness and locally available tools —such as public databases and energy accreditation schemes— naturally increased the complexity in identifying and defining relevant DBL stakeholders in the different practical environments. Ultimately, the survey questionnaire and the focus group sessions collectively identified 38 stakeholders that are used in the subsequent mapping exercise, as can be seen in Table 7.

Table 7: 38 stakeholder groups identified from the qualitative studies.

Stakeholder groups	
Owner-occupiers	Real estate agents
Architects and designers	Investors
Construction contractors	IT providers
Certifiers	Local authorities
Public authorities	Public waste agency
Policy makers	Guarantee bodies
Facility and/or building managers	Distribution network operators (DNOs)
Landlords and/or leasers	Renovation advice providers
Researchers	Service designers (UX)
Banks and/or insurers	Retrofit service providers
Utility providers	Social housing providers
Building material suppliers	Surveyors
Developers (real-estate)	Building safety regulator
Tenants (Lessees)	Competent person schemes
Valuers	Energy data providers
Demolition contractors	Funding party
Energy experts	Maintenance contractor
Energy suppliers	International organizations
Data companies (inventories and registering)	Lawyers, solicitors and/or conveyancers

DBL stakeholder mapping exercise

With the information gathered from desk research and preliminary analysis, a mapping template was created to then compare the results with how the stakeholder roles are perceived in practice. Figure 4 (see next page) provides an overview of the combined findings of the four focus group sessions that was later validated by Demo-BLog experts.

The approach to DBL stakeholder mapping proposed in this paper builds on the approach undertaken by European Commission et al. (2021; see Figure 1). Similarly, the mapping of the DBL stakeholders in this exercise is likewise two-part: (1) distinguishing the role of a stakeholder as a data user and/or provider along the horizontal axis of the matrix and (2) identifying the phase(s) in the building lifecycle in which the stakeholder is engaged along the vertical axis.

On the former, a data user is defined to be benefiters of the building data stored in the DBL. These stakeholders need specific data to address specific needs. For instance, an owner-occupier may want to observe data on the energy consumption of the home to identify energy-saving practices. A data provider on the other hand represents the supply side of the spectrum. Taking the same example, the utility providers may either provide monthly statements on the energy consumption of the home or install meters that provide real-time information on the energy use. The significance of stakeholders of both categories is also measured in the matrix to distinguish the levels of value and contribution per stakeholder.

The phases in the building lifecycle were first defined through desk research on the preceding European studies as was discussed in earlier sections of this paper, refined later with the inputs from the focus group studies on the operational aspects of the DBLs in the regions studied.

Every stakeholder analyzed in this mapping exercise is tagged to the relevant region to provide context on the difference in stakeholder roles and significance in the locations studied. As indicated in the legend, the stakeholders raised in the discussion in Flanders (Belgium), are tagged in orange, France in blue, Germany in green and the United Kingdom in pink.

Several ambiguities were raised in the validation stage of the stakeholder mapping. Conclusions largely orbit around a common observation: there is too little overlap between the DBL active regions than was hypothesized. A number of causes were identified here:

1. Inconsistencies in the representation of the stakeholder groups during the focus group sessions;
2. The inclusion of personal aspirations rather than reflect how the DBLs are used at volume; and
3. Gaps in terminology due to differences in processes and language.

However, the exercise itself proved that a critical reflection into the AECO sectors as well as the legal and operational systems in specific regional contexts is fundamental when defining the stakeholders of the DBL.

For instance, the roles of the competent person schemes, guarantee bodies and existing renovation advice and/or service providers in the renovation phase of the building are unique in the United Kingdom, while these roles are most likely present in different entities in other regions. In Flanders on the other hand, energy experts and the public waste agency exhibited great interest in the available DBL data in both the pre-use and the end-use phases of the building to identify how the (newly) available resources can be redistributed.

Being able to identify how the roles are dispersed across unique legal systems and the various steps involved has shown to help create discourses around the DBL and the possibilities of data sharing among stakeholders of different domains. Such discourses are crucial to close the gap in understanding among these stakeholders to not only increase efficiency in the collection of building data, but also the awareness on the significance of their roles in DBLs and the larger ecosystem to ensure that the collaboration between the sectors is present. The common

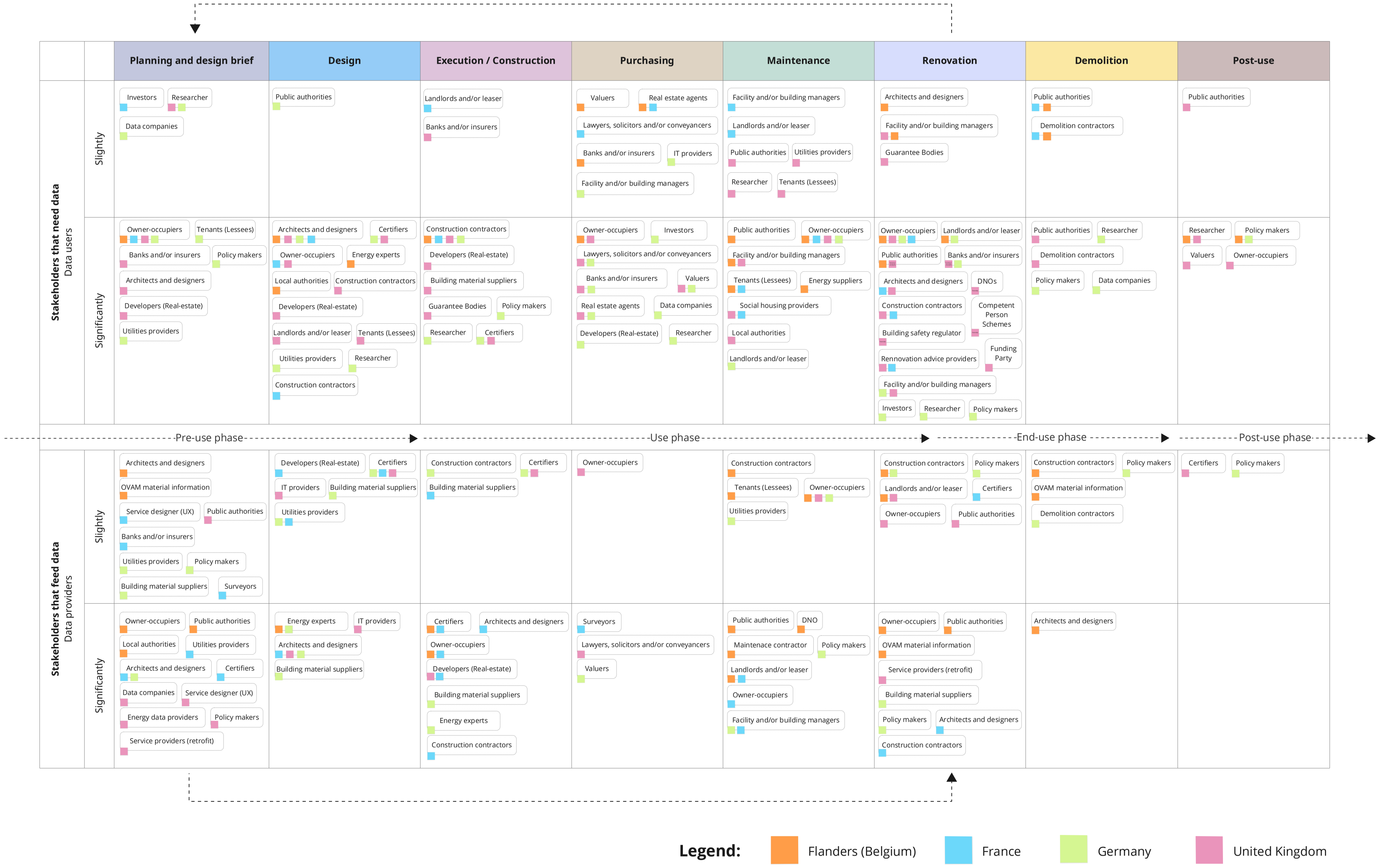


Figure 4: Combined findings of the stakeholder mapping with key stakeholders in the four regions studied, validated with DBL experts (Original work created for Demo-BLog, 2023)

consensus established will then positively open opportunities to further develop a centralized data collection method, a holistic quality scheme for the data collected, as well as the interoperability of the DBL that provides a secured platform for data sharing.

Conclusions

This paper discusses the challenge of stakeholder understanding in the context of digital building logbooks and calls for a wider discourse around the DBL for the cross-sectoral involvement in data collection and sharing. This is a crucial and challenging topic because the DBL and the EU policies around it are still under development. Previous studies indicated that a lifecycle approach to identifying DBL stakeholders is key, but the regional context must likewise be adequately anticipated to yield improved stakeholder engagement.

The findings of this paper indicate that the existing European approaches to DBL stakeholder mapping require a more critical approach. This paper therefore proposes a more comprehensive methodology which involves co-creation between stakeholders of various domains across the construction market value chain and a mapping matrix to visually represent the DBL stakeholder ecosystem that is unique to specific contexts. A key takeaway from this exercise is to place emphasis on the fact that the engagement of the stakeholder to the DBL is two-way, and stakeholders should be viewed as both data users and providers. This additional step helps provide an indication to the overlaps in the inflow and outflow of DBL data among stakeholders, therefore demonstrating the relevance of building data for all identified stakeholders.

Acknowledgments

This research has received funding from the European Union's Horizon Europe research and innovation programme, under the Demo-BLog project (Grant agreement ID: 101091749).

Notes

[1] For more information on the DG-GROW project, see https://www.ecorys.com/case-studies/technical-study-for-the-development-and-implementation-of-digital-building-logbooks-in-the-eu/.

[2] See https://demo-blog.eu/ for more information on the Demo-BLog project.

References

Blum, A. (2009) Documentation, assessment and labelling of building quality the German 'building passport' issue. Sustainable Urban Development, 3.

Ecorys, TNO, Arcadis & Contecht (2023) Technical guidelines for digital building logbooks: Guidelines to the Member States on setting up and operationalising digital building logbooks under a common EU framework. DG-GROW.

Centraal Bureau voor de Statistiek (2023a) Hoogste aantal nieuwbouwwoningen in afgelopen decennium. Available at https://www.cbs.nl/nl-nl/nieuws/2023/05/hoogste-aantal-nieuwbouwwoningen-in-afgelopen-decennium (Accessed: 5 December 2023).

Centraal Bureau voor de Statistiek (2023b) Voorraad woningen; eigendom, type verhuurder, bewoning, regio. Available at https://www.cbs.nl/nl-nl/cijfers/detail/82900NED (Accessed: 5 December 2023).

European Commissions (2020) A Renovation Wave for Europe - greening our buildings, creating jobs, improving lives (White Paper 52020DC0662). Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020DC0662

European Commission, Executive Agency for Small and Medium-sized Enterprises, Dourlens-Quaranta, S., Carbonar, G., De Groote, M. et al. (2021) Study on the development of a European Union framework for digital building logbooks: final report. Publications Office. Available at: https://data.europa.eu/doi/10.2826/659006

European Commission, Executive Agency for Small and Medium-sized Enterprises, Volt, J., Toth, Z., Glicker, J. et al. (2020) Definition of the digital building logbook: report 1 of the study on the development of a European Union framework for buildings' digital logbook. Publications Office. Available at: https://data.europa.eu/doi/10.2826/480977

Formosa S., & Formosa Pace, J., (2022). Digitisation, Digitalisation, Digital Transformation: The Maltese Spatial Encounter in S., Formosa, J., Formosa Pace, & E., Sciberras, (Eds) Virtualis: Social, Spatial and Technological Spaces in Real and Virtual Domains - SpatialTrain III. Planning Authority & Kite Group, pp. 305-317. ISBN 978-9918-23-097-6

Gómez-Gil, M., Espinosa-Fernández, A. and López-Mesa, B., 2022. Review and analysis of models for a european digital building logbook. Energies, 15(6), p.1994.

Housing Europe Observatory, 2023. State of Housing in Europe 2023. Housing Europe. Available at: https://www.housingeurope.eu/resource-1825/the-state-of-the-housing-in-europe-2023

Monzón, M. and López-Mesa, B., 2018. Buildings performance indicators to prioritise multi-family housing renovations. Sustainable Cities and Society, 38, pp.109-122.

Talamo, C. and Bonanomi, M., 2015. Knowledge management and information tools for building maintenance and facility management (Vol. 3). Berlin/Heidelberg, Germany: Springer International Publishing.

ACQUIRING DIGITAL COMPETENCES FOR CONSTRUCTION PROJECT MANAGEMENT

Zoran Pučko[1], and Nataša Šuman[1]
[1] Faculty of Civil Engineering, Transportation Engineering and Architecture, University of Maribor, Slovenia

Abstract

The construction industry is currently undergoing a profound transformation in project management, led by the integration of Building Information Modeling (BIM) and Industry 4.0. BIM is central to modern construction project management and is at the heart of this change. Slovenian universities have adapted civil engineering curricula to equip students with essential digital skills. The paper describes in detail the experience in teaching BIM and advanced technologies, and thus describes the objectives of the subject, competences acquired, teaching methods and outcomes. The article emphasizes the importance of cutting-edge training and looks at BIM-based and Industry 4.0 courses. By integrating advanced methods into civil engineering education, graduates will be equipped with the digital skills they need for their future careers.

Introduction

Project management is a scientific discipline that is present in all sectors of the economy. Over the last decade, this field has changed rapidly due to the development of advanced technologies. In the construction industry, however, the situation is unfortunately not yet as advanced in terms of the introduction of digitalization (Manyika et al., 2015). Therefore, current efforts are focused on research and examples of best practice, emphasizing a responsible and lifelong approach to the planning and execution of construction projects. At the forefront of these efforts is the implementation of advanced technologies and methods (Xavier and Jacob, 2018). A major challenge is the transformation of traditional business models, shifting from material-centric to information- and digital-centric products, data-driven approaches and intellectual business models (Klinc and Turk, 2019). In general, this means the digitalization of business in the construction sector.

The introduction of Industry 4.0 principles, technologies and methods is enabling transformative advances in the construction industry today. To close the digitalization gap and drive the construction industry forward, it must keep pace with the trends of Industry 4.0 (Klinc and Turk, 2019). The latest research findings highlight trends such as Building Information Modeling (BIM), robotics, project management, advanced materials, offsite construction, occupational safety, monitoring, green building, 3D printing and networked construction sites ("Top 10 Industry 4.0 Trends in 2024 | StartUs Insights," n.d.). These innovations are recognized in both academic and professional circles (Hasan and Sacks, 2021).

The construction sector faces numerous challenges in the digital transformation of management processes. Modern and advanced technologies are gradually being integrated into the entire project life cycle, from planning to execution and post-occupancy. It is crucial that future and current engineers acquire digital skills.

Due to the rapid adoption of BIM, many construction management (CM) and civil engineering programs have added BIM courses to the curriculum or are integrating BIM topics into various courses. According to 2013 data, 54% of construction programs had included BIM courses in their curricula, and 52% of BIM materials were embedded in traditional courses (Wu and Issa, 2014). In the USA, for example, in 2022/23 there are almost 120 universities that are fully accredited and offer degree programs that offer BIM courses as part of the curriculum (Morganti et al., 2023).

Studies dealing with the use of BIM in higher education (Wang et al., 2020) provide a comprehensive review of the current state of BIM education literature and list important journals and conference proceedings in which research on BIM education is published. An analysis of the bibliometric study in the Scopus database was performed (Chihib et al., 2019). For the citations »BIM in University« it was found that the development of publications in the period from 2003 to 2018 is relatively weak compared to BIM worldwide (only 6.4% of articles refer to BIM in universities); the distribution by country and institution shows that the most productive institutions are the USA, the UK and China (46% of all articles).

Many studies that investigated pedagogical strategies for BIM courses show that the two main teaching approaches are team-based learning and project-based learning, for ACE students (Anderson et al., 2020; Jin et al., 2018; Olowa et al., 2022) and for CEM students (Wang et al., 2022; Zhang et al., 2018). Some studies also propose the T-shaped model as a concept for BIM education (Partl et al., 2023).

Two Slovenian faculties that train civil engineers have also recognized the importance of incorporating modern content into the study process and have successfully integrated it by active learning methods (Lassen et al., 2018).

At the Faculty of Civil Engineering, Transportation Engineering, and Architecture (FGPA), University of Maribor ("UM FGPA-a. 2023," n.d.), six undergraduate programs, four postgraduate programs and three doctoral programs are offered in the field of the built environment. The undergraduate and postgraduate programs in civil engineering and industrial engineering, which are crucial for the development of the construction sector, are gradually introducing the principles of BIM approach and Industry 4.0. In this, introduction of Industry 4.0 methods is particularly important for the business and project management aspects of construction projects.

At the Faculty of Civil Engineering and Geodesy of the University of Ljubljana, BIM-related content has been part of the study process since 2008 (Šuman et al., 2023). After 2010, the first courses dealing exclusively with modelling and processes within the BIM approach were accredited. The faculty incorporates the topics of digitization and automation into various courses in the different degree programs. Teaching is implemented with integrated problem-based learning (Computer Integrated Construction degree course, 2nd degree) and a workshop approach, which is primarily used in the undergraduate degree courses. Since the 2019/20, the faculty has also been introducing the international Master's degree course BIM A+ (BIM A+, n.d.).

This paper aims to provide an overview of the formal and nonformal acquisition of digital competences in the field of construction management at the FGPA of the University of Maribor. It describes how students acquire digital competences during the learning process, especially in subjects based on the BIM approach and Construction 4.0 technologies. Informal methods for acquiring digital skills are also outlined. The paper concludes with insights into lifelong learning, including a brief description of the buildingSMART Professional Certification Program ("buildingSMART Professional Certification," n.d.) and the presentation of the Platform 5.0 project as part of the Recovery and Resilience Plan (RRP) ("University of Maribor, Pilotni projekti NOO - UM.si," n.d.), which aims to train workers in the field of digitalization and acquire micro-certificates. In this paper, experiences from teaching digital skills acquisition in construction management are shared with the international community and discussions and suggestions for improvement are encouraged.

Formal acquisition of digital competences through the study process

At the FGPA of the University of Maribor, the importance of introducing methods and technologies for digitalization and automation was already recognized in the 1990s, initially within the framework of the Chair of Construction Informatics. With the development of ICT methods and applications, the importance of their use expanded to other areas within the AEC. Significant benefits were recognised in construction management, which led to the integration of digitalization content into the courses offered by the Chair of Construction Management, Technology and Economics. These topics have been gradually introduced over the last 10 years and are taught by competent teachers who keep up to date with the latest trends and research findings. Digitalization is integrated in most professional subjects and in all elective subjects of the 3rd doctoral level.

Since the 2015/16 academic year, the FGPA UM has been offering a BIM (Building Information Modeling) subject at 2nd level. This subject introduces the general concept of BIM, its role and technologies. Specifically for the area of construction management, methods and tools of digitization are used as a basis for construction processes in further courses at the Chair of Construction Management, Technology and Economics e.g. Project management in construction, which is carried out in final year master students with a background in civil and industrial engineering. In addition to BIM-based content, topics relating to Industry 4.0 methods and technologies have increasingly been covered in recent years. Therefore, in the past two years, the chair offers electives such as Computer modeling of construction objects in the 1st level of study as well as BIM approach in operational construction in the 2nd level. In addition, most of the chair's courses cover topics related to digitization and automation, for example in Safety at construction work, Spatial arrangement and advance planning, Maintenance and rehabilitation of structures and Rehabilitation of civil engineering structures. Figure 1 shows the scheme of courses and degrees related to BIM and Industry 4.0 at FGPA UM. All courses, with the exception of the general BIM course, which is a stand-alone course, benefit from the approach of implementing BIM into existing courses (Ghanem, 2022) and include basic and advanced BIM applications. Electives are labelled (e) and courses in the module are labelled (m).

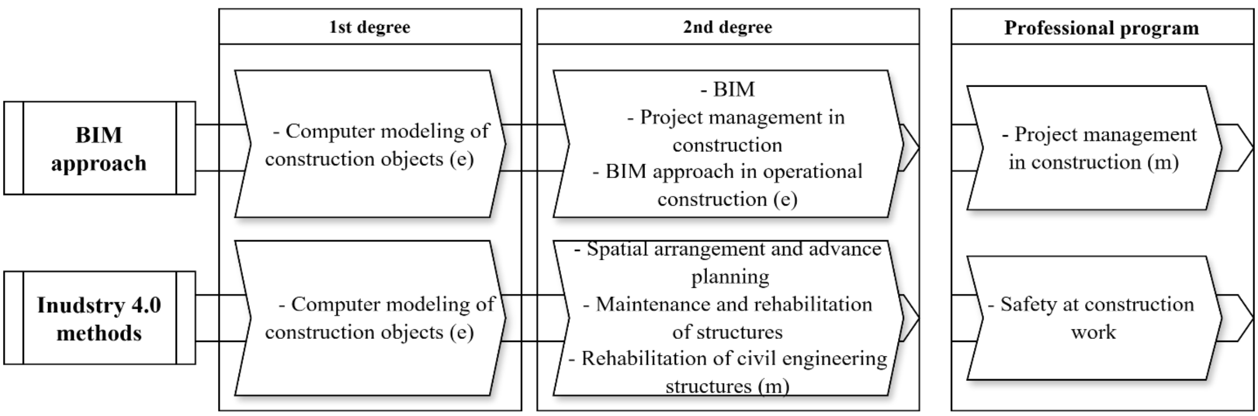


Figure 1: Scheme of courses and degrees with topics on BIM and Industry 4.0

Members of the chair also supervise student theses on digitalization topics. Through courses and theses, students acquire advanced digital skills, roughly divided into a) BIM-supported learning processes, b) learning processes with Industry 4.0 methods and technologies and c) theses on digitization in construction. The following sections present the FGPA's efforts to integrate BIM-based and Industry 4.0 topics into the curricula of the 1st and 2nd degree programs.

Integrating the BIM-based approach

The integration of the BIM-based approach into the study process has been underway at FGPA for many years. The first working example of BIM was introduced specifically for the management of construction projects in the 2015/16 academic year. These concepts were integrated into the subject Project management in construction in the 2nd degree program, which is offered jointly in two study programs: Civil Engineering and Industrial Engineering. This year, a 4D and 5D BIM model for a garage building in Skopje was created as a working example. In following years, the students expanded their knowledge by creating 4D and 5D BIM models for various building construction projects, including a school sports hall, a single-family house with associated infrastructure, a commercial office building, a two-family house, an apartment building, a single-family house with outdoor area and a high-bay warehouse. Since the 2019/20, this approach has also been used at the 1st degree professional study program Civil Engineering, subject of the Operational construction module.

Project work with students on the Project management in construction course follows a structured procedure: students first familiarize themselves with a basic 3D BIM model for a real object and learn about the importance of 3D model information and its application. They then extend the basic model to 4D (construction schedule) and 5D (construction costs) BIM models. Theoretical knowledge about higher BIM dimension levels, i.e. 4D, 5D and 6D BIM models, is taught in lectures. The project work itself is carried out through tutorials and computer-based tasks, with groups of up to three students working together. Selected software tools, including the modules of the Trimble Vico Office™ software (“Vico Office,” n.d.), the local tool 4BUILD for cost estimation and MS Project (“Project Management Software | Microsoft Project,” n.d.) for scheduling, are used to create 4D and 5D BIM models. A more detailed description and a diagram of the BIM creation process during the study at the FGPA of the University of Maribor can be found in the articles in (Pučko et al., 2019; Pučko and Šuman, 2024).

In the 2022/2023 academic year, the students created a highly detailed BIM model for a single-family house that encompasses both the 4D and 5D dimensions. The input data came from a conceptual design that was developed using Allplan BIM software. The original 3D BIM model was previously created and served as an input model for working with students. Figure 2 shows the 3D BIM model of the single-family house and illustrates the level of detail achieved in this comprehensive BIM model.

Figure 2: 3D BIM model of a single-family house

The Vico Office software was used to create the 4D and 5D BIM models, which does not allow direct import from the Allplan program. This led to particular challenges when using the basic 3D BIM model, which had to be converted via the IFC format. Despite these challenges, this was an additional learning opportunity for the students to effectively customize the 3D BIM model and familiarize themselves with the IFC format. For this reason, the subsequent detailed work focused exclusively on the basement of the building.

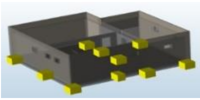
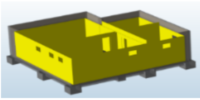
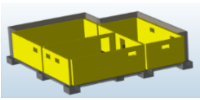
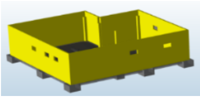
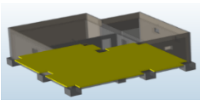
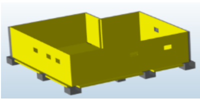
The creation of 4D and 5D BIM models was carried out in groups of two students each, so that a total of six groups were formed. Each group was assigned a specific set of building elements, which are listed in Table 1. First, the groups thoroughly analyzed the elements assigned to them, taking into account dimensions, material, technology, quantity, etc. They collected input data important for estimating unit costs and scheduling, including market prices for materials, equipment and machinery, as well as average wages and standards for labor.

A 5D BIM model was then created using the Cost Planner module. This software offers a significant advantage as it enables the direct and bidirectional linking of cost information with the 3D BIM model elements. This means that every change to the geometry of the 3D BIM model automatically updates the associated costs. The activities and their sequence were then defined as the basis for planning the construction work.

The software facilitates time analysis by creating a 4D BIM model in the Task Manager module. When analysing time, the elements of the 3D BIM model are directly linked to the activities in the schedule. This approach takes into account the links between the activities and the allocated resources and provides a comprehensive understanding of the time aspects of the project.

Table 1: Students group number, assigned project tasks and visualization of building elements

#	The topic of the project task	Visualization (set in Vico Office)

1	001 Reinforced concrete point foundations 010 Reinforced concrete stair foundation	
2	007 Drywall walls 017 Plastering of reinforced concrete basement walls	
3	008 Reinforced concrete basement walls	
4	009 Reinforced concrete columns-ties 015 Facade	
5	004 Reinforced concrete base plate	
6	003 Gravel filling 005 Waterproofing 006 Thermal insulation XPS 016 Facade waterproofing	

In December 2023, a final presentation of the groups' project work took place at the faculty. Each group presented the 4D and 5D BIMs they had created (Figure 3). At the end, the overall result for the execution of the basement was also presented to the students as a comprehensive 3D, 4D and 5D BIM model (Figure 4). A simulation of the construction process was also created by the course teacher. This representation of common 4D and 5D BIM models gives students an insight into the overall cost estimate and schedule.

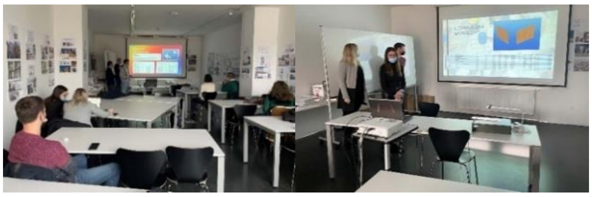


Figure 3: Final presentation of project work

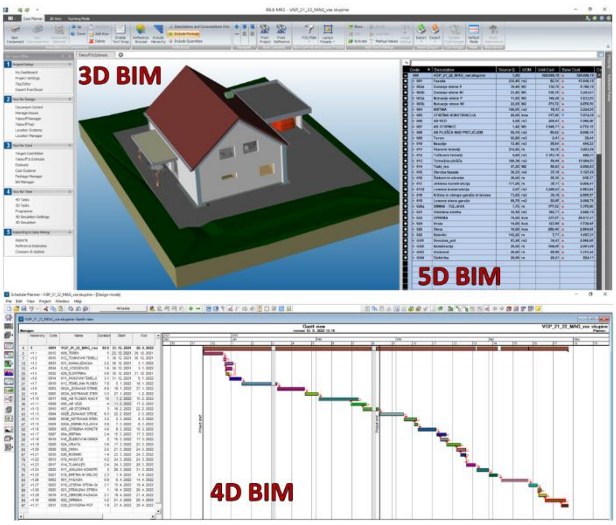


Figure 4: Comprehensive 3D, 4D and 5D BIM model

Integrating methods and technologies of Construction 4.0

Although there is no uniform definition for Construction 4.0, as can be seen from the literature review (Forcael et al., 2020; Jazzar et al., 2020; Perrier et al., 2020), it can be summarized that there is a consensus that the introduction of Construction 4.0 will not only change the construction process, but also the organizational and project structures, transforming the dispersed construction industry into an integrated industry. Individual disruptive technologies that are increasingly being used in the construction industry stand out. In recent years, students have become familiar with the theoretical starting points of Construction 4.0 and acquired individual knowledge and practical skills. For example, in numerous subjects, both in the 1st and 2nd levels of study, digital content is obtained using the technology of data acquisition of the environment using a 3D laser scanner. This technology is used in the following subjects Safety at construction work (1st degree), Spatial arrangement and advance planning (2nd degree), and Maintenance and rehabilitation of structures (2nd degree). The use of a 3D laser scanner is possible for a variety of purposes, such as capturing the as-built condition of the environment, ensuring site safety, determining work progress and creating a 4D as-built BIM model, etc. As part of the subjects listed, students are shown how to use a 3D laser scanner and the method of collecting point cloud data. Figure 5 shows the point cloud data acquired by the students for the selected site as part of the Spatial arrangement and advanced planning course on 2nd degree of study program Industrial Engineering.

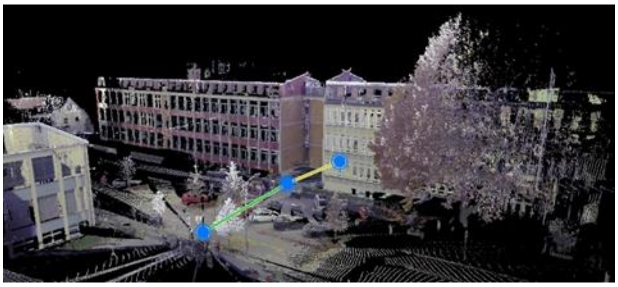


Figure 5: Acquired point cloud data by the students

Students also gain from the modern digital content of Industry 4.0 through advanced augmented reality technologies. With these methods, they acquire knowledge and skills to connect a real environment with digital elements that enable a 1:1 visualization of the built environment. This technology is still relatively new in the construction sector and therefore only a few applications can currently be demonstrated in practice. However, we believe that this technology has great potential and it is expected that further innovations will be developed in the coming years. Figure 6 shows an example of the use of advanced mixed reality in Spatial arrangement and advanced planning course on 2nd degree of study program Industrial Engineering in the 2022/23 academic year.

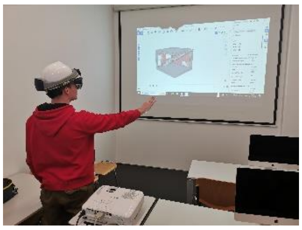


Figure 6: Use of advanced mixed reality during the study process

Nonformal acquisition of digital competences

Many other nonformal learning activities on the topic of digitalization and automation in the construction industry are also carried out at the FGPA. These activities include student participation in various professional events, conferences, trade fairs, excursions, and site visits.

In terms of student participation in various events such as conferences and trade fairs, we would like to highlight a 2-day professional excursion in 2018 with the participation of 50 students at the BIM WORLD Munich Congress (Figure 7-a) (Pučko et al., 2019). In 2023, a visit to the 33rd International Fair for Construction (MEGRA) in Slovenia was organized for students ("Megra 2023," n.d.), where the students took part in an event entitled An educational day in the Slovenian construction industry. Presentations of the latest digital technologies were held at the thematic lectures ("Chamber of Commerce and Industry of Slovenia," n.d.).

In addition, the students took part in an open day in May 2023 at the invitation of the Protim Ržišnik Perc design office. They were introduced to the latest software for working in the BIM environment, as used in the office's design practice (Figure 7-b) ("UM FGPA-b. 2023.," n.d.).

In 2023, various excursions were organised for students with visits to construction sites where the BIM approach was used for detailed design and during construction. We visited the following construction sites with the students, among others Second track Divača-Koper railway line (Figure 7-c) ("UM FGPA-c. 2023.," n.d.) and short visits to construction site Center Rotovž in Maribor (Figure 7-d) ("UM FGPA-d. 2023.," n.d.).

(a)

(b)

(c)

(d)

Figure 7: a) BIM WORLD Munich Congress, b) open day of the project bureau Protim, c) Construction site Second track Divača-Koper railway and d) Construction site Center Rotovž Maribor

All these activities help students to broaden their horizons and gain a closer insight into how modern techniques and methods are applied in the construction process.

Activities of lifelong learning at FGPA

Lifelong learning for working graduates on the topic of acquiring digital competences has been implemented at the FGPA since 2022. This involves running the Fundamentals of BIM course, which is offered under the auspices of the bSI Professional Certification Program ("buildingSMART Professional Certification," n.d.). After completing the course, participants can take an exam to obtain a certificate for basic BIM knowledge, the so-called buildingSMART Professional Certificate, which internationally certifies that professionals are highly qualified and recognisably trained in digital working methods. FGPA is designated as the official training center for the implementation of this program. Civil engineers, architects, electrical engineers and some students (civil engineering and architecture) have successfully completed the course so far.

From July 2022, a pilot project introducing advances digital technologies is also being implemented at the FGPA. The project is being implemented as part of the Recovery and Resilience Plan (RRP) ("University of Maribor, Pilotni projekti NOO - UM.si," n.d.). The University of Maribor is carrying out a total of 23 such pilot projects, which are intended to enable a green and resilient transition to Society 5.0. At the FGPA, the project is called Platforma 5.0 FGPA. The aim of the platform is to identify the knowledge and skills required to promote sustainable development and digitization in study programs over a period of three years. One of the aims of the project is also to create modules for lifelong learning for professionals who will acquire micro-certificates after completing the courses. At the same time, digital technologies for informatization of education will be recognized and implemented, including BIM-based and Industry 4.0 content. Upon completion of the project, the study programs will be aligned with the needs of the economy and society, and students will acquire digital skills and competencies for the green transition and digitization in addition to the competencies dictated by the labor market.

Lessons learned

Each academic year, FGPA UM conducts a course evaluation survey to evaluate the teacher's grade point average and student workload. In the 2022/23 academic year, teachers grades for courses with a BIM approach and Industry 4.0 content were above average at 1.66 (grade range from -2.00 to +2.00). In general, no 1-year or 2-year deviations were identified in terms of student workload. The deviation was only found in the subject Computer modeling of construction objects (+-20%), where students felt that there were too many lectures and tutorials, but students had more independent work. The course was run for the first time this academic year and the results of the survey provide a good basis for the direction of the work in the future.

During several years of running courses on the integration of BIM and two years of implementing Industry 4.0 methods in existing courses, we have come to the following conclusions:

- Working with the BIM approach enables students to better understand the entire construction process, in particular the specific integration and upgrade of 4D BIM (time planning techniques) and 5D BIM (cost estimation approaches)
- The use of interdisciplinary teamwork helps students to develop their collaboration and problem-solving skills and increase student engagement to achieve better results.
- Teachers who actively implement project- and team-based learning together with their students generally ensure a higher quality of teaching.

Based on the knowledge gained, organizational and structural changes are planned:

- A new computer classroom should be set up for practical exercises, equipped with high-quality computers (for conducting BIM courses, modeling and simulation) and the necessary equipment for 3D printing, VR/AR/MR equipment and robots in the construction industry (Industry 4.0).
- As part of the renovation of study programs at the 1st level, a stand-alone BIM course should be introduced to teach the BIM concept and specific modeling skills. The course could replace the current CAD course or be offered as an elective.
- BIM software training should be offered to students and teachers, preferably in collaboration with local industry, to give students hands-on experience.
- Collaboration between teachers on BIM and Industry 4.0 courses should be encouraged and their education made possible.

Conclusions

Despite the rapid development of Industry 4.0 methods, the construction industry is lagging far behind other sectors in terms of digitalization and automation. Project management in the construction industry has changed significantly with the introduction of modern tools, which requires rapid adaptation of learning processes and the integration of BIM and Industry 4.0 into university curricula. This paper outlines the approach taken at the FGPA of the University of Maribor to teach formal and acquiring nonformal digital competences, with a focus on BIM and advanced technologies of Construction 4.0. The emphasis is on providing a comprehensive understanding of BIM principles, mastery of the software and practical application through project work. Nonformal learning, including conferences, trades, and site visits, plays a crucial role in students' education, especially in the context of the accelerating global digital transformation. The implementation of lifelong learning at FGPA through courses such as Fundamentals of BIM and the FGPA Platform 5.0 project ensures continuous education and professional buildingSMART certification for students and professionals alike.

Acknowledgments

The project is co-financed by the Republic of Slovenia, Ministry of Higher Education, Science and Innovation, and the European Union – NextGenerationEU. The project is implemented in accordance with the Smart, Sustainable and Inclusive Growth development area, Strengthening of Competencies component, especially digital competencies and those required by new professions and the green transition (C3 K5), for the investment measure Investment F. Implementation of pilot projects, the results of which will serve as a basis for the preparation of grounds for the reform of higher education for a green and resilient transition to Society 5.0: the project Pilot projects for the Reform of Higher Education for a Green and Resilient Transition.

References

Anderson, A., Dossick, C.S., Osburn, L., 2020. Curriculum to Prepare AEC Students for BIM-Enabled Globally Distributed Projects. Int J Constr Educ Res 16, 270–289. https://doi.org/10.1080/15578771.2019.1654569

BIM A+, n.d. BIM A+ European Master in BIM - Building Information Modelling [WWW Document]. URL https://bimaplus.org/ (accessed 4.2.24). buildingSMART Professional Certification [WWW Document], n.d. URL https://education.buildingsmart.org/ (accessed 12.16.23).

Chamber of Commerce and Industry of Slovenia [WWW Document], n.d. . Izobraževalni dan slovenskega gradbeništva 2023 na sejmu MEGRA 2023. URL https://www.gzs.si/Dogodki/10-03-2023/izobrazevalni-dan-slovenskega-gradbenistva-2023-na-sejmu-megra-2023 (accessed 12.16.23).

Chihib, M., Salmerón-Manzano, E., Novas, N., Manzano-Agugliaro, F., 2019. Bibliometric maps of BIM and BIM in universities: A comparative analysis. Sustainability (Switzerland) 11. https://doi.org/10.3390/su11164398

Forcael, E., Ferrari, I., Opazo-Vega, A., Pulido-Arcas, J.A., 2020. Construction 4.0: A literature review. Sustainability (Switzerland). https://doi.org/10.3390/su12229755

Ghanem, S.Y., 2022. IMPLEMENTING VIRTUAL REALITY - BUILDING INFORMATION MODELING IN THE CONSTRUCTION MANAGEMENT CURRICULUM. Journal of Information Technology in Construction 27, 48– 69. https://doi.org/10.36680/j.itcon.2022.003

Hasan, S., Sacks, R., 2021. Building Information Modeling (BIM) and construction tech integration for construction operations: state of the art, in: Proceedings of the 2021 European Conference on Computing in Construction. pp. 358–365. https://doi.org/10.35490/EC3.2021.179

Jazzar, M. El, Urban, H., Schranz, C., Nassereddine, H., 2020. Construction 4.0: A roadmap to shaping the future of construction, in: Proceedings of the 37th International Symposium on Automation and Robotics in Construction, ISARC 2020: From Demonstration to Practical Use - To New Stage of Construction Robot. International Association on Automation and Robotics in Construction (IAARC), pp. 1314–1321. https://doi.org/10.22260/isarc2020/0180

Jin, R., Yang, T., Piroozfar, P., Kang, B.G., Wanatowski, D., Hancock, C.M., Tang, L., 2018. Project-based pedagogy in interdisciplinary building design adopting BIM. Engineering, Construction and Architectural Management 25, 1376–1397. https://doi.org/10.1108/ECAM-07-2017-0119

Klinc, R., Turk, Ž., 2019. Construction 4.0 – Digital Transformation of One of the Oldest Industries. Economic and Business Review 21. https://doi.org/10.15458/ebr.92

Lassen, A.K., Hjelseth, E., Tollnes, T., 2018. Enhancing learning outcomes by introducing bim in civil engineering studies – experiences from a university college in Norway. International Journal of Sustainable Development and Planning 13, 62– 72. https://doi.org/10.2495/SDP-V13-N1-62-72

Manyika, J., Ramaswamy, S., Khanna, S., Sarrazin, H., Pinkus, G., Sethupathy, G., Yaffe, A., 2015. Digital America: A Tale of the Haves and Have-Mores.

Megra 2023 [WWW Document], n.d. URL https://www.megra.pomurski-sejem.si/en/ (accessed 12.16.23).

Morganti, C., Coraglia, U.M., Bragadin, M.A., Rissolo, D., Witt, E., Kähkönen, K., 2023. Teaching BIM: a comparison between actual and future perspectives, in: Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference. https://doi.org/10.35490/EC3.2023.212

Olowa, T., Witt, E., Morganti, C., Teittinen, T., Lill, I., 2022. Defining a BIM-Enabled Learning Environment —An Adaptive Structuration Theory Perspective. Buildings 12. https://doi.org/10.3390/buildings12030292

Partl, R., Miya, S., Hengel, F., Heschl, C., 2023. T-shaped model design for further education in BIM, in: Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference. https://doi.org/10.35490/EC3.2023.181

Perrier, N., Bled, A., Bourgault, M., Cousin, N., Danjou, C., Pellerin, R., Roland, T., 2020. Construction 4.0: A survey of research trends. Journal of Information Technology in Construction. https://doi.org/10.36680/J.ITCON.2020.024

Project Management Software | Microsoft Project [WWW Document], n.d. URL https://www.microsoft.com/en-us/microsoft-365/project/project-management-software (accessed 12.16.23).

Pučko, Z., Šuman, N., 2024. Advanced Construction Project Management: The Meaning and Importance of the New Learning Paradigm, in: Vrečko, I., Gajšek, B. (Eds.), The Future of Project Management: Adapting to Modern Needs. Cambridge Scholars Publishing , Newcastle, UK, pp. 119–149.

Pučko, Z., Šuman, N., Klanšek, U., 2019. Teaching and Lifelong Education on BIM at FCETEA in Maribor, Slovenia. In: OTMC Conference 2019 Proceedings. Croatian Association for Construction Management: Zagreb, Croatia; n.d.; pp. 465–473. Zagreb.

Šuman, N., Pučko, Z., Srdić, A., Klinc, R., 2023. Education for the Use of BIM on the Example of Slovenian Higher Education, in: Proceedings of the 15thAdvances in Civil Engineering Congress. pp. 147–163.

Top 10 Industry 4.0 Trends in 2024 | StartUs Insights [WWW Document], n.d. URL https://www.startus-insights.com/innovators-guide/top-10-industry-4-0-trends-innovations-in-2021/ (accessed 12.16.23).

UM FGPA-a. 2023 [WWW Document], n.d. URL https://www.fgpa.um.si/en/ (accessed 12.16.23).

UM FGPA-b. 2023. [WWW Document], n.d. . Dan odprtih vrat Protim Ržišnik Perc, Senčur - Fakulteta za gradbeništvo, prometno inženirstvo in arhitekturo Univerza v Mariboru. URL https://www.fgpa.um.si/dan-odprtih-vrat-protim-rzisnik-perc-sencur/ (accessed 12.16.23).

UM FGPA-c. 2023. [WWW Document], n.d. . Ekskurzija študentov na gradbišča drugega tiraželezniške proge Divača – Koper, 17.5.2023 - Fakulteta za gradbeništvo, prometno inženirstvo in arhitekturo Univerza v Mariboru. URL https://www.fgpa.um.si/ekskurzija-studentov-nagradbisca-drugega-tira-zelezniske-proge-divaca koper-17-5-2023/ (accessed 12.16.23).

UM FGPA-d. 2023. [WWW Document], n.d. . Predavanje iz gospodarstva in ogled gradbišča Center Rotovž - Fakulteta za gradbeništvo, prometno inženirstvo in arhitekturo Univerza v Mariboru. URL https://www.fgpa.um.si/predavanje-iz-gospodarstva-in-ogled-gradbisca-center-rotovz/ (accessed 12.16.23).

University of Maribor, Pilotni projekti NOO - UM.si [WWW Document], n.d. URL https://www.um.si/o-univerzi/pilotni-projekti-noo/ (accessed 12.16.23).

Vico Office [WWW Document], n.d. URL https://vicooffice.dk/en/ (accessed 12.16.23).

Wang, L., Huang, M., Zhang, X., Jin, R., Yang, T., 2020. Review of BIM Adoption in the Higher Education of AEC Disciplines. Journal of Civil Engineering Education 146. https://doi.org/10.1061/(asce)ei.2643-9115.0000018

Wang, L., Huang, M., Zhang, X., Yan, X., Jin, R., Wanatowski, D., Cheshmehzangi, A., Chohan, N., 2022. Incorporating BIM into the upper-division curriculum of construction engineering and management. European Journal of Engineering Education. https://doi.org/10.1080/03043797.2022.2112150

Wu, W., Issa, R.R.A., 2014. BIM education and recruiting: Survey-based comparative analysis of issues, perceptions, and collaboration opportunities. Journal of Professional Issues in Engineering Education and Practice 140. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000186

Xavier, D.K., Jacob, R.M., 2018. Digitalization of the Construction Industry: The Revolution is Underway [WWW Document]. Oliver Wyman. URL https://www.oliverwyman.com/our-expertise/insights/2018/sep/digitalization-of-the-construction-industry.html

Zhang, J., Wu, W., Li, H., 2018. Enhancing Building Information Modeling Competency among Civil Engineering and Management Students with Team-Based Learning. Journal of Professional Issues in Engineering Education and Practice 144. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000356

ELECTRONIC PERFORMANCE MONITORING AT THE WORKPLACE: A SURVEY TO DIAGNOSE ENGINEER'S AWARENESS

Diego Calvetti[1], Miguel Chichorro Gonçalves[1], Pedro Mêda[1], and Hipólito de Sousa[1]
[1]CONSTRUCT/Gequaltec, University of Porto, Porto, Portugal

Abstract

The electronic performance monitoring (EPM) of human resources using new technologies increases current control methods while introducing ethics and moral concerns. In chasing EPM deployments, a survey was conducted to diagnose the perception of workers and students. The engineers were elected as core participants in EPM due to their knowledge of both the technology and the field of application. Surprisingly, 69% of the 261 respondents were unaware of the topic. Results indicate a 56% acceptance level (95% confidence level and 6.3% error). The research points out the most accepted and rejected aspects of EPM, illustrating scenarios and alternatives for implementation.

Introduction

The industry has been applying numerous innovations in electronic devices and wearables to monitor workers' performance. Institutions seek to use electronic monitoring of human resources to increase their financial results (Alder, 2001). Electronic Performance Monitoring is about the use of hardware, software, and techniques to access and evaluate human resources job performance (Alder & Tompkins, 1997; Lund, 1992; Oz et al., 1999; Panina & Aiello, 2005). The great challenge to electronic monitoring is related to trust and security in sharing data between workers and employers (Calvetti et al., 2020; Calvetti, Mêda, et al., 2021; European Commission. Directorate General for Research and Innovation., 2021; Kelly, 2018; *Use of Wearables in the Workplace Is Halted by Lack of Trust – PwC Research - Press Room*, n.d.). Perceiving the people's reaction to monitoring is a fundamental part of achieving success in implementing technological innovation projects (Bhave, 2013; Holland et al., 2015; Panina & Aiello, 2005; Stanton & Julian, 2002). Industrial sectors, such as construction, whose financial result is strongly linked to the physical performance of human resources, seek to improve their control tools in safety, health, productivity and quality of services (Calvetti et al., 2022; Calvetti, Goncalves, et al., 2021). A worker away from work due to accidents or work-related illnesses has a negative impact on the performance of companies. Work activities with a great need for human physical effort and little repeatability have significant variability in productivity rates and demand numerous processes to control the quality of services.

In addition to current control methods such as using computers (email, internet, etc.), access control by badges, and monitoring the use of equipment and machines. Electronic monitoring advances in analysing location, movements, gestures, and the physiological state of workers. Tracking the location of each worker in near real-time provides, in addition to the analysis of their productive occupation, the diagnosis of permanence in areas with safety hazards and routes with the risk of being run over by machinery and equipment (Bangaru, Wang, & Aghazadeh, 2020; Duan et al., 2022; Igwe et al., 2022; Jung & Chi, 2020; Ryu et al., 2020). The detailed assessment of workers' movements and gestures allows for a precise ergonometric analysis, which helps analyse occupational health, safety and productivity (Bangaru, Wang, Zhou, et al., 2020; Yang et al., 2019; Zhang et al., 2019). Collecting data regarding the workers' physical state (cerebral electrical activity, temperature, pressure, heart rate, etc.) makes it possible to assess the physiological state (tiredness, fatigue, stress, tension, etc.), which also allows for the assessment of performance in health, safety and productivity (Ahn et al., 2019; Choi et al., 2019; Costa et al., 2019; Jebelli et al., 2018; Wang et al., 2019).

However, current technological advances allow for a new and vast field of investigation regarding human resources performance. Nevertheless, legal paradigms and restrictions regarding electronic monitoring must be confronted. Realising workers' reactions to electronic monitoring is a crucial element in achieving successful technological innovations in the workplace (Ahn et al., 2019; Maurice et al., 2019; Ryu et al., 2019; Yan et al., 2018). Because of that, a survey was conducted to understand better peoples' reactions to electronic monitoring at work.

Survey sample

A total of 8914 people received the questionnaire, which 8478 - Students of the Faculty of Engineering of the University of Porto (FEUP), 16 – Professors of FEUP, and 420 - People from the authors' network (mainly from LinkedIn social network). A total of 261 people answered the questionnaire, which determined a response rate of 2.93%. Based on the population (8478) and sample (261), within a confidence level of

95% (P = 0.5), the accuracy of the responses reached an index of ± 6.3% (Dupont et al., 2017).
The collected data were tabulated in MS Excel for analysis and graphical development. Furthermore, the IBM SPSS system was used for the univariate and multivariate statistical analysis. Finally, the results were discussed by comparing similar research conducted by the insurance company AIG (American International Group, Inc.) in nine countries with 400 employees and 250 entrepreneurs in the year 2016 (Mike Abbott, 2017).

Young Portuguese students gave the most significant number of answers from Engineering in Information Technology and Computers, Civil Engineering, Mechanics, Metallurgy and Materials Engineering. Concerning the education level, 67% of respondents are students, and 13% study and work; in effect, 80% of the sample is currently a student. It is worth noting that the vast majority of these, 58%, attend the integrated master's degree, followed by 16% to develop degrees, 15% for doctorates and 10% for master's degrees.

Regarding the field of activity (academic/professional), it was identified that most of the respondents are from technology and computing fields, with 43%, followed by 24% of civil engineering, mechanics, metallurgical, and materials.

As for gender, there was 64% of males, 35% of females, and 1% of respondents chose not to be identified by a specific gender. Moreover, as seen previously, the more representative public sampled are students who are predominantly young, with 34% between 16 and 21 years old and 33% between 22 and 25 years old. However, the sample represents the perception of future engineers who are going to enter the labour market in the coming years. Finally, 87% of the respondents live in Portugal, and 10% are Brazilian residents, among other countries with 3%.

Results and discussions

In order to conduct the respondents for real situations about the application of electronic performance monitoring at work, fourteen situational questions were applied. Moreover, two questions sat the respondents in the position of the employer and twelve as workers. Besides, interviewees' personal view-positioning was observed in two questions about believing that monitoring their own work is essential to improve performance and if the government should regulate EPM. Finally, to identify whether the respondents already had any knowledge about the EPM subject, one question was asked. The results of these seventeen questions are presented and discussed in the exact order developed in the survey.
Questions 1 and 3-8 were the types of yes/no answers with room for alternative/free opinions. Question 2 presents different options to agree/disagree as well as with room to write a distinguished option.

Question 1: Do you believe that monitoring the progress of activities developed during the workday is important for improving your performance?
Result: Yes 209 (80%), No 52 (20%). It is noticed that the great majority understand that monitoring is a relevant tool for the improvement of their own performance.
Question 2: By placing yourself in the position of EMPLOYER, you would want to use technological innovations to monitor. Alternatives: The safety of workers; The health of workers; The productivity of workers; The quality of services performed by workers; Other (field for text insertion).
Results: As shown in Figure 1, it is perceived that respondents would be more interested in monitoring order: Safety of workers 200 (77%); The productivity of workers 185 (71%); The quality of services performed by employees 182 (70%); Workers' health 155 (59%); Do not monitor 5 (2%). The values in safety, productivity, and quality are close, and health monitoring is significantly below, with still fewer responses contrary to any form of monitoring. Compared to the survey carried out by AIG (Mike Abbott, 2017), it is observed that the population responding has higher interest levels, as for security (77%) compared to (59% AIG), productivity (71%) versus (55% AIG), and health (59%) versus (55% AIG).

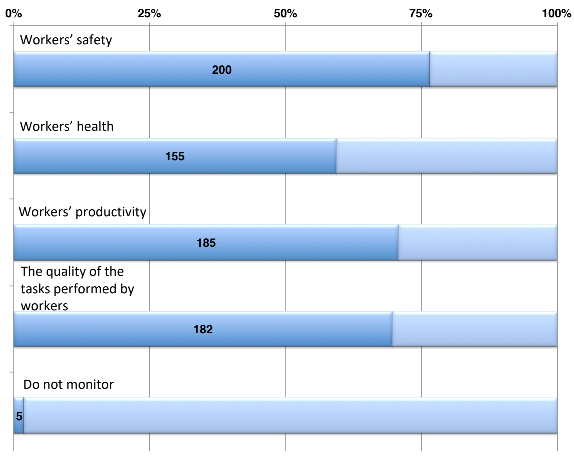


Figure 1: Subjects of interest to be monitored

Question 3: By placing yourself in the position of EMPLOYER, do you believe that in the case of applying the monitoring, you should share this data collected with the employees?
Result: No (this information must be exclusive to the company) 47 (18%), Yes (I must allow workers access to information) 214 (82%). It is noticed that in the vast majority, people understand it necessary to share the data collected with the workers. The study conducted by AIG (Mike Abbott, 2017), evaluated the intention to share data of employees and employers, where per country:

- Japan - employees (80%), employers (65%)

- United States of America - employees (75%), employers (65%)
- Australia - employees (74%), employers (65%)
- Singapore - employees (76%), employers (69%)
- United Kingdom - employees (74%), employers (68%)
- Italy - employees (87%), employers (85%)
- Germany - employees (70%), employers (70%)
- China - employees (93%), employers (97%)
- France - employees (66%), employers (76%)

There is an alignment of the sample surveyed regarding the sharing of data, mainly countries with higher indexes of intention to share, such as Japan and Italy.

Question 4: Do you believe that the GOVERNMENT should regulate electronic monitoring at work?

Result: Yes, 141 (54%), No 107 (41%), Depends on the level of monitoring 4 (2%), I have no opinion 4 (2%), Other responses 5 (2%). Compared with the results of AIG's survey (Mike Abbott, 2017), which indicated that 89% of employers and 87% of employees understand that there should be legal regulation of monitoring (Mike Abbott, 2017), the rate of only 54% seems low. Because the survey sample is predominantly of young students (still out of the job market), two hypotheses can be assumed for such a low value. Firstly, students still do not know/fear the possible persuasive use of EPM. Secondly, young people accept electronic monitoring more naturally than older generations (employees and employers).

Question 5: By placing yourself in the position of EMPLOYEE, would you allow the mapping of your location in real-time by the employer?

Result: Yes, 111 (43%), No 150 (57%). At this point in the survey, respondents are confronted with more frontal questions about electronic monitoring methods at work. Of effect, it can perceive some greater reactivity from the respondents regarding EPM, in this specific case of allowing the mapping location in real-time.

Question 6: By placing yourself in the position of EMPLOYEE, would you allow the mapping of your movements/motions (trunk, legs, arms, and hands) by the employer?

Result: Yes, 86 (33%), No 175 (67%). Within the same analysis performed previously, and specifically regarding the monitoring of the movements and gestures, the respondents' significant reactivity was identified. This sentiment is aligned, for example, with the journalistic material published on CNN Tech, regarding the controversial patent application by the Amazon company with the intention of monitoring the workers' arms motions (Kelly, 2018).

Question 7: By placing yourself in the position of EMPLOYEE, would you allow the mapping of your physical conditions (electrical activity of the brain, temperature, pressure, heart rate, etc.) by the employer?

Result: Yes, 138 (53%), No 123 (47%). Given the two previous answers, the monitoring of physical conditions reached a higher rate of acceptance. One possible hypothesis is that the respondents connect the monitoring of physical conditions to health issues, and in the opposite, the two before productivity issues. However, it should be noted that monitoring these physiological aspects could undoubtedly be used to assess productivity.

Question 8: By placing yourself in the position of EMPLOYEE, would you feel safe sharing the data collected by wearable electronic devices with the employer?

Result: Yes, 143 (55%), No 118 (45%). It is identified that more than half of the respondents would trust their data to the employer. However, the survey carried out by AIG regarding the employees' intention to share the data indicates higher indexes of the order of 66% to 93% (Mike Abbott, 2017).

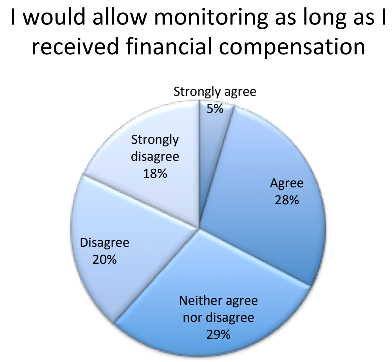
Questions 9, 10, and 11 requested a situational opinion, as employees, if they were prone to allow the monitoring to receive some rewards. Meanwhile, questions 12 and 13 try to perceive awareness about safety and health in the workplace. Smartphone data collection/analysis is a current issue. Questions 14 and 15 asked if they would share data from their own device and in case of that device was from the company. Question 16 puts respondents against the wall, asking if their job depends on the acceptance of EPM if they would allow it. Finally, in the last question, 17 asked if they knew, before that survey, about those EPM methods at the workplace.

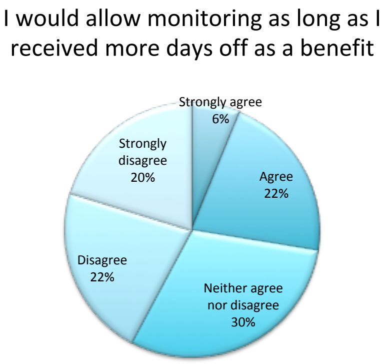
Question 9: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring as long as I received financial compensation.".

Question 10: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring as long as I received more days off as a benefit.".

Question 11: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring as long as this would help my hierarchical progression in the company.".

It is noticed that the respondents value more the professional career concerning other rewards such as financial benefits and days off. The hypothesis that can be inferred is that the young public of the sample aims to enter the job market and to have professional success, not yet realise the relevance of the financial benefits and times of rest; see Figure 2, where it is highlighted: 33% agree, or strongly agree to allow monitoring through financial compensation; 28% agree, or strongly agree to allow monitoring by receiving more days off; 44% agree, or strongly agree to allow monitoring through a hierarchical progression in the company.





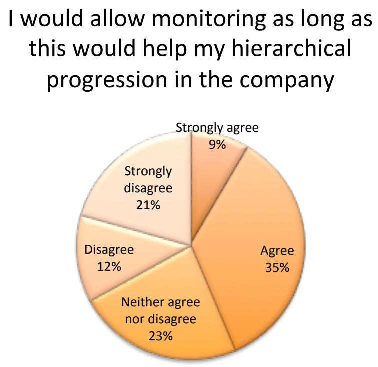
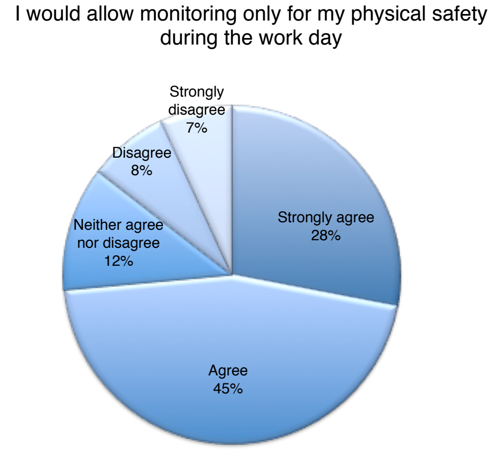


Figure 2: Situation of enabling monitoring in exchange for benefits

Question 12: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring only for my physical safety during the workday.”.

Question 13: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring only in the concern to improve my physical health during the workday.".

Figure 3 presents the results. It can be identified that the respondents have great motivation to accept monitoring for their own safety and health, where 74% agree or strongly agree to allow monitoring in the pursuit of improving physical safety during the workday; 68% agree or strongly agree to allow monitoring in the quest to improve physical health during the workday.



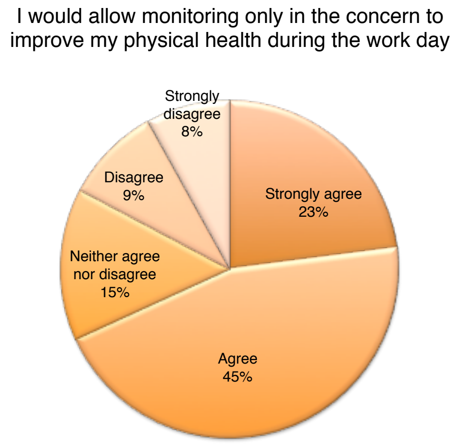


Figure 3: Situation of allowing monitoring for greater safety and health during the workday

Question 14: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would allow monitoring of data from my own smartphone.".

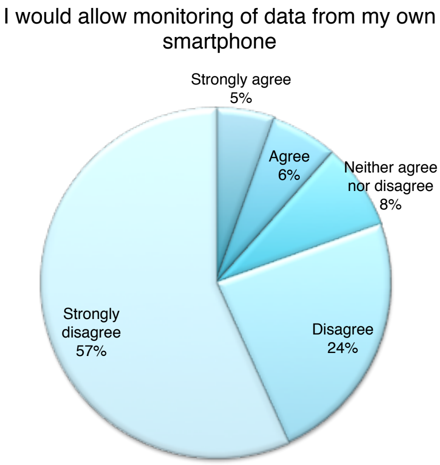
Question 15: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would only allow monitoring data from a smartphone provided by the company.".

Figure 4 presents the results. It is noticed that the respondents strongly reject the idea of allowing the availability of data on their own smartphones. However, when the hypothesis of the smartphone was provided by the employer, there was a significant tendency of acceptance to share data, where it is highlighted that 11% agree or strongly agree to allow monitoring of data from the smartphone itself; 50% agree, or strongly agree to allow the monitoring of smartphone data provided by the company.

The study carried out by AIG (Mike Abbott, 2017) highlights the favourable position of employees in using a device provided by the employer, by country as forward presented:

- Germany (29%)
- Japan (36%)
- United States of America and Australia (38%)
- United Kingdom and France (40%)
- Italy, Singapore and China (56%)

The indices identified in this survey of 50% are above the position of several countries, being close to Italy, Singapore, and China (56%), countries with more favourable indexes.



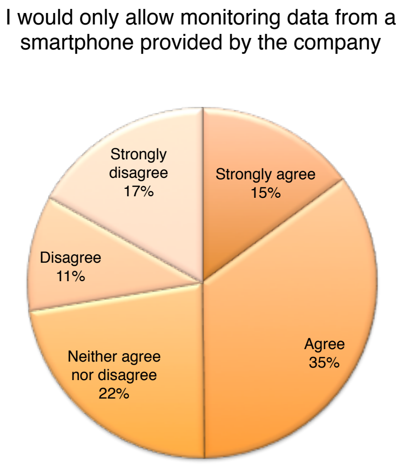


Figure 4: Situation of allowing monitoring for greater safety and health during the workday

Question 16: Regarding the monitoring by means of electronic devices and wearables, by placing yourself in the position of EMPLOYEE in these situations: "I would never allow monitoring, even if that would be a prerogative to get the job.".

It should be noted that 55% disagree or strongly disagree with the proposed statement, which indicates that these respondents in the postulated case would accept the monitoring to access/guarantee a job vacancy. Only 14% responded by indicating that they would withdraw from the job if there were the prerogative of monitoring, where the result points out that: Strongly agree (7%); Agree (7%); Neither agree nor disagree (31%); Disagree (28%); Strongly disagree (26%).

Question 17: Have you had contact or knowledge of the use of electronic monitoring systems in the workplace?

Surprisingly, a large part of the respondents (69%) only has known about the subject of electronic monitoring at work through this survey. In practice, only 7% had real-life experience with EPM, and 24% knew some piece of information from some news. Where the options presented were: No. “I did not know the matter.” 180 (69%); Yes. “I've followed or participated in projects of this type.” 19 (7%); Yes. “I had already obtained information on the subject (internet, videos, lectures, etc.).” 62 (24%)

An EPM level of acceptance is possible by quantifying the answers. The questions with answers of type "yes" or "no" were quantified on a binary scale, been by unfavourable added zero (0), and most favourable to the monitoring one (1), where: Question: 1; Yes (1), No (0); Question 3; No (1), Yes (0); Question 4; Yes (0), No (1), Depends on the level of monitoring (0), I have no opinion (0), Other answers (0); Questions 5-8; Yes (1), No (0).

As a criterion for the question of multiple choices (question 2), one value (1) was established for each selected item. Given a value of zero (0) to the indications of not monitoring. In this question, the values have variables from zero (not monitor) to four (selection of all items), where the safety of workers (1), Workers' health (1), The productivity of workers (1), The quality of services performed by workers (1); Do not monitor (0).

For situational questions with five alternatives (questions 9-15), according to the favourable or unfavourable sense of monitoring, the following values were assigned: Strongly agree (1); Agree (0,5); Neither agree nor disagree (0); Disagree (-0.5); Strongly disagree (-1).

Specifically for question 16, which has an otherwise sense of agreement, the values were: Strongly agree (-1); Agree (-0,5); Neither agree nor disagree (0); Disagree (0.5); Strongly disagree (1).

With this, based on the above-detailed quantification, the maximum attainable value (all favourable responses) is “19 values”; in counterpoint, the lowest possible value (all unfavourable responses) is “-8 values”. Based on the responses and values, a value level was determined for each respondent, see Figure 5.

Proportionally, the total spectrum was divided into four quartiles. The first is with an index between 0-25% (Q0), the second between 25-50% (Q25), the third between 50-75% (Q50), and finally, the fourth quartile between 75-100% (Q75).

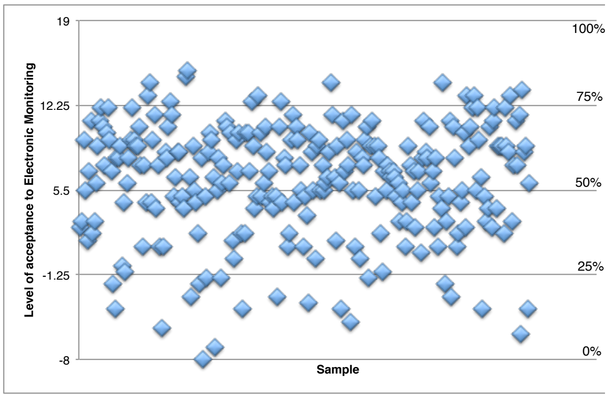


Figure 5: Acceptance level for electronic sample monitoring

It is observed that 60% of the sample (156 respondents) is in the Q50 (between 50-75%) of acceptance to the EPM. In the top quadrant, Q75, with an acceptable level greater than 75%, there are 14 respondents (5%). In this sense, it is highlighted a trend of sixty-five per cent favourable to EPM by these populations under analysis. Within Q25 (25-50%), there are 71 respondents (27%). Furthermore, at the lower end, positions totally contrary to EPM, only 20 respondents (8%) were found.

Employing the SPSS software, the values and histogram graph, see Figure 6, of the total sample were determined statistically. It is observed that the average (6.3180) and the median (7.0000) have approximate values. The kurtosis value (0.3630) is below 0.5, indicating the normality of the distribution. However, the marked asymmetry (-0.7500) and the mode value (9.0000) accept the values' distribution as expected. The value that most represents the population under analysis is the median (7.0000). Other values are Standard deviation (4.4411), Variance (19.7230), Minimum (-8,0000), and Maximum (15,0000).

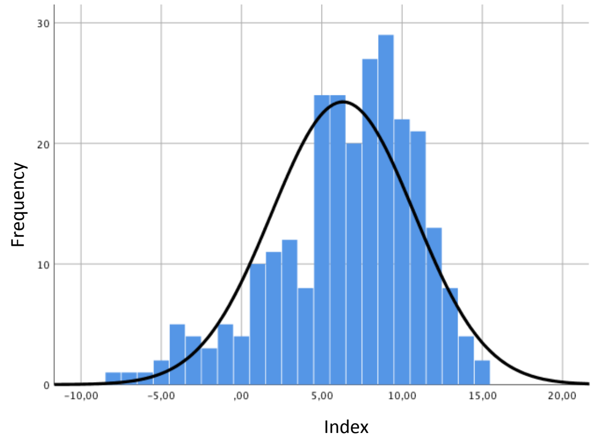


Figure 6: Samples' histogram distribution of the statistical analysis

With this, a level of 56% acceptance is determined for the electronic monitoring of the population under analysis. Within a 95% confidence level (P = 0.5), the accuracy of the responses reached an index of ± 6.3% (Dupont et al., 2017).

Accounting for the quartiles (Q0 - Q25 - Q50 - Q75) and the proportion of their population, Table 1, which indicates a higher or lower level of acceptance to EPM, it is possible to identify whether parameters for the sample characteristics (e.g., age, gender, etc.) can guide clusters of individuals. In this sense, multivariate statistical analysis was performed through SPSS-IBM to cluster the individuals. However, no characteristic of the sample was identified that was guiding the formation of groups of individuals, "more or less" favourable to the monitoring. The results and cross-analysis of sample characteristics and responses within the four quartiles (Q0 - Q25 - Q50 - Q75) are presented below. It should be noted that all values were determined to vary according to the statistical error of ± 6.3%; see exemplification in Table 1.

Table 1: Acceptance levels for electronic performance monitoring

Quartile	Total	%	Lower limit	Upper limit
Q0	20	8%	1%	14%
Q25	71	27%	21%	34%
Q50	156	60%	53%	66%
Q75	14	5%	-1%	12%
Total	261			

Conclusions

The applied survey allowed a very detailed view of a population predominantly of engineers (students and workers) regarding EPM at work. Surprisingly 69% of the respondents first knew about EPM through the survey. The population under analysis is more predisposed to accept monitoring to increase safety and evaluate health. Also, reward systems (financial, days off, and hierarchical advancement) are detected as factors that may facilitate acceptance of EPM. Raising people's awareness that monitoring makes it possible to improve performance can be a facilitating method for implementing EPM at work. However, the EPM projects' success needs to guarantee data sharing security and provide devices for the workers.

An expressive level of 56% EPM acceptance was determined. Within a 95% confidence level (P=0.5), the accuracy of the responses reached an index of ±6.3%, thus delimiting the determined level in the range of 49.7% to 62.3%. It is observed that 60% of the sample (156 respondents) is in Q50 (between 50% and 75%) from reception to monitoring. And that still is in the maximum quadrant, Q75, with an acceptance level greater than 75%, 14 respondents (5%). In this sense, there is a tendency (65%) in favour of electronic monitoring by the population under analysis. By how much, at lower levels, within Q25 (25% to 50%), there are 71 respondents (27%). And, with a completely opposite position (a level below 25%) to electronic monitoring, only 20 respondents (8%) were monitored.

Most respondents, 180 (69%), had, for the first time, knowledge about electronic monitoring at work through the survey in question. The population under analysis is more predisposed to accept monitoring when it aims to increase safety and assess health. Also, benefit systems (financial and days off) and opportunities for hierarchical advancement in companies are factors that facilitate the monitoring of electronic monitoring.

Raising people's awareness that monitoring enables performance improvement can be a facilitating method for including electronic monitoring at work. However, it is accompanied by the success of implementing transparent information monitoring and investment in processes that transmit security to workers regarding data sharing.

Acknowledgments

This work was financially supported by: Base Funding - UIDB/04708/2020 with DOI 10.54499/UIDB/04708/2020 (https://doi.org/10.54499/UIDB/04708/2020) of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC).

References

Ahn, C. R., Lee, S., Sun, C., Jebelli, H., Yang, K., & Choi, B. (2019). Wearable Sensing Technology Applications in Construction Safety and Health. Journal of Construction Engineering and Management, 145(11), 03119007. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001708

Alder, G. S. (2001). Employee reactions to electronic performance monitoring: A consequence of organizational culture. The Journal of High Technology Management Research, 12(2), 323–342. https://doi.org/10.1016/S1047-8310(01)00042-6

Alder, G. S., & Tompkins, P. K. (1997). Electronic performance monitoring. Management Communication Quarterly, 10(3).

Bangaru, S. S., Wang, C., & Aghazadeh, F. (2020). Data Quality and Reliability Assessment of Wearable EMG and IMU Sensor for Construction Activity Recognition. Sensors, 20(18), 5264. https://doi.org/10.3390/s20185264

Bangaru, S. S., Wang, C., Zhou, X., Jeon, H. W., & Li, Y. (2020). Gesture Recognition–Based Smart Training Assistant System for Construction Worker Earplug-Wearing Training. Journal of Construction Engineering and Management, 146(12), 04020144. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001941

Bhave, D. P. (2013). The Invisible Eye? Electronic Performance Monitoring and Employee Job Performance: PERSONNEL PSYCHOLOGY. Personnel Psychology, n/a-n/a. https://doi.org/10.1111/peps.12046

Calvetti, D., Goncalves, M., Vahl, F., Meda, P., & Sousa, H. de. (2021). Labour productivity as a means for assessing environmental impact in the construction industry. Environmental Engineering and Management Journal, 20(5), 781–790. https://doi.org/10.30638/eemj.2021.073

Calvetti, D., Magalhães, P. N. M., Sujan, S. F., Gonçalves, M. C., & Campos de Sousa, H. J. (2020). Challenges of upgrading craft workforce into Construction 4.0: Framework and agreements. Proceedings of the Institution of Civil Engineers - Management, Procurement and Law, 173(4), 158–165. https://doi.org/10.1680/jmapl.20.00004

Calvetti, D., Mêda, P., Sousa, H., & Chichorro Gonçalves, M. (2021). Human Data Interaction in Sensored Sites, Challenges of the Craft Workforce Dimension. 2021 European Conference on Computing in Construction, 173–180. https://doi.org/10.35490/EC3.2021.171

Calvetti, D., Sanhudo, L., Mêda, P., Martins, J. P., Gonçalves, M. C., & Sousa, H. (2022). Construction Tasks Electronic Process Monitoring: Laboratory Circuit-Based Simulation Deployment. Buildings, 12(8), 1174. https://doi.org/10.3390/buildings12081174

Choi, B., Jebelli, H., & Lee, S. (2019). Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk. Safety Science, 115, 110–120. https://doi.org/10.1016/j.ssci.2019.01.022

Costa, A., Rincon, J. A., Carrascosa, C., Julian, V., & Novais, P. (2019). Emotions detection on an ambient intelligent system using wearable devices. Future Generation Computer Systems, 92, 479–489. https://doi.org/10.1016/j.future.2018.03.038

Duan, P., Zhou, J., & Tao, S. (2022). Risk events recognition using smartphone and machine learning in construction workers' material handling tasks. Engineering, Construction and Architectural Management. https://doi.org/10.1108/ECAM-10-2021-0937

Dupont, Q. F. M., Chua, D. K. H., Tashrif, A., & Abbott, E. L. S. (2017). Potential Applications of UAV along the Construction's Value Chain. Procedia Engineering, 182, 165–173. https://doi.org/10.1016/j.proeng.2017.03.155

European Commission. Directorate General for Research and Innovation. (2021). Industry 5.0: Towards a sustainable, human centric and resilient European industry. Publications Office. https://data.europa.eu/doi/10.2777/308407

Holland, P. J., Cooper, B., & Hecker, R. (2015). Electronic monitoring and surveillance in the workplace: The effects on trust in management, and the moderating role of occupational type. Personnel Review, 44(1), 161–175. https://doi.org/10.1108/PR-11-2013-0211

Igwe, C., Nasiri, F., & Hammad, A. (2022). Construction workspace management: Critical review and roadmap. International Journal of Construction Management, 22(10), 1960–1973. https://doi.org/10.1080/15623599.2020.1756028

Jebelli, H., Hwang, S., & Lee, S. (2018). EEG-based workers' stress recognition at construction sites. Automation in Construction, 93, 315–324. https://doi.org/10.1016/j.autcon.2018.05.027

Jung, M., & Chi, S. (2020). Human activity classification based on sound recognition and residual convolutional neural network. Automation in Construction, 114, 103177. https://doi.org/10.1016/j.autcon.2020.103177

Kelly, H. (2018, February 2). Amazon's idea for employee-tracking wearables raises concerns. CNNMoney. https://money.cnn.com/2018/02/02/technology/amazon-employee-tracker/index.html

Lund, J. (1992). Electronic performance monitoring: A review of research issues. Applied Ergonomics, 23(1), 54–58. https://doi.org/10.1016/0003-6870(92)90011-J

Maurice, P., Malaisé, A., Amiot, C., Paris, N., Richard, G.-J., Rochel, O., & Ivaldi, S. (2019). Human movement and ergonomics: An industry-oriented dataset for collaborative robotics. The International Journal of Robotics Research, 38(14), 1529–1537. https://doi.org/10.1177/0278364919882089

Mike Abbott. (2017). The Data Sharing Economy: Quantifying Tradeoffs that Power New Business Models. American International Group, Inc. (AIG).

Oz, E., Glass, R., & Behling, R. (1999). Electronic workplace monitoring: What employees think. Omega, 27(2), 167–177. https://doi.org/10.1016/S0305-0483(98)00037-1

Panina, D., & Aiello, J. R. (2005). Acceptance of electronic monitoring and its consequences in different cultural contexts: A conceptual model. Journal of International Management, 11(2), 269–292. https://doi.org/10.1016/j.intman.2005.03.009

Ryu, J., McFarland, T., Banting, B., Haas, C. T., & Abdel-Rahman, E. (2020). Health and productivity impact of semi-automated work systems in construction. Automation in Construction, 120, 103396. https://doi.org/10.1016/j.autcon.2020.103396

Ryu, J., Seo, J., Jebelli, H., & Lee, S. (2019). Automated Action Recognition Using an Accelerometer-Embedded Wristband-Type Activity Tracker. Journal of Construction Engineering and Management, 145(1), 04018114. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001579

Stanton, J. M., & Julian, A. L. (2002). The impact of electronic monitoring on quality and quantity of performance. Computers in Human Behavior, 18(1), 85–101. https://doi.org/10.1016/S0747-5632(01)00029-2

Use of wearables in the workplace is halted by lack of trust – PwC research—Press room. (n.d.). Retrieved 26 June 2023, from https://pwc.blogs.com/press_room/2016/06/use-of-wearables-in-the-workplace-is-halted-by-lack-of-trust-pwc-research.html

Wang, D., Li, H., & Chen, J. (2019). Detecting and measuring construction workers' vigilance through hybrid kinematic-EEG signals. Automation in Construction, 100, 11–23. https://doi.org/10.1016/j.autcon.2018.12.018

Yan, X., Li, H., Zhang, H., & Rose, T. M. (2018). Personalized method for self-management of trunk postural ergonomic hazards in construction rebar ironwork. Advanced Engineering Informatics, 37, 31–41. https://doi.org/10.1016/j.aei.2018.04.013

Yang, Z., Yuan, Y., Zhang, M., Zhao, X., & Tian, B. (2019). Assessment of Construction Workers' Labor Intensity Based on Wearable Smartphone System. Journal of Construction Engineering and Management, 145(7), 04019039. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001666

Zhang, H., Yan, X., Li, H., Jin, R., & Fu, H. (2019). Real-Time Alarming, Monitoring, and Locating for Non-Hard-Hat Use in Construction. Journal of Construction Engineering and Management, 145(3), 04019006. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001629

THE CYCLICAL DEVELOPMENT OF BIM AUTHORING TOOLS AND WORKFLOW METHODOLOGIES INTO THE MULTI-VARSITY CURRICULA OF ARCHITECTURAL, ENGINEERING AND CONSTRUCTION PROGRAMMES

Gordon Chisholm[1], Joao Pereira de Sá[2], Geoff Olner[3], Frances Robertson[3], Paweł Hawryszków[4], Diogo Rodrigo Ribeiro[5]

[1]South East Technological University (SETU), Ireland

[2]Copenhagen School of Design and Technology (KEA), Denmark

[3]Sheffield Hallam University (SHU), Sheffield, England

[4] Wrocław University of Science and Technology (WUST), Poland

[5]CONSTRUCT-LESE, School of Engineering of Polytechnic of Porto (ISEP), Portugal .

Abstract

This paper explores how the Trivarsity initiative addresses misalignment between AEC education and industry needs and its tangible results in showing the benefits and advantages for both staff and students of incorporating an interdisciplinary, multi-varsity curricular development into their European AEC programmes. Running annually since 2014, the collaborative BIM project has been an action research vehicle for an inter-varsity, interdisciplinary, collaborative design workshop, applying current and nascent digital tools and technologies. The constant re-alignment of AEC sector education to new, effective ways of working may ensure future proofing of educational provision.

Introduction

Background

This paper explores the benefits and advantages for both staff and students of incorporating an interdisciplinary multi-varsity curricular into their European AEC programmes. Effective and conclusive research generally provides evidence of what the curriculum in Higher Education Institutions (HEIs) should be (Emmitt, 2006). Through active participation in the research process, professional educators and students become clearer of the kinds of competence and attitudes which are critical for current and future effective professional performance (Schein, 1972). Businesses require graduates who not only have the skills to do the job, but who can aid the organisation to evolve in the face of continuous and rapid economic and technological change (CBI, 2012; Government of Ireland, 2019). Construction 2025 (Department for Innovation and Skills, 2013) reports on the future aspirations of the Architecture, Engineering, Construction and Owner/operations (AEC) sector, to include it being a world-leader in research and innovation which embraces Information Communication Technologies (ICTs) and smart construction. The NBS BIM report surveys the international adoption of digital tools and technologies (NBS, 2013; NBS, 2023), providing evidence of AEC sector practice now moving from the digital to the machine age. In digital construction, information is required at the level of data across the whole project eco-system. The radical change to the AEC sector disciplinary educational model, is in the application of the basal skills in practice of implementing digital tools, machine learning, and AI, necessitating the ability to design, manipulate and modify the (software) programming of architectural processes (Scheer, 2014, p.124): education needs to explore these innovative and disruptive developments in practice (nascent and emerging fields) to remain future-proof.

Methodological Approach

Research that aims to make projections about the future requires a *prospective* methodological approach (as per Ratcliffe, 2008). Through on-going, collaborative research between academia, industry and the professions, the future (r)evolutionary application of ICTs to professional practice across the disciplines in the AEC sector may be determined (as per *Industry 4.0* (Deloitte 2020)). Technical knowledge and skill are not enough in practice, the participants in a building project must have the necessary social skills to work together effectively and efficiently (Emmitt, 2010). The implementation of an emergent general plan for the integration of Building Information Modelling/Management (BIM/M) authoring tools and workflow methodologies, and soft skills into the curriculum of AEC sector education is challenging - it requires a methodological approach to systematically review forces for, and barriers to, effective change in terms of potential points of entry or intervention into the present educational system (Schein, 1972, p.90).

Action research: the active and interested participation by the researchers in the issue and processes being investigated so that they can identify, appraise and conjecture potential solutions is proposed. The intention of action research is to effect a change: knowledge gained through reflection on action is used to instigate ongoing evolutionary change, and to create knowledge about the process of change, the consequences of this change, and about the nature of the change itself (Fellows and Liu, 2008). Action research is complex and is appropriate for the study of future collaborative practice to inform education.

In the digital age, AEC sector educational evolution may mimic effective development in the I.T. industry, i.e., a *Dev Ops* cycle which is more like a figure of eight 'cycle', with continuous improvement of the product through feedback, reflection, adjustment, and concomitant improvement through the process itself (Kim et al, 2016; Kim, Behr and Spafford, 2018). At any point in time the product is deliverable but in a constant state of flux in terms of the understanding of its development. This is predicated on an approach to learning like 'transformative reflection' (Biggs and Tang, 2011 p.46) which is a multi-stage collaborative process of: reflect-plan-apply-evaluate which is then fed back into the iterative evolutionary development process. In vocational education, a renewed focus on the HEI developer ('Dev') co-creating course devisal/delivery and effective learning mechanisms collaboratively *with* industry operators ('Ops'), presents a symbiotic learning opportunity (see Figure 1 (Robertson, 2022)). All contributors, practitioners, educators, learners, and industry innovators may collectively test out the possible answers to the real world-relevant problems, and interdependently propose solutions, ensuring the alignment and interconnectedness of 'Dev' and 'Ops'. This adventurous type of learning better incubates and develops in students the skills they will need to apply professionally, involving activities with an element of professional, social, and emotional risk. Feedback from activities allow error to be used constructively (Biggs and Tang, 2011 p.65) following Emmitt and Ruikar's (2013) recommendation for through-project and post-project reflection and learning, thus, creating an academic community of reflective practitioners (Schön, 1983) and agile and socio-emotionally intelligent graduates.

HEI's current curricular development is often predicated on a protracted feedback loop. However, AEC sector in-house R&D is 'live' action research - an iterative feedback loop over a continuous two-yearly cycle but with opportunities for shorter term change implementation. Research-informed change in HEIs is challenged by bureaucratic structures and processes, and the reluctance to embrace change. A re-framing of change management in HEIs to the Dev Ops approach successfully adopted by innovatory practices, instigates a revolutionary change in professional education enabling it to remain relevant and travel in alignment with best practice.

The researchers demonstrate an analogous DevOps approach in the re-framing of the quality assurance (QA) mechanisms in their HEIs: the current emphasis on declarative (subject-specific) knowledge and threshold standards is shifted to include achieving 'threshold concepts', i.e., functioning (procedural) knowledge which informs appropriate action by the learner as 'performances of understanding' (Biggs and Tang, 2011, p.81), in order to propose 'satisficing' solutions (Simon, 1996) to industry-relevant problems. Practice problems often involve disciplinary professionals functioning in a 'silo' of specialisation, the research collective aim to bring conflicting 'silos' together to develop collaboratively a solution to the complex problems presented.

This DevOps approach is how that change may be most effectively accomplished in professional AEC sector education: as a partnership between active-research-informed practitioners and up skilled educators. To verify and validate HEI provision the learning mechanisms incorporate how people work, practice in action, patterns in practice, effective methods of action, interpersonal/ inter-professional relationships, and societal influences/events (like Covid-19) as per Seymour and Rooke (1995); Seymour et al (1997).

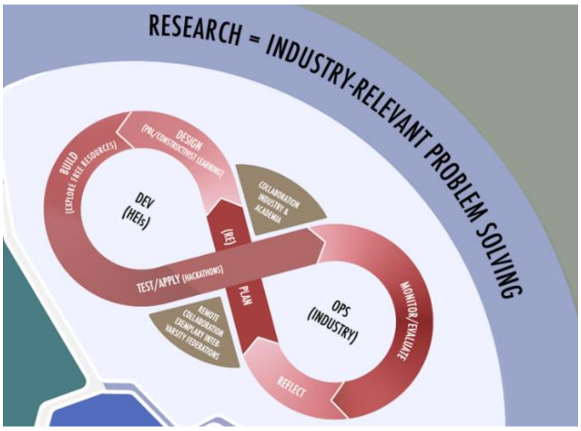


Figure 1: (Produced by Irina Vasiliu) DevOps Approach to Course Development: An on-going, cyclical process of collaborative research and development of the implementation of ICTs between HEIs and the industry (Robertson, 2022).

Reflection on the Cyclical Development of the Workshops (2014-2023)

The workshops started as a vehicle for staff and students to learn about building information modelling that was being mandated in UK, Ireland and Denmark. As educators in AEC, it needed to be addressed, because the curriculum was lagging and falling behind governmental policy. The collaboration outlined below started in 2014 with like-minded *'change agents'* (Schein, p.72, 1972) from academia identifying a *'specific change target'*; a collaborative BIM design project. They then instigated a change process to simulate 21st century, interdisciplinary, professional practice within academia. An inter-varsity, interdisciplinary, collaborative design exercise using industry-standard BIM/M software has evolved within a community of academic experts and practitioners in international HEIs.

In managing this evolutionary change process, the ultimate goals are continuously re-examined: it is a dynamic and cyclical process reinforced by the change agents finding a *'change role model'* (Schein, p.78, 1972). Autodesk, as thought leaders in the evolution of digital technologies in the AEC sector, are an apt change role model. This is a symbiotic relationship between academia and software developers involved in planned change, and the learners are a crucial group in the change-agent collective: their feedback is pivotal to the success of the cyclical development process.

The workshop collaborators from industry/practice (architectural, structural, environmental and MEP), professional education (multi-disciplinary academics and students) and software design and development (Autodesk) collectively review each workshop to identify driving and restraining forces and key barriers to its successful adoption into the existing educational system.

Trivarsity Workshop Development

The genesis for the workshop came through networking during the International Congress on Architectural Technology (ICAT) 2013, hosted at Sheffield Hallam University (SHU). Educators from SHU, England, VIA University, Denmark and Waterford Institute of Technology, Ireland (WIT now SETU Est. 2022). A common goal, with the onset of BIM mandates/legislation across the EU was to collaborate and mimic industry best practice utilising BIM methods and methodologies, upskilling staff and students in understanding and implementation of BIM tools across disciplines. Academia-Industry Partnerships (Thomas, 2013) have been fostered throughout each workshop with guidance, presentations and sponsorship. The first five workshops involved the original three institutes with a mixture of project types, selected to learn software and processes. As the workshop progressed student numbers grew as did the disciplines.

2015 SHU workshop explored authoring the 3D model to ease the downstream pricing/scheduling. The student groups involved Architectural Technology SHU, Derby and Huddersfield and Architectural & BIM Technology WIT.

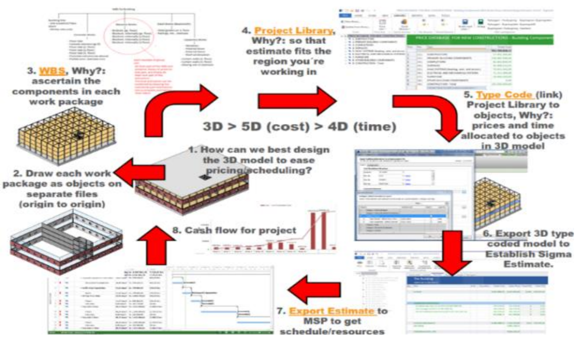


Figure 2: Infographic of Workshop workflow (Muller, 2014)

2015 VIA Horsens workshop introduced the discipline of "Architectural Technology and Construction Management" through the VIA students. This workshop used the same base models as the first workshop and introduced some remodelling aspects to optimise the daylight to the building. Students explored costing, external envelope design and daylight analysis.

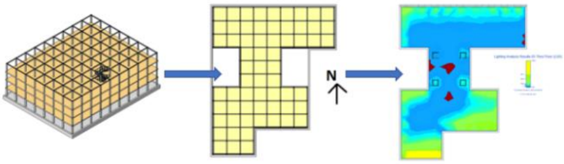


Figure 3: Group 6: Design Development and Analysis (Chisholm, 2024)

2016 WIT workshop introduced two new student disciplines from WIT; Sustainable Energy Engineering who return in 2023 and Quantity Surveying who rejoined in 2019 and have continued to be involved with Trivarsity. A key aspect of the workshops was introduced at this point. "The Prelude" is a period of 2-3 weeks prior to the physical workshop where the project and group members are introduced. The groups set up the CDE, BEP, MIDP, file naming conventions, project programming & roles; communicating through collaborative platforms; at this point it was BIM 360 and later workshops MS Teams. This has since developed to include shared online classes run simultaneously across all participating institutes. The workshop looked at the design of eight mixed use tall buildings to create a new district on Waterford Quay focusing on cost, programming, wind & daylight analysis, structure, architecture and MEP models. Each group had one representative who formed a master planning group to coordinate all groups. Figure 4 below shows them in discussion and the decisions made on the whiteboard.

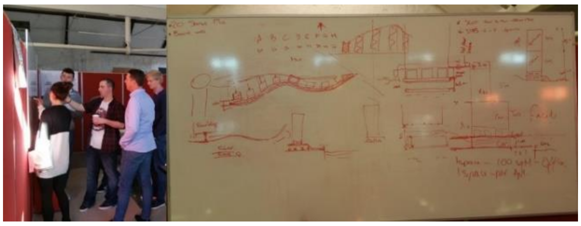


Figure 4: Master Planning Group (Chisholm, 2016)

2017 SHU workshop integrated industry practitioners as a group into the workshop. The principle objective was to explore, gain experience of and optimise the co-ordination of information between the disciplines on a simulated interdisciplinary refurbishment project utilising BIM360 and Autodesk Glue. Preparation work was undertaken during the Prelude. This workshop introduced clash detection and resolution to the students.

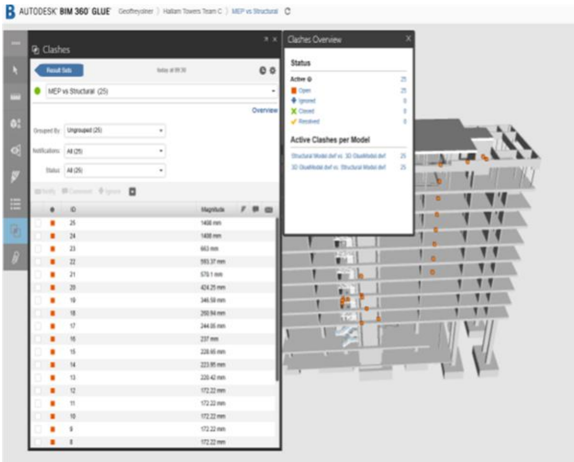


Figure 5: Team C: BIM360 clash analysis (Chisholm, 2024))

2018 VIA Aarhus workshop asked the groups to act as the design team associated with a design & build contractor in a bidding competition against the other teams for the design of a new student union building on the campus. A brief was set for the client requirements, a project programme was to be delivered along with costing packages and overall cost. The assessment of the workshop shifted from formal presentations to an exhibition style which saw each team produce an audio-visual presentation and two A1 posters, displayed in a common area. Refreshments were provided and each student and staff member then marked the projects based on the Most Economically Advantageous Tender criterion set out in the brief. This process has proven to be very successful and has carried on throughout all subsequent workshops. The benefits are seen in the students discussing and reviewing each others work within a social environment at the end of the workshop.

Figure 6: Exhibition & Marking, VIA Aarhus (Chisholm, 2017)

2019 WIT workshop saw VIA exit and Copenhagen School of Design and Technology (KEA) join Trivarsity with their Architectural Technology and Construction Management students. The WIT quantity surveying students took part again. The workshop built upon previous experiences with a design project based on the design and build tender for a tall gateway tower for Waterford Quay. Key aspects of the Trivarsity workshops were now in place: the prelude, multi-disciplinary (to a point) teams, working in the cloud to BIM standards, industry experts presenting, student social event with exhibition and marking to cap the workshop. The collaboration software had matured and the academic staff and students' knowledge had reached a point where the early day set up issues were eradicated.

Throughout all the workshops students could work remotely. This facilitated those who could not travel and added a vital aspect of the cloud collaboration experience.

Figure 7: Group Web call at WIT (Chisholm, 2019)

2020 SHU, the Covid workshop. With the collective experience gained through six Trivarsity workshops in cloud-based collaboration running a workshop during national lockdowns was not an issue. Tools such as Microsoft Teams and Autodesk BIM360 were now integrated into the workshops and our individual programme teaching. The entire genesis of Trivarsity workshops to work to industry best practice was tested with great success. The project dealt with a real SHU campus project, the proposed "Skills Factory". For the first time shared classes ran during the prelude and the workshop. New topics were MEP authoring and integrating Dynamo to size ducts. Both SHU & WIT were in lockdown, KEA were able to attend the classroom which resulted in some interesting dynamics with KEA students working in the one space and connected to the wider student/staff body via MS Teams.

Figure 8: KEA workshop and on the main screen Trivarsity Staff and group members reviewing project work (Sá, 2020)

2021 KEA. Structural engineering staff and students from Instituto Superior de Engenharia do Porto (ISEP) joined the workshop. This made a significant impact on the structural designs and the use of Autodesk Robot within collaboration process. The project was for the design of a new mixed-use district at Refshaleoen and to include building integrated renewables. SHU attended the workshop remotely due to travel restrictions. This was the largest Trivarsity to date with 16 No. groups comprising of 166 students taking part. This was the first Trivarsity workshop to be funded via the Erasmus+ Blended Intensive Programme (BIP). The BIP requirements matched the existing structure of the workshop with a mix of online classes and physical workshop. The length of the workshop increased from three to five days to include cultural aspects such as a river cruise and site visit.

Figure 9: Group 04, poster presentation (Chisholm, 2024)

2022 ISEP, BIP funded. Structural engineering staff and students joined Trivarsity from WUST, bringing more balance to the student disciplines. A comprehensive mix of cultural and workshop events were organised across the five days. The project was the development of a new urban area on the banks of the Douro River. Deliverables were consistent with recent workshops with highlights being the organised cultural events such as the river tour and bridge climb.

Figure 10: Arrábida Bridge arch climb (Chisholm, 2022)

2023 SETU, BIP funded. A return to Waterford and of SETU sustainable energy engineering staff & students, the introduction of Frederick University (FU) architecture & engineering students, and staff from ISCTE School of Technology and Architecture University Institute of Lisbon observing the workshop with a view to joining for the next iteration. The workshop consisted of seven institutes covering eight programmes/disciplines. The project was a revisit to the 2016 project with additional requirements to investigate cross laminated and glulam structure, environmental and embodied carbon analysis with Autodesk Forma and Insight tools. Again, a comprehensive mix of cultural and workshop events were organised.

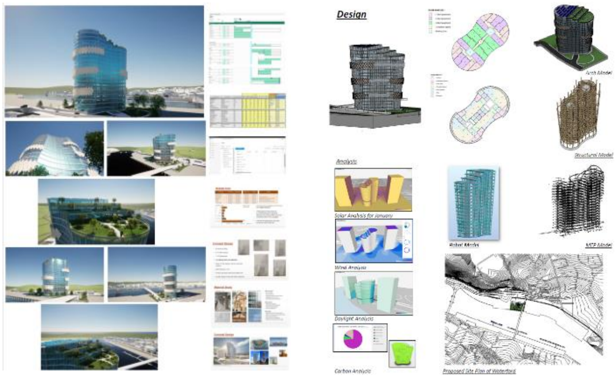


Figure 11: Group 03, poster presentation (Chisholm, 2024)

Trivarsity Impact on Programmes

Across all participating institutions the Trivarsity Workshops have had a positive effect on curriculum and teaching methods. Taking the three longest running members KEA, SHU & SETU, the impact can be seen in:

- the early adoption of BIM authoring tools, management and clash detection tools and the integration of these across years and programmes.
- support by software developers & vendors
- applied knowledge of international standards embedded into curriculum and programmes

- development of multi-disciplinary modules, workshops with their own institutes and locally.
- Senior years mentoring during workshops
- Development of BIM/Digital Construction focused modules building on the knowledge gained through Trivarsity

The impact of participation in Trivarsity for ISEP & WUST institutes who joined later, is similar but enhanced as they have benefitted from the preceding years development and body knowledge gained. Key effects are:

- Rapid amendments to curriculum across bachelor and master's programmes in areas such as, digital modelling & technologies, BIM fundamentals, management, standards & policies, digital surveying, AR/VR, digital twins
- Modifications to course delivery to include collaborative aspects improving teamwork and leadership skills and an uplift in the quality of student work.

Examples of didactical effects achieved in 2023 are presented in Figure 12 from WUST, utilising the software and methodologies from Trivarsity in their Bridge Engineering programme.

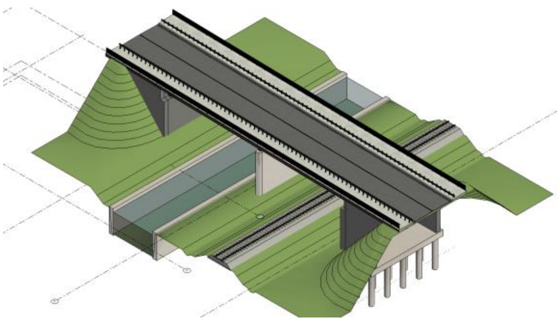


Figure 12: BIM model of a 2-span composite bridge (Hawryszków and Janicki, 2023)

Conclusions and Future Development

The ongoing success of the Trivarsity Workshops demonstrates the importance of interdisciplinary collaborative role-based teaching to allow students practice their discipline specific skillsets whilst gaining soft skills in teamwork and collaboration. Throughout the workshops skills are transferred between students across institutions and disciplines.

Trivarsity students are succeeding through engagement with WorldSkills International Digital Construction competitions, with SHU and SETU students winning medals at both national and international levels.

Staff are collaborating on research projects, studying PhDs at partner universities and the group have been asked to join the recently formed "I Build Sustainable and Smart" group. A European wide collective of educators with aims to broaden collaboration between institutes. One of the aims within the group is to disseminate the findings and knowledge gained through the Trivarsity Workshops to assist others to collaborate in a similar way.

Acknowledgments

The authors would like to Acknowledge the contribution of Academic staff involved with each Trivarsity Workshop.

2015_SHU:

Ernest Muller VIA, Frances Robertson SHU, Geoff Olner SHU, Boris Ceranic UoD, Charles Hippisley-Cox UoH, Brian Dempsey WIT, Robin Stubbs WIT, Gordon Chisholm WIT

2015_VIA Horsens:

Ernest Muller VIA, Frances Robertson SHU, Geoff Olner SHU, Liane Duxbury SHU, Jacob Ware SHU, Mark Botham SHU, Brian Dempsey WIT, Gordon Chisholm WIT

2016_WIT:

Ernest Muller VIA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Robin Stubbs WIT, Gordon Chisholm WIT, Tom O'Brien WIT, John Mernagh WIT

2017_SHU:

Ernest Muller VIA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Robin Stubbs WIT, Gordon Chisholm WIT

2018_VIA Aarhus:

Ernest Muller VIA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Gordon Chisholm WIT, Tiberius Gruia Autodesk

2019_WIT:

Joao Pereira de Sá KEA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Robin Stubbs WIT, Gordon Chisholm WIT, Tom O'Brien WIT, John Mernagh WIT, Tiberius Gruia Autodesk

2020_SHU:

Joao Pereira de Sá KEA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Gordon Chisholm WIT, John Mernagh WIT

2021_KEA:

Joao Pereira de Sá KEA, Gabija Kaltenyte KEA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey WIT, Robin Stubbs WIT, Gordon Chisholm WIT, Diogo Rodrigo Ribeiro ISEP, Ricardo Pereira Santos ISEP

2022_ISEP:

Joao Pereira de Sá KEA, Benedetto Alessi KEA, Kristian Lisičak KEA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey SETU, Robin Stubbs SETU, Gordon Chisholm SETU, Diogo Rodrigo Ribeiro ISEP, Ricardo Pereira Santos ISEP, Marta Knawa-Hawryszków WUST, Paweł Hawryszków WUST

2023_SETU:

Joao Pereira de Sá KEA, Nikola Dimovski KEA, Frances Robertson SHU, Geoff Olner SHU, Brian Dempsey SETU, Gordon Chisholm SETU, Tom O'Brien SETU, John Mernagh SETU, Jurgen Bauer SETU, Diogo Rodrigo Ribeiro ISEP, Ricardo Pereira Santos ISEP, Paweł Hawryszków WUST, Marta Knawa-Hawryszków WUST, Ricardo Pontes Resende ISCTE

List of Institute Abbreviations:

Copenhagen School of Design & Technology (KEA)

Frederick University (FU)

Instituto Superior de Engenharia do Porto (ISEP)

School of Technology and Architecture University Institute of Lisbon (ISCTE)

Sheffield Hallam University (SHU)

South East Technological University (SETU) formerly Waterford Institute of Technology (WIT)

University of Derby (UoD)

University of Huddersfield (UoH)

VIA University College (VIA)

Wroclaw University of Science & Technology (WUST)

Diogo Rodrigo Ribeiro was financially supported by: Base Funding - UIDB/04708/2020 with DOI 10.54499/UIDB/04708/2020 and Programmatic Funding - UIDP/04708/2020 with DOI 10.54499/UIDP/04708/2020 of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC).

References

Biggs, J. & Tang, C. 2011, 'Teaching for Quality Learning at University', 4th ed., England.

Confederation of British Industry (CBI) 2012, Boosting Employability Skills. Available at: http://www.cbi.org.uk/business-issues/education-and-skills/in-focus/employability/ (Accessed: 18th December 2014).

Deloitte 2020, Industry 4.0 [Homepage of Deloitte.com], [Online]. Available at: https://www2.deloitte.com/uk/en/insights/focus/industry-4-0.html (Accessed: 10th December 2023)

Department for Business, Innovation and Skills 2013, Construction 2025: industrial strategy for construction - government and industry in partnership., HM Government.

Emmitt, S. 2010, Managing interdisciplinary projects: a primer for architecture, engineering, and construction, Spon, London.

Emmitt, S. 2006, 'Investigating the Synergy between Teaching and Research in a Teaching-led University: The Case of an Architectural Technology Undergraduate Programme', Architectural Engineering and Design Management, vol. 2, no. 1, pp. p.62.

Emmitt, S. & Ruikar, K. 2013, Collaborative design management, Routledge, London.

Fellows, R., & Liu, A. 2008, Research methods for construction, 3rd ed., Wiley-Blackwell, Chichester; Oxford.

Government of Ireland 2019, Future Jobs Ireland, Government of Ireland, Available at: https://www.enterprise.gov.ie/en/Publications/Publication-files/Future-Jobs-Ireland-2019.pdf (Accessed: 10th December, 2023).

Kim, G., Behr, K. & Spafford, G. 2018, The Phoenix Project: A Novel about IT, DevOps, and Helping Your Business Win, IT Revolution.

Kim, G., Humble, J., Debois, P. & Willis, J. 2016, The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Technology Organisations, IT Revolution.

NBS 2023, Digital Construction Report 2023, NBS & Glenigan, Available at: https://www.thenbs.com/digital-construction-report-2023/ (Accessed: 7th December 2023).

NBS 2013, NBS International BIM Report, NBS, Available at: https://www.thenbs.com/knowledge/nbs-international-bim-report-2013 (Accessed: 10th December 2023).

Ratcliffe, J. 2008, 'Built environment futures research: the need for foresight and scenario learning' in Advanced Research Methods in the Built Environment, eds. A. Knight & L. Ruddock, 1st edn, Wiley-Blackwell, pp. 216-229(p.217).

Robertson, F. J. 2022, The Future Practice of Architectural Technology; Informing Practice and Education, Vol. 1, PhD Thesis, Available at: https://researchportal.bath.ac.uk/en/studentTheses/the-future-practice-of-architectural-technology-informing-practic (Accessed: 10th December, 2023)

Scheer, D.R. 2014, The death of drawing: architecture in the age of simulation, Routledge.

Schein, E.H. 1972, Professional Education: Some New Directions, McGraw Hill, New York.

Sch n, D.A., 1983, The Reflective practitioner: how professionals think in action, Basic Books.

Seymour, D. & Rooke, J. 1995, 'The culture of the industry and the culture of research', Construction Management and Economics, vol. 13, no. 6, pp. 511-523.

Seymour, D., Rooke, J. & Crook, J. 1997, 'The role of theory in construction management: A call for debate', Construction Management and Economics, vol. 15, no. 1, pp. 117-119.

Simon, H.A. 1996, The Sciences of the Artificial, 3rd edition edn, MIT Press, Cambridge, Massachusetts.

Thomas, K 2013, Collaborative BIM Learning via an Academia-Industry Partnership, Construction IT Alliance BIM Gathering 2013, Dublin.

Hawryszków, P. & Janicki M. 2023, 'Building Information Modelling of bridge structures over the railways' (in Polish), Proceeding of the 18th Wrocław Bridge Days Conference, 23-24 November 2023, pp. 411-418, Wrocław, Poland

DIGITAL PRODUCT PASSPORTS IN CONSTRUCTION – BARRIERS AND OPPORTUNITIES AT PEOPLE, PROCESS AND TECHNOLOGY DIMENSIONS

Pedro Mêda[1], Eilif Hjelseth[2], Diego Calvetti[1], Hipólito Sousa[1]
[1] CONSTRUCT – Gequaltec, Porto University Faculty of Engineering, Porto, Portugal
[2] Department of Civil and Environmental Engineering, NTNU - Norwegian University of Science and Technology Trondheim, Norway

Abstract

Digital Product Passports (DPP) will be inevitable as they pursue objectives aligned with the Sustainable Development Goals (SDGs). DPP is grounded in digital data, enabling integration and traceability throughout life cycles and value chains. DPP can play a central role in accomplishing twin transitions in construction. A People, Process and Technology (PPT) dimensions study aimed to fill the knowledge gap between strategies and sector visions—the mixed methods integrated approach collected data focusing on barriers and opportunities at each dimension. Focus groups analysed and clustered the results. Despite relevant barriers, construction stakeholders see opportunities and recognise the need, demonstrating eagerness towards adoption.

Introduction

Digital Product Passport (DPP) means a set of data specific to a product that includes, at minimum, relevant information for performance, environment and waste dimensions and is accessible via electronic means through a data carrier defined under the EU Ecodesign Regulation (European Parliament and European Union Council, 2022). EU legislation on the Green Deal and Data Act will make DPP a reality for different areas of activity in Europe and worldwide (Commission, 2022).

The construction sector is often labelled as highly resistant to change (Lines *et al.*, 2015). However, when awareness and a concerted change effort are made, the sector can behave and perform as one of the best. Stakeholders' realisation and perception of barriers and opportunities are crucial to anticipate implementation problems, better tune the support actions, and provide valuable and aligned contributions, fostering smooth transitions. This work stems from these assumptions, collecting impressions to forecast and support streamlined DPP implementation in the construction sector.

Despite some uncertainties on how the passports will be mandated, the objectives they aim to achieve are known, especially those concerning environmental targets and Sustainable Development Goals (SDGs) (United Nations, no date). According to (Munaro and Tavares, 2021), product passports are generally seen as a mechanism to influence consumer behaviour concerning sustainable purchasing and responsible product ownership. However, to accomplish this objective, many other aspects must be handled, enabling a broad range of services, businesses and possibilities and establishing a relevant ecosystem that should be anticipated (King, *et al.*, 2023).

Although recent, the DPP results from continuous work that has been ongoing for decades. The stakeholders' capacity for new requirements, more digital methods, and abilities related to products have been tested with EU-scale developments in areas such as energy efficiency and chemicals. However, there is the intuition that product passports might raise some fear across sectors due to the introduction of significant changes in how trade is processed and how product information is managed among all involved in the different value chains.

Several researchers have devoted their efforts to detailing and anticipating implementation challenges surrounding DPPs or associated concepts (Byers and De Wolf, 2023; Ducuing and Reich, 2023; King, *et al.*, 2023; van Capelleveen *et al.*, 2023). However, none have approached the topic from a People, Processes and Technology (PPT) dimension and grounded in visions from industry stakeholders. The motivation for this work is based on previous studies where the same research design was used to generate an improved understanding of the positioning of specific stakeholders regarding the use of construction technologies.

The research topic is the DPP for Construction Products. A survey is used as the ground for a mixed methods approach collecting quantitative and qualitative data from respondents in PPT dimensions. As part of the qualitative approach, specific questions for each dimension focus on identifying barriers and opportunities. A focus group involving the authors and experts analysed and clustered the answers and thoughts shared via open questions. The results aim to answer the following Research Questions:

- From the different PPT dimensions, how mature is the construction sector for DPPs?
- What ideas do stakeholders have regarding the opportunities and barriers raised by DPPs, considering the same dimensions?

This research comprises this Introduction followed by the research design and methods presentation. A brief overview is drawn on the origin and motivations for product passports, ranging from several sector perspectives to the specific features of the subject under investigation: Construction DPP. The survey structure and an overview of the quantitative data are presented. To end, the analysis of the qualitative results and overall discussion point to the conclusions section, where answers to the research questions are provided and summarised, endorsing the contributions, limitations, and paths for future works.

Research Design and Methods

The design for this research is inspired by previous studies (Mêda *et al.*, 2020). The reason for using PPT dimensions is associated with the imperatives for Integrated Design and Delivery Solutions (IDDS), where the use of collaborative work processes and enhanced skills, with integrated data, information, and knowledge management, aim to minimise structural and process inefficiencies and to enhance the value delivered during design, build, and operation, and across projects (Owen, 2013). As it will be detailed, this definition aligns with what Digital Product Passports aim to achieve. Complementary to this is the intuition that DPPs' accomplishment contributes to IDDS goals.

The integrated-use approach for mixed methods, according to (Cresswell and Poth, 2018), was chosen to perform simultaneously the qualitative and quantitative data collection using a survey.

After receiving the answers (the sample), the authors gathered a focus group formed by the authors and four experts to discuss the results, assess maturities, and analyse the most relevant barriers and opportunities identified from the PPT dimensions. Figure 1 illustrates the research design detailed above.

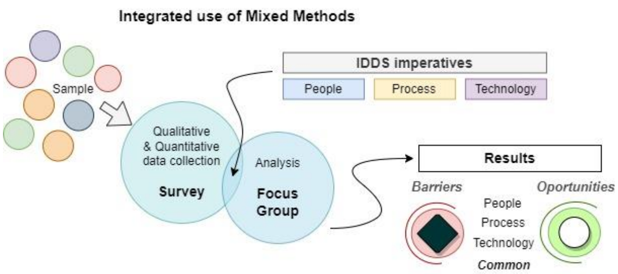


Figure 1: Research Design and method adopted in the study

Why Passport for Products?

When addressing a passport, most people understand it as a travel document issued by a determined country containing relevant information about the person who owns it. That information is secured and curated through time by authorities and individuals. This experience is based on a global standard that facilitates the understanding of passport information by all countries (Torpey, 2018).

Free trade and a single market of products and services constitute the EU's foundational principles. Over the last decades, rules and standards were developed to break down barriers to open trade across member states. Focusing on the free trade of goods, one of the key elements was the definition of a common framework of principles and procedures for marketing products (Official Journal of the European Communities, 2008). This set of rules was summarised on the CE conformity marking or "CE mark", indicating that a product meets the applicable requirements and has undergone the relevant conformity assessment procedures. This label has been working as a sort of passport "facilitating the understanding" between manufacturers, distributors, importers, and other relevant stakeholders across the EU and worldwide.

Performance characteristics were prioritised above all others. This was essential for confidence across markets and not to overcomplicate, ensuring feasibility and widespread implementation. Despite some shortcomings, in the construction sector, the process can be recognised as a successful (Ecorys, 2018), creating the needed background and opening paths to new interests and requirements.

Sustainability and environmental concerns command most 21st-century agendas (United Nations, no date). The product's characterisation imperatives regarding chemicals and energy consumption led the EU to increase the requirements and develop tools to disclose energy efficiency labels and track and disclose chemicals and substances of high concern (Agency, no date; Energy, no date). In this respect, and to improve the legal framework, the EU proposed a new Ecodesign regulation (European Parliament and European Union Council, 2022). The objective is to repeal rules currently in force, mainly concentrated on energy-related products, broadening the scope to a broader variety of products and making them more durable, reusable, reparable, upgradable, recyclable and less harmful to the environment (European Parliament and European Union Council, 2022). The proposal establishes performance and information requirements, supported by a digital instrument to streamline and facilitate data access, management, exchange, and traceability throughout the product lifecycle: the DPP. The proposal sets the link to specific product delegate acts to determine which information should be included and the accessibility rules to different stakeholders (European Parliament and European Union Council, 2022). For the case of most products used in construction, the Delegate Act will be the new version of the Construction Products Regulation (CPR) (Commission, 2022).

One of the most challenging aspects of this process is balancing the requirements with a successful and widespread implementation. In this respect and following the work developed by (King, *et al.*, 2023), it becomes clear for the European Commission that DPP objectives must be made more explicit, the administration costs for businesses must be anticipated, supported or limited, and requirements must be generated on a product-by-product basis, being determined in the construction products case, by harmonised standards.

DPPs must be perceived as an industry symbiosis enabler, given the amount of data that can be collected and managed and the viewpoints and stakeholders that can engage. (King, *et al.*, 2023) envisages an ecosystem for DPPs detailing the abovementioned aspects without focusing on a specific sector or delegated act. Among the many challenges identified is the awareness that such an ecosystem will operate within constraints (such as commercial interests, data quality and data ownership, a variety of sustainability metrics, privacy concerns, legacy systems, cost, skills, and current capacity) to achieve the sustainability values and goals of societal stakeholders (King, *et al.*, 2023).

In what respects the construction sector, DPPs should be additionally envisaged as Digital Twin (DTw) enablers at building scale (Mêda *et al.*, 2021), being a relevant part of the information structures being prepared to characterise built entities (Mêda *et al.*, 2022). Other initiatives originating from data-driven and waste perspectives, such as data templates, ISO standards, and material Passport initiatives, increase the complexity of this already delicate ecosystem (Honic, *et al.*, 2024). Formalising the discerning concept of passports is crucial, supporting decision makers' understanding and stakeholders' awareness of the full potential and realising the benefits together with the surrounding challenges (van Capelleveen *et al.*, 2023).

The first passport was designed to introduce changes in people's identification for the purpose of travelling across borders. By doing that, many services and business opportunities opened, most of which were not realised initially. DPPs are envisaged to pursue similar intentions and, most importantly, to enable a transformation in the product value chain supported by digital tools. The challenges for this endeavour in the construction industry are as broad as the opportunities, therefore justifying a multidimensional approach such as PPT (Derenzi *et al.*, 2009).

Development

Structuring the Survey

The survey was structured to suit a wide range of respondents. The initial objective was to share it via LinkedIn social media. Due to the number of lectures given to postgraduate students in different universities on DPPs' origin, main characteristics, and objectives according to the regulations, it was included as part of the teaching material. Google Forms was used as the support tool. An account subscribed to by the University of Porto was used to fulfil data protection requirements. Fifteen answers were obtained from the lectures that engaged a universe of sixty professionals. In this respect, respondents represent several areas of the construction value chain, from the primary sector (raw materials) to real estate and facility management, including architecture and engineering practice, manufacturing, wholesale, and owners. The answers range from six different European countries, and, in terms of professional practices, one-third has between 5 and 10 years, 26,7 and 20% has between 10 and 15, and 15 and 20 years, respectively. Lastly, 13,3% answered more than 20 years of practice. The survey is organised into five sections, promoting a brief explanation and capturing different types of respondents regarding DPP knowledge.

The Introduction and the Data management, positioned at the beginning and end, frame the study objectives, and collect relevant information regarding the respondents' professional experience, background, and geographic location. The three middle sections focus, as previously framed, on the PPT dimensions, setting a series of open questions with short answers and multiple-choice questions, where personal evaluations are to be provided considering four maturity levels, "Emerging", "Basic", "Advanced" and "Expert", with variations on the definitions depending on the dimension being questioned.

Table 1 presents the Process dimension questions as an example, as the changes are minor for the People and Technology dimensions.

Table 1: "Process" dimension questions used in the survey and presented as example

Question	Type of answer
How do you envisage the PROCESSES involving DPPs? (provide an answer based on a meaningful example from your perspective: awareness on how DPPs are to be structured and how they should be used as data providers for different outcomes throughout the construction life cycle; providing U-value of products for energy performance certificate issuance, among others)	open question with short answer
Can you provide an example of a process or outcome/deliverable that could benefit from the data set in a DPP?	open question with short answer
In your opinion, what might be the biggest CHALLENGES/BARRIERS at the PROCESS level regarding the implementation of DPPs?	open question with short answer
What might be, in your opinion, the biggest OPPORTUNITIES at the PROCESS level regarding the implementation of DPPs?	open question with short answer
How do you assess the existing PROCESSES for STRUCTURING and DELIVERING DATA for DPPs.	multiple-choice
How do you assess the routines' maturity for the DPPs to STREAMLINE PROCESSES during the Design, Construction and Use phases?	multiple-choice
How do you assess the routines' maturity for the DPPs to support PROCESSES contributing to the SUSTAINABLE GOALS (LCA, Level(s), others)?	multiple-choice

Quantitative Results

Different multiple-choice questions were prepared for each dimension to understand better and assess the respondents' maturity and their perceptions of daily practices.

Starting with the People dimension, when asked about their impressions of the level of knowledge regarding DPP, 42,9% pointed it as "Emerging", 50% as "Basic", and 7,1% as "Advanced". The same results were achieved for the questions focusing on awareness of the changes in EU Regulations that will make DPPs mandatory and how different stakeholders realise DPPs' benefits for the value chain.

A different landscape of answers was obtained for the Process dimension. When asked to provide a personal assessment of the existing processes for structuring and delivering data into DPPs, 42,9% classified them as "Emerging", 35,7% as "Basic", and 21,4% as "Advanced". Regarding the maturity of the routines to streamline DPP processes during the Design, Construction and Use phases, 50% considered these as "Emerging", 28,6% as "Basic", and 21,4% as "Advanced". When the question focused on the routines' maturity to support sustainability assessments and goals, 35,7% considered them "Emerging". The same result was obtained for "Basic", and 28,6% classified it as "Advanced". These results demonstrate higher maturity at the Process level if the scope is sustainability-related deliverables.

Concerning the Technology dimension, and when asked to provide a maturity assessment for existing tools to support and deploy DPPs, 21,4% considered these as "Emerging", 35,7% as "Basic", and 42,9% as "Advanced". Regarding the existing technology frameworks enabling information exchange for DPPs, 35,7% considered them as "Emerging", 35,7% as "Basic", and 28,6% as "Advanced". The last question impacted the data ownership and security topic, namely assessing the extent to which these concerns are being anticipated. 42,9% of the respondents considered them "Emerging" and the same percentage as "Basic". 14,3% considered them to be "Advanced".

From the overall view of the results for different dimensions, it is possible to state that higher maturity was achieved for the questions under the Technology dimension. This is not surprising compared to previous PPT studies (Mêda *et al.*, 2020). It is relevant to stress that interesting maturity levels are observed for some questions under the Process dimension.

Analysis and Discussion

Quantitative Analysis

Focus group participants were involved in previous PPT studies, meaning there is a knowledgeable but eventually biased view of the results. Given the topic's novelty and, more importantly, the absence of final decisions on its framework and contents, it was already expected to have low or 0% answers in all dimensions for the "Expert" maturity level.

It is worth noting that the lowest result obtained for "Advanced" in the Technology dimension is associated with data ownership and security, revealing stakeholders' concerns about how these issues are being safeguarded and explained. Still, in this dimension, it is interesting to observe the differences between the results of existing tools supporting and deploying DPPs and those enabling the exchange of DPP information. Respondents recognise that improvements must be made on the second.

From the dimension results, it is impossible to clearly state that Processes and People are falling far behind. When questioned on the routines' maturity to provide outcomes at the sustainability level, a relevant percentage of inputs was given for "Advanced". This situation may indicate that respondents recognise how the processes are aligned to provide these outcomes. In contrast, it might raise concerns about the inability to see how processes will handle others. Associated with these results might be the knowledge of systems and processes to deliver this type of outcome and how they need to be adjusted to accomplish the aimed integration. The people dimension lies behind all questions, namely if we look at the percentages obtained for the "Advanced" level. However, this confirms how awareness is needed. Complementing the legal picture of DPPs, such actions should focus on increasing the stakeholders' maturity where the lowest results were obtained.

Focus Group - Qualitative PPT Analysis

The first exercise consisted of grouping and clustering the answers for each dimension according to barriers and opportunities. At the Process and Technology levels, additional questions were made to improve how respondents see DPP implementation outcomes. These will be described as additional elements.

Starting with the People dimension, the lack of skills is identified as a constraint, confirming results from quantitative data collection. The absence of clear guidelines, reaction to change, the perception of the resources needed to systematise data, and concerns about data overflow can be highlighted as relevant bottlenecks. On the other hand, respondents clearly perceive how productivity can be raised with DPPs. As well as opportunities for business and innovation with a strong focus on sustainability. It is interesting to note that, from a professional perspective, DPP skills are envisaged as a relevant knowledge asset.

Looking at the Process dimension from the barrier's perspective, several concerns are raised regarding the cost of deploying and curating DPPs. Aside from the cost of technology (development/acquisition), respondents expressed concerns about the cost of changes at the process level and the human resources competencies needed. The requirements in terms of new processes, compliance with legacy systems and innovations and processes to ensure data quality and validity are also pointed out. On the opportunities side, the productivity gains, the timing for integrating processes, and the vision of how transactions and comparability can be streamlined are the topics respondents demonstrate higher confidence.

From the Technology dimension, the investment costs needed for new tools, the adaptation, when possible, of legacy systems and the management and update over time raise concerns. Interoperability, data security, and ownership are also aspects pinpointed as hurdles. From the opportunity's perspective, the data access, continuity, and the recognition of a data-driven approach as the background for digitalisation are the most recognised. Some respondents already forecast Artificial Intelligence (AI) developments to bring added value. In contrast, others strengthen how this can constitute an unparalleled opportunity to improve and widespread information management systems in construction.

These results prove that some barriers and opportunities range from more than one dimension; in some cases, they are common to the three. This vision steered the focus group discussions to produce the answers to the research questions.

Discussion on Barriers

Figure 2 clusters the main barriers framed within the dimensions and their overlap. From the core to the boundaries, the main concerns expressed by respondents, ranging from the PPT dimensions, are those associated with the costs, data collection and quality assurance. This aligns with some of the issues raised by EU entities regarding the final text for the regulations, where improvements are needed. However, their implementation should be feasible, scalable, and widespread at all levels. When looking at the overlap between People and Process, the perception of DPPs' value and the development of guidelines, standards, and supporting documentation are identified as the main barriers. In this respect, it is recognised that awareness actions and real use cases are needed as demonstrators and facilitators of the potential contributions of using DPPs. The idea is that these must, in the first moment, concentrate on the already identified goals, namely the ones associated with sustainability.

Notwithstanding, other topics should be presented, namely those that use the same data. This will provide an improved perception of the DPP value, not just for new requirements but also for the long existing ones, such as structural resistance or thermal behaviour. When looking at the identified potential constraints targeting the Process and Technology dimensions, the systems curation through time, hosting resources, update requirements, lifelong operation, and integration with legacy systems stand out. As all respondents have a professional practice, this result may derive from past experiences with systems replacement. This was pointed out in some answers. Establishing the framework and systems for DPPs must consider the objectives and all the relevant legacy tools and processes to accommodate the integration needs as best as possible. In addition, the framework should assume an incremental vision to accommodate more requirements and evolutions through time.

When looking at the aspects set in a single dimension and starting with Technology, interoperability, security, and ownership management constitute the main barriers. Respondents expressed the need to consider all the existing knowledge and systems architecture devoted to interoperability issues. When looking at DPPs from a multi-sector perspective, this aspect raises many concerns

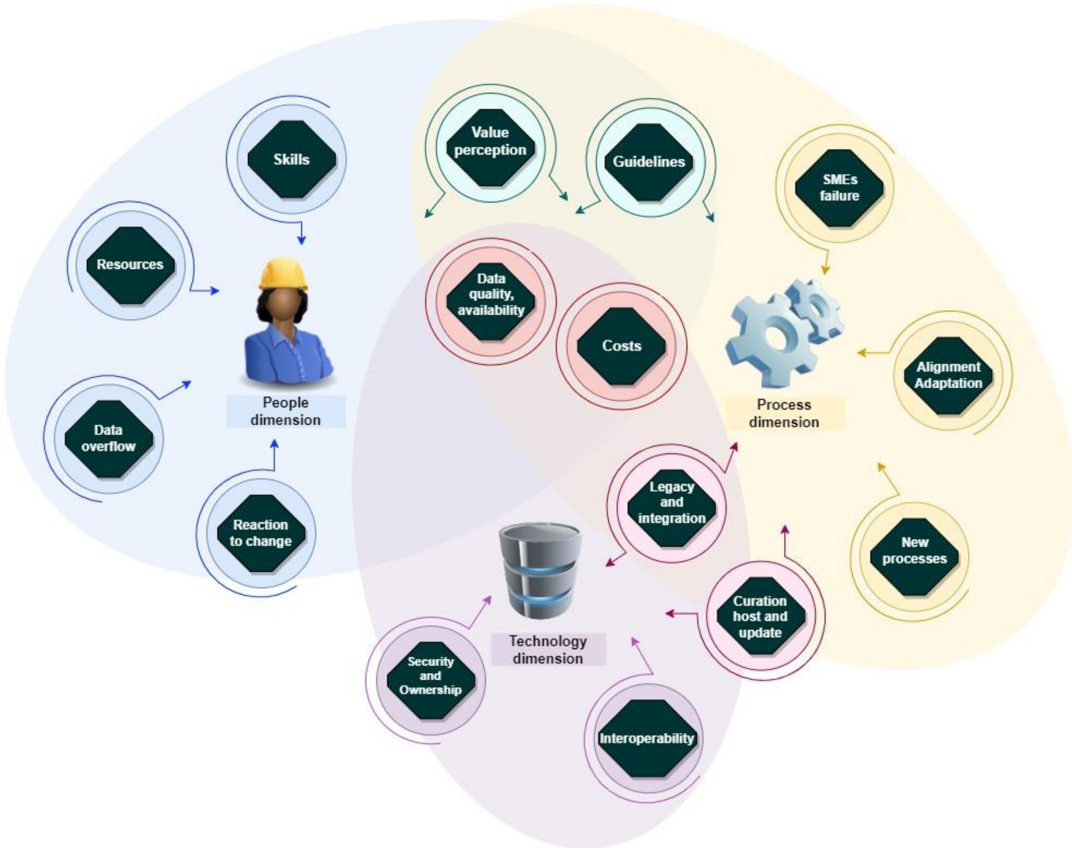


Figure 2: Main barriers clustered at People, Process and Technology dimensions

as several stakeholders are developing tools in silos. In construction, the topic is more safeguarded due to Building Information Modelling (BIM), meaning that developments can be built from the prescriptions set in BIM standards. This link and consensus favours overcoming the ownership and security of data issues. However, the DPP case will need new stakeholders, profiles, and restrictions. From a Process dimension, respondents pointed out three main barriers associated with the need to establish new processes or adapt existing ones, as well as the potential for failure from SMEs, given the dimension and demand underlying this transformation. The capacity to transform, keeping as much as possible or introducing effortless innovations, is a critical challenge common to most endeavours. In this case, and considering that most construction product manufacturers are SMEs, where some are very small, it is necessary to be aware of ways to support these structures to deliver successfully.

Finally, when looking at the People dimension, respondents identify some barriers when discussing innovation, namely reaction to change and lack of skills. The data overflow concern, the incapacity to collect and manage all properties for various products, is part of a data challenge that deserves in-depth research. This topic is associated with the lack of resources. In the present practice, many resources are wasted searching for data and keeping it in the head for some time to avoid getting lost during the processes. This somewhat "mental traceability" must be shifted by improving the ability to search faster and knowing where to access data or where it is stored in larger quantities. This is to say that, despite data overflow being a reality and an issue to prevent, misjudgements can, to some extent, blur this topic.

Discussion on Opportunities

Following the approach used for barriers, Figure 3 presents the landscape of DPP opportunities. Data integration, trust, and transparency stand out by ranging all dimensions under discussion. From the answers, it is observed that respondents understand relatively well the role of DPP and, more importantly, the transformations that it can bring to the construction sector if well implemented. Aligned with this are productivity gains and efficiency, also mentioned as joint opportunities. Only two aspects were identified as sharing two dimensions, one for People and Technology and the other for Technology and Process. Starting with the first case, an ample opportunity is imagined for the sector from data-driven developments, with many answers detailing the importance in having tools that manage and exchange data with a common understanding. Somewhat related is the

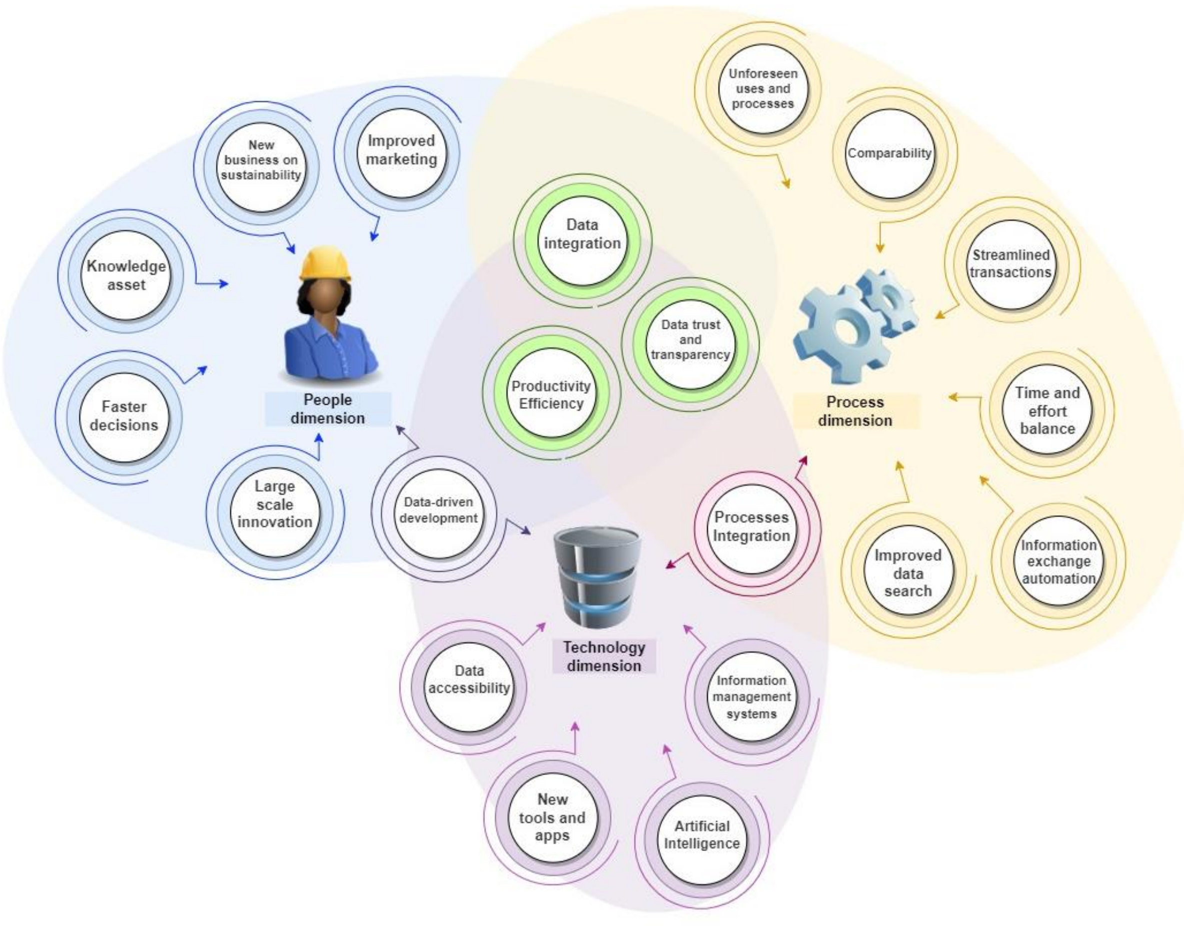


Figure 3: Main opportunities clustered at People, Process and Technology dimensions

other aspect, where respondents highlight the process integration as a game changer.

From a single-dimension vision, the identified opportunities are related to developing new tools and applications, namely information management systems that can support and efficiently administrate rich data environments. In this respect, many respondents pointed out the prospects in AI as essential to speed up and support data management. The last aspect mentioned by respondents reinforces the topic by highlighting the opportunities deriving from access to more organised and broader amounts of data. In the Process dimension, many answers focus on how the DPP-associated processes constitute an opportunity to find data better, increase comparability capacity, streamline transactions, and automate information exchanges. The results are very interesting because examples supporting these topics are provided, including all the construction process phases and the life cycle value chain. The expressed changes envisaged for the comparability opportunity from design to construction or during the bidding process were highlighted. The other topic mentioned is the unforeseen uses and processes. These answers reveal, to some extent, how the DPPs motivate and engage stakeholders. This enthusiasm must be used to tackle barriers previously addressed in the People dimension. Respondents state that improved marketing and faster decisions can occur in this respect. Some pointed out that new businesses can be built based on sustainability. Large-scale innovation is an opportunity deriving from the perception of how DPPs can work in terms of data integration. Finally, more than one respondent identified DPPs as a knowledge asset and an opportunity, supported by the idea that proficiency in this subject will become a competitive advantage in the short term, allowing more competitive professional possibilities.

Conclusions

From applying mixed methods and the People, Process, and Technology dimensions, it can be concluded that regarding Construction DPPs, the Technology dimension is more mature. It is also possible to note that a significant push should be made at the People dimension to provide awareness, guidelines, use cases and training to tackle erroneous value perceptions and reactions to change. DPPs are not a complete novelty. So, in the first moment, there should be a focus on what already exists and how DPPs can be built from it. Working on the outcomes deriving from the existing data is critical, making it part of DPP's information structure. The next step will be understanding the new requirements and the data needed for a comprehensive accomplishment. There are still unknown details on the passport framework and functionalities. The new regulations will bring about clarity. Despite the infancy of the concept and the absence of a full realisation, a solid background should be built using the known facts, fostering the conditions for a concerted change effort for the transformation to be consistent.

The research outcomes constitute relevant contributions, namely on how professionals engaged in the construction processes perceive DPPs and how their thoughts align with the strategies. When detailing the barriers, relevant bottlenecks were identified, targeting a single dimension from PPT, or ranging more than one, with cases comprising the three. The landscape of clustered barriers presents topics that are well-known and labelled as transversal to most innovation initiatives, whereas others can be recognised as specific from the DPP. As presented in Figure 2, the vision of the barriers can be paramount to tackle the main concerns in the short-medium time. Regarding the opportunities presented in Figure 3, it is relevant to stress that DPP seems to motivate stakeholders given the wide range of issues pinpointed and the associated comments. Relevant advantages can be taken from this enthusiasm, namely by establishing strategies to ensure a positive environment for the transition, engaging a high number of stakeholders and working on the most meaningful outcomes for each role. From this overview, it is relevant to highlight the similarity, in terms of pattern, achieved for the barriers and the opportunities.

The research is limited to the questions set in the survey and the sample, which can be classified as above average concerning DPP knowledge. The number of respondents could be more significant. However, it can be recognised that, due to the infancy of Construction DPP, few professionals can provide inputs such as the ones obtained. One of the objectives was to fill the gap in the existing studies by providing an overall perspective on the passport ecosystem for the construction sector when collecting insights from a multidimensional perspective. With the mentioned limitations, this was accomplished. Future research will focus on amplifying the sample and extending the number of questions and topics. A deep understanding of the motivations and reasons for the identified barriers and opportunities is also part of future actions. From the DPP perspective, future research will incorporate new elements as they become public, namely, the EU Regulations, when published, and discussions on the roles and implementation of Construction DPPs.

Digital Product Passports are inevitable in the EU as they are the cornerstone instrument of the digital and green transitions supporting industrial and carbon-neutral strategies. Despite the novelty, stakeholders identify the benefits and opportunities, expressing positive impressions towards a widespread implementation. Many answers state that what DPPs aim to accomplish was long expected, given the inefficiencies derived from manual work that is still mostly needed when using digital tools. This means the novelty is more on the concept than the features and enablers. However, given the push from the sustainability side, concerns are raised regarding the data overflow and the data priorities for the aimed goals. The situation is particularly relevant to SMEs as more support may be required to keep pace. To conclude, the impressions point to an easy understanding of the DPP and to a positive environment for its adoption in the construction industry.

The Construction DPP seems well-positioned to create a concerted change effort across the industry. Therefore, the capacity to set clear guidelines for this to occur is critical.

Acknowledgments

This work was financially supported by: Base Funding - UIDB/04708/2020 with DOI 10.54499/UIDB/04708/2020 (https://doi.org/10.54499/UIDB/04708/2020) of the CONSTRUCT - Instituto de I&D em Estruturas e Construções - funded by national funds through the FCT/MCTES (PIDDAC) and European Economic Area (EEA) Financial Mechanism 2014-2021, Bilateral Relations, FBR_OC2_46, Innovation for Circular Competency - Accelerating GrowingCircle with Digital Product Passports Awareness.

References

Agency, EC (no date) ECHA - SCIP database. Available at: https://echa.europa.eu/scip (11 December 2023).

Byers, B.S. and De Wolf, C. (2023) 'QR Code-Based Material Passports for Component Reuse Across Life Cycle Stages in Small-Scale Construction', Circular Economy, 1(2), pp. 1–16. Available at: https://doi.org/10.55845/iweb6031.

van Capelleveen, G. et al. (2023) 'The anatomy of a passport for the circular economy: a conceptual definition, vision and structured literature review', Resources, Conservation and Recycling Advances, 17(May 2022), p. 200131. Available at: https://doi.org/10.1016/j.rcradv.2023.200131.

Commission, E. (2022) EU Construction Products Regulation proposal, COM(2022) 144 final. Brussels.

Cresswell, J.W. and Poth, C.N. (2018) Qualitative Inquiry and Research Design: Choosing among five approaches. 4th edn. Thousand Oaks: SAGE Pub Ltd.

Derenzi, D. et al. (2009) 'A successful approach in integrating people, process, and technology inside collaborative environments: A practical view of challenges and lessons learned', Digital Energy Conference and Exhibition 2009, pp. 268–277. Available at: https://doi.org/10.2118/123287-ms.

Ducuing, C. and Reich, R.H. (2023) 'Data governance: Digital product passports as a case study', Competition and Regulation in Network Industries, 24(1), pp. 3–23. Av. at: https://doi.org/10.1177/17835917231152799.

Ecorys (2018) Survey on users' need for information on construction products. Brussels. Available at: https://doi.org/10.2873/87907.

Energy, D.-G. EPREL - European Product Registry for Energy Labelling: (11 December 2023) https://eprel.ec.europa.eu/screen/home.

European Parliament and European Union Council (2022) Ecodesign requirements for sustainable products, COM(2022) 142 final. Brussels.

Honic, M., Magalhães, P.M. and Bosch, P. Van Den (2024) 'From Data Templates to Material Passports and Digital Product Passports', in C. De Wolf, (eds) A Circular Built Environment in the Digital Age. Springer. Springer Cham, pp. 79–94. Available at: https://doi.org/10.1007/978-3-031-39675-5_5.

King, M.R.N., Timms, P.D. and Mountney, S. (2023) 'A proposed universal definition of a Digital Product Passport Ecosystem (DPPE): Worldviews, discrete capabilities, stakeholder requirements and concerns', Journal of Cleaner Production, 384(Nov. 2022),. Av. at: https://doi.org/10.1016/j.jclepro.2022.135538.

Lines, B.C. et al. (2015) 'Overcoming resistance to change in engineering and construction: Change management factors for owner organizations', International Journal of Project Management, 33(5), pp. 1170–1179. Available at: https://doi.org/10.1016/j.ijproman.2015.01.008.

Mêda, P. et al. (2020) 'People, Process, Technology in Construction 4.0 - Balancing Knowledge, Distrust and Motivations', in S. Toledo, Eduardo (ed.) 37th CIB W78 Information Technology for Construction Conference, São Paulo, Brazil, pp.218–231. Av. at: https://doi.org/10.46421/2706-6568.37.2020.paper016.

Mêda, P. et al. (2021) 'Incremental Digital Twin Conceptualisations Targeting Data-Driven Circular Construction', Buildings, 11(11), p. 554. Available at: https://doi.org/10.3390/buildings11110554.

Mêda, P. et al. (2022) 'A Process-Based Framework for Digital Building Logbooks', in EC3 (ed.) 2022 European Conference on Computing in Construction. Ixia: EC3, p. 8. Available at: https://doi.org/10.35490/EC3.2022.183.

Munaro, M.R. and Tavares, S.F. (2021) 'Materials passport's review: challenges and opportunities toward a circular economy building sector', Built Environment Project and Asset Management, 11(4), pp. 767–782. Available at: https://doi.org/10.1108/BEPAM-02-2020-0027.

Official Journal of the European Communities (2008) 'Decsion 768/2008/EC: common framework for the marketing of products', OJEU, p. 768/2008/EC Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32008D0768.

Owen, R. (2013) Integrated Design & Delivery Solutions (IDDS). Rotterdam.

Torpey, J.C. (2018) The Invention of the Passport: Surveillance, Citizenship and the State. 2nd edn. Edited by CS in L. and Society. Cambridge: Cambridge University Press.

United Nations (no date) U.N. Sustainable Development Goals. Available at: https://www.un.org/sustainabledevelopment/sustainable-development-goals/ (Accessed: 17 May 2022).

GAMIFICATION OF SAFETY TRAINING IN CONSTRUCTION: A UK PROFESSIONAL PERSPECTIVE

Charbel Ajaka[1], SeyedReza RazaviAlavi[1], Amit Kant Kaushik[1], Pablo Martinez[1]
[1]Northumbria University, Newcastle upon Tyne, United Kingdom

ABSTRACT

Most practitioners consider safety training in construction as a must-have but tedious and repetitive task. Novel technologies and learning approaches, such as gamification, have the potential to increase worker engagement and motivation towards safety training. This paper reports on the development of a game for safety training in construction and the game evaluation by a group of construction professionals. The study aims to evaluate the capacity of gamification to impact safety training. Based on the results, changing current practice to a more gamified approach to safety training and learning has a significant positive impact on safety awareness.

INTRODUCTION

The use of information and communication technologies (ICTs) have increased exponentially in the past decades, and educational institutions and their classrooms are no different. ICTs have been slowly but uninterruptedly introduced in the classroom to facilitate learning and acquiring knowledge, skills, and positive values. With the recent pandemic, it has become a fundamental pillar in improving students' attitudes towards learning (Lazar and Panisoara 2018; McGovern et al. 2020). Among ICTs, digital virtual platforms are revolutionizing educational experiences by providing simulated environments in which users can interact with the environment in an apparent real-time manner. These virtual environments can recreate classrooms, ensure safe experimentation of otherwise risky learning experiences, and often, a more viable and inclusive solution to overcome logistical and organizational challenges and expose learners to certain professional scenarios (Tzanavari and Tsapatsoulis 2010).

For construction education, for example, it facilitates bringing numerous students to a construction site at once, transforming the logistics and safety problems related to having students in a hazardous environment into a technological problem. From a technical aspect, virtual platforms are usually created to be accessed through various means. Hardware selection (e.g., flat screens vs. virtual reality) and the delivery method integrated into the software greatly impact the learning experience and outcomes. On that last part, an interesting approach that has been gaining popularity in the past years is serious gaming. This concept refers to a mix of gamification with an additional educational value that pushes the learner to use game mechanics in non-gaming contexts. It promotes learning, challenging performance, engagement, and personal initiative (Caponetto et al., 2014).

The rationale behind the gamification of training is, at its core, the need to change something within the learning process that is not working properly or adapt it to newer generation trainees. While health and safety in construction is a heavily regulated practice, with specific regulations at national and international levels, construction operations remain hazardous and high-risk workplaces. Gamification can be an approach to provide information at individualized levels (for operators, supervisors, technicians, etc.) an accurate and contextualized learning experiences (Benito Rodriguez et al., 2021). However, users' perspective on such games has not been recorded in literature, specifically from active professionals. This paper presents the evaluation of a safety training game by UK construction professionals, aiming to understand the benefits and drawbacks of gamified learning experiences.

LITERATURE REVIEW

Several studies have reported the success of gaming approaches for learning and skill development (Gee, 2006; Cavalcanti et al., 2021). A game environment presents an immersive environment full of stimuli that makes for an excellent sandbox to improve learning processes. Games are flexible in content and goals and can be used in a wide range of locations and engagement. This is especially important for flexible learning approaches in cases of worker availability disruptions, as occurred with the recent COVID-19 pandemic.

In a previous study, a framework for the development of games for construction activities was proposed (Le et al., 2015). Three modules are required for any game: a knowledge dissemination phase, a knowledge reflection phase, and a knowledge assessment phase. This approach has been used in the literature, and the results showed improvements in hazard identification (Bernardes et al., 2015). It also has proven applications in risk management to help learn activity interpretation and cause-consequence analysis through simulated scenarios (Barot et al., 2013).

Also, current literature highlights the dynamic presentation capabilities of gamified learning processes for safety training as a key element that supports higher immersion and learning outcomes. This is in line with learning theory regarding attention and visual cues. In fact, when delivered using virtual reality methods, games provide an additional level of dynamism and an improved learning experience (Harichandran et al., 2021). In this line, a large number of studies have been done over several learning training processes to address construction education. For instance, Kazar et al. (2021) developed a serious game for safety training and assessed its effectiveness with senior civil engineering students. Their results showed that the serious game is an effective training approach, and its effectiveness is not influenced by the game experience of the learners. In another study, Gallerati et al. (2017) assessed students' learning retention and the cost associated with training activities through serious gaming and VR. In their experiment, students used a simulated immersive VR that emulated a drilling land rig, the results showed improvement in learning retention and a reduction in safety risks and training costs. For flood safety training in the urban built environment, D'Amico et al. (2023) developed a non-immersive VR game, which could make a significant improvement in the self-efficacy and safety knowledge of the learners. The use of digital twin, which can dynamically update the digital models with real-time data, was studied by Harichandran et al. (2021) for developing dynamic VR games. They developed a framework for creating safety training scenarios from the project intent information, project status knowledge, safety regulations, and historical knowledge provided by the digital twin. Despite the novelty of the proposed approach, manual creation of the scenarios is reported as its drawback to be addressed in future.

This literature review has highlighted that the current implementation of serious gaming in AEC education has a lot of potential for education and training. Safety training has been identified as a key area for gamification. However, most of the existing studies focused on experiential learning for students, disregarding professionals who may have a more pragmatic approach to learning approaches (Abotaleb et al., 2023).

This paper explores how industry professionals perceive the introduction of serious gaming for safety training and how it compares to their current approaches. The population is a sample of AEC professionals who are actively involved in the construction industry, not necessarily around safety. Based on this, the following research questions are proposed:

1. What is the opinion of industry professionals regarding the gamification of safety training?
2. How does gamification compare to current approaches for safety training?

METHODOLOGY & GAME DESCRIPTION

The research is based on a quasi-experiment due to the lack of randomization in the testing group. The professionals involved in this experiment were selected based on experience and availability. A total of 28 professionals participated in the experiment in four different sessions due to space and availability constraints. However, a single group is created to check and evaluate the methods and learning outcomes presented through the experiment (see Figure 1). The learning activity is designed to introduce the population to two different concepts: site traffic safety from a trucker perspective, and crane safety from the crane operator and ground safety personnel. The first activity uses serious gaming to present different hazard profiles and safety procedures of driving across a congested and operational construction site while driving a loaded truck. Similarly, the second activity focuses on crane operations from two points of view, one while performing a module lifting operation from the cabin and the other one while doing a routine ground inspection around the crane lifting area. After the experiment, lasting around 40-45 minutes, a post-assessment point is given to allow the population to provide feedback and assess their learning experience.

Figure 1. Serious gaming for safety training activity.

To kickstart the learning activity, the participants are given a small tutorial prior to immersion in the game. The tutorial lasts about 20 minutes and covers the basics of the game experience regarding controls, main menu, and game structure and goals. This phase is meant to facilitate navigation within the game to all participants. Any participants who had difficulty with navigation or any game-related issue were provided with support and additional training.

After that, the participants start the game. They are thrown in first person on a construction site environment where they can interact with the main menu of the game (see Figure 2). This menu allows players to select the different levels available in the game. Each training topic/theme is structured as a set of levels, starting with a tutorial phase where participants are taught using visual cues and written statements regarding safety rules on specific scenarios that the participants walk into.

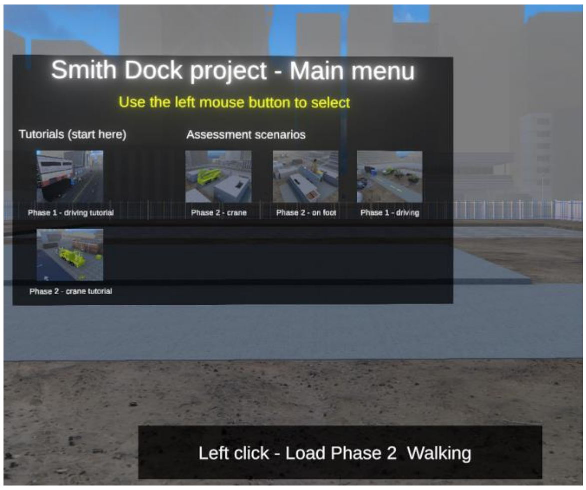


Figure 2. Screenshot of the main menu of the serious game.

Upon completion of the tutorial, an assessment level on the theme just covered is unlocked. This assessment level is similar to the tutorial level in terms of tasks and environment; however, no cues or support is given. A scoring system is introduced to automatically assess the results of the participant over their behavior and motion across the level. Participants receive their scores and feedback on how to improve the sections in which they lose points. The level structure is illustrated in Figure 3. Participants are free to navigate through the different tutorial levels at any given time, as there is no order to those levels.

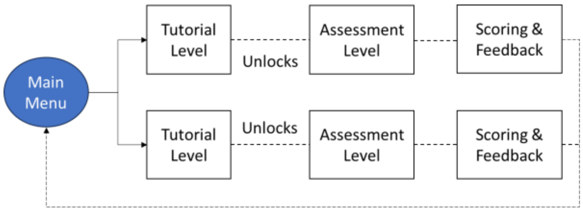


Figure 3. Game flow and level structure used.

The game brings the participants across different situations where safety training can be taught and tested. Within the same construction environment, that can be exported directly from BIM for additional user immersion, a large variety of scenarios can be proposed to imitate even the less plausible hazard profiles. While the safety training focused on traffic and crane operations, additional safety training elements of more generic nature can be added. For example, as shown in Figure 4, signage awareness or construction debris removal can be incorporated.

After the completion of an assessment phase, the participants are scored based on their performance, behavior, and interactions with the built environment. Performance is scored based on timing, behavior is scored around whether a set of safety tasks are done or not, and the interactions with the built environment are built as penalties (i.e., ignoring signage results in a loss of points). A final score is provided to the participant, with some feedback regarding potential improvement or where the participant was penalized. Then, the score is uploaded onto the leaderboard.

Once the participants are done playing with the game, about 40 to 45 minutes, a post-test questionnaire is given to each participant to fill. This questionnaire focuses not just on the learning experience, the game development, or the effectiveness and drawbacks of using serious gaming for safety training but also on prior engagement with safety training methods and current practice.

Figure 5. Signage included in the construction scenario for safety training.

RESULTS

This study and its protocol have received full ethical approval from the Northumbria University ethics online system (project ID: 5567). All participants gave permission to use the collected data. Initially, a personal assessment of the participants' professional experience and prior experience with safety training is provided. It allows us to showcase a better understanding of the participants and potential biases introduced by the population subconsciously. Figure 5 illustrates those results (left to right) for work experience, participant professional active field or background, and their prior experience with video games and safety training methods respectively. Note that the population is slightly biased towards early career professionals, with a current role in architecture or engineering. This will be addressed in

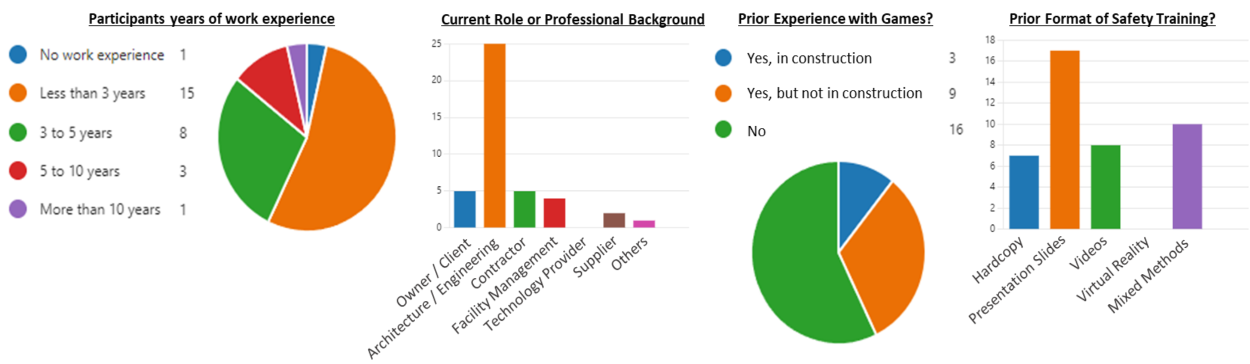


Figure 4. Experiment population data.

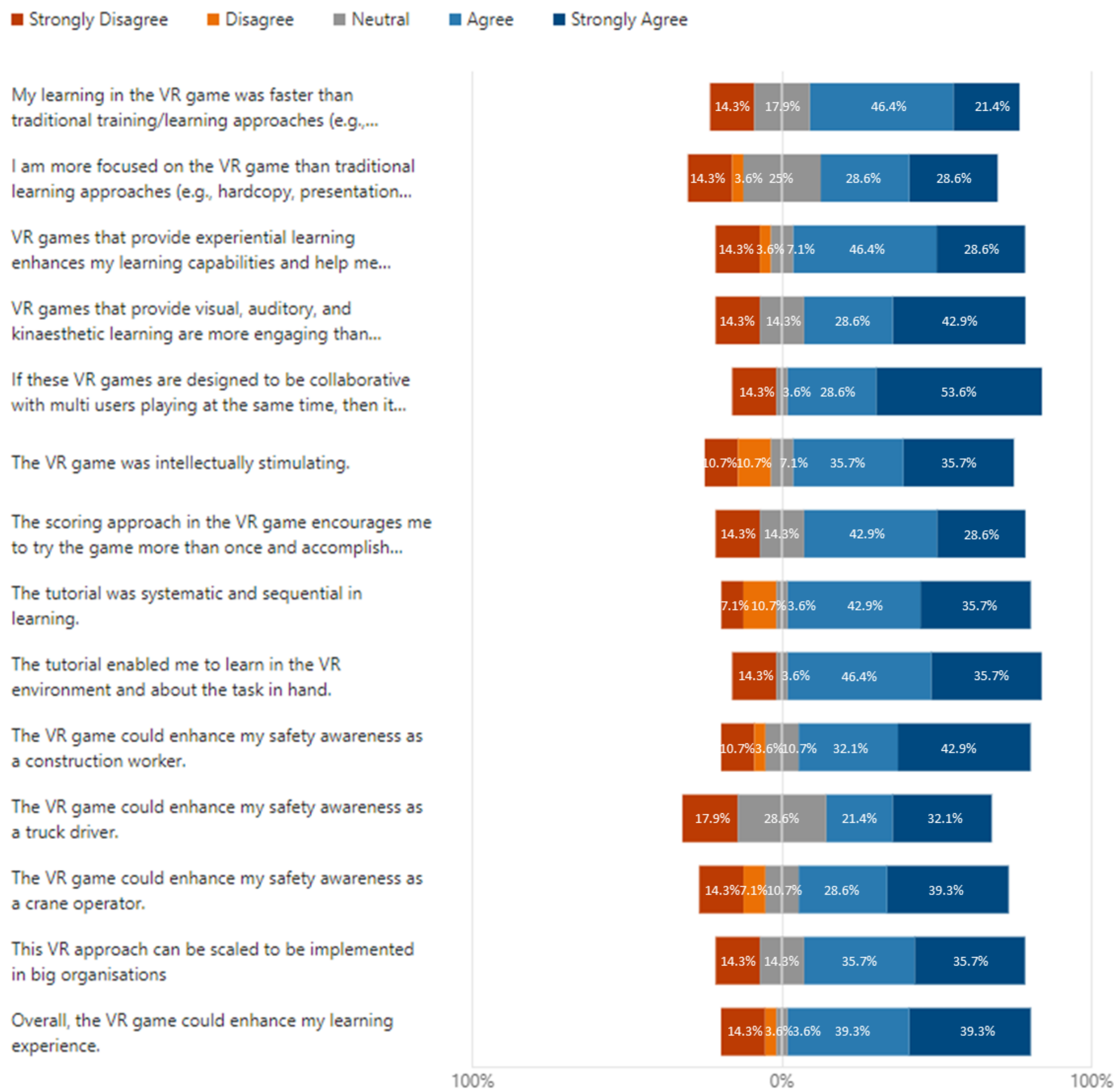


Figure 6. Participants' answers to the final questionnaire.

further experimental testing to provide a more homogenous population.

The results from the quasi-experiment can be found in Figure 6. Participants answered a series of questions using a 5-step Likert scale (1 – Strongly Disagree; 2 – Disagree; 3 – Neutral; 4 – Agree; 5 – Strongly Agree). Overall, the participants reacted positively to safety training through gaming instead of more traditional learning methods. Table 1 showcases the statistical analysis of the results obtained per question, in the order shown in Figure 6.

Table 1. Statistics of the questionnaire results.

	Questionnaire Results	
	Median	**Std. Deviation**
Question 1	4	1.235
Question 2	4	1.322
Question 3	4	1.305
Question 4	4	1.355
Question 5	5	1.361
Question 6	4	1.326
Question 7	4	1.278
Question 8	4	1.205
Question 9	4	1.291
Question 10	4	1.280
Question 11	4	1.402
Question 12	4	1.411
Question 13	4	1.319
Question 14	4	1.82

Based on the results obtained, on average, the learning experience of the professionals questioned is positive. The gamification approach is seen as an improvement of current approaches and can potentially increase safety awareness and training capabilities for operators/workers. The questionnaire results highlight the value given by the users of the collaborative aspect of the game (question 5), the immersion (question 4), and the increase in safety awareness when compared to traditional training delivery (question 10). Note that most professionals agree that the technology is scalable to an organization level.

However, more divisive results are obtained for certain questions: a more average neutral stance is given towards the efficacy of engagement and focus during the training of the game users compared to current approaches and the overall view of the game from a learning perspective is relatively neutral. Regarding those points, the participants could further elaborate on. Their feedback on possible improvements is listed below:

- Considering adding a story mode that simulates certification.
- Incorporating multiplayer.
- Improved tutorials for a broader range of users (especially non-gamers).
- Clearer instructions and transparency on the scoring system.
- Providing different camera angles instead of only first-person point of view.
- Better details and graphics for additional immersion, especially sound.

Separating the comments on improved technical performance, some of the desired improvements are of a complex nature. Developing certifications or learning assessments in a multiplayer environment would be an interesting challenge. While it challenges the idea of individual knowledge testing, it would probably be a more realistic representation of safety considerations and efforts on construction sites. As safety on sites is not an individual responsibility, could training be more effective in collaborative environments? Further research to explore feedback provided and perform more extensive evaluations on the potential of gamification will be pursued.

STUDY LIMITATIONS

The showcased study is a preliminary work to understand the impact of serious gaming on safety perceptions and training around hazards on construction sites, however the results obtained have to taken with the consideration of certain limitations in the scope and untested biases that can be accounted for.

For example, how the novelty of the serious gaming experience is affecting the perception of the usefulness of the approach by the users is uncertain. Or how the safety perception given by the gaming experience is translate onto the physical world. In the end, even the gaming optics of the criteria used to develop the training scenarios or even the scenarios themselves may had an impact on the experience outcomes that is uncertain.

Overall, multiple variables that would need to be accounted for are not considered in this study and should definitely be included in the experiment planning and posterior analysis to come up to more decisive conclusions.

CONCLUSIONS

The study's investigation into the use of serious gaming for construction safety training presents innovative findings. It demonstrates that gamification significantly enhances professionals' engagement, motivation, and safety awareness, offering a novel approach to addressing the complexities and hazards inherent in construction environments. The positive reception of the game's immersive and interactive features encourages a pivotal shift towards more dynamic training methodologies. This research enriches the academic discourse on educational technology by applying gamification within a novel context. It also provides evidence of its potential to revolutionize traditional safety training practices in construction. It underscores the efficacy of serious gaming in improving learning outcomes, thereby advocating for its integration into standard training protocols to better meet the evolving demands of the construction industry. The study outlines several pathways for future research, including enhancing game functionalities, exploring gamification's long-term impact on safety practices, and its scalability and cost-effectiveness. These potential avenues demonstrate further the role of serious gaming in professional education and its potential to transform safety training in construction and other sectors.

REFERENCES

Abotaleb, I., Hosny, O., Nassar, K., Bader, S., Elrifaee, M., Ibrahim, S., ... & Sherif, M. (2023). An interactive virtual reality model for enhancing safety training in construction education. Computer Applications in Engineering Education, 31(2), 324-345.

Barot, C., Lourdeaux, D., Burkhardt, J. M., Amokrane, K., & Lenne, D. (2013). V3S: A virtual environment for risk-management training based on human-activity models. Presence, 22(1), 1-19.

Benito Rodríguez, A., Ángel Casanova, M., Rodríguez Jiménez, M. E., Gutiérrez Canete, M. T., & Rodríguez-Martín, M. (2021, October). Based-on gamification activity for training in occupational risk prevention in the context of the works at height. In Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'21) (pp. 183-188).

Bernardes, S. M. F., Rebelo, F., Vilar, E., Noriega, P., & Borges, T. (2015). Methodological approaches for use virtual reality to develop emergency evacuation simulations for training, in emergency situations. Procedia Manufacturing, 3, 6313-6320.

Caponetto, I., Earp, J., & Ott, M. (2014, October). Gamification and education: A literature review. In European Conference on Games Based Learning (Vol. 1, p. 50). Academic Conferences International Limited.

Cavalcanti, J., Valls, V., Contero, M., & Fonseca, D. (2021). Gamification and hazard communication in virtual reality: A qualitative study. Sensors, 21(14), 4663.

D'Amico, A., Bernardini, G., Lovreglio, R., & Quagliarini, E. (2023). A non-immersive virtual reality serious game application for flood safety training. International journal of disaster risk reduction, 96, 103940.

Gallerati, P., Bagnato, S., Casciaro, D., Conte, A., & Maisano, M. (2017, March). Use of serious gaming and virtual reality applications improves students' learning retention and reduces safety risks and costs associated with training activities. In Offshore mediterranean conference and exhibition (pp. OMC-2017). OMC.

Gee, J. P. (2006). Are video games good for learning?. Nordic Journal of Digital Literacy, 1(3), 172-183.

Harichandran, A., Johansen, K. W., Jacobsen, E. L., &

Kazar, G., & Comu, S. (2021). Effectiveness of serious games for safety training: A mixed method study. Journal of Construction Engineering and Management, 147(8), 04021091.

Harichandran, A., Johansen, K. W., Jacobsen, E. L., & Teizer, J. (2021). A conceptual framework for construction safety training using dynamic virtual reality games and digital twins. In ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction (Vol. 38, pp. 621-628). IAARC Publications.

Teizer, J. (2021). A conceptual framework for construction safety training using dynamic virtual reality games and digital twins. In ISARC. Proceedings of the International Symposi4um on Automation and Robotics in Construction (Vol. 38, pp. 621-628). IAARC Publications.

Lazar, I., and I. O. Panisoara. (2018). "Understanding the role of modern technologies in education: A scoping review protocol." Psychreg Journal of Psychology, 2: 74–86.

Le, Q. T., Pedro, A. K. E. E. M., Lim, C. R., Park, H. T., Park, C. S., & Kim, H. K. (2015). A framework for using mobile based virtual reality and augmented reality for experiential construction safety education. International Journal of Engineering Education, 31(3), 713-725.

McGovern, E., G. Moreira, and C. Luna-Nevarez. (2020). "An application of virtual reality in education: Can this technology enhance the quality of students' learning experience?" Journal of Education for Business, 95 (7): 490–496.

Tzanavari, A., and N. Tsapatsoulis. (2010). Affective, interactive and cognitive methods for e-learning design: Creating an optimal education experience: Creating an optimal education experience. IGI Global.

PRELIMINARY EXPERIMENTS WITH A HANDS-ON INTRODUCTION TO GENERATIVE DESIGN TECHNIQUES USING TANGIBLE OBJECTS

Divyanshu Bajpai[1], Vishal Singh[1], and Aradhya Singh[1]
[1]Department of Design and Manufacturing, Indian Institute of Science, Bengaluru, India.

Abstract

Generative Design (GD) is typically introduced to students through computer-aided design (CAD) courses. Diverging from this approach, this research uses hands-on experiments to introduce GD concepts to students. Shape Grammar (SG) and Cellular Automata (CA) techniques are explored using a tangible toolkit, distinct from CAD sessions. In an experimental setup, students use SG and CA techniques in separate sessions, following a think-aloud protocol. Data is captured using video recording. The study aims to uncover patterns in students' thought processes and assess if tangible exercises can instil confidence and interest in GD. The paper outlines the experimental setup, presents preliminary results, and discusses implications for teaching and GD tool development.

Introduction

Generative design (GD) is an iterative design process widely used in architecture to support human designers and help in design exploration. Various generative design techniques, including cellular automata (CA) (Herr & Ford, 2016; Von Neumann, J., 2017) shape grammar (SG) (Stiny, 1980), L-system (LS) (Lindenmayer, 1968), genetic algorithm (GA) (Jo & Gero, 1998) and swarm intelligence (SI) (Bonabeau et al., 1999) are reviewed based on preliminary literature review (Singh & Gu, 2012). In this paper, CA and SG have been introduced through tangible hands-on exercises.

Cellular automata (CA) is a collection of cells arranged in a grid of specified shape, and each cell changes state as a discrete time step, according to a set of rules depending on the state of neighbouring cells (Zaitsev, 2017). The essential components of CA are Cell, Cell States, and Neighborhood. Two kinds of cell neighbourhoods are usually considered in a 2D cellular automata theory. They are Von Neumann and Moore. Von Neumann's neighbourhood is considered in this work.

Fig. 1: Classical neighborhood in 2D CA a) von Neumann's neighborhood b) Moore's neighborhood (Zaitsev, 2017)

CA systems are widely considered in architectural design for their spatial abilities and form generation (Herr & Kvan, 2007; Krawczyk, 2002). CA produces unpredictable results after every iteration, which helps designers push their capabilities of creativity and imagination (Wolfram & Gad-el-Hak, 2003).

Shape grammar (SG) is a set of shape rules that can be applied to generate a set or language of design (Knight, 1999). According to Stiny (1980), the essential components of SG are a finite set of shapes, a finite set of rules, a finite set of symbols, and an initial shape. SGs are widely used to recreate architectural and design styles (Stiny & Mitchell, 1980; Poon & Maher, 1997; Stiny & Mitchell, 1998). The speciality of SG is to generate several design solutions using a finite number of shapes and rules. SG is widely used for design exploration and visual compositions (Tapia, 1999). The distinct feature of SGs is that they can be used as an analysis and synthesis tool to decompose complex shapes and generate complicated forms starting from a simple initial shape, respectively (Fasoulaki, 2008).

In this article, design experiments are performed with the help of a tangible kit relevant to solving the given problems using GD techniques. The design experiments assist in assessing and developing knowledge about the subjects' learning process and may provide a medium to support subjects' learning (Cobb et al., 2003).

Initially, the GD (CA & SG) technique is briefly discussed, followed by a background on related work. Then, the methodology is detailed, covering experiment design, subjects, data collection, and students' thought processes. Next, the experiment framework is outlined, including the problem description, constraints, rules, and setup details. Finally, results and discussions are presented, along with future work directions.

Related Work

Several studies have shown that incorporating GD techniques into design education can enhance the students' ability to explore design possibilities and improve their problem-solving skills (Fischer & Herr, 2001). These authors argue that design teaching should include more focus on techniques and skills in GD, especially in the initial stages of learning. Eris (2003) highlights the cognitive paradigm associated with the type of questions designers ask during the design process, which represents a convergent-thinking distinction. These studies also highlight the importance of hands-on experience in the learning process (Cassim, 2013). They present the findings from a case study of design students' reflections on their project experience. The reflections highlight the importance of hands-on experience with real-world problems in advancing a student's design knowledge and skills. Pusca et al. (2017) use case studies to show that hands-on activities can enhance students' knowledge and skills in design, creativity, communication, and modelling. They stress the importance of not only hands-on experiences but also minds-on to promote deeper thinking and understanding.

Junginger (2007) introduces the concept of "heart, hand and mind" as essential tools for design inquiry. However, most of these studies have focused on introducing GD through CAD systems (Krish, 2011 proposed a CAD-based design exploration method**;** Khan et al., 2019; Khan & Awan, 2018). CAD-based GD may not be accessible or appealing to all students initially because they need to learn to use the CAD tools and, perhaps, have adequate confidence in programming or coding scripts.

Hence, this research explores the effectiveness of incorporating hands-on exercises in teaching GD techniques, specifically using CA and SG. The goal is to determine if these tangible exercises can foster student interest and confidence in GD while improving their problem-solving and design exploration skills. The study also aims to analyse the students' thought processes during the exercises.

Numerous studies have explored the use of tangible tools for teaching computational and design thinking (Want et al., 2014; Lin et al., 2020; Matthee & Turpin, 2019). These studies highlight the benefits of using tangible tools in enhancing students' understanding of computational concepts and promoting hands-on learning. Soleimani et al. (2016) present a hands-on activity for the Computer Organization course in Computer Science using low-cost single-board computers. The results show that 90% of the students prefer hands-on activities to improve their learning and understanding of computational concepts. However, these studies have mostly focused on teaching programming concepts, not specifically GD. Marshall (2007) presents T-Maze, a tangible programming tool for children aged 5–9 to build computer programs in maze games by placing wooden blocks. Therefore, the current study aims to fill this gap by exploring the use of tangible exercises to teach GD techniques.

Research Methodology -

Experiment Design

The research on GD systems is often focused on reduced time and cost to achieve design alternatives, optimization, accuracy, consistency, and so on. Problem formulation and choice of the GD technique and representation play an equally important role, even if areas like optimisation steal the limelight. In this context, the experiment was designed to examine whether the tangible kit (discussed in a later section) can help students analyze the patterns in their thought processes while applying a GD technique to solve a given problem.

The experiments include three sub-experiments. Separate tangible kits are designed for each sub-experiment, as explained later in the framework section.

Subjects

The subjects for this experiment are postgraduate students in product design or undergraduates in architecture. Six subjects were selected, four from design and two from architecture. Although all the subjects have a background in design and architecture, they had yet to gain any knowledge of GD before this activity.

Experiment modules

The experiment sessions are divided into three modules. Each module has a demonstration video and reference manual explaining how to perform the activities. A moderator is still available during the experiment to clarify doubts (Toyong et al., 2020) and to ensure that the quality of recordings and data capture is acceptable. For reference, a demonstration video of rules and instructions is displayed on the screen throughout the activities.

Data Collection

Think-aloud protocol (Fonteyn et al., 1993) is used in the design sessions of approx. two hours, with approximately 10 minutes of break in between. The data is video recorded on tape, capturing both the verbal data and the evolution of the design through a camera directed towards the worktable with the tangible kit (Goldschmidt, 1994). A 10-minute post-session semi-structured interview was conducted with each subject, followed by open-ended feedback from the subjects about the whole experience. The interview (Patton, 1990) questions are outlined below -

Interview Questionnaire (Semi-Structured)

- *Which technique did you find easier to perform? Why?*
- *Any tangible thing or object you felt was missing from the experiment apparatus that'd have helped you during the experiment? Any suggestions?*
- *Which technique granted you more freedom while designing? Rank all three techniques.*
- *Did any part of the material or the experiment make you want to learn more about it?*
- *What part of the experiment did you find the easiest/clearest? What parts were unclear/more difficult?*

Feedback Questionnaire

- *What did you like about the experiment and its modules? What did you not like?*
- *How can we improve the delivery of content for better and easier understanding?*
- *Which content delivery method (manuals, videos, etc.) was most useful?*
- *Was there any ambiguity in any part of the experiment that halted your progress?*
- *Was the time allocated for reading manuals, watching videos and performing the experiment enough?*
- *Did you like the hands-on approach to learning Generative Design?*

A five-point Likert scale, a standard measurement method in educational research (Boone & Boone, 2012), is used. A specific questionnaire for each activity is developed. Data is collected from the subjects at the end of the activity. Each question has a declarative statement that measures one trait, such as measuring the challenges associated with a specific technique. Numerically, the scale is assumed to vary from -2 to 2.

- *The experiment has certainly helped me learn the Cellular Automata technique.* ***(SA | A | NAND | D | SD)***
- *The experiment has certainly helped me learn the Shape Grammar technique.* ***(SA | A | NAND | D | SD)***

- *The experiments conducted today:*
  Helped me learn more | Made no difference to how I learn | Were detrimental to my learning process
- *The Cellular Automata module of the experiment is challenging.* ***(SA | A | NAND | D | SD)***
- *I am comfortable with the activities performed in the Cellular Automata module.* ***(SA | A | NAND | D | SD)***
- *The Shape Grammar module of the experiment is challenging.* ***(SA | A | NAND | D | SD)***
- *I am comfortable with the activities performed in the Shape Grammar module.* ***(SA | A | NAND | D | SD)***
- *The Tangible Free Design module of the experiment is challenging.* ***(SA | A | NAND | D | SD)***
- *I am comfortable with the activities performed in the Tangible Free Design module.* ***(SA | A | NAND | D | SD)***
- *I'd like to have such interactive exercises in my courses.(* ***SA | A | NAND | D | SD)***

SA= *Strongly Agree |* **A**= *Agree |* **NAND**= *Neither Agree Nor Disagree |* **D**=*Disagree |* **SD**=*Strongly Disagree*

Data Analysis

The recorded data was transcribed (Tilley & Powick, 2002) and cleaned by an experienced transcriber. The transcribed data was subjected to thematic analysis (Maguire & Delahunt, 2017; Braun et al., 2016). Thematic analysis is an effective tool for identifying, analysing, and reporting patterns or themes within the data (Braun & Clarke, 2006). Two researchers independently conducted the thematic analysis for the whole transcript. The identified themes were reviewed again, and cross-checking against the transcription was done together. The questionnaire-survey results are analysed using descriptive statistics.

Experiment Framework

An exhibition layout problem was given for all three sub-experiments. The description of the problem is as follows.

"To design an exhibition space layout to display paintings such that there are four partitions/zones with similar spatial capacity. The paintings are displayed on stands such that there are an adequate number of surfaces to show the paintings. Adequate space must be provided inside the zones for visitors to walk and view the paintings. A 2D layout is expected (top view)".

Requirements and constraints on the display stand:

A display stand combines flat panels as surfaces for hanging the paintings. A panel is represented by a tangible object (matchstick), and a display stand is essentially a set of multiple connected matchsticks. The viewing angle should be between 90 and 180 degrees. These panels are to be placed so that the two sides of the panel – front and back – are easily visible to the viewers.

Experiment type and description

This work involves three sub-experiments with respective kits. The details of each experiment are as follows:

Sub-experiment A: Free tangible (FT) without rules

There are no predefined rules and constraints. A free, unconstrained space made of cardboard is provided. Tangible objects, including matchsticks, clay, thread, and clips, are provided. There is no explicit connection or reference to GD for this part of the experiment. The free-form thinking of the students is captured. Figure 2 shows an example of the layout created by one of the subjects. -

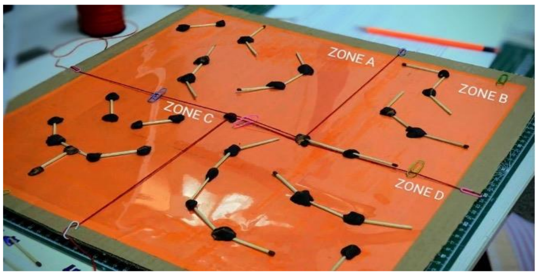


Fig. 2: Display structure layouts designed by a subject for an exhibition hall with four zones.

Sub-experiment B: Cellular automata tangible (CAT)

This activity addresses the challenge of designing the display structures using a cellular automaton (CA) kit while considering structural rigidity constraints. The subjects are constrained to use no more than five panels to ensure stability, and a closed-loop structure is excluded due to ensure dual-sided visibility. A grid outline serves as the canvas for CA operations. The tangible kit employs matchsticks for creating display stands distributed in a 2*2 grid. Subjects engage in two sub-activities, focusing on display structure and layout planning. The goal is to assess the tangible kit's effectiveness in instilling fundamental computational thinking for GD while fostering confidence and interest. Decision-making is pivotal as subjects adhere to predefined rules, with their choices and justifications contributing to the learning outcomes.

CA rule formation and application –

CA usually uses a grid, and state changes are made according to the neighbourhood conditions driven by rules. The rule formation is discussed below:

Two cell states are considered for this activity, named A and B (Fig. 3). The state of the cells in the next iteration is governed by the state of the neighbouring cells in the previous iteration. Von Neumann's neighbourhood is considered for this activity.

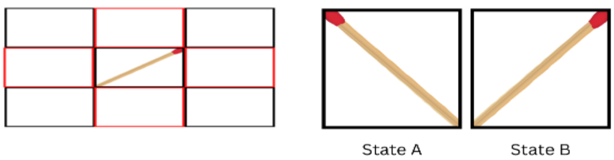


Fig. 3: Cell Von Neumann neighborhood

The following rules govern the cell state in the "New" iteration:

Rule #1: Assign cell state B if it has more neighbouring cells with state A; otherwise, assign cell state A if it has fewer neighbouring cells with state B.
Rule #2: If the number of neighbouring cells with state A is equal to the number of neighbouring cells with state B for some arbitrary cell position, then the cell state (in the "New" iteration) for that specific position is declared as:

Rule #2.1: State B (if the cell number of that arbitrary cell position is odd)

Rule #2.2: State A (if the cell number of that arbitrary cell position is even).

The three specified rules guide manual CA iterations. After three iterations, the subject evaluates the resulting display structures. Selection is based on logical assessment, requiring structures to maintain a continuous connection of panels (or matchsticks) without breaks for consideration. Figure 4 shows examples of some of the display structures generated by the subjects –

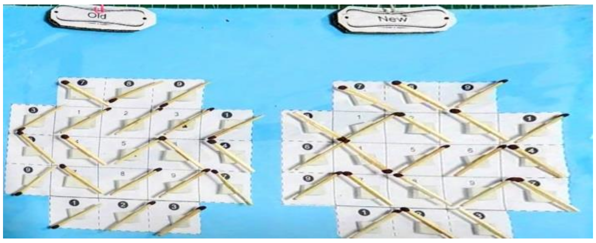


Fig. 4: Display structures obtained by a subject after performing an iteration of CA rules.

Once a display structure is selected, it is ready for arrangement in the four zones, marking the completion of the first part. Subsequently, three rules govern layout planning iterations, outlined below.

Arrangement of Tokens in the 4 Zones:

The chosen structures form four tokens through altered orientation and connection to a token's corner. Subjects can freely connect joints, ensuring the structure fits the grid. Design decisions and orientations must be explained, fostering robust decision-making. Figure 5 shows examples of tokens created by the subjects.

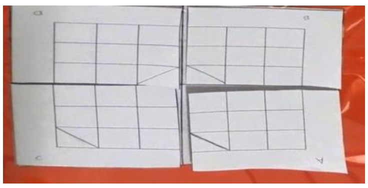
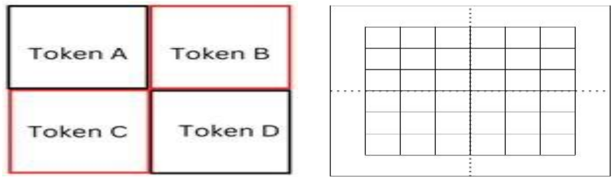


Fig. 5: Four tokens created by a subject.

Following token creation, these states represent CA, contingent on the subjects' creativity. A 2x2 grid, forming four exhibition zones, is generated for placing tokens A, B, C, and D. The grid has a computationally generated initial configuration, and tokens follow a 3x3 structure with walking space around the periphery, exemplified for exhibition formation. Figure 6 shows an example of a token configuration to form an exhibition-



(a) **(b)**

*Fig. 6 (a) 2*2 grid for placing tokens (b) Tokens arranged to form an exhibition.*

The newly created 2x2 matrix is used as the basis for the next set of 3 iterations in adherence to the set of grid rules given below:
Rule A: If token D neighbours either token C or B (or both) in the "Old" iteration, the token in the "New" iteration at that position will be token A.
Rule B: If token B neighbours either token A or D (or both) in the "Old" iteration, the token in the "New" iteration at that position will be token C.
Rule C: If token A neighbours either token C or token D (or both) in the "Old" iteration, the token in the "New" iteration at that position will be token B.
Rule D: If token C, with an odd cell label number, neighbours either token A or token B (or both) in the "Old" iteration, the token in the "New" iteration at that position will be token D.
Rule E: If token C, with an even cell label number, neighbours either token A or token B (or both) in the "Old" iteration, the token in the "New" iteration at that position will remain token C.

Default Rule: If none of the above rules are applicable, copy the "Old" state into the "New" state without any changes.

After performing three iterations, subjects have three different layouts and can use their decision-making skills and the requirement to select or reject a layout. Figure 7 shows an example of one complete layout made by one of the subjects.

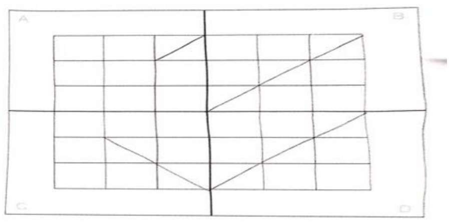


Fig. 7: A token layout obtained by a subject after performing the 3rd iteration on the 2x2 grid.

Sub-experiment C: Shape grammar tangible (SGT) –

SG has predefined shape rules, starting from the initial shape. This activity aims to analyze the subject's thinking pattern while applying this technique to solve a problem and whether the tangible SG kit can help subjects recognise the fundamental computational thinking approaches required in GD.

Two sets of rules are formed to perform this activity: the traditional SG rule to make geometric entities and the extension of SG called parametric SG (Angelo et al.,

2013) to make or play with the layout. Below are the rules for performing the SG activity –

Rule to Make Display Structure:

× - **Marker**

This rule is used to make display structures with the help of clay and matchsticks. There are two constraints that the designer should keep in mind when designing display structures:

- The angle between any two panels should be between 90 and 180 degrees.
- The number of panels should be less than or equal to five to handle structural rigidity.

Based on these constraints, a designer can make any number of display structures and name them entity 1, entity 2, entity 3 and so on. These entities will further help in parametric design rules. Figure 8 shows a few display structures made by subjects according to the above rule –

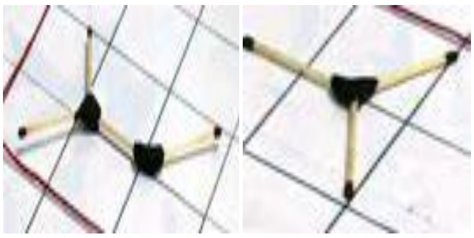
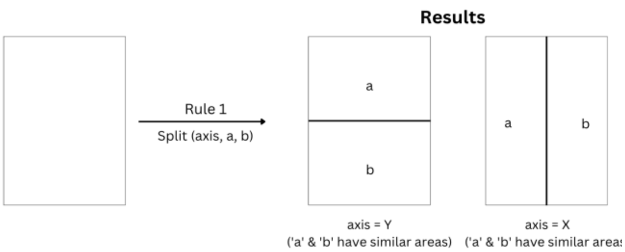


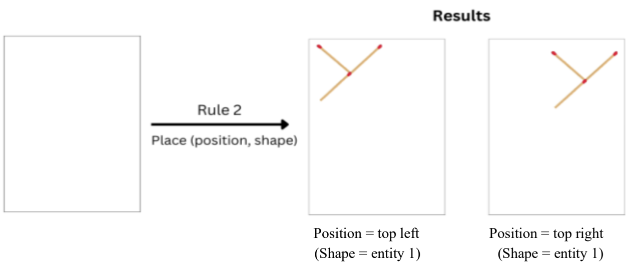
Fig 8. Display structures made by the subject

Rule to Design Exhibition Layout

A 10x10 grid with proper scaling and thread will be provided to the designer. The two parametric rules which will be used for the layout design are mentioned below-

Rule 1 –





- Entity1

Figure 9 shows an example of the layout created by the subjects with the help of the given rule and objects-

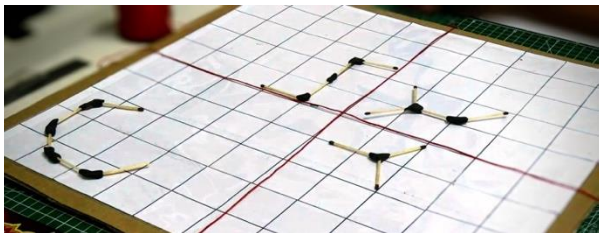


Fig. 9: Complete layout obtained using the SG rules

Table 1. summarizes the experimental setup.

S. No.	**Experiment (E) / Activities (A)**	**Timeline (minutes)**		
		Sub experiment A- FTD	**Sub experiment B – CAT**	**Sub experiment B – SGT**
A)	Exp Handout	COMMON (5 MIN)		
B)	Video Briefing	✓ 2MIN	✓ 15MIN	✓ 5 MIN
C)	Doubt Session	✓ 5 MIN	✓ 5 MIN	✓ 5 MIN
D)	Ref Manual	×	✓ 5 MIN	✓ 5 MIN
E)	Exp Activity	✓ 20 MIN	✓ 20 MIN	✓ 20 MIN
F)	Feedback	COMMON (10 MIN)		
G)	Questionnaires	COMMON (5 MIN)		

Result and Discussion

Four major themes emerged through thematic analysis of the transcribed data (Table 2). A detailed discussion of the themes is presented below –

User/Visitor Experience

This theme represents the importance of user flow in exhibition space layout. User flow is an essential aspect of spatial layout problems. Four out of six subjects spoke about the walking space arrangement of the panel, the field of view and the angle between panels (matchsticks), etc., to improve the visitor experience. For example, the experience of one subject is –

"*Even if it is rigid 90 degrees, the corner space will not be well utilized. Here it provides a larger space for the people to have a standard view, and even the angle of view will be better. So, this will be straightforward, and there will be the utilization of space. If it is rigid 90 degrees, the corner will not be very convenient for the flow.*"

– (Subject 4)

Here, the subject is describing the user flow along with the angle between panels when placing the display structure (entities in the case of SG) on free space.

Subjects' Perception of the Problems

Although each activity involved enough supporting media and small doubt clarification sessions, each subject perceived the situation differently. The response of one subject is –

“Just want to clarify? So we may have to make four zones, right? Similarly spatial capacity!. So what do you mean by similar like? Is it like the number of displays should be the same for this one?”

– (Subject 1)

Response to GD Techniques

This theme includes and represents the intuitiveness while performing the GD technique. Almost all subjects spoke about the ease of applying CA and SG, the ease/difficulty of following the rules, constraints, and the freedom while designing the layout. How the computational thinking approach is recognised while applying the GD technique is also a part of this theme. Below is an example of the interview session –

“Interviewer - which technique did you find easier to perform and why?

Subject 2 - If we talk in terms of performing, the last exercise was much easier because we have more options to do, I mean more things to do and with fewer complications, so we can just very freely work.”

This extract talked about intuitiveness in performing the GD technique. The subject found SG (the last exercise) easy to perform due to the freedom of design.

How the Kit Relates to Real-Life Problems

The tangible object and chosen problem statement make the whole experience like a real-life problem. Every subject visualized and tried to make the experience like designing an exhibition layout based on their understanding. Here is the experience of visualization of one subject while creating the layout plan –

“So this one is for having different and hear another theme. And similarly, this one like combines kind of these two, so one inclusive space kind and then have different themes on the outside”

– (Subject3)

Using codes for thematic analysis can help ensure consistency and rigour in the analysis process. By establishing a clear set of codes that define the themes of interest, the analysis can be more objective and less prone to bias. Combining codes to create broader themes can help synthesize the data and provide a more holistic understanding of the subject's thought processes.

A qualitative survey was collected from each subject based on the Likert scale to understand the subjects’ experience and whether such a simple tangible exercise can generate confidence and interest in subjects’ toward GD. The result of the study is presented in a graph (fig.10) below. The question is divided into three categories: FT design, CA, and SG. The responses of all six subjects based on the Likert scale (-2 to 2) are collected in three categories.

Table 2. Major Themes, along with a few associated Codes

Theme 1– User/Visitor Experience	**Theme 2–** Subjects Perception of the Problems.	**Theme 3-** Response to GD Techniques	**Theme 4-** How the Kit relates to Real Life Problem
Codes	**Codes**	**Codes**	**Codes**
Space for walking	Exploration of design space	Understanding of CA and SG rules	Geometry specific knowledge
Panel arrangement	Reinforcing problem objectives	Intuitiveness to perform	The relative size of the panel
Gallery theme experience	Doubt clarification	Design Freedom within and between techniques	Available zone space
Field of view	Constraint understanding	Computational thinking	Spatial Capacity
Angle arrangement between panels			

The incorporation of tangible kits into GD education has demonstrated positive outcomes in enhancing learning and comprehension. Subjects responded favourably to the tangible kit's effectiveness in teaching GD techniques, leading to improved understanding and comfort in applying each method. The kit facilitated the application of structured thinking, enabling subjects to explore a broader solution space—an essential aspect of GD. This hands-on approach allowed subjects to grasp computational thinking, offering insights not as apparent through traditional teaching methods.

However, challenges surfaced when subjects engaged in free-thinking activities, where the absence of constraints and ambiguity in design tasks led to negative responses. This suggests that open-ended design tasks requiring high levels of creativity and dealing with ambiguity can pose difficulties for some subjects. Despite these challenges, subjects did not express negativity regarding the comfort of performing each technique, highlighting the engaging and user-friendly nature of the designed experiment kit.

The study's findings emphasize subjects' positive responses to the tangible kit and its associated activities, particularly for understanding and comprehension. The contrast between activities with well-defined rules and constraints and the free-thinking task indicates subjects' potential struggles in open-ended design scenarios, suggesting the need for additional support to develop problem-solving skills in such contexts.

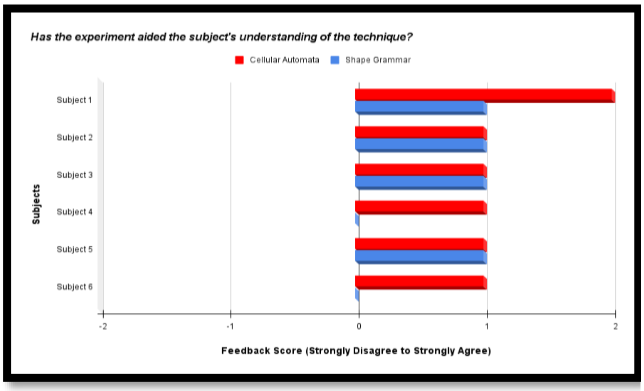
Subjects exhibited positive responses towards all the techniques in the kit, indicating their overall engagement and ease of use. This suggests that tangible exercises can be effective in teaching GD techniques and enhancing student engagement. Thematic analysis and code combinations were employed to delve deeper into subjects' thought processes, providing valuable insights for future design education practices.

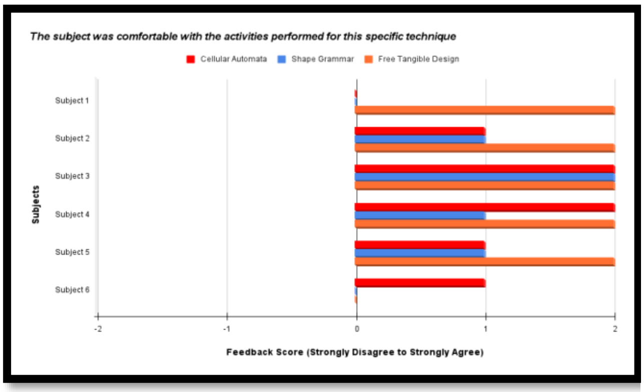
The study's methodology, including subject selection and activity time allocation, was carefully considered based on a pilot study. The pilot aimed to refine tangible exercises, ensuring feasibility and effectiveness, addressing potential issues, and guiding adjustments to the experiment procedures. While the sample size may be perceived as small, it aligns with the study's focus on exploring tangible exercise effectiveness. The comprehensive exploration offered valuable insights for educators and designers considering hands-on learning in generative design teaching.

The inclusion of diverse subjects with varying design education levels and backgrounds enriched the study's perspective. The multi-session approach contributed to an in-depth understanding of subjects' experiences, enhancing the study's qualitative analysis.

This research's significance lies in its potential to elevate GD education. The study offers insights into improved student engagement with complex design concepts by assessing the effectiveness of tangible exercises in teaching CA and SG. Addressing a literature gap, it specifically focuses on tangible exercises for teaching GD, potentially influencing GD education and tools development.

In conclusion, this research advances GD practices and underscores the transformative impact of hands-on learning in design education.





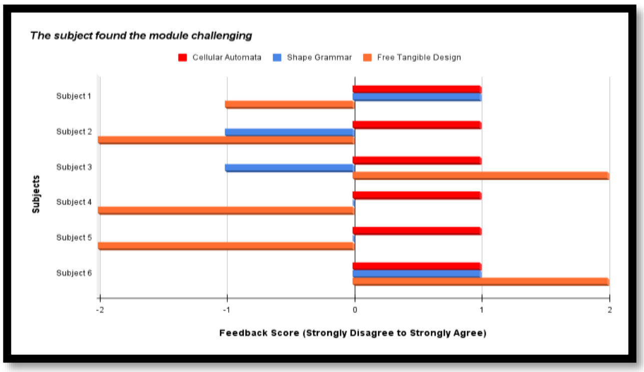


Fig.10 Results of subjects' responses based on the Likert data

Conclusion and future work –

Generative Design (GD) systems in architecture often employ various techniques, with two notable ones being Cellular Automata (CA) and Shape Grammar (SG). This study introduces tangible experiments using kits to enhance students' understanding of computational thinking in GD. The experiment aims to reveal cognitive processes and patterns in students' problem-solving approaches. Thematic analysis unveils insights into domain-specific knowledge and its role in design exploration. The findings underscore the importance of tangible exercises in fostering confidence and interest in GD, particularly in developing divergent thinking skills. The study acknowledges limitations, such as challenges in the free-thinking category, prompting the need for refined teaching methods balancing freedom and structure and the need for further experiments with more subjects having diverse backgrounds, which cover more diversity of student experiences and learning outcomes. Valuable insights into the effectiveness of tangible exercises using CA and SG emerge, suggesting potential extensions with additional GD techniques like L-system and Genetic Algorithm. Future work may include quantitative assessments of the experimental kit's efficacy and more detailed transcription analyses to identify computational thinking support. The future work could involve developing a taxonomy of knowledge based on the understanding of designers' and students' thought processes, identifying fundamental knowledge crucial in the generative design process. This research contributes significantly to design education, paving the way for future developments and exploration in the field.

References

Bonabeau, E., Dorigo, M. and Theraulaz, G., 1999. *Swarm intelligence: from natural to artificial systems*. Oxford university press.

Boone Jr, H.N. and Boone, D.A., 2012. Analyzing likert data. *The Journal of extension*, *50*(2), p.48.

Braun, V. and Clarke, V., 2006. Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), pp.77-101.

Braun, V., Clarke, V. and Weate, P., 2016. Using thematic analysis in sport and exercise research. In *Routledge handbook of qualitative research in sport and exercise* (pp. 213-227). Routledge.

Cassim, F., 2013. Hands on, hearts on, minds on: design thinking within an education context. *International Journal of Art & Design Education*, *32*(2), pp.190-202.

Cobb, P., Confrey, J., DiSessa, A., Lehrer, R. and Schauble, L., 2003. Design experiments in educational research. *Educational researcher*, *32*(1), pp.9-13.

Di Angelo, M., Ferschin, P. and Paskaleva, G., 2013, June. Shape grammars for architectural heritage. International Conference on Architecture and Urban Design.

Eris, Ö. (2003). Asking generative design questions: a fundamental cognitive mechanism in design thinking. In *DS 31: Proceedings of ICED 03, the 14th International Conference on Engineering Design, Stockholm* (pp. 587-588).

Fasoulaki, E., 2008. *Integrated design: A generative multi-performative design approach* (Doctoral dissertation, Massachusetts Institute of Technology).

Fischer, T. and Herr, C.M., 2001, December. Teaching generative design. In *Proceedings of the 4th Conference on Generative Art* (pp. 147-160). Politechnico di Milano University Milan.

Fonteyn, M.E., Kuipers, B. and Grobe, S.J., 1993. A description of think aloud method and protocol analysis. *Qualitative health research*, *3*(4), pp.430-441.

Goldschmidt, G., 1994. On visual design thinking: the vis kids of architecture. *Design studies*, *15*(2), pp.158-174.

Herr, C.M. and Ford, R.C. (2016), “Cellular automata in architectural design: From generic systems to specific design tools”, *Automation in Construction*, Elsevier BV, Vol. 72, pp. 39–45.

Herr, C.M. and Kvan, T., 2007. Adapting cellular automata to support the architectural design process. *Automation in Construction*, *16*(1), pp.61-69.

Jo, J.H. and Gero, J.S., 1998. Space layout planning using an evolutionary approach. *Artificial intelligence in Engineering*, *12*(3), pp.149-162.

Junginger, S., 2007. Learning to design: giving purpose to heart, hand and mind. *Journal of Business Strategy*, *28*(4), pp.59-65.

Khan, S. and Awan, M.J., 2018. A generative design technique for exploring shape variations. *Advanced Engineering Informatics*, *38*, pp.712-724.

Khan, S., Gunpinar, E. and Sener, B., 2019. GenYacht: An interactive generative design system for computer-aided yacht hull design. *Ocean Engineering*, *191*, p.106462.

Knight, T., 1999, April. Applications in architectural design, and education and practice. In *NSF/MIT Workshop on Shape Computation* (Vol. 67).

Krawczyk, R.J., 2002. Experiments in architectural form generation using cellular automata. *eCAADe20. Education in Computer Aided Architectural Design in Europe. K. Koszewski, S. Wrona (Ed.), Warsaw, Poland*, pp.552-555.

Krish, S., 2011. A practical generative design method. *Computer-Aided Design*, *43*(1), pp.88-100.

Lee, S.J., Jung, A., Park, J. and Yun, M., 2020, August. Cost-efficient hands-on learning design for computer organization course. In *2020 15th International Conference on Computer Science & Education (ICCSE)* (pp. 150-155). IEEE.

Lin, S.Y., Chien, S.Y., Hsiao, C.L., Hsia, C.H. and Chao, K.M., 2020. Enhancing computational thinking capability of preschool children by game-based smart toys. *Electronic Commerce Research and Applications*, *44*, p.101011.

Lindenmayer, A., 1968. Mathematical models for cellular interactions in development I. Filaments with one-sided inputs. *Journal of theoretical biology*, *18*(3), pp.280-299.

Maguire, M. and Delahunt, B., 2017. Doing a thematic analysis: A practical, step-by-step guide for learning and teaching scholars. *All Ireland journal of higher education*, *9*(3).

Marshall, P., 2007, February. Do tangible interfaces enhance learning?. In *Proceedings of the 1st international conference on Tangible and embedded interaction* (pp. 163-170).

Matthee, M. and Turpin, M., 2019. Teaching critical thinking, problem solving, and design thinking: Preparing IS students for the future. *Journal of Information Systems Education*, *30*(4), pp.242-252.

McNerney, T.S., 1999. *Tangible programming bricks: An approach to making programming accessible to everyone*(Doctoral dissertation, Massachusetts Institute of Technology).

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications, inc.

Poon, J. and Maher, M.L., 1997. Co-evolution and emergence in design. *Artificial Intelligence in Engineering*, *11*(3), pp.319-327.

Pusca, D., Bowers, R.J. and Northwood, D.O., 2017. Hands-on experiences in engineering classes: the need, the implementation and the results. *World Trans. on Engng. and Technol. Educ*, *15*(1), pp.12-18.

Singh, V. and Gu, N., 2012. Towards an integrated generative design framework. *Design studies*, *33*(2), pp.185-207.

Soleimani, A., Green, K.E., Herro, D. and Walker, I.D., 2016, June. A tangible, story-construction process employing spatial, computational-thinking. In *Proceedings of the the 15th international conference on interaction design and children*(pp. 157-166).

Stiny, G. and Mitchell, W.J., 1978. The palladian grammar. *Environment and planning B: Planning and design*, *5*(1), pp.5-18.

Stiny, G. and Mitchell, W.J., 1980. The grammar of paradise: on the generation of Mughul gardens. *Environment and planning B: Planning and design*, *7*(2), pp.209-226.

Stiny, G., 1980. Introduction to shape and shape grammars. *Environment and planning B: planning and design*, *7*(3), pp.343-351.

Tapia, M., 1999. A visual implementation of a shape grammar system. *Environment and Planning B: Planning and Design*, *26*(1), pp.59-73.

Tilley, S.A. and Powick, K.D., 2002. Distanced data: Transcribing other people's research tapes. *Canadian Journal of Education/Revue canadienne de l'éducation*, pp.291-310.

Toyong, N.M.P., Abidin, S.Z. and Anwar, R., 2020. Design Focus Group as a Controlled-Experiment Setting. *Environment-Behaviour Proceedings Journal*, *5*(SI3), pp.79-85.

Von Neumann, J., 2017. The general and logical theory of automata. In *Systems research for behavioral science* (pp. 97-107). Routledge.

Wang, D., Wang, T. and Liu, Z., 2014. A tangible programming tool for children to cultivate computational thinking. *The Scientific World Journal*, *2014*.

Wolfram, S. and Gad-el-Hak, M., 2003. A new kind of science. *Appl. Mech. Rev.*, *56*(2), pp.B18-B19.

Zaitsev, D.A., 2017. A generalized neighborhood for cellular automata. *Theoretical Computer Science*, *666*, pp.21-35.

AN APPROACH TO THE ESTIMATION OF DIGITAL TWINS TECHNOLOGICAL COMPONENTS

Giovanni Bortolin[1], Nicola Moretti[2], Karim Farghaly[2], and Weiwei Chen[2]
[1]Savile Brown Associates, London, United Kingdom
[2]Bartlett School of Sustainable Construction, University College London, London, UK

Abstract

The Digital Twin (DT) concept in the Architecture Engineering Construction and Operation (AECO) sector is gaining momentum. Various studies and case examples highlight its benefits. However, for broader implementation within the sector, there is a need for a standardised method for estimating the cost of DTs. This paper proposes an extension of the UK New Rules of Measurement 2 (NRM2) for estimating the technological components of DT systems based on the desired capabilities. The DT capabilities are identified through the analysis of the theoretical framework, which, combined with the analysis of current industry standards, enables the development of the NRM2 extension. The need for this extension has also been validated through a set of semi-structured interviews with industry experts in the UK. This research contributes to reducing ambiguity in the identification and estimation of technological components of DTs in buildings.

Introduction and background

The benefits of the digitalisation of the construction sector (still poorly digitised) are many and include addressing major challenges such as carbon reduction, resources and energy efficiency, labour shortage and competitiveness (European construction sector observatory, 2021). The digitalisation trend has been accelerated in recent years by the introduction of the DT concept, which can be broadly defined as "*A digital twin is a virtual model of an object, a system, or a process. It is connected to its real-world counterpart by a 2-way flow of right-time data, meaning it mimics it in all aspects. This helps us test decisions before we make them and understand how different actions might affect the real world.*" (UK Department for Business & Trade, n.d.). Though it is known that the DT concept has been investigated in several other sectors as aerospace and manufacturing (Grieves, 2016), there is no agreement on the components that should be included in a DT implementation (Shahzad *et al.*, 2022), thus it is hard to drive the digital transformation in the construction industry, with nothing more than a concept and a number of few custom applications.

In fact, the digital transformation in Architecture, Construction, Engineering and Operations (AECO) is not confined to the economic cost for the deployment of a new set of technologies. In the case of DTs, these costs can include for examples the change management activities required at the ecosystem level, training and workforce upskill, security, together with the hardware and software costs.

In the construction industry, innovation dynamics and therefore also the digital innovation, are often realised through projects (Xue *et al.*, 2014) and is well known that the success of a project is related to the careful estimation of the time, cost and quality of the desired outcomes (Association for Project Management, 2019). Within this context and with a focus on the economic dimension of projects, a careful cost estimation is key to understand in depth the required efforts to deliver the desired outputs and realise the client requirements and objectives. This is crucial, especially in the pre-construction phase in order to avoid cost overruns which could deviate completely the project scope.

The Bill of Quantities (BoQ) techniques are fundamental tools in the construction sector, used to control the project costs. To develop the BoQ it is essential to involve methods of estimations which respects standards in the way that different stakeholders will be able to understand it. One of the widely acknowledged standards is the NRM2 from the Royal Institution of Chartered Surveyors (RICS, 2021), which is used for detailed measurements of building works.

This paper proposes a method to estimate the technological components of DTs, required to deliver a finite set of general capabilities, identified from existing DT implementations. The method of estimation of DT in this paper is considered as a case of digital innovation in AECO which has not been completely explored and therefore requires further investigation. The paper also proposes a research roadmap on the study of the cost associated with the digital transformation in AECO.

The rest of this paper is organised as follows: the method section identifies the research methodology and tools used, the analysis of the theoretical framework analyse and identifies capabilities and technologies for DT systems, the NRM2 extension section showcases the developed standard for estimation and the preliminary validation through expert interviews, the discussion and conclusion refine what has been done against the identified research gap, and inform on the research limitations and gives direction on future steps.

Method and tools

The methodology used for the study is related to the design science research methodology (Peffers *et al.*, 2007), which is used for information systems as this allows the design of a method for the estimation for DT systems. The research methodology consists of six activities which are developed in sequence, the first activity, entails the identification of the problem statement, the second activity relates to the definition of the objective for a solution, the third consists in the design and development, the fourth is the demonstration, the fifth is for the evaluation, and the sixth consists in the communication.

Of the above-mentioned activities, the research focuses on the first three activities with the latter three and an outline of the next steps being discussed in the discussion and conclusion section.

The first activity is presented in the analysis of the theoretical framework, and results with the need for the development of a method of estimation for DT systems, which is needed for solving the identified research gap.

The second activity is identified through an understanding of existing standards of estimation and the relative implementation of estimation of DT systems, as currently, the research did not identify standards of estimation, this would accomplish a better identification of the DT systems in the phase of design and their related capabilities, and better define the DT requirements needed in the project.

The third activity includes the design and development stage, which is centred towards a developed method of estimation for DT systems for buildings, this is developed in the analysis chapter through the design of a method of estimation with the identified DT technologies capabilities.

Furthermore, a validation of the proposed method of estimation was undertaken through structured interviews in order to understand the experiences of the various construction experts involved in digital construction and in specific DTs.

NRM 2 to estimate DT systems

There are different guidance and standards used for estimation, and one of those standards, which is accepted internationally, is the New Rules of Measurement (NRM) collection, formed by NRM1, NRM2 and NRM3 and utilised as a guidance for estimating the building works throughout the lifecycle.

Furthermore, NRM can be taken into account in legal proceedings (RICS, 2021) and can be used to produce BoQs which follow the standard methods of estimation and can be attached as a contractual document in standard form of contracts.

The NRM1 is used for initial cost planning, the NRM3 is used for the estimation of building maintenance works, and the NRM2 is a standard for the estimation of building works during the pre-phase of construction and can also be used as a contract document which allows to obtain a total price for completing building works. This allows to evaluate the project variations and to support the management of the cash flow, implementing value engineering tasks. The NRM2 can be used as a tool to extend methods of estimation of DT, since it contains a method of measurement for quantifying the related systems and for the exclusion of the components already quantified using other parts of the standard (RICS 2021). Therefore, the proposed NRM2 extension aims at linking the DT capabilities with the technological implementation of both the hardware and software parts of DTs in buildings.

Analysis of the theoretical framework

To define the current costs associated with DT systems and their related technological components, in this section, we identify the current approaches that deliver the capabilities of systems, as described in the ISO 23247 series (ISO, 2021a). This standard provides a guidance for the development of a DT reference architecture and its correlation to the functional entities. Likewise, the DT application in building systems, assets and whole infrastructure involves the use of different capabilities that are summarised in Table 1. It has to be noted that the application level does not cover the use of DT systems for the benefit of the construction site.

The data acquisition capability (ISO, 2021b) is the functional entity for data acquisition and might also contain an Internet of Things (IoT) network, which links sensors, actuators and the collected data, supporting automation. Furthermore, data can be retrieved from historical building systems, as for example the Building Automation Systems (BAS) and Building Management Systems (BMS) (Lu et al., 2020).

The data modelling reflects the need to create a digital representation of the building and its features. It includes different data formats and their requirements (ISO, 2021c), which can enable data, process and system interoperability. For example, in Xu *et al.* (2023), the Industry Foundation Classes (IFC) standard and the semantic information in JSON file format are combined and converted into the 3D tiles standard for its interoperability.

The data analytics capability allows to gain insight or prediction from the data provided in the digital representation, as for example, in the case of Hosamo *et al.*, (2022), who describe the use of "the algorithms Artificial Neural Network (ANN), support vector machine (SVM), and decision trees to predict severe Air Handling Units (AHU) faults".

The actuation capability allows the use of the outcomes of the data analytics to automate processes and plan maintenance. This capability can be enabled through advanced functionalities of the IoT network which can enhance applications, as for instance the control of the Heating Ventilation and Air Conditioning (HVAC) system based on room occupancy as outlined in García-Monge *et al.* (2023). The curation of the DT system involves the maintenance of the whole system, this is done through the creation of appropriate data storage and its maintenance.

Table 1: Case studies and related capabilities

Application level	Purpose	Data Acquisition	Data modelling	Data Analytics	Actuation	Curation	Author
Asset	Human robot collaboration in Facility Management	Robot sensing data, asset data capture.	RDF and entity relationship diagram	Control flow for anomaly detection	Physical actuators and robot motion	Cloud based server, DB, web platform and BIM for visualisation	(Lu *et al.*, 2023)
Asset, system, building	Real-time monitoring of the hospital facility	Data on facility users, data on facility and processes	Web service, JSON Schema, Connection adapters, MOM	OLAP,HDFS, OLTP	Process control, defect detection, alarm rate of monitored points	BIM for visualisation, WebGL and JSON DB	(Han *et al.*, 2023)
System	Predictive Maintenance AHU	IoT sensor network, environmental sensing, and system monitoring	Brick ontology based on COBie data and IFC	ML, ANN, SVM and decision trees	Planning of maintenance	Sensor data DB and BIM for visualisation	(Hosamo *et al.*, 2022)
System	Anomaly detection of centrifugal pumps	Tags, environmental sensing, system monitoring, historic building data	Multi-source data through IFC and COBie	CUSUM and BOCPD	Anomaly detection and reduction of false alarms	BIM of the building and web based DT platform	(Lu *et al.*, 2020)
System	Fault detection and diagnosis of HVAC	Environmental sensing, and system monitoring	JSON file brick model with fault tags to form a hierarchical structure	Goodness-of-Fit test with Kullback–Leibler divergence	Unsupervised and supervised fault detection	Adaptive City Platform digital twin platform, which integrate brick schema	(Xie *et al.*, 2023)
System	Fan coils units fault detection for HVAC	Environmental sensing, system monitoring, IoT sensor network	Integrated data using BIM platform and NPM modules	Integrated anomaly detection flow chart	Fault detection	BIM and web based program	(Villa *et al.*, 2021)
Systems, spaces	Energy optimisation of HVAC	Environmental sensing, system monitoring, tags, data on facility users, IoT network	Integration of data through proprietary developed application	Quantitative data analysis	Decision through data analysis to improve efficiency	Web user interface	(García-Monge *et al.*, 2023)
Spaces	Monitoring of room occupancy	Image recognition and historical data	SQL database coupled with BIM model	Defined thresholds	Decision-making through a cost model	SQL Cloud server and BIM for visualisation	(Mannino *et al.*, 2019)
Building	Structural Health monitoring of building	Structural sensing, environmental sensing, computer vision	BIM with IFC and JSON for data exchange and 3D tile standard	Dynamic and static mechanical analysis	Building Structural health monitoring	SQL server to store data and 3D engine for visualisation	(Xu *et al.*, 2023)
Building, system, spaces	Indoor environmental quality, energy consumption, estimation of unobservable elements	Environmental sensing, systems monitoring, structural sensing	MATLAB backend software for data exchange between local software and the simulated building	CFD and ODEs for building zone dynamics	Detection of air contamination and test of fault diagnosis to improve protocols	Desktop application, cloud platform for remote access and as is BIM model	(Hadjidemetriou *et al.*, 2023)

This also involves the use of security systems for the collected and stored data, for example in a SQL database and a Building Information Modelling (BIM) model as showcased in Mannino *et al.* (2019).

Table 1 summarises not only the applications of the DT concept to several assets levels (i.e., assets, systems, spaces, and buildings), but also allows to identify the technologies implemented for each individual capability in the analysed DT case studies. This forms the knowledge basis used for the development of the NRM2 extension. The identified technologies are in fact generalised and indicated as possible cost which must be considered in the design phase, for the estimation of built asset DTs.

The NRM2 extension

The developed method of estimation is outlined in Figure 1 which showcases the different types of components needed for estimation and avails of the current format of the NRM2 estimation standard which also includes a section including items already incorporated in the NRM2. The estimation method designed includes sections for components which are identified through Table 1. The multiple technologies needed for each capability are differentiated between hardware components and software components. The costs related to the hardware technological components is estimated as their prime costs as the supply rate for goods (RICS, 2021). On the other hand, for software technological components, the cost is related to time-related variables.

Data acquisition

The data acquisition includes within sub-component 1.1 the cost of installation of the data acquisition source and its associated prime costs, these data acquisition sources can be and are not limited to sensors of different kinds and QR Codes tags. These sensing technologies should be considered in addition to the BMS/BAS parts and are typical of DT implementations (e.g., occupancy monitoring sensors, air quality, indoor comfort etc.)

As this requires the installation of a physical element the quantities will be enumerated (nr) and for systems which could include multiple data acquisition sources, the quantities might be itemised (item). These data acquisition technologies have all different costs associated with their installation and require separate analysis.

A time related charge is allowed for the process related to its installation and the duration that the data acquisition source will have to stay in position during the construction phase which could include the protection to other type of works undertaken on site.

Furthermore, by following the process of installation of the data acquisition technologies, there is the need for installation and temporary maintenance of communication technologies. Thus, the sub-component 1.2 in the data acquisition section could require batteries or cabling - such as an ethernet cable as seen in the ISO 23247-:2021(ISO, 2021c) - directly to the data acquisition source, hence quantities might be itemised for systems or numbered for single priced components with a fixed charge or time-related charge. Furthermore, the NRM2 extension allows the possibility of integration of historical building data (sub-component 1.3), as outlined previously. For example, through BMS/BAS data, quantification can be itemised and have a fixed charge which would describe the type of data and the related cost for provision.

Data engineering

The data engineering includes the capabilities identified in the data modelling and the curation detected in the analysis of theoretical framework. The sub-component 2.1 includes the preparation of a digital model for the target assets processes and their interrelations. For this purpose, the BIM approaches can be used as well as other Modelling and Simulation (M&S) methods which should be identified and quantified as items. BIM, in fact, is an information management technique that follows the project life cycle. For this reason, a complete data management platform needs to be developed and costed, in support of the DT functionalities which need to be curated along the DT lifecycle. The included sub-component 2.2 includes the data storage and its related maintenance. Sub-component 2.3 regards the implementation of security systems and their quantification is itemised. Sub-component 2.4 regards the data access and interoperability which reflects the need for producing and using structured information in DTs. This allows systems, processes, and data interoperability. Since multiple methods could be used, quantities need to be itemised with each system separately, to identify the associated costs and pricing method (e.g., fixed charge or a time-related charge).

Data analytics

As there are multiple approaches for data analytics which could be developed, with different systems and which needs distinct correlated works, each individual data analytics techniques might identified (as outlined in Table 1). The data analytics method should be stated, and the quantities separately itemised and quantified to identify the associated costs and pricing method (e.g., fixed charge or a time-related charge).

Actuation

The actuation includes the sub-component 4.1: the installation of actuators of the building systems which facilitate the building control and automation of the spaces and equipment. These components should be considered in addition to the BMS/BAS. Furthermore, the included sub-component 4.2 will have to include the installation and temporary maintenance of communication technologies, which could require installations with batteries or cabling such as ethernet cables as in examples outlined in the ISO23247:2021(ISO, 2021c).

Digital Twin System

Component	Included (Sub-Component)	Unit	Pricing method	Excluded
1 Data acquisition.	The type of data collection source to be provided should be stated, with each type separately quantified: **1.1** Installation charge of data collection source.	item/nr	**1** Fixed charge **2** Time-related charge	**1** Operation and maintenance manuals (included in B.12 Contractor's cost items: completion and post-completion requirements). **2** Preparation of handover plans (included in B.13 Contractor's cost items: completion and post-completion requirements). **3** Inspection (included in B.13 Contractor's cost items: completion and post-completion requirements). **4** Training of building user's staff in the operation and maintenance of the building engineering services systems (included in B.13 Contractor's cost items: completion and post-completion requirements). **5** Information model and information management for BIM (included in A.9 Employer's requirements: building information modelling (BIM)). **6** Mechanical services system (included in Work section 38: Mechanical services). **7** Electrical services system (included in Work section 39: Electrical services).
	1.2 Installation and temporary maintenance of communication technologies.			
	1.3 Provision of historical building data.			
2 Data engineering.	The type of digital representation should be stated and if the representation avails of a BIM model: **2.1** Preparation of digital asset/process representation.	item	**1** Fixed charge **2** Time-related charge	
	2.2 Data storage and maintenance.			
	2.3 Implementation of data security system.			
	The method used should be stated with each method separately quantified: **2.4** Data access and interoperability.			
3 Data analytics.	The method used should be stated with each method separately quantified: **3.1** Work related to data analytics with supplied data and prediction algorithms.		**1** Fixed charge **2** Time-related charge	
4 Actuation.	The type of actuator to be provided should be stated, with each type separately quantified: **4.1** Installation charge of actuators.	item/nr	**1** Fixed charge **2** Time-related charge	
	4.2 Installation and temporary maintenance of communication technologies.			

Figure 1: Proposed method of estimation for DT systems based on the NRM2 structure.

Excluded components

Other tables and existing components already included in the NRM2, which may overlap, are excluded for clarity purposes. These are the operation and maintenance manuals (included in B.12 contractor's cost items: completion and post-completion requirements), the preparation of handover plans (included in B.13 contractor's cost items: completion and post-completion requirements), the inspection (included in B.13 contractor's cost items: completion and post-completion requirements) and the training of building user's staff in the operation and maintenance of the building engineering services systems (included in B.13 contractor's cost items: completion and post-completion requirements), the information model and information management for BIM (included in A.9 employer's requirements: building information modelling - BIM), the mechanical services system (included in Work section 38: Mechanical services), the electrical services system (included in Work section 39: Electrical services).

Validation of DT Systems designed method of estimation

The next stage of the research concerns the validation of the proposed NRM2 extension. For achieving this, a series of interviews with four experts was carried out. The participants were high-profile professionals in the fields of digital construction and have been selected based on their standing in the UK industry.

A total of eight open questions were asked in semi-structured interviews and are summarised in Table 2.

Table 2: Interview questions.

Question	
1	What is your current profession and your expertise in the AEC industry?
2	How would you define your experience in DTs generally? Also, do you have any experience in estimation for building works?
3	Do you think that DTs are beneficial?
4	What is your current opinion about DTs system for operation and maintenance?
5	Do you think the current proposed section can be used in your profession and why?
6	Can you suggest any improvements to the proposed new section of the NRM2?
7	Would you adopt this section for future implementations of DTs?
8	Do you think the current section could help other professionals in future implementation of the DT systems in the built environment?

The experts' profiles are summarised in Table 3.

Table 3: Summary of interviewed experts

Expert	Profession and expertise
P1	Planning and project management
P2	Works with a built environment professional body withing the standards team
P3	Chartered quantity surveyor working in academia
P4	Specialised in DTs and strategies

The analysis of the interviews is presented in the following section and focuses on the validation of the NRM2 extension while identifying possible further developments of the proposed work.

Validation and benefit of the NRM2 extension for DT systems

The experts P1, P2, P3 expressed the possibility of using the proposed NRM2 extension in their work. There are several reasons for that. For example, the interviewee P1 confirms that using a standard such as the proposed designed NRM2 extended method of estimation for DT systems allows a better understanding and breakdown of possible costs and this would increase the transparency of the DT design and development process.

P3 explains that the proposed approach gives the professionals involved in the cost process the possibility to better understand and quantify what should be covered in DT implementation. Furthermore, P2 explains that implementing standards for estimation for DT systems would be beneficial, and professionals should make sure that this is on their agenda when people are looking at creating new assets, so the DT components would be inserted in the programme and subsequentially costed from day one.

Possible Improvements of the NRM2 extension for DT systems

The experts expressed different suggestions for improving the NRM2 extension too. For example, one participant suggests to expand each sub-component, since there are many things needed to define certain aspects of the DTs. However, for effective pricing of these components it is understandable to have a broader description.

Other participants explains that the proposed extension has covered few or most grounds of the DT concept and this work should not be the end of the elaboration of the designed method and the proposed extension should be taken further and take into consideration DT systems generally and identify the parameters that can be measured. Also, the methods to effectively measure them should be studied.

Additionally, the NRM2 extension interlaps with methods of estimation already existing on BMS systems. However, since a DT system main value is to track data that is currently not tracked, the proposed extension should focus on the data interoperability aspect, the information and where this information needs to be stored as an essential part of the DT system. Finally, understanding the system architecture that needs to be implemented and how to link all the data from different sources should be taken into account, as well as focusing on the two essential aspects of DTs: fidelity and frequency.

Discussion and conclusions

The key objective of this research is to develop a method of estimation for DT for aiding the digital transformation process, and this sheds light on how digital technologies which fit within the DT concept can be effectively costed. In fact, the DT implementations play nowadays a central role in the construction industry digitalisation and, given the wide breadth of applications, represent a significant case of digitalisation in the construction industry.

To address the research gap, we looked into current standards and approaches utilised for DT systems implementation and through the analysis of the theoretical framework we identify some key capabilities: data acquisition, data modelling, data analysis, actuation and curation. These capabilities are analysed from the technological point of view and the enabling items to be estimated are identified. For this purpose, the NRM2 is used, and this aids the development of a method of estimation for the DT system during the phase of design which can be used for tendering. This research has been developed using a mixed evidence-based method and a literature analysis approach.

One of the main benefits is that the proposed method of estimation contributes to creating a standard for the estimation of the DT technological components, which streamlines the budgeting activities and improves the design and value engineering of the system.

Further studies should explore in more depth the DT capabilities and identify the detailed costs of DT systems proposed in the research within the NRM2 extension which includes the unit of measure and the applied pricing method. For example, the cost functions for each of the DT components have not been defined, as well as the total cost function. This would form part of the future development of this research.

Moreover, the last three steps of the design research method should be addressed, these include demonstration, evaluation and communication. Accordingly, the NRM2 extension should be implemented in a pilot design process, so that the results would be evaluated on a real testbed. The validation phase (an additional one to the preliminary validation carried out through interviews with the industry experts) can be carried out using a comparative evidence-based approach, assessing the effectiveness of the proposed method as opposed to the traditional estimation methodologies and through an additional round of interviews.

Also, further studies should explore the possibilities of life cycle costing for the implementation of the system and its financial/economic benefits throughout the life cycle of the building. This would contribute to a better understanding of the costs related to digital transformation.

References

Association for Project Management (2019) *APM body of knowledge*. Seventh edition. Princes Risborough, United Kingdom: Association for Project Management.

European construction sector observatory. (2021). Digitalisation in the Construction Sector. In *Digitalisation in the construction sector. Analytical Report* (Issue April). https://ec.europa.eu/docsroom/documents/45547/attachments/1/translations/en/renditions/native

García-Monge, M. *et al.* (2023) 'Is IoT monitoring key to improve building energy efficiency? Case study of a smart campus in Spain', *Energy and Buildings*, 285, p. 112882. Available at: https://doi.org/10.1016/j.enbuild.2023.112882.

Grieves, M. (2016) 'Origins of the Digital Twin Concept'. Available at: https://doi.org/10.13140/RG.2.2.26367.61609.

Hadjidemetriou, L. *et al.* (2023) 'A digital twin architecture for real-time and offline high granularity analysis in smart buildings', *Sustainable Cities and Society*, 98, p. 104795. Available at: https://doi.org/10.1016/j.scs.2023.104795.

Han, Y. *et al.* (2023) 'Digital twinning for smart hospital operations: Framework and proof of concept', *Technology in Society*, 74, p. 102317. Available at: https://doi.org/10.1016/j.techsoc.2023.102317.

Hosamo, H.H. *et al.* (2022) 'A Digital Twin predictive maintenance framework of air handling units based on automatic fault detection and diagnostics', *Energy and Buildings*, 261, p. 111988. Available at: https://doi.org/10.1016/j.enbuild.2022.111988.

ISO (2021a) 'Automation systems and integration — Digital twin framework for manufacturing, Part 1: Overview and general principles BS ISO 23247-1:2021'. International Organization for Standardization.

ISO (2021b) 'Automation systems and integration — Digital twin framework for manufacturing Part 2: Reference architecture BS ISO 23247-2:2021 Licenced'. International Organization for Standardization.

ISO (2021c) 'Automation systems and integration — Digital twin framework for manufacturing Part 4: Information exchange BS ISO 23247-4:2021 Licenced'. International Organization for Standardization.

Lu, Q. *et al.* (2020) 'Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance', *Automation in Construction*, 118, p. 103277. Available at: https://doi.org/10.1016/j.autcon.2020.103277.

Lu, W. *et al.* (2023) 'Digital twin-enabled human-robot collaborative teaming towards sustainable and healthy built environments', *Journal of Cleaner Production*, 412, p. 137412. Available at: https://doi.org/10.1016/j.jclepro.2023.137412.

Mannino, A. *et al.* (2019) 'Office building occupancy monitoring through image recognition sensors', *International Journal of Safety and Security Engineering*, 9(3), pp. 371–380. Available at: https://doi.org/10.2495/SAFE-V9-N4-371-380.

Peffers, K. *et al.* (2007) 'A Design Science Research Methodology for Information Systems Research', *Journal of Management Information Systems*, 24(3), pp. 45–77. Available at: https://doi.org/10.2753/MIS0742-1222240302.

RICS (2021) 'New Rules of Measurement 2'. RICS.

Shahzad, M. *et al.* (2022) 'Digital Twins in Built Environments: An Investigation of the Characteristics, Applications, and Challenges', *Buildings*, 12(2), p. 120. Available at: https://doi.org/10.3390/buildings12020120.

UK Department for Business & Trade. (n.d.). *What a digital twin is and how you can contribute.* Retrieved December 5, 2023, from https://www.gov.uk/government/publications/what-a-digital-twin-is-and-how-you-can-contribute/what-a-digital-twin-is-and-how-you-can-contribute

Villa, V. *et al.* (2021) 'IoT Open-Source Architecture for the Maintenance of Building Facilities', *Applied Sciences*, 11(12), p. 5374. Available at: https://doi.org/10.3390/app11125374.

Xie, X. *et al.* (2023) 'Digital twin enabled fault detection and diagnosis process for building HVAC systems', *Automation in Construction*, 146, p. 104695. Available at: https://doi.org/10.1016/j.autcon.2022.104695.

Xu, Jinghai *et al.* (2023) 'Developing a digital twin model for monitoring building structural health by combining a building information model and a real-scene 3D model', *Measurement*, 217, p. 112955. Available at: https://doi.org/10.1016/j.measurement.2023.112955.

Xue, X. *et al.* (2014) 'Innovation in Construction: A Critical Review and Future Research', *International Journal of Innovation Science*, 6(2), pp. 111–126. Available at: https://doi.org/10.1260/1757-2223.6.2.111.

Energy Modelling and Monitoring

ANALYZING THE ENVIROMENTAL SUSTAINABILITY OF SMART BUILDINGS

Sara Aouichaoui[1], Farina Koller[1] and Niels Bartels[1]
[1]University of Applied Sciences, Cologne, Germany

Abstract

Smart buildings are becoming increasingly important for the construction industry. They are associated with optimization potential for efficiency and therefore ecological sustainability. However, there are currently no studies that can scientifically prove this thesis. Through a comprehensive methodology involving literature reviews, analysis of smart buildings, and expert interviews, the study explores the complexities of achieving environmental sustainability in smart buildings. Emphasizing the importance of standardized definitions, data availability, and collaboration among stakeholders, the paper presents use cases and highlights their impact on the sustainability of a smart building. Furthermore, information for evaluating the environmental sustainability of smart buildings is defined.

Introduction

According to the European Copernicus Climate Change Service 2023 will be the hottest year in the last 125.000 years. Oktober 2023 even sets a record as the warmest month ever recorded, with a global average temperature 1.7 degrees above pre-industrial levels (Schmid, 2023). This underscores the urgency of adhering to the goals of the Paris Climate Agreement, not only in limiting global temperature rise but also in reducing emissions and adapting to climate change (Europäische Union, 2016), (Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, 2023). The building sector, responsible for 40 % of final energy consumption in Europe (United Nations Environment Programme, 2022), plays a crucial role. In Germany, 35 % of total energy consumption goes to buildings, contributing to 30 % of annual CO_2 emissions, with 75 % of a building's carbon emissions occurring during operation (SmartScore, 2021).

To address this, there is a need to reduce energy consumption in existing buildings and enhance the sustainability of new ones. Innovative construction methods and materials, such as 3D printing, wood, clay, and CO_2-reduced concrete, are being explored (Bäuerle & Lohmann, 2021), (Neef, et al., 2020), (Haist, et al., 2023). Simultaneously, the adoption of technological solutions, especially smart building applications, is increasing (Mordor Intelligence, 2023).

Despite the growing interest, there is a lack of comprehensive scientific studies on the ecological sustainability of smart buildings, particularly in terms of user and facility management practices. This paper aims to fill this gap by scientifically analyzing the environmental sustainability of smart buildings.

- Which applications and associated data are necessary for evaluating the ecological sustainability of a smart building?
- What data is available for use in a smart building that goes beyond a classic building atomization and what additional data is required?

This case study was carried out exclusively in the Smart office area, because smart homes have too individual a design, as well as other data protection difficulties. Furthermore, the smartness of other types of buildings has not yet been sufficiently further developed. It will assess optimization potential, consider stakeholders (project developers, owners, manufacturers, and facility managers), and scrutinize data flow within smart buildings, identifying potential interface problems and data losses. The paper concludes with recommendations on calculating the environmental sustainability of smart buildings based on practical experiences and expert insights.

Literature Review and related work

Definition of a smart building

A clear distinction between buildings with classic building automation and smart buildings was drawn up with the help of a literature review. The literature and practical applications lack a uniform definition of a smart building, attributed partly to differing perspectives among project participants throughout a building's life cycle (SmartScore, 2021), (Monterioa Froufe, et al., 2020). As outlined in literature developers emphasize user experience for increased property value, while facility management (FM) focuses on safety and energy savings (Qolomany, et al., 2019), (Li, et al., 2020), (Pašek & Sojková, 2018). To sustain a smart building efficiently and ecologically, there is a need to network stakeholders and establish a unified definition.

Current consensus, based on literature and project analysis, describes a smart building as one equipped with diverse technologies (actuators, sensors, microchips) for data collection, storage, and analysis (Bosch & Deckert, 2023), (Plageras, et al., 2018). According to the (Fraunhofer-Allianz Bau, 2022), smart buildings use automation to enhance physical properties through holistic control concepts and suitable technologies. They integrate existing building automation systems or use autonomous sensor technology installed solely for smart building solutions (Graveto, et al., 2022).

There exist different perspectives on smart building applications. (Bosch & Deckert, 2023) focus on automated, energy-efficient, and sustainable building networking that positively impacts user experience, climate protection, and environmental conservation. (Moretti, et al., 2020) emphasizes optimizing maintenance services in smart buildings.

A comprehensive definition emerges from these perspectives that is shown in Figure 1: A smart building is a digitally connected building with diverse technologies and systems. It aims to enhance user experience and positively impact the three pillars of sustainability. Smart buildings continuously collect real-time data through sensors of the facilities, systems and users. This data is transparently stored, processed, and evaluated, ideally in a cloud and made available to a building management platform. Considering data over the building's life cycle supports optimization, acting as a foundation for new building planning.

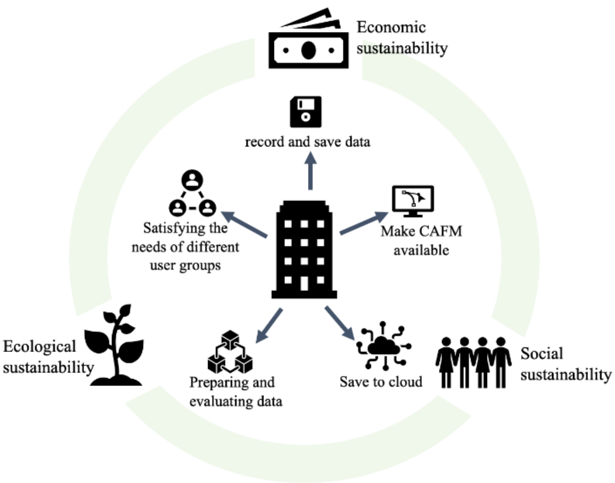


Figure 1: Definition of a smart building

Influence of smart buildings on environmental sustainability

Research in the field of smart building is becoming increasingly important, especially in relation to environmental sustainability. Figure 2 shows the number of scientific publications on the keywords "smart building" and "sustainability" and a continuous increase on Scopus and google scholar between 2018-2023.

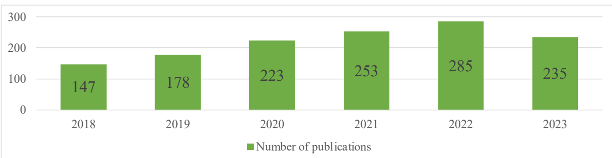


Figure 2: Publication in the area of smart buildings and environmental sustainability published in the years 2018-2023

With the Smart Readiness Indicator (SRI), the European Union provides a basis for assessing the readiness to implement smart components in buildings. However, this does not provide a basis for assessing the sustainability of the building (The European Commission, 2020). Rising energy costs and the need to meet the Paris climate targets (Koch, 2023) mean that companies want to reduce their consumption in office space through energy monitoring (Nulsch, 2023). ISO 52120-1:2021 deals with the influence of building automation and management on energy efficiency (DIN e.V., 2019). This shows a correspondence with the analyzed use cases of a smart building. The standard uses a factor-based method to estimate the influence of energy consumption based on the degree of automation of various functions. Improving the energy efficiency standard according to the standard can lead to significant savings, although previous standards have listed similar factors. In literature it is stated, that smart buildings could save 10 % - 25 % in energy (Fraunhofer-Allianz Bau, 2022), (Fraunhofer-Allianz Bau, 2023), (Jacob & Kukovec, 2022). Annex 81 addresses the fact that the potential of digitalization for the energy efficiency of buildings has not yet been fully exploited. among other things, the project aims to develop software applications with the help of available cost-efficient data in buildings. Semantic standards are defined and an open data platform is required (International Energy Agency, 2023).

Manufacturers promote smart buildings as a driver for environmental sustainability in the building sector, focusing on reduced energy consumption, cost reduction and increased comfort through intelligent optimization systems that adapt energy consumption to occupancy and use. More efficient use of space saves resources and maximizes stakeholders' profits (Siemens Schweiz AG, 2020). In order to evaluate the savings and optimizations of environmental sustainability by smart buildings, an analysis of 23 manufacturers have been conducted in addition to the literature review. The information was taken from product catalogs, white papers, proof of concepts or the companies' websites. Eight of 23 companies indicate potential savings, that vary from 10 % up to 40 %. Some providers regard the energy in general, others differentiate between heating and lighting energy. A few put the savings in heating costs or operating costs at between 30 % and 40 %. One claims to achieve area savings of 30 %. The other 15 companies do not provide any information on savings. The analysis showed 5 major problems in defining the environmental sustainability for smart buildings:

1. No information is provided on the initial values (e.g., planning or previous years of operation) or the basis for calculating the savings,
2. predominantly no information on whether the savings relate to thermal or electrical energy,
3. the number given cannot be substantiated, as generalized statements about CO_2 savings are often made or an increase in the users' sense of comfort is suggested, as well as falling personnel costs due to a healthier environment,
4. the focus is on long-term savings potential rather than short-term added value and
5. changes in the use or operation of the property, deviating conditions in facility management and other relevant factors on the technical side are not mentioned.

None of these problems will be solved by special solutions in showcase buildings. In order to make the building sector more sustainable, solutions are needed for the building sector as a whole.

Methodology

Based on the literature review (Literature Review and related work) on "smart building" and "environmental sustainability" keywords was conducted using databases like

Google Scholar and Scopus, the revision of the research gaps is carried out in three steps, that are shown in Figure 3.

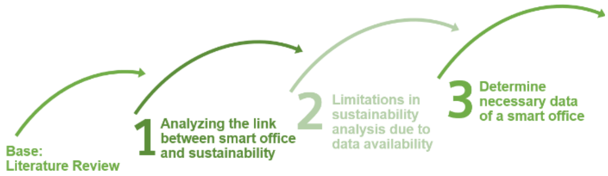


Figure 3: Methodology of this paper

Step 1 analyzed the current link between sustainability and smart buildings in Practice. Therefore, firstly 15 Case Studies of Smart Buildings in Central Europe were analyzed. Afterwards, the results were validated by guided interviews with 67 experts provided qualitative and quantitative insights. By doing so, the use cases, that affect the ecological sustainability of smart buildings, were evaluated.

Step 2 outlines limitations in sustainability analysis due to data availability. Hereby, the challenges highlighted in the expert interviews (e.g., data availability) were first summarized and discussed. Based on that, the data of 4 smart buildings, that were provided by the Facility Management and the building owner, were analyzed. In addition to that, the loss of data throughout the data processing from data collection to data usage is evaluated based on the case studies and further expert interviews.

Due to the lack of complete data, that could proof the sustainable impact of smart buildings, recommendations for necessary smart building data are made in a final step (Step 3).

Step 1: Analysis of the environmental sustainability of smart buildings in practice

In order to evaluate the environmental sustainability 15 Case Studies were analyzed. Furthermore, 25 guided interviews were conducted with participants from 67 requested companies in asset and facility management (AM & FM), general contracting (GC), hardware and software manufacturers (manufactures) in building services, engineering, planning (P), project development (PD), and owners (O) (Figure 4) in order to validate the results of the Case Study analysis.

The companies were selected based on literature, project analysis, and smart building experience. Only 25 companies could provide insights into the link between smart buildings and sustainability, resulting in a 37 % response rate. Interviews took place between April 17 and June 30, 2023, either via Teams or on-site, with four companies also submitting written information.

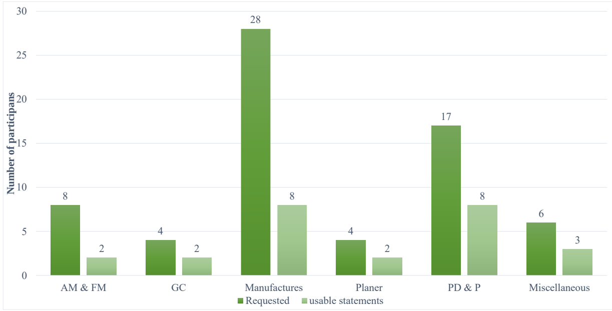


Figure 4: Participants' structure of the expert interviews

The expert interviews were based on the analysis from the basic analysis and the use cases of smart buildings, which were evaluated in workshops with experts. Drawing on existing knowledge, freely available content from implemented smart buildings, planning documents, and input from Central European smart building planners, use cases were clustered and evaluated for their environmental sustainability impact. A total of 64 use cases were evaluated.

Evaluation of the use cases of smart buildings based on Case Studies and interviews

The analysis of 15 Smart Offices in Germany and Central Europe identified 64 use cases. Due to expert interviews those can be categorized into seven main areas. In addition, the potential influence of these on the three pillars of sustainability was identified. Examining the implementation rates of these use cases in smart offices, Figure 5 illustrates their distribution. The further out the use cases are placed, the more frequently they are implemented (Percentages in Figure 5). Notably, 34 % of all use cases impact environmental sustainability (shown in bright green), while 52 % influence economic sustainability, and 72 % impact social sustainability (shown in dark green).

The dominance of energy supply (82 % of use cases) and water supply (100 % of use cases) is evident in their substantial influence on the environmental sustainability of buildings. Only 36 % of the environmentally sustainable use cases were implemented in at least 50 % of the analyzed buildings, challenging the notion that smart systems are primarily implemented for environmental reasons. Most use cases affecting environmental sustainability (72 %) were implemented in less than 33 % of the smart buildings. Further examination highlights that use cases related to information management, user experience, and building security significantly impact social sustainability and user comfort. Over half (56 %) of these use cases were integrated into more than 33 % of the buildings, indicating a predominant focus on user satisfaction and experience rather than increasing environmental sustainability.

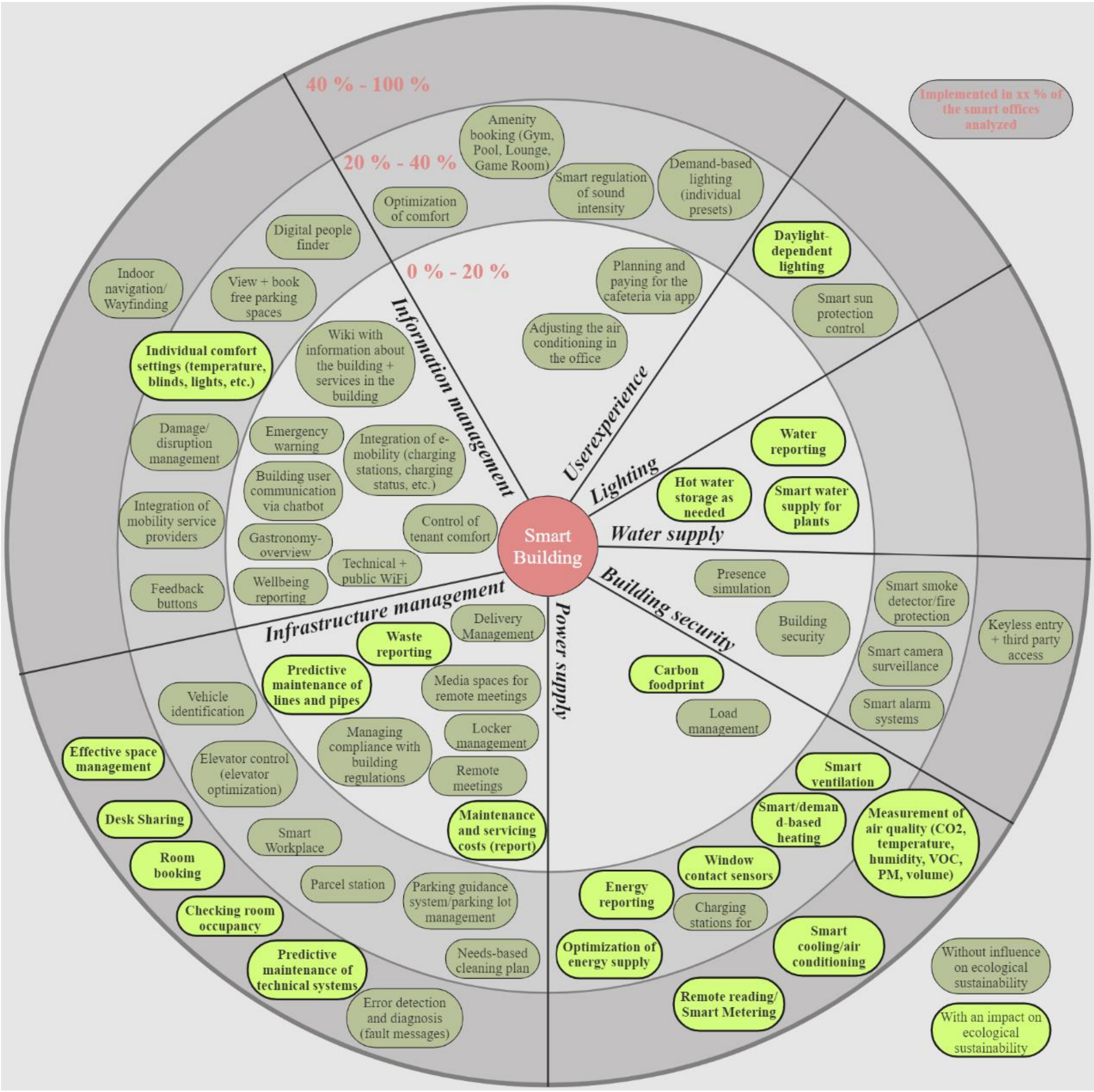


Figure 5: Illustration of the use cases of a smart building depending on the areas of application and illustration of the influence on environmental sustainability, as well as the frequency of use cases in realized smart buildings

Detailed expert interviews demonstrate that environmental sustainability is regularly not the primary objective of smart buildings but emerges as a by-product. Figure 6 the distribution of use cases impacting environmental sustainability, as well as their combinations with social and economic sustainability. Only 6 % of use cases exclusively impact environmental sustainability, while 61 % influence both environmental and economic sustainability. This dual impact is often due to the intention to operate buildings more cost-effectively, with reduced energy consumption correlating with improved ecological sustainability. Additionally, 33 % of use cases impact both social and environmental sustainability, reflecting the interconnected nature of these factors. This reiterates that the influence on ecological sustainability is typically a secondary outcome of smart building implementations.

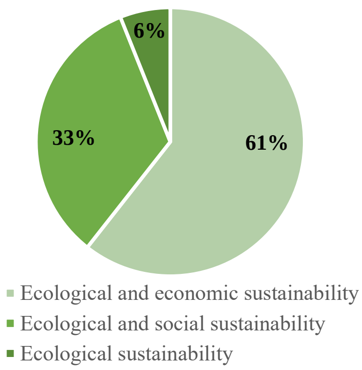


Figure 6: Percentage distribution of use cases with influence on the three pillars of sustainability

Summary of the results and critical discussion

Based on the statements of the experts, 4 trends can be identified in which the use of smart buildings can lead to an optimization of ecological sustainability:

1. several experts confirm that the increase in mobile working that occurred during the Corona pandemic resulted in a high vacancy rate. This increased the importance of and interest in use cases that optimize the use of space, such as workplace booking or space management. This allows space savings to be achieved, which can reduce both grey energy and energy in operation, as space is used more optimally.

2. At the same time, these are aimed at environmental sustainability through side effects such as needs-based heating or air conditioning.

3. Sustainable management of buildings can also be achieved through the shared use of resources, such as meeting rooms, cars, workstations, etc.

4. an analysis of consumption values and the presence of people can contribute to optimized energy use (e.g., by reducing the use of space in the company).

Expert interviews reveal that smart building use cases can enhance environmental sustainability, but there is disagreement on whether smart functionalities are essential for sustainable construction and operation. Some argue that sustainable practices can be achieved through optimal design and material choices alone, while others emphasize the importance of needs-based regulation and suggest additional energy savings through sensor and hardware utilization. The following section provides detailed insights into these perspectives. The experts also assert that a building's operation can be made more ecologically sustainable through analyses and subsequent optimization of energy consumption or space utilization (Figure 7). This definition of the functionality levels can also be found in the definition of the SRI. However, insufficient information is available on the data basis for calculating these savings, with only one interviewee from the manufacturer group mentioning the starting point as the planned energy demand of the building. It is unanimously agreed among the experts that smart functionalities serve as optimizations rather than remedies for planning errors.

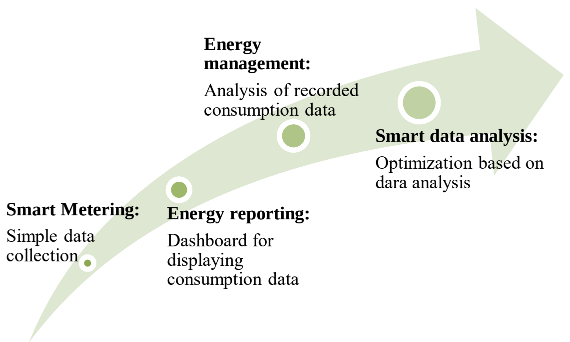


Figure 7: Exemplary optimization levels of the use cases

Step 2: Analyzing limitations of the sustainable smart building due to data availability

This section summarizes expert assessments of the initial situation regarding data availability in the real estate sector. Afterwards, the data basis used by the experts in their own companies will be examines. Finally relevant building data from the study that can be used to measure and assess the environmental sustainability of a building is listed.

Case Studies

6 % of the requested experts were able to provide data from their smart buildings. The remaining respondents referred to data protection or were unable to provide data for other reasons. The received data sets were incomplete or not meaningful. For two buildings, the data could only be provided monthly and for one building only annually. It was not possible to clearly identify the system boundaries to which the data refers. In addition, in three cases, information on the building was missing. Therefore, it was not possible to analyze the sustainable operation of an actually implemented smart building. As a result, the savings claimed by the experts cannot be validated. For this reason, data and their collection interval, which are necessary for the evaluation of the building operation, were identified and summarized in Table 1.

Initial situation of the real estate sector regarding data availability according to the experts

A smart building requires the effective use of data stored in the cloud to promote environmental sustainability. It is necessary to capture data from the building management system, integrated use cases and the entire life cycle of the building. All data (e.g., building model, consumption data, attendance data, FM data, BMS data) from the building must be available over the entire life cycle in order to be able to derive optimization potential from the data. 20 % of experts emphasize that large FM companies often prioritize functionality and user well-being, while building sustainability is neglected. 50 % of manufacturer experts report impairments due to manual interventions in automated processes. Furthermore, two experts noted that the operators often do not have the necessary in-depth knowledge of all the processes and functions of the building and also do not have the capacity to operate a building in a technically professional manner. Another expert criticized the fact that the owners have no need to make the management of the building more sustainable, as the tenants bear the energy costs. Continuous optimization for sustainable operation moves into the background. Some owners have only shown an interest in sustainable building operation since the changes in the real estate industry, as tenants now bear the energy costs.

Current data basis of the experts

Figure 8 shows the results of the expert interviews on the use of the data collected in smart buildings and highlights the data flows and losses. It was interrogated whether data on building operation is recorded and if this data is used

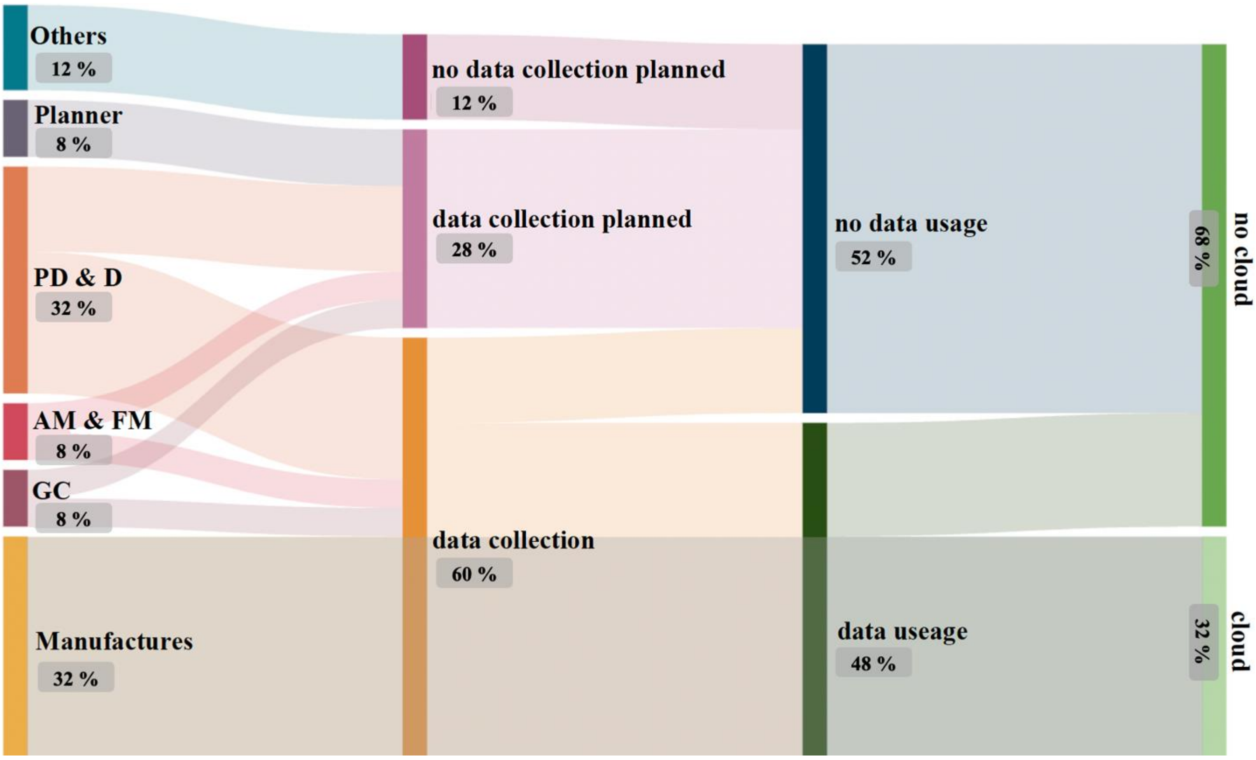


Figure 8: Illustration of the current data basis based on the expert interviews

for further purposes. The status of cloud solutions for data storage was also determined.

60 % of the companies surveyed collect building operating data (marked in orange), either directly by owners, FM or third-party providers. 28 % plan to collect data, while 12 % have no plans to do so.

Of the companies collecting data, 20 % do not use the data further and the storage locations are unclear. The remaining 80 % use the data for FM support, plant optimization or sustainable operational control. This corresponds to 48 % of all companies surveyed, with 75 % of these data-collecting companies being manufacturers. Only 62 % of the project developers and owners surveyed receive data from manufacturers, which is merged in PowerBI dashboards. A proprietary building operating system is rarely used. Over 50 % of experts do not use any building technology or energy-related data for optimization in their companies. Only 32 % of interviewed companies store data centrally in a cloud. However, they all belong to the group of manufacturers who sell the clous as part of their portfolios This means that it is not possible to draw any conclusions as to whether the relevant players (FM, building owners, etc.) use the data centrally for long-term analyses and evaluations of optimization scenarios.

Difficulties in data availability

The aim of this paper is to compile relevant data for the uniform assessment of smart buildings in terms of environmental sustainability. The following section presents the necessary data basis. In the validation phase following the interviews, only four companies provided data on their smart buildings, which was largely implausible and incomplete. Due to a lack of data quality, no conclusions could be drawn about the operation and sustainability of the buildings. Data protection issues and concerns, particularly in relation to the General Data Protection Regulation (GDPR), also prevented 73 % of experts from releasing data. Works councils were against use cases that collect personal data, such as access controls, due to concerns about employees' privacy rights. The debate on data protection versus sustainability remains open. Two experts criticized the lack of availability of data for optimization and research purposes, while one manufacturer considered anonymized provision of customer data in the future but pointed to limited resources.

Step 3: Definition of Relevant Data for assessing environmental sustainability

The results of the expert interviews and available data show the current problem in the building sector: data availability. Data is regularly not recorded and stored sufficiently over a long time by the building management system. Data failures go unnoticed and even existing data is not used optimally. A sufficient data basis is crucial so that the collected data can be automatically processed and made available to users, asset managers and other stakeholders in a clear and concise manner. Only in this way can the data be used for optimization and have a positive impact on environmental sustainability. The building management system often records various data over a period of time, including supply and return temperatures, control signals and status messages from technical components. This data can already indicate major errors in the regulation and control system.

To assess the sustainability of smart buildings, the experts listed implemented use cases and provided consumption data for heating, cooling and electricity at quarter-hour intervals. The inclusion of occupancy data, such as the number and type of occupants, was considered crucial. The information requested included building area, type of use, number of employees, occupancy data, metering concepts, floor plans, heating/cooling systems, ventilation and sensor infrastructure. Based on these criteria, relevant data for the environmental sustainability analysis in smart buildings is compiled in Table 1.

Table 1: Data basis for the sustainability assessment of a smart building

Nature of the data	**Description**
Energy consumption data (heating, cooling and electricity) of the building	
Consumption data for the entire building, individual parts of the building or rental sections - General, technical and tenant electricity - Heating/cooling consumption of the respective rental areas, if applicable meeting and conference rooms recorded individually	Ideally, consumption data should be recorded on a quarter-hourly basis and broken down into individual building sections. Smaller-scale recording enables more precise monitoring and must be selected according to the type of use of the building. *[kWh]*
Emission factor of the respective energies	CO_2-equivalent data for different energy sources according to the Law on Energy Efficiency. To compare the type of energy supplier in different buildings. *[gCO_2e/ kWh]*
Outdoor temperature	To classify the amount of energy consumed. Ideally every 15 minutes for consumption *[°C]*
Planning documents	
Average demand of Energy	Comparative value for checking the operation. This must not be exceeded. *[kWh/ m²a]*
Building area or individual rental space	For consumption data of individual sections, in each case with details of the gross floor area. Allocation and comparison of the different areas. *[m²]*
Type of use	List of the types of use in the building (office, canteen, daycare center, fitness studio, etc.). Relevant for assessing the energy consumption of the smart building as such. As well as for the comparison of other smart buildings by type of use.
Heating and cooling generation systems, air conditioning systems	Information on the type and output of the systems. To classify the energy requirement depending on the respective energy source. As well as for checking the control behavior. *[e.g. kW]*
Building automation and smart building	
Data points of the different systems	Type and number of data points of the technical building systems installed in the building
Individual room control	Zone size incl. number and type of connected data points
Fault message forwarding	Information on the process chain in the event of a fault message from a system or system component. Not directly for comparison or evaluation of ecological sustainability, but decisive for optimization potential of the systems.
Metering concept	Necessary for the allocation of consumption data.
List of implemented use cases	Incl. description, for assessment of smartization.
Data storage	Information on the duration of data storage and storage location to ensure data availability for the analysis of the building.
Sensor infrastructure	Type and number of sensors or concepts regarding the installed sensors (brightness, occupancy, temperature, air quality such as CO_2, Bluetooth beacons), to assess smartization.
Operation	
Occupancy data	Information on m² per person, as well as time-related information on the use of the building area. Is required to assess the energy consumption. *[m²/n; n/h]*
Number of employees	For comparison with other smart buildings *[n]*
Evaluation of the use cases	Evaluation of the frequency of use of different use cases. To identify the optimization potential regarding the use of the building's Smartness.

The identified essential data for evaluating building operation provides a basis for analyzing smart buildings in terms of their sustainability. This database must be validated in further studies.

Conclusion and recommendations for action

The analysis relies on the standardized definition of a smart building established in this paper across different building types. The clustering of use cases serves as a common foundation for assessing potential impacts on environmental, economic, and social sustainability. Currently, ecological sustainability is often considered a by-product of improved economic or social sustainability. To enhance ecological sustainability, more dedicated ecological use cases need implementation. It's important to note that technological advancements may introduce new use cases for smart buildings.

However, the analysis reveals the incompleteness and poor traceability of provided data, particularly concerning energy consumption. Technical and data protection reasons hindered obtaining further information on data classification, occupancy figures, outdoor temperatures, or measurement concepts, rendering the database unreliable. Consequently, results were supplemented with expert interviews, and strategies for future data collection were developed.

The study's limitations include the exclusive use of smart offices from Central Europe and a lack of representation from companies outside Europe. Additionally, the short lifespan of the smart buildings studied suggests that operators may lack a holistic understanding of building functions and processes. The study does not address the embodied carbon of components like sensors, actuators, and cables, posing a need for comprehensive assessments. The study highlights potential for further investigations, particularly regarding the embodied carbon of data storage throughout a building's life cycle.

The identified essential data for evaluating building operation provides a basis for analyzing smart buildings in terms of their sustainability. To validate the defined data basis, real buildings must be monitored over an extended period. Potential increase in CO_2 emissions from additional hardware and data storage should be considered. A comparison with similar buildings lacking smart implementations can provide valuable insights.

References

Bäuerle, H. & Lohmann, M.-T., 2021. *Ökologische Materialien in der Baubranche - Eine Übersicht der Möglichkeiten und Innovationen.* Stuttgart: Springer Vieweg.

Bosch, M. & Deckert, R., 2023. *Digitalisierung und Smart Building - Ein kritischer Erfolgsfaktor für nachhaltige Entwicklung.* Wiesbaden: Springer Gabler.

Bundesministerium des Innern, für Bau und Heimat (BMI), 2019. *Leitfaden Nachhaltiges Bauen - Zukunftsfähiges Planen, Bauen und Betreiben von Gebäuden.* Berlin: s.n.

Bundesministerium für wirtschaftliche Zusammenarbeit und Entwicklung, 2023. *Klimaabkommen von Paris.* [Online] Available at: https://www.bmz.de/de/service/lexikon/klimaabkommen-von-paris-14602, [Accessed 03.09.2023].

DIN e.V., 2019. *DIN EN ISO 52120-1 Energieeffizienz von Gebäuden - Einfluss von Gebäudeautomation und Gebäudemanagement - Teil 1.* Berlin: Beuth Verlag.

Europäische Union, 2016. *Übereinkommen von Paris.* Paris: Amtsblatt der Europäischen Union.

Fraunhofer-Allianz Bau, 2022. *Positionspapier der Task-Force "Bauen der Zukunft".* Valley: s.n.

Fraunhofer-Allianz Bau, 2023. *Smart Building.* [Online] Available at: https://www.bau.fraunhofer.de/de/forschungsbereiche/Gebaeudeautomation.html, [Accessed 26.09.2023].

Graveto, V., Cruz, T. & Simöes, P., 2022. Security of Building Automation and Control Systems: Survey and future research directions. *Computer & Security,* Volume 112.

Haist, M. et al., 2023. Nachhalitger Betonbau Vom CO2- und ressourceneffizienten Beton und Tragwerk zur nachhaltigen Konstruktion. In: *2023 Bauphysik Kalender: Nachhaltigkeit.* Berlin: Ernst & Sohn GmbH, pp. 259-363.

Herkströter, C., Jilge, B. & Beckers, C., 2023. ESG-Konformität - Entwicklung bei der Regulierung im Immobiliensektor. In: *Rating von Industrieimmobilien.* Wiesbaden: Springer Gabler; Frankfurt School Verlag, pp. 167-185.

International Energy Agency, 2023. *ANNEX 81 - Data-Driven Smart Buildings: State-of-the-Art Review - Energy in Builindgs and Communities Technology Collaboration Programme.* Newcastle : CSIRO.

Jacob, C. & Kukovec, S., 2022. *Auf dem Weg zu einer nachhaltigen, effizienten und profitablen Wertschöpfung von Gebäuden.* Wiesbaden: Springer Vieweg.

Koch, D., 2023. Immobilien- und Baurecht. In: *Innovative Rechtberatung.* Stuttgart: Schäffer-Poeschel Verlag, pp. 87-96.

Kropp, A., 2019. *Grundlagen der nachhaltigen Entwicklung - Handlungsmöglichkeiten und Strategien zur Umsetzung.* São Paulo: Springer Gabler.

Li, Z. et al., 2020. A Review of Smart Design Based on Interactive Experience in Building Systems. *Sustainability,* 12(17).

Monterioa Froufe, M. et al., 2020. Smart Buildings: Systems and Drivers. *Buildings 2020,* 10(9).

Mordor Intelligence, 2023. *Analyse der Marktgröße und des Anteils von Smart Building - Wachstum und Prognose (2023-2028),* Hyderabad: Mordor Intelligence.

Moretti, N. et al., 2020. Maintenance service optimizatio in smart buildings through ultrasonic snesors network. *Intelligent Buildings International,* 13(1), pp. 4-16.

Neef, T., Müller, S. & Mechtcherina, V., 2020. 3D-Druck mit Carbonbeton: technologien und die ersten Untersuchungsergebnisse. *Beton- und Stahlbeton,* 115(12), pp. 943-951.

Nulsch, N., 2023. Die Energiekrise 2022 - Eine Zeitenwende für den Gaasmarkt. *Handelsblatt*, 20 April, pp. [Online] Available at: https://live.handelsblatt.com/die-energiekrise-2022-eine-zeitenwende-fuer-den-gasmarkt/.

Pašek, J. & Sojková, V., 2018. Facility Management of Smart Buildings. *International Review of Applied Sciences and Engineering,* 9(2), pp. 181-187.

Plageras, A. et al., 2018. Efficient IoT-based sensor BIG Data collection-processing and analysis in smart buildings. *Future Generation Computer Systems,* Volume 82, pp. 349-357.

Qolomany, B. et al., 2019. *Leveraging Machine Learning and Big Data for Smart Buildings: A Comprehensive Survey.* Kearney (USA): IEEE Access.

Schmid, K., 2023. *So heiß wie seit mindestens 125.000 Jahren nicht.* [Online] Available at: https://www.tagesschau.de/ausland/klimawandel-hitze-extremwetter-100.html#:~:text=Auch%20im%20Oktober%20gab%20es,w%C3%A4rmste%20seit%20125.000%20Jahren%20werden., [Accessed 09.11.2023].

Siemens Schweiz AG, 2020. *Whitepaper Smrt Office: Das intelligente Büro der Zukunft.* Zürich: Siemens.

SmartScore, 2021. *Smart Buildings. Die Zukunft ist smart - SmartScore White Paper.* s.l.:SmartScore.

The European Commission, 2020. *COMMISSION DELEGATED REGULATION (EU) 2020/2155 - supplementing Directive (EU) 2010/31/EU of the European Parliament and of the Council by establishing an optional common European Union scheme for rating the smart readiness of buildings.* Brussels: s.n.

United Nations Environment Programme, 2022. *2022 Golbal Status Report for Buildings and Construction.* Nairobi: s.n.

BETTER URBAN MANAGEMENT: A SYSTEMATIC REVIEW OF MULTI-SCALE DIGITAL MODELLING

Feng Guo[1], Ling Ma[1]
[1] Huazhong University of Science and Technology, Wuhan, China

Abstract

Digital modelling has significantly progressed. Many scholars have reviewed and summarized the research of digital modelling. However, the reviewer needs to consider the practical applications of digital modelling at different scales. Multi-scale digital modelling (MSDM) refers to digital models that contain different levels of geometric or semantic details. This paper presents a systematic literature review (SLR) of MSDM. Distributed energy generation is one of the most worthwhile research directions. In the future, distributed energy generation systems - energy storage systems - and buildings or urban infrastructure should be interconnected based on MSDM.

Introduction

Digital modelling of buildings and cities is becoming an important research topic. Digital models of buildings, usually referred to as Building Information Models (BIM) or 3D building models, are detailed, accurate, and visual representations of buildings. BIM can be utilized across all stages of a building's complete lifecycle, including design, construction, operation, and maintenance (Qiu et al., 2021). In the design phase, BIM facilitates a more profound understanding among all stakeholders of the building's details and overall design compared to traditional 2D drawing models, allowing for the early identification of potential design issues such as clashes in piping or structural elements (Huh et al., 2023); in the construction phase, BIM aids in visualizing construction progress and enhancing on-site safety management (Tu et al., 2021); in the operation and maintenance phase, the energy consumption (Kamel and Kazemian, 2023) and carbon emissions (Lu and Deng, 2023) generated during the use phase of a building can be estimated from the site data collected by BIM and devices like sensors. A key benefit of BIM and equivalent technologies is their object-oriented nature, which ensures that models accurately represent physical and functional characteristics and encompass objectified relationships among building components. Despite its benefits, BIM is mainly only used in the design phase (Tu et al., 2021). The main reason is that BIM is required to carry different information for different phases, and too much useless information will reduce users' willingness (Wang et al., 2022). Therefore, it is necessary to develop BIMs containing different information or to automate the addition or subtraction of BIM information for different usage requirements.

Digital models of cities usually referred to as City Information Models (CIM) or 3D city models, which need to consider topographical data such as mountains, rivers and buildings and infrastructure (Xia et al., 2022). Souza and Bueno (2022) have shown that CIM contains Geographic Information System (GIS) data and BIM data, as well as data collected by the Internet of Things (IoT) and sensors. It enables automatic management of urban information and predictive reasoning about urban problems. For example, as different regions and buildings in a city have different energy demands and consumption (Camero and Alba, 2019), CIM can help predict energy-associated problems ((Xia et al., 2022) (Weil et al., 2023)). Although CIM can present a wealth of information, CIM faces a similar problem as BIM: they both carry a wealth of information, which makes users who only need a specific piece of information reluctant to invest additional time and resources to get the information they need from it. In other words, the development of BIM/CIM with different information or the ability to automatically add or subtract information from BIM/CIM is a real need for most users. In BIM/CIM, 3D geometric models are critical primary data. To improve the user experience and meet different business requirements and application scenarios, developing digital models with different geometric levels or graphical simplification may be an effective way.

The same Digital models may have different levels of detail (LoD) in different application scenarios. LoD refers to the representation of a 3D model with different data information, mainly concerned with the geometric details of the graphics. In BIM/CIM, there is a similar concept of LoD, but it is not limited to graphical rendering but also includes other information (Deng et al., 2016). In the commonly used CIM data exchange format City Geography Markup Language (CityGML), LoD is not only the complexity of the geometrical level of the city model but also includes the complexity of the semantic information (Kolbe et al., 2005). The concept of LoD also exists in BIM. However, the current level of development (LOD) is more widely used than LoD, which was developed by the American Institute of Architects (AIA) as an improvement of LoD and started with five levels from LOD 100 to LOD 500 (Abualdenien and Borrmann, 2022), corresponding to the five phases of conceptualization, approximate components, exact components, construction, and completion. The BIMForum has further improved the AIA definition by developing a new LOD350 level and publishing the LOD Specification. In addition to defining LoDs with expert knowledge, it is also possible to adjust the LoDs of the model using some software, such as Blender and Unity, which can modify graphics details. Fig. 1 shows the result of representing the exact Monkey in Blender with a different number of surfaces.

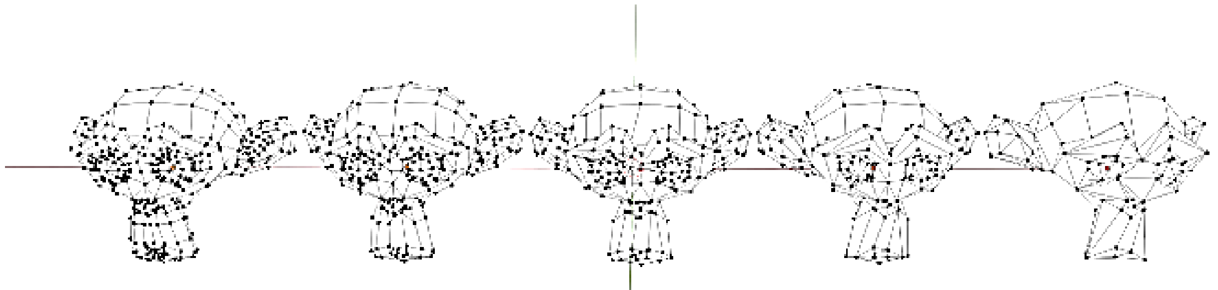


Figure 1: Blender generates a Monkey with different numbers of faces

Although some standards for LoDs exist, there still needs to be more clarity in understanding the information required for LoDs in different countries and organizations, especially concerning geometry. The main reason is that semantic information can usually be normalized into a set of attributes, but systematically examining the geometric details of a model, developing appropriate criteria to classify the hierarchy of geometries, and applying them to specific models is very difficult (Abualdenien and Borrmann, 2022). In addition, integrating the models of different LoDs into a city model is also a challenge. Table. 1 shows the LoD of BIM defined by AIA and the definition of LoD by CityGML

Table 1: LoD Definitions for BIM and GIS

LoD of BIM	LoD of GIS
LoD 100: Information on building height, area, and volume for conceptual design	LoD 0: Digital Elevation Model (DEM) and Digital Orthophoto Map (DOM) for regional landscape representation scale
LoD 200: Approximate information on generic elements for preliminary design	LoD 1: Block models at the city and regional representation scale
LoD 300: Detailed information on specific components for detailed design	LoD 2: Basic models at the district and street representation scale
LoD 400: Complete construction and installation information for the construction design	LoD 3: Standard models for representative building exteriors
LoD 500: Information on structural components corresponding to the as-built design	LoD 4: Advanced models for representative building interiors

Similar to this review, many scholars have reviewed and summarized the literature on BIM, GIS, Urban Digital Twin (UDT), and other themes related to digital modelling. Table. 2 shows some of the representative literature.

Table 2: A representative collection of literature reviews related to BIM/GIS/UDT

Articles	Main Contents
Wang et al., (2019)	It provided a review of BIM-GIS integration in sustainable built environments.
Jiang et al., (2021)	It identified recent advances in DT in civil engineering and made recommendations for future DT developments.
Khan et al., (2022)	It highlighted the challenges and future directions of multiscale modelling for megacities and intelligent cities.
Kamra et al., (2023)	It summarised techniques for lightweight city reconstruction and indicated future research directions.
Weil et al., (2023)	It summarised the main challenges and unresolved issues related to the implementation of UDT.

Compared to the representative literature review in Table. 2, this SLR focuses on the MSDM for buildings and cities. With the rapid development of information technology and the further expansion of urban boundaries, the demand for MSDM will skyrocket(Abualdenien and Borrmann, 2022). Therefore, an SLR is needed to inform scholars of the current research gaps on MSDM and more accurately capture future research priorities.

The rest of the paper follows: Chapter 2 introduces the methodology applied, describing the SLR protocol.

Chapter 3 shows the results obtained by SLR. Chapter 4 discusses the inspirations gained and points out future research priorities. Chapter 5 summarises the whole paper and further highlights the contributions made in this SLR.

Methodology

Searching database and keywords

The literature covered in this SLR is at the intersection of civil engineering and computer graphics. Since computer technology requires high timeliness and many scholars will publish their papers at conferences, this SLR needs to cover conference and journal papers. Four academic databases, ACM Digital Library, IEEE Xplore, Scopus, and Web of Science (WOS), were selected for searching. The ACM Digital Library and IEEE Xplore are critical academic databases for computer science, containing many conference papers and some journal papers related to computer graphics (Cheng et al., 2022). Scopus and WOS are more comprehensive in their coverage of the subject areas and contain more abundant journal papers (Srinivasan and Yadav, 2023) and, therefore, serve as a supplement to the ACM Digital Library and IEEE Xplore.

After determining the search database, this section will identify the search keywords. Considering that the core research question of this SLR is "Research on Building/City MSDM," this SLR divides the required keywords into two parts. Fig. 2 shows the complete search formula (four databases use a similar search logic).

Building modeling

(“ifc” OR “industry foundation class*” OR “bim” OR “building information model*” OR “building model*” OR “building digital twin*”)
AND
(“Multiscale” OR “lod” OR “level of detail*” OR “log” OR “level of geometr*” OR “loi” OR “level of information”)

City modeling

(“citygml” OR “city geography markup language” OR “cim” OR “city information model*” OR “cit* model*” OR “urban* model*” OR “town* model*” OR “cit* digital twin*” OR “urban* digital twin*” OR “town* digital twin*” OR “3D GIS” OR “3D geoinformation”)
AND
(“Multiscale” OR “lod” OR “level of detail*” OR “log” OR “level of geometr*” OR “loi” OR “level of information”)

Figure 2: Schematic diagram of the used search formula

Literature screening criteria

After reading the title-abstract-keywords of each piece of literature, literature duplicated or did not meet the research objectives was excluded, and the detailed literature screening criteria were as follows.

(1) Removing literature whose form of publication is not an article or conference paper (published as a book, magazine, etc.); (2) Removing literature whose language of publication is not English; (3) Removing literature in apparently unrelated subject areas (e.g., Nursing, Veterinary, Dentistry, etc.); (4) Removing literature that is purely theoretical, purely conceptual and not focused on digital modelling; (5) Removing duplicates. After obtaining all the powerfully relevant literature, further reading the full text and delineating the literature at both cities/buildings’ scales. Define the building scale as a single building or a component within a building and the city scale as multiple buildings or the integration of BIM with GIS.

Extracted information

After obtaining the required literature, further reading of the full text extracted more detailed information. The extracted information included the primary statistical information data of the literature. The following 12 categories of information were extracted: authors, affiliation of the first author, country of the first author's affiliation, date of publication, name of the journal or conference, keywords, purpose of the study, software or tool used, source of the data, format of the data, methodology of the study, and results of the study.

Results

Overview

Firstly, the initial search results were IEEE Xplore (123), ACM Digital Library (1316), Scopus (1655), and Web of Science (787), with a total of 3881 pieces of literature retrieved. Secondly, based on the screening criteria, 3881 pieces of literature were refined, resulting in 452. Then, the full text of 452 pieces of literature was carefully read, resulting in 45. Finally, the literature was divided according to the definitions of building scale and city scale. Twenty-four pieces of literature were obtained for the building, 21 for the city.

Research themes

Fig. 3 shows the application scenarios for city/building MSDM. In City management, there are main eight application scenarios. For Urban Renewal Synchronization, Qin, (2014) detected changes in the LoD2 building model using satellite-based stereo images from different dates. For Energy Demand Simulation, Nouvel et al., (2017) investigated the impact of data quality, such as geometry and semantics, on SimStadt's estimation of heating demand. The LoD1 model can reliably estimate the heating demand of an area, and the LoD2 model can reliably estimate the heating demand of a single building. De Jaeger et al., (2018) quantified the geometric and energy characteristics differences between the five variants of LoD1 and LoD2. It investigated the availability of LoD1 and LoD2 and the impact of building geometry in regional energy modelling. For Energy Consumption Simulation, Johari et al., (2022) analyzed the impact of LoD on energy use in different types of multifamily residential buildings. The effect of LoD was negligible in heating seasons, whereas it was significant throughout the year. The deviation between LoD1 and LoD3 was as high as 9%. For Solar Energy Simulation, Nahon et al., (2013) argued that using adaptive LoD speeds up calculations and ensures accuracy. Besuievksy et al., (2014) introduced a flexible LoD system that allows the user to specify the geometric hierarchy of geometric models for solar simulation. Besuievksy et al., (2014) estimated the annual solar irradiance of a building's roof using Monte Carlo simulation and conducted uncertainty analyses in different LoD models. Besuievsky et al., (2018) proposed a new LoD strategy that automatically detects and preserves the geometries (e.g., roofs) that affect the solar simulation. It simplified the rest of the geometries. Machete et al., (2018) evaluated the impact of the urban environment on the exposure of building roofs and facades to solar radiation by combining a 3D GIS with a solar radiation tool and using two different modelling approaches (2.5D and 3D methods). Peronato et al., (2018) simulated the solar irradiance of 109 buildings at different spatial and temporal granularities using the LoD2 model. Saretta et al., (2020) defined the LoD2.5 model to estimate the BIPV potential of a city containing facades. Han et al., (2022) predicted the regional power potential by analyzing the variation of shading effects between the LoD1 and LoD2 models. Krapf et al., (2022) extracted the roof superstructure from aerial imagery using deep learning, which can improve the solar simulation's accuracy. Yan et al., (2023) reconstructed three-dimensional buildings from high-resolution satellite imagery using a detail-oriented deep learning method and estimated the solar power potential. For Climate Simulation, Chen et al., (2020) cited 16 LoD models defined by Biljecki et al., (2016) for climate simulation.

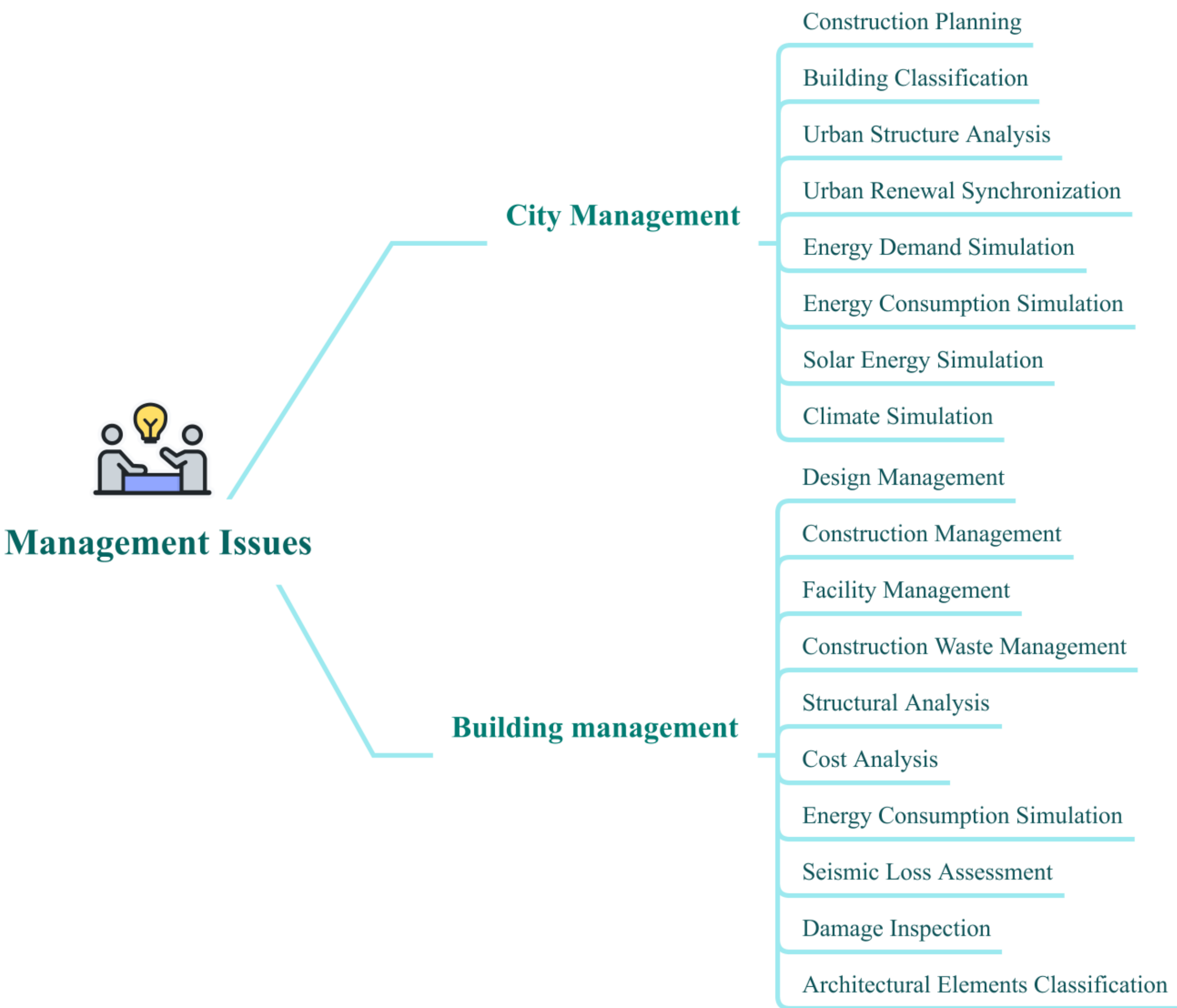


Figure 3: Application scenarios for city/building MSDM

In Building management, there are main ten application scenarios. For Design Management, Leite et al., (2011) evaluated the effort required to generate BIMs with different LoDs and the impact of LoDs on design coordination, such as electrical and plumbing (MEP). Huang et al., (2022) extended IFC to develop BIM models of metro stations with multiple LoDs, which can cope with different scenarios. For Construction Management, Han et al., (2015) enhanced the 4D BIM with lower LoDs to provide visual construction progress assessment for some projects. For Facility Management, Dias and Ergan, (2016) developed LoDs for facilities management. For Construction Waste Management, Sanchez et al., (2021) developed methods to automatically determine parameters that can be disassembled and defined BIM models and parameters for disassembly. For Structural Analysis, Briggen et al., (2009) used Computational Fluid Dynamics to predict Wind-driven rain volumes, concluding that façade geometric detailing significantly affects localization. Gonizzi Barsanti et al., (2023) obtained appropriate geometric detailing level models from the point cloud data, which can be used directly in finite element calculations. For Cost Analysis, Wood et al., (2014) enabled the automatic calculation of design costs by matching building material data with cost data on different LoDs. For Energy Consumption Simulation, Geyer et al., (2018) proposed a multi-LoD modelling approach combined with machine learning algorithms to predict energy performance. Malhotra et al., (2019) compared different LoDs and determined the LoDs required for heating load simulation. Singh and Geyer, (2020) analyzed the extent to which design parameters influence energy consumption prediction. It identified the information required for each level in a multi-LoD. Mediavilla et al., (2023) developed a methodology based on graph techniques for using multiscale BIM for energy analysis. Jung et al., (2023) proposed that LoD350 was suitable for modelling the energy performance of actual buildings. Xu et al., (2023) proposed a new LoD2 (LoD2ES) for energy simulation. For Seismic Loss Assessment, Xu et al., (2019) evaluated earthquake damage using BIM with different LoDs. For Damage Inspection, Pantoja-Rosero et al., (2023) constructed LoD3 BIM models for building damage assessment. For Architectural Elements Classification, Koo et al., (2021) classified IfcDoor and IfcWall using Multi-view CNN and PointNet. Abualdenien and Borrmann, (2022) proposed a framework for analyzing and checking the LoG of BIM.

In City/Building Management, energy issues are hot research themes. Scholars try to use different LoD models, combine different frameworks, and use different methods to predict the energy demand, consumption, and solar power generation potential. Scholars have already obtained valuable research results. However, many issues still need to be fully resolved due to the excessive number of factors and the complexity of the problem. There will remain a topic worthy of in-depth research.

Implication

Current researches

MSDM is for a better solution to practical problems. By this paper, it is clear that MSDM can serve all stages of the whole life cycle of a building, solving problems in all aspects of urban construction and maintenance, which contribute to the sustainable development of cities. However, there are still the following limitations in the current researches.

Much of the current literature on the application of MSDM to city management focuses on problems related to building groups. Although the problems addressed are relevant to individuals and societies, the boundaries of research should be further expanded. Scholars need to consider infrastructure and management issues further. MSDM is a substantial database for solving all kinds of problems in the city and society.

Future directions

MSDM has to serve practical problems. The above limitation is a direction that deserve research in the future, but the current most worthwhile research is to use MSDM to solve urban energy-related problems.

Excessive use of non-renewable energy sources (such as fossil fuels) may cause environmental and social problems. Widespread use of renewable energy in cities will contribute to sustainable urban development. Scholars have estimated the energy demand, energy consumption, and solar power generation potential of cities/buildings through various means of simulation. It contributes to energy management in cities and the development of urban distributed generation. However, there needs to be research on unifying these results in a single system or considering the optimal arrangement of solar power generation systems, energy storage systems, and the connection to buildings because most of the distributed energy generation in cities tends to be self-generated and self-consumed. Urban planners need to consider how to manage the excess renewable energy to respond flexibly to different situations. The potential for urban distributed energy generation should be combined with regional energy demand consumption to achieve a rational distribution of renewable energy sources. Fig. 4 shows the framework for future research on energy issues.

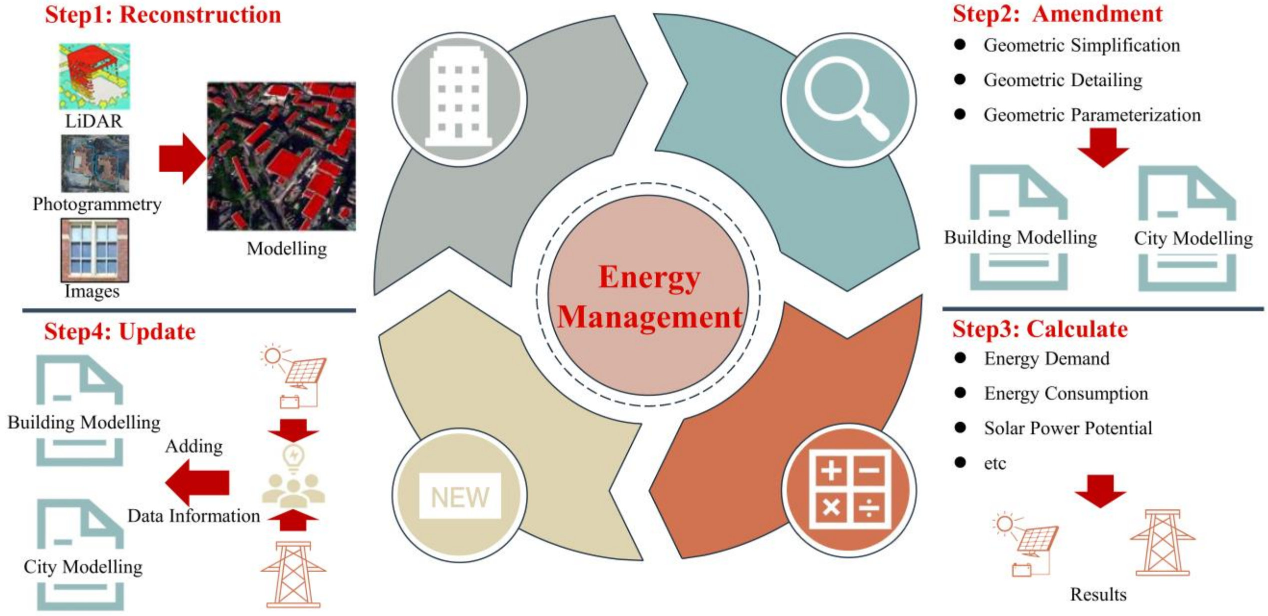


Figure 4: Research framework for future city energy management

Firstly, using LiDAR/Photogrammetry/Images to reconstruct the MSDM of the actual city. Secondly, using Geometric Simplification/Detailing/Parameterization to amend the MSDM. Then, based on the MSDM, calculate parameters such as the city's Energy Demand/Consumption and Solar Power Potential. Based on these results, developing algorithms to rationally plan the distributed energy generation system, energy storage system and the specific allocation of energy use in the region. Finally, the obtained information is fed back to the MSDM to update the MSDM to achieve sustainable urban energy management.

For researchers, LoD definitions for different scenarios and the level of detail provided can help designers identify potential problems early in the project, reduce errors and design modifications, and thus improve the design quality. It will help designers better understand and plan their projects, which will help optimise the use of materials and construction techniques and further reduce costs.

The conversion of models at different levels of detail can effectively facilitate collaboration between project participants by providing a shared, multi-scale project model. Different teams can work on the same model and view and update information in real time, significantly improving project collaboration's efficiency and effectiveness.

Conclusions

This paper presents a comparative analysis of the current advances in MSDM research from both city/building perspectives. Compared to previous literature reviews, this review focuses on the different scales of digital models, especially their geometric aspects. This paper identifies the research themes of MSDM for cities/buildings, analyses the findings of the existing literature by theme, and points out the hot research themes in "City/Engineering Management" and the limitations of each research theme. Regional energy management will be one of the hot research directions, based on MSDM, to unify distributed energy generation system - energy storage system - building/infrastructure in one system will achieve the optimal distribution of energy, give full play to the potential of renewable energy, and promote the sustainable development of the city and the society.

Furthermore, this paper finds that scholars in Africa and South America are not concerned with MSDM research, much less using MSDM for energy management. The equatorial region is extremely rich in solar energy resources. From a global energy perspective, it is necessary to apply the research results of MSDM in equatorial countries to enhance global energy sustainability further.

However, this paper still has some limitations. Only a representative portion of the literature selected for this paper and other valuable literature may have yet to be analysed. Meanwhile, the search terms may need to be more comprehensive, and some may need to be complete.

Acknowledgments

The authors would like to acknowledge the financial support provided by the National Natural Science Foundation of China (Grant No. U21A20151).

References

Abualdenien, J., Borrmann, A., 2022. Ensemble-learning approach for the classification of Levels Of Geometry (LOG) of building elements. Adv. Eng. Inform. 51, 101497.

Besuievksy, G., Barroso, S., Beckers, B., Patow, G., 2014. A configurable LoD for procedural urban models intended for daylight simulation 19–24.

Besuievsky, G., Beckers, B., Patow, G., 2018. Skyline-based geometric simplification for urban solar analysis. Graph. Models 95, 42–50.

Biljecki, F., Ledoux, H., Stoter, J., 2016. An improved LOD specification for 3D building models. Comput. Environ. Urban Syst. 59, 25–37.

Briggen, P.M., Blocken, B., Schellen, H.L., 2009. Wind-driven rain on the facade of a monumental tower: Numerical simulation, full-scale validation and sensitivity analysis. Build. Environ. 44, 1675–1690.

Camero, A., Alba, E., 2019. Smart City and information technology: A review. Cities 93, 84–94.

Chen, S., Zhang, W., Wong, N.H., Ignatius, M., 2020. Combining CityGML files and data-driven models for microclimate simulations in a tropical city. Build. Environ. 185.

Cheng, Y.L., Lim, M.H., Hui, K.H., 2022. Impact of internet of things paradigm towards energy consumption prediction: A systematic literature review. Sustain. Cities Soc. 78, 103624.

De Jaeger, I., Reynders, G., Ma, Y., Saelens, D., 2018. Impact of building geometry description within district energy simulations. Energy 158, 1060–1069.

Deng, Y., Cheng, J.C.P., Anumba, C., 2016. Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. Autom. Constr. 67, 1–21.

Dias, P., Ergan, S., 2016. The Need for Representing Facility Information with Customized LOD for Specific FM Tasks 2563–2572.

Geyer, P., Singh, M.M., Singaravel, S., 2018. Component-Based Machine Learning for Energy Performance Prediction by MultiLOD Models in the Early Phases of Building Design, in: Smith, I.F.C., Domer, B. (Eds.), Advanced Computing Strategies for Engineering, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 516–534.

Gonizzi Barsanti, S., Guagliano, M., Rossi, A., 2023. DIGITAL (RE)CONSTRUCTION FOR STRUCTURAL ANALYSIS. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XLVIII-M-2–2023, 685–692.

Han, J.-Y., Chen, Y.-C., Li, S.-Y., 2022. Utilising high-fidelity 3D building model for analysing the rooftop solar photovoltaic potential in urban areas. Sol. Energy 235, 187–199.

Han, K., Cline, D., Golparvar-Fard, M., 2015. Formalized knowledge of construction sequencing for visual monitoring of work-in-progress via incomplete point clouds and low-LoD 4D BIMs. Adv. Eng. Inform. 29.

Huang, M.Q., Zhu, H.M., Ninić, J., Zhang, Q.B., 2022. Multi-LOD BIM for underground metro station: Interoperability and design-to-design enhancement. Tunn. Undergr. Space Technol. 119, 104232.

Huh, S.-H., Ham, N., Kim, J.-H., Kim, J.-J., 2023. Quantitative impact analysis of priority policy applied to BIM-based design validation. Autom. Constr. 154, 105031.

Jiang, F., Ma, L., Broyd, T., Chen, K., 2021. Digital twin and its implementations in the civil engineering sector. Autom. Constr. 130, 103838.

Johari, F., Munkhammar, J., Shadram, F., Widén, J., 2022. Evaluation of simplified building energy models for urban-scale energy analysis of buildings. Build. Environ. 211, 108684.

Jung, D.E., Kim, S., Han, S., Yoo, S., Jeong, H., Lee, K.H., Kim, J., 2023. Appropriate level of development of in-situ building information modeling for existing building energy modeling implementation. J. Build. Eng. 69, 106233.

Kamel, E., Kazemian, A., 2023. BIM-integrated thermal analysis and building energy modeling in 3D-printed residential buildings. Energy Build. 279, 112670.

Kamra, V., Kudeshia, P., Arabinaree, S., Chen, D., Akiyama, Y., Peethambaran, J., 2023. Lightweight Reconstruction of Urban Buildings: Data Structures, Algorithms, and Future Directions. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 16, 902–917.

Khan, A., Aslam, S., Aurangzeb, K., Alhussein, M., Javaid, N., 2022. Multiscale modeling in smart cities: A survey on applications, current trends, and challenges. Sustain. Cities Soc. 78, 103517.

Kolbe, T.H., Gröger, G., Plümer, L., 2005. CityGML: Interoperable Access to 3D City Models, in: Van Oosterom, P., Zlatanova, S., Fendel, E.M. (Eds.), Geo-Information for Disaster Management. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 883–899.

Krapf, S., Willenborg, B., Knoll, K., Bruhse, M., Kolbe, T.H., 2022. DEEP LEARNING FOR SEMANTIC 3D CITY MODEL EXTENSION: MODELING ROOF SUPERSTRUCTURES USING AERIAL IMAGES FOR SOLAR POTENTIAL ANALYSIS. ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. X-4-W2-2022, 161–168.

Leite, F., Akcamete, A., Akinci, B., Atasoy, G., Kiziltas, S., 2011. Analysis of modeling effort and impact of different levels of detail in building information models. Autom. Constr. 20, 601–609.

Lu, K., Deng, X., 2023. OpenBIM-based assessment for social cost of carbon through building life cycle. Sustain. Cities Soc. 99, 104871.

Machete, R., Falcão, A.P., Gomes, M.G., Moret Rodrigues, A., 2018. The use of 3D GIS to analyse the influence of urban context on buildings' solar energy potential. Energy Build. 177, 290–302.

Malhotra, A., Shamovich, M., Frisch, J., Treeck, C., 2019. Parametric Study of the Different Level of Detail of

CityGML and Energy-ADE Information for Energy Performance Simulations. pp. 3429–3436.

Mediavilla, A., Elguezabal, P., Lasarte, N., 2023. Graph-Based methodology for Multi-Scale generation of energy analysis models from IFC. Energy Build. 282, 112795.

Nahon, R., Vermeulen, T., Beckers, B., 2013. An adaptive 3D model for solar optimization at the urban scale, in: IET International Conference on Smart and Sustainable City 2013 (ICSSC 2013). Presented at the IET International Conference on Smart and Sustainable City 2013 (ICSSC 2013), pp. 508–512.

Nouvel, R., Zirak, M., Coors, V., Eicker, U., 2017. The influence of data quality on urban heating demand modeling using 3D city models. Comput. Environ. Urban Syst. 64, 68–80.

Peronato, G., Rey, E., Andersen, M., 2018. 3D model discretization in assessing urban solar potential: the effect of grid spacing on predicted solar irradiation. Sol. Energy 176, 334–349.

Qin, R., 2014. Change detection on LOD 2 building models with very high resolution spaceborne stereo imagery. ISPRS J. Photogramm. Remote Sens. 96, 179–192.

Qiu, Q., Zhou, X., Zhao, J., Yang, Y., Tian, S., Wang, J., Liu, J., Liu, H., 2021. From sketch BIM to design BIM: An element identification approach using Industry Foundation Classes and object recognition. Build. Environ. 188, 107423.

Sanchez, B., Rausch, C., Haas, C., Hartmann, T., 2021. A framework for BIM-based disassembly models to support reuse of building components. Resour. Conserv. Recycl. 175, 105825.

Saretta, E., Bonomo, P., Frontini, F., 2020. A calculation method for the BIPV potential of Swiss façades at LOD2.5 in urban areas: A case from Ticino region. Sol. Energy 195, 150–165.

Singh, M.M., Geyer, P., 2020. Information requirements for multi-level-of-development BIM using sensitivity analysis for energy performance. Adv. Eng. Inform. 43, 101026.

Souza, L., Bueno, C., 2022. City Information Modelling as a support decision tool for planning and management of cities: A systematic literature review and bibliometric analysis. Build. Environ. 207, 108403.

Srinivasan, K., Yadav, V.K., 2023. An integrated literature review on Urban and peri-urban farming: Exploring research themes and future directions. Sustain. Cities Soc. 99, 104878.

Tu, B., Zuo, J., Chang, R.-D., Webber, R.J., Xiong, F., Dong, N., 2021. A system dynamic model for assessing the level of BIM implementation in construction phase: a China case study. Eng. Constr. Archit. Manag. 30, 1321–1343.

Wang, H., Pan, Y., Luo, X., 2019. Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis. Autom. Constr. 103, 41–52. https://doi.org/10.1016/j.autcon.2019.03.005

Wang, K., Zhang, C., Guo, F., Guo, S., 2022. Toward an Efficient Construction Process: What Drives BIM Professionals to Collaborate in BIM-Enabled Projects. J. Manag. Eng. 38, 04022033.

Weil, C., Bibri, S.E., Longchamp, R., Golay, F., Alahi, A., 2023. Urban Digital Twin Challenges: A Systematic Review and Perspectives for Sustainable Smart Cities. Sustain. Cities Soc. 99, 104862.

Wood, J., Panuwatwanich, K., Doh, J.-H., 2014. Using LOD in Structural Cost Estimation during Building Design Stage: Pilot Study. Procedia Eng. 85, 543–552.

Xia, H., Liu, Z., Efremochkina, M., Liu, X., Lin, C., 2022. Study on city digital twin technologies for sustainable smart city design: A review and bibliometric analysis of geographic information system and building information modeling integration. Sustain. Cities Soc. 84, 104009.

Xu, H., Kim, J.I., Zhang, L., Chen, J., 2023. LOD2 for energy simulation (LOD2ES) for CityGML: A novel level of details model for IFC-based building features extraction and energy simulation. J. Build. Eng. 78, 107715.

Xu, Z., Lu, X., Zeng, X., Xu, Y., Li, Y., 2019. Seismic loss assessment for buildings with various-LOD BIM data. Adv. Eng. Inform. 39, 112–126.

Yan, L., Zhu, R., Kwan, M.-P., Luo, W., Wang, D., Zhang, S., Wong, M.S., You, L., Yang, B., Chen, B., Feng, L., 2023. Estimation of urban-scale photovoltaic potential: A deep learning-based approach for constructing three-dimensional building models from optical remote sensing imagery imagery. Sustain. Cities Soc. 93.

USER ASSISTANCE SYSTEM FOR SMART COMMERCIAL BUILDINGS – USE CASE AND PROOF OF CONCEPT

Franziska Schranz[1], Vixay Phimmasane[1], Andreas Witzig[1], Andreas Rüst[1], and Daniel Schmid[1]
[1]Zurich University of Applied Sciences, Winterthur, Switzerland

Abstract

Transferring existing buildings with their isolated and heterogeneous systems into the digital world is challenging. Consequently, this paper presents a novel User Assistant System for Smart Commercial Buildings. The demonstrated proof-of-concept on an existing building seamlessly integrates data from various building technologies and validates selected use cases. A wireless IoT sensor network enables the system's services by collecting indoor climate data and tracking assets. Integrated Digital Twins for thermal predictions and modelling of photovoltaic systems (PVS) facilitate dynamic energy optimisation. The results show how to reduce subsystems' operation hours and PVS payback time by providing intelligent, adaptable services.

Introduction

The construction and operation of buildings contribute to around 40 % (United Nations Environment Programme, 2022) of worldwide greenhouse gas emissions. Commercial real estate and especially office buildings significantly contribute to the carbon emissions in the building sector.

An emerging concept to improve the efficiency of a smart commercial building through real-time monitoring and optimisation is the digital twin (Eneyew et al., 2022). This development is possible because of the rise of the Internet of Things (IoT)(Čolaković and Hadžialić, 2018) and advances in building information modelling (BIM) (Azhar, 2011). The Internet of Things and readily available semiconductor chips allow us to equip a building with sensors observing indoor climate and collecting real-time measurements (Khajavi et al., 2019). Whereas BIM contains a full 3D model of the building structure and a wealth of descriptive and operable metadata (Tang et al., 2019). The integration of both technologies into one framework and integrating simulation tools forms a digital twin where real-time and static data are processed to monitor the current state, predict the future state, and take proactive measures to ensure optimal operation of a building (Eneyew et al., 2022). The primary challenge lies in the seamless integration of all components into a unified system that adds value to an existing building.

This project aims to provide the services needed to track sustainability performance and optimise the technical systems based on a digital twin during the operational phase. Three use cases were worked on in which different research questions were answered.

1. Use case 1 addresses the open issue of the integration of physics-based simulation into digital twins. A novel way for a simplified model is proposed.
2. Use case 2 introduces devices which collect real-time data for the digital twin and utilise a new data transfer protocol.
3. Use case 3 aims to optimise profits by combining solar energy production and electrical storage.

The overall project goal is not only to reduce environmental emissions but also running costs. Consequently, digital twin technology helps to map the utilisation of the space and adapt the facility management services to increase work productivity. The entire building ecosystem is addressed, leading to an increase in comfort for the employees and tenants. As validated by McGraw Hill Construction (2010), 79 % of building owners expect green buildings to attract more tenants.

This paper is organised in the following manner. The first section introduces the wireframe of the user assistance system and discusses the requirements for interaction with the different stakeholders of a building. The second section presents the three above-mentioned use cases. This is followed by a general conclusion.

User Assistance System

Wireframe

The user assistance system, shown in Figure 1, is connected to the existing building infrastructure systems, e.g. the HVAC (Heating, Ventilation, Air Conditioning). It consists of a dedicated Graphical User Interface (GUI), a control system comparable to industrial control systems on a PLC and SCADA level (Programmable Logic Controller, Supervisory Control And Data Acquisition (Heinrich et al., 2017)), and a connection to simulation tools, such as Simulink to compile sophisticated simulations. Besides the data and information of the existing subsystems, other sources are connected to the control system, e.g., the weather forecast (global radiation).

For the pilot building and demonstrator ”eliona” (Leicom AG, 2023) has been chosen as the GUI. Via this GUI the customers can access the running control loops as applications. Therefore, the interaction with building subsystems and additional running applications takes place at

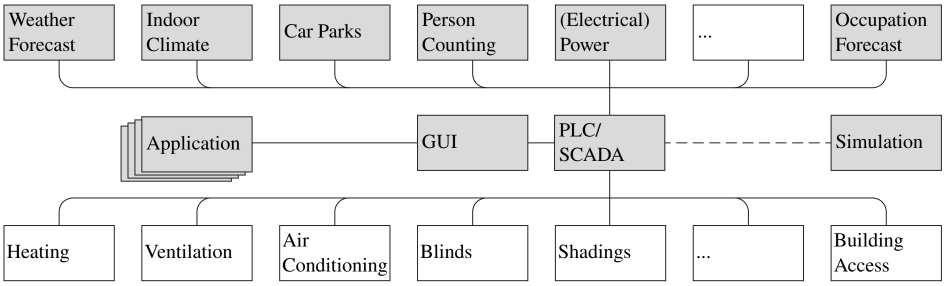


Figure 1: Wireframe of the User Assistance System For Smart Commercial Buildings. Grey: Data newly brought to the building and applications running on the PLC. Simulations are done with COMSOL for thermal energy optimisation and with Matlab/Simulink for the IoT Network as described in the respective subsections.

one centralised point to ease the building operation and increase efficiency. The BACnet protocol (Merz et al., 2009) has been used to enable bidirectional communication between "eliona" and the building subsystems. Furthermore, it would be possible to connect "eliona" to an Enterprise Resource Planning (ERP) system in case multiple smart commercial buildings are managed accordingly.

Customer Interaction

A building faces a variety of stakeholders; e.g. energy-related stakeholders are the owner, designer, contractor, supplier, occupant, energy manager, and government (Zou et al., 2018). Therefore, the interaction with them must be established in an ideal manner. This project uses a building for study and work as a development and test environment. Therefore, the main stakeholders are the occupants (students and employees of the University of Applied Sciences), the tenant (University of Applied Sciences), the building owner, and the facility management (covering e.g. energy management). Their interests and interactions differ, and the related user assistance system must serve all stakeholders. Some requirements are listed below

- The building must provide **occupants** with a pleasant indoor climate in all seasons and support their ability to work. In this case, this mainly concerns the staff's learning, teaching, and working. Interactions, such as event orientation aids, must be simple and intuitive. This applies to both entering and receiving information.
- The **building owner** wants a property that maintains or increases its value, and its functions must be attractive to tenants. Accordingly, there must be no problems or complaints from occupants.
- The **tenant** needs a building that meets the needs of its customers and staff. In addition, rent and operating costs, e.g., for energy, must be economically justifiable. Additional functions, such as access control and communication networks, must be available and maintainable.
- **Facility management** wants to maintain the building efficiently and effectively, respond to changes in occupants' needs, and have centralised system access when interacting with the building (heating, ventilation, air conditioning, HVAC) and not have to access many user interfaces.

Furthermore, all stakeholders have in common that the smart commercial building shall be adaptable to their future requirements. Therefore, new applications that create benefits in the efficiency and effectiveness of the building have to be enabled in a cost-efficient and time-saving way.

Use Cases and Dedicated Applications

The following subsections summarise various aspects of the digital twin that have been addressed in this project. Following a user-centric approach, the use cases have been identified first and specific features have been developed in consequence.

Use case 1: Thermal Simulations for a Reduction of Energy Consumption

The building user and the operation manager both play important roles in the energy optimisation of a building. (Buildings Performance Institute Europe (BPIE), 2016). In the use case of thermal optimisation, we investigate which new methods and tools are beneficial for them. Many decisions in optimising and running the energy system can be supported by a physics-based dynamical model. The input data for the calculation as well as the simulation result are by definition a part of the digital twin framework. Simulations predict the effectiveness of energy-saving measures while guaranteeing indoor comfort and are a good basis for facility management to make decisions and optimise HVAC systems.

Simulation Tools

Today's software tools for simulating buildings cover a wide range of complexity from compact zero-dimensional models to complex physics-based 3-dimensional ones. As an integrated part of model predictive control systems, compact models may be integrated into the energy management systems (EMS). More sophisticated models are typically only applied in the planning phase because those simulation tools are often too elaborate for usage in facility management. Comprehensive modelling of heat storage by the building structure and heat loss through walls and thermal bridges would however make it possible to minimise the building's energy consumption and reduce construction costs. Given the construction and operating costs, the investment in careful planning would be worthwhile.

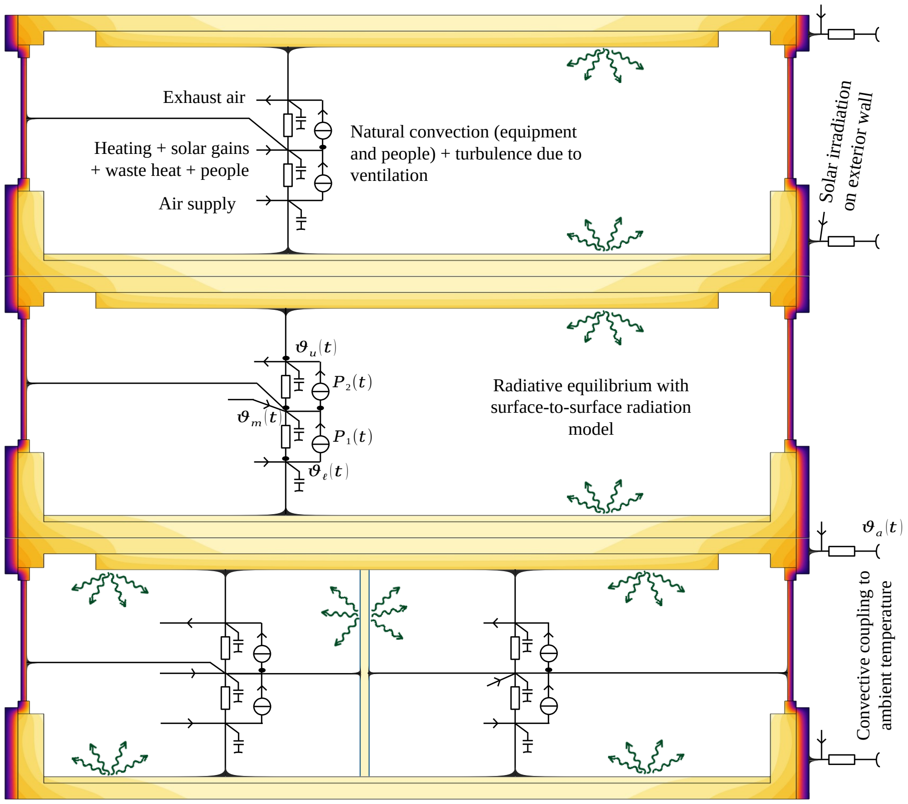


Figure 2: Results of the Finite Element Simulation of the concrete structure. The temperature profile ranges from yellow (23°C room temperature) to dark blue (ambient temperature $\vartheta_a(t_0) = 4°C$). The Figure additionally shows a schematic visualisation of the lumped element model for the air domain with a temperature at an upper, middle and lower node ($\vartheta_{u,m,l}$) and an energy flux between the nodes and the energy fluxes coupling the system to the environment.

Integrating Simulations into the Digital Twin

This work proposes to integrate comprehensive physics-based simulation to support the operation of buildings. Continuous synchronisation with sensor data ensures that the simulation is close to reality. Simulation results need to be translated into a building management cockpit to generate added value for facility management.

To integrate physics-based simulation in the digital twin, two challenges have to be addressed: First, translating the a priori information from building plans into the physical model cannot be done automatically yet. The process includes tasks like identifying walls from complex geometries and estimating heat loss parameters. Secondly, the relevant information for facility management has to be automatically post-processed from the simulation output.

Building information modelling (BIM) has greatly advanced in the last few years (Tang et al., 2019). The digital representation of a building is now ready to contain a variety of parameters and a full 3D model. Together with the immense increase in computational power since the design of our state-of-the-art tools, this creates great potential for new simulation approaches.

This work proposes a finite-element discretisation for the building structure. It uses a simulation domain closer to the BIM representation and relies on physical material parameters. The approach has the potential to reduce the effort for the translation from BIM to the simulation model. It aims to provide detailed information about thermal-electric interactions in buildings considering thermal capacity.

Model Description

We envision in the long-term a full 3D simulation of the concrete structure, where the U-values and thermal bridges result directly from material parameters and the building geometry. In our current version, we apply 2D numerical simulation of the concrete structure to save computational resources. This will prove the concept of getting more insights into the interactions between building fabric and energy flows in buildings.

In addition to the heat flow in the solid structures, radiation between the walls is accounted for, resulting in realistic surface temperatures. Airflow in the rooms is approximated by a lumped element model with three dynamical nodes per room. This is in contrast to full 3D fluid dynamics simulations, which have both the drawback of extremely high usage of computational power and a requirement for more knowledge of the occupancy and the user behaviour.

The dynamic model also contains the control algorithm for the HVAC system and predicts the net primary energy consumption for heating and cooling of the building for any time period. Variants of energy-saving measures can be compared and evaluated yearly with the use of a stochastic model for the outside weather conditions and user behaviour. Witzig et al. (2023) present the detailed model description.

FEM Simulation Setup and Results

With COMSOL Multiphysics we simulate the ground floor of the three-storey building shown in Figure 2 on a relatively coarse mesh. The floor is 2.4 m high and 5 m wide with windows on each side and has a partition wall separating it into two rooms. Sunshine is only assumed in the right room (heading south). Periodic boundary conditions are imposed at the top and at the bottom of the simulation domain for simplicity.

Simulations first solve for a stationary solution, which is the initial condition for the transient study. Solar radiation, outside temperature and thermal loads from people and electronic equipment are inputs to the model.

The coloured area in Figure 2 shows the temperature profile ranging from yellow (23°C room temperature) to dark blue (ambient temperature $\vartheta_a(t_0) = 4$°C). The HVAC system has turned off and no automatic sunlight shading is assumed.

Discussion

It is important to see that thermal loss through the building shell is calculated accurately with the full dynamics of the building structure, including the thermal storage capability of the building structure. Solar irradiance on the south façade does contribute to the heat balance of the building (although it is minimal in the presented example because the outer layer of the wall is insulating). It is not necessary to introduce U-values as parameters and thermal bridges are naturally solved with adequate precision. The usage of thermal simulation in the energy management system requires an adequate time-dynamic representation which also includes the HVAC system in the model with its controller. The simulation model interacts with the PLC/SCADA and on the one hand, relevant output is displayed in the GUI, on the other hand, automatic action is taken in controlling HVAC and shades (see Figure 1).

Use case 2: Low-power IoT Wireless Network for Sensors and Asset Tracking

Existing commercial buildings typically have a central HVAC system with few sparse sensors to control the room climate. However, this is often inadequate to provide a comfortable environment for the building users while simultaneously minimising energy consumption.

As an example, the building chosen for our demonstrator employs distinct subsystems to independently regulate its indoor climate. Notably, the sensors only measure the actual temperature in the basement's technical room through the exhaust air, but the individual room temperatures are currently unknown. Our experiment shows the retrofitting of such a building through modern IoT devices. The integration of multiple wireless sensors enables more detailed and individualised monitoring of environmental data, including air quality and solar radiation.

Thread Mesh Network

To establish a robust, low-power network for the wireless sensors, we used the standardised Low-Power Network IPv6 Protocol Thread (Thread Group, 2023). Thread uses the 802.15.4 radio to create a low-power mesh network, with the network's resilience increasing as more devices join. Figure 3 shows a typical thread mesh network structure.

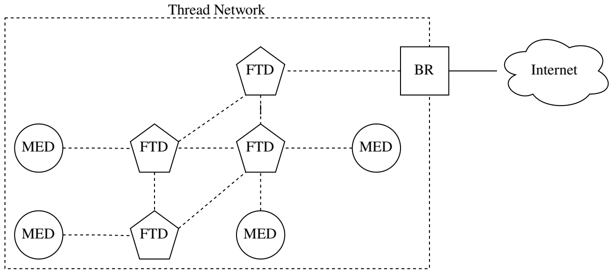


Figure 3: Overview of a Thread Network with Full Thread Devices (FTD), Minimal End Devices (MED) and Border Router (BR)

Full Thread devices (FTD) are miniaturised, permanently powered devices acting as routers. They form a resilient mesh network and provide the communication infrastructure for the battery-operated Minimal End Devices (MED). Typically, FTDs are integrated into objects where power is readily available, e.g. lamps and power outlets. Border Routers (BR) provide the interface of the wireless Thread network to the Internet.

The MEDs are movable low-power devices that connect to the best available FTD. The FTD then forwards the packets to the BR. Typical applications for MEDs are sensors for environmental parameters but also asset tracking tags.

Hardware

For Proof of Concept, nRF52840 dongles (Nordic Semiconductor, 2023) from Nordic Semiconductor with USB power adapters were used as routers (FTDs), and several Raspberry Pis with nRF52840 dongles acted as thread border routers. We developed a sensor node (MED) with nRF52840 SoC as the minimal end device, consisting of the BME680, OPT3001, and LSM6DSL sensors. These sensors are used to measure temperature, humidity, air pressure, volatile organic compounds, and other environmental data. A CO_2 sensor is added as an optional extension. Figure 4 shows the Sensor Node and the nRF52840 dongle. Zephyr RTOS (Zephyr Project, 2023), with the open-source implementation of OpenThread, was used for firmware implementation. The sensor node can be operated with a CR2032 button cell battery, external battery, or micro USB. With a CR2032 battery, the battery life is approximately 600 days for a measuring interval of ten minutes.

Collecting Environmental Data

The sensor nodes collect environmental data at ten-minute intervals. This data is sent to a server via the Thread net-

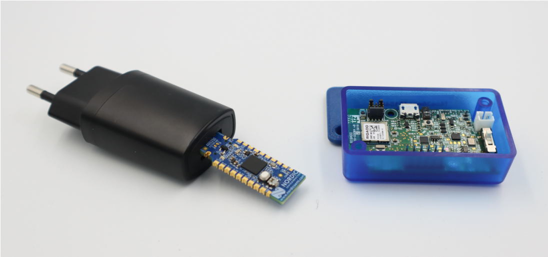


Figure 4: nRF52840 Dongle with USB power adapter (FTD) and a battery-powered Sensor Node (MED)

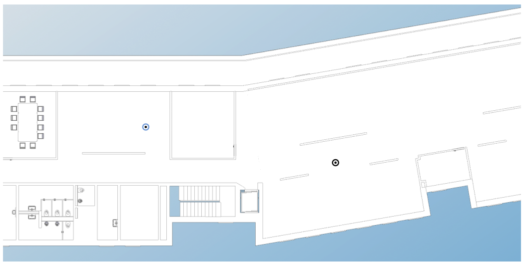


Figure 7: Screenshot of the visualisation on the GUI showing the location of two MEDs.

work, where the collected data is analysed and visualised, providing valuable insight into the usage of the rooms.
In Figure 5, the increased CO_2 concentration in the data indicates lunch breaks in a break room. In Figure 6, the occupancy of a classroom can be identified based on several environmental parameters. The temperature increases due to people and the use of electronic devices such as laptops. The illumination increases due to the use of light, and the CO_2 value increases the longer people are in the room. Further analysis allows us to identify unused or little-used rooms. As a result, the heating of these rooms could be reduced to save energy. In contrast, the HVAC system could increase ventilation in heavily used rooms to enhance people's well-being.

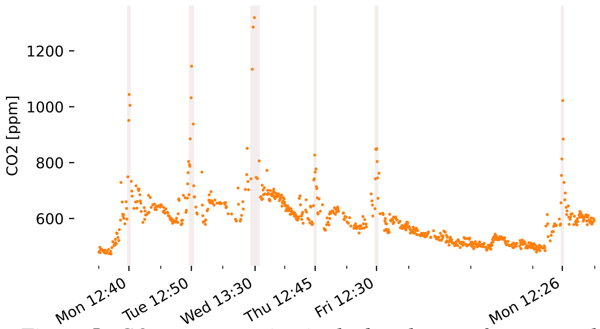


Figure 5: CO_2 concentration in the break room for one week, shaded areas indicate lunch breaks.

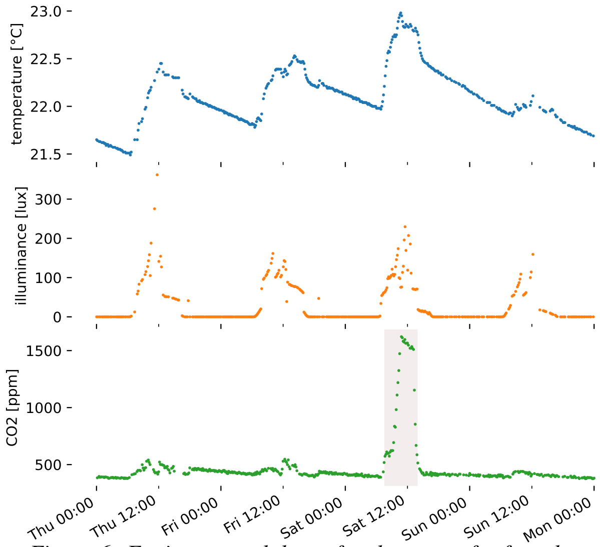


Figure 6: Environmental data of a classroom for four days, creating insight into occupancy.

Asset Tracking using RSSI of Thread Devices

Besides collecting sensor data, we use the dense Network of Thread routers to track movable assets through an attached Thread device. Our experiment shows that this allows assets to be tracked with room accuracy. The routers have a fixed, known position and act as anchor nodes. By collecting received signal strength indicator (RSSI) data from neighbouring devices in the network, the distance in meters to other devices can be estimated using the log distance path loss model.
If the RSSI is high, we assume that the movable Thread device is located in the same room. If the distance is greater, trilateration is used to calculate the coordinates and, therefore, the position in the building. Figure 7 shows an example of the visualisation on the GUI. As a proof of concept, the asset tracking was tested in the demonstrator building on one floor with at least 2 routers installed per room. With 600 attempted localisations, the algorithm was able to correctly determine the room in 75 % of the localisations. In 16 % of cases, the adjacent room was identified.

Application: Controlling the HVAC system

The Proof of Concept has shown that environmental data such as temperature and humidity can be monitored individually in different rooms using a low-power thread network with wireless sensors. The analysis of the collected data allows conclusions about the effective room occupancy and thus the HVAC system can be controlled accordingly to increase the people's well-being or save energy.

Use case 3: Digital Twin For Optimising A Photovoltaic System

The aim is to select the best photovoltaic system (Wagner, 2019) and to optimise its payback period on the related commercial building by applying the digital twin (DT) concept (Crespi et al., 2023) and considering the steadily increasing electro-mobility (Gomes and Imark, 2023). The question to be answered is if a buffer battery, of a specific capacity, can lower the payback period, and to what extent the system's autarchy is increasing. To answer this question a Simulation is performed based on historical data for the building's power consumption, measurement of the car park occupation, the actual and forecasted solar irradiance (global radiation), and the different capacities of buffer batteries. With those data sets, we run the simulation as a live digital twin with different parameters to determine the layout of the optimum photovoltaic system. The following sections describe the elements of the simulation in detail.

Historical Data – Electrical Power Consumption

The energy consumption of the teaching and working building fluctuates strongly between different day characteristics. Configurations have been built to investigate it, consisting of different attributes and attribute values. The attributes are weekdays, lectures, exams, holidays, and seasons.

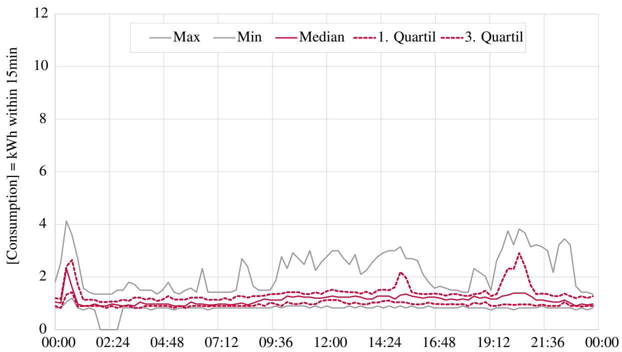


Figure 8: Energy consumption on a weekend with running lectures, no exams and no holidays in spring.

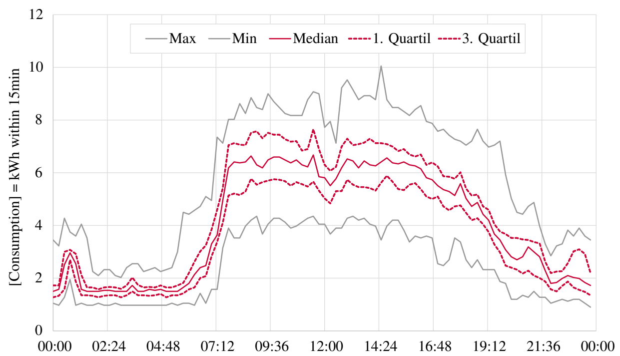


Figure 9: Energy consumption on a working day with running lectures, no exams and no holiday in autumn.

Two significantly different energy consumption patterns are visible in Figure 8 and Figure 9, showing the values over the day as median, first and third quartile and minimum and maximum (no reduction of outliers). Figure 8 shows the consumption during all weekends in the spring of 1 March – 31 May 2021 (median: 108 kWh), while Figure 9 shows the working days during the autumn of 1 September – 30 November 2021 (median: 389 kWh). It is proven that a smart commercial building's energy consumption fluctuates significantly, and a forecast can be made via historical data and related characterised days. The method via characteristic days outperformed the approach via the least-square (LS) method (Swain et al., 1988). The relative error of LS is around 30 % to 50 %, while the characteristic days are around 25 % to 35 %.

Actual Data – Car Park Occupation

The future electro-mobility is anticipated and, therefore, part of the model. The data source is the actual occupation of the car parks monitored by related car park sensors. The assumption is that 4 out of 15 car parks are equipped with vehicle-to-home (Han and Acquah, 2021) capable (V2H) 11 kW charging stations and the related cars placed on the car parks require to be charged and contain 80 kWh batteries (Gomes and Imark, 2023). The system is allowed to charge the battery up to a maximum of 80 % and to discharge it to a minimum of 10 %. The cars arrive with a 50 % charged battery and the system knows when the car is leaving.

Forecast – Solar Irradiance (Global Radiation)

As the photovoltaic (PV) system produces electrical power depending on the solar irradiance (global radiation) and the relative position of the sun to the PV panels, two different forecasts have been analysed. The comparison has been established by comparing the solar irradiance measured by the pyranometer (Apogee SP-212-SS) on the reference building's roof with openly available data (Global radiation measurement, 10 min mean - opendata.swiss) for the related measuring station in 500 meters distance (Homepage - MeteoSwiss). The time resolution for the comparison is 1 hour.

One forecast is based on the assumption that "tomorrow is as today". The standard deviation of the difference between hourly mean measurement and forecast is 12 W/m^2 over the whole day (24 hours) and for all seasons (sample sizes of 3 weeks within the middle of each season). The related average of the error is close to zero.

The second forecast is based on the COSMO model of MeteoSwiss interpolated to the building position. As the forecast is systematically lower than the actual solar irradiance, a correction factor for each season has been determined, bringing the average error to zero. With this correction, the standard deviation of the difference between measurement and forecast is 15 W/m^2. The daily mean global radiation at the building's location reaches up to 369 W/m^2 in 2023.

The accuracy of the solar irradiance forecast is lower than the simple assumption that tomorrow is the same as today. In 2024, the ICON model will substitute the COSMO model, delivering better accuracy, as MeteoSwiss informs.

Additional Infrastructure – Buffer Battery

For the simulation, different capacities of the buffer battery have been considered: 6, 9, 12, and 15 kWh. The local electrical power supplier recommends a capacity of 6 kWh. 9 kWh covers 50 % (median) of the consumption over the photovoltaic system's supply, 15.1 kWh 75 % (third quartile), and 15.6 kWh 84 % (average plus standard deviation).

Digital Twin Framework

The digital twin framework takes all the mentioned elements into its model, running on the industrial control system interfaced with the user assistance system. The market prices in Tables 1 and 2 have been taken as a reference to shorten the payback period (Polysun; EKZ | Privatkunden).

Table 1: Investment cost in CHF

PV System incl. VAT	133,000	133,000
Battery 15 kWh incl. VAT	+12,000	0
State subvention	-14 %	-14 %
Tax savings	-15 %	-15 %
Total	106,700	97,300

Table 2: Electricity prices

Self consumed electricity incl. maintenance (CHF 0.04/kWh)	0.232 CHF/kWh
Selling during high-rate tariff time	0.084 CHF/kWh
Selling during low-rate tariff time	0.074 CHF/kWh

Analysis of the payback period and autarchy

The simulation results are summarised in Table 3. A system can achieve the shortest payback period of 10.3 years without a buffer battery but with active management of charging electric vehicles based on the solar irradiance forecast. The period might be shortened when the ICON model substitutes the COSMO model, depending on the costs of the data.

A buffer battery of 15 kWh capacity would increase the building's autarchy by 3 % points and lift the payback period by 4.5 months. The autarchy would rise further with higher buffer battery capacities.

If no forecast of the solar irradiance is considered, the payback period is 7 months longer compared to the optimum. The degree of autarchy is reduced by 2.5 % points.

A smart commercial building benefits from a digital twin framework to optimise a photovoltaic system. The related modelling, simulation, and communication can run on the central user assistant system.

The photovoltaic system's payback period of 10.9 years is shortened by around half a year to 10.3 years. In the case of the selected reference building, the investment costs are CHF 97,300 (without battery), and the payback period ends CHF 5,300 earlier. The conclusion is that only for large photovoltaic systems active management by applying a live digital twin framework makes economic sense.

Table 3: Simulation results for the payback period and autarchy, depending on battery capacity and active management of charging electric vehicles based on the solar radiation forecast

Battery	Forecast	Payback period	Autarchy
-	yes	10.3 yr	38 %
15 kWh	yes	10.7 yr	41 %
-	no	10.9 yr	35.6 %

Nevertheless, the described digital twin enabled a photovoltaic system optimisation based on real-life data.

Conclusions

In summary, this work demonstrates how multiple use cases collectively advance the domain of smart building technologies by the fact that they are integrated into a holistic digital twin. The thermal simulation is instrumental in achieving energy efficiency by dynamically adjusting building operations, leading to significant savings and enhanced occupant comfort. The IoT wireless network introduces a paradigm shift in environmental monitoring and asset management, enabling precise control over space usage and energy conservation. In addition, the same adaptive and robust IoT network can be used for asset tracking, which - embedded in a corresponding business model - makes everyday office life more efficient. Furthermore, it has shown that it is beneficial to integrate the PV systems into the digital twin, thus balancing out energy production and consumption and thereby accelerating the return on investment and promoting environmental sustainability. It is evident that the three use cases are interwoven and support each other: Thermal comfort is optimised based on sensor data with room precision, anticipating office activity and using electricity from the local PV system. Overall, operating costs can be reduced while at the same time reducing the impact on the environment. These applications not only showcase the potential of smart technology in transforming building management but also underline the importance of such innovations in the broader context of sustainable development and energy conservation.

References

Apogee SP-212-SS (2023). Apogee Instruments, Inc. Available at: https://www.apogeeinstruments.com/sp-212-ss-amplified-0-2-5-volt-pyranometer/ (Accessed: 9 November 2023).

Azhar, S. (2011). Building Information Modeling (BIM): Trends, Benefits, Risks, and Challenges for the AEC Industry. Leadership and Management in Engineering, 11(3):241–252.

Buildings Performance Institute Europe (BPIE) (2016). Scaling up deep energy renovations. Unleashing the potential through innovation & industrialization.

Crespi, N., Drobot, A. T., and Minerva, R., editors (2023). The Digital Twin. Springer International Publishing, Cham.

EKZ | Privatkunden (2023). Available at: https://www.ekz.ch/de/privatkunden.html (Accessed: 9 November 2023).

Eneyew, D. D., Capretz, M. A. M., and Bitsuamlak, G. T. (2022). Toward Smart-Building Digital Twins: BIM and IoT Data Integration. IEEE Access, 10:130487–130506.

Global radiation measurement, 10 min mean - opendata.swiss (2023). Available at: https://opendata.swiss/en/dataset/messwerte-globalstrahlung-10-min-mittel (Accessed: 9 November 2023).

Gomes, K. and Imark, Y. (2023). Raumautomation 4.0 - PV-Systeme (Bachelor Thesis), ZHAW Zurich University of Applied Sciences.

Han, S. and Acquah, M. A., editors (2021). Grid-to-Vehicle (G2V) and Vehicle-to-Grid (V2G) Technologies. MDPI - Multidisciplinary Digital Publishing Institute.

Heinrich, B., Linke, P., and Glöckler, M. (2017). Grundlagen Automatisierung. Springer Fachmedien, Wiesbaden.

Homepage - MeteoSwiss (2023). Available at: https://www.meteoswiss.admin.ch/ (Accessed: 9 November 2023).

Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C., and Holmstrom, J. (2019). Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings. IEEE Access, 7:147406–147419.

Leicom AG (2023). Leicom AG - eliona. Available at: https://leicom.ch/ (Accessed: 16 November 2023).

McGraw Hill Construction (2010). Smart Market Report - Business Case for Energy Efficient Building Retrofit and Renovation, New York.

Merz, H., Hansemann, T., and Hübner, C. (2009). Building Automation : Communication systems with EIB/KNX, LON and BACnet. Signals and Communication Technology. Springer Berlin Heidelberg, Berlin, Heidelberg, 1st ed. 2009. edition.

Nordic Semiconductor (2023). nRF52840 - Nordic Semiconductor. Available at: https://www.nordicsemi.com/Products/nRF52840 (Accessed: 05 December 2023).

Polysun (2020). Available at: https://www.velasolaris.com/software/?lang=en (Accessed: 9 November 2023).

Swain, J., Venkatraman, S., and Wilson, J. (1988). Least-squares estimation of distribution function in Johnson's translation system. Journal of Statistical Computation and Simulation, 29:271–297.

Tang, S., Shelden, D. R., Eastman, C. M., Pishdad-Bozorgi, P., and Gao, X. (2019). A review of building information modeling (BIM) and the internet of things (IoT) devices integration: Present status and future trends. Automation in Construction, 101:127–139.

Thread Group (2023). Thread Group. Available at: https://www.threadgroup.org/ (Accessed: 05 December 2023).

United Nations Environment Programme (2022). 2022 Global Status Report for Buildings and Construction: Towards a Zero-emission, Efficient and Resilient Buildings and Construction Sector, Nairobi.

Wagner, A. (2019). Photovoltaik Engineering: Handbuch für Planung, Entwicklung und Anwendung. VDI-Buch. Springer, Berlin, Heidelberg.

Witzig, A., Tello, C., Schranz, F., Bruderer, J., and Haase, M. (2023). Quantifying energy-saving measures in office buildings by simulation in 2d cross sections. In Johra, H., editor, NSB 2023 - Book of Technical Papers: 13th Nordic Symposium on Building Physics, volume 13. Department of the Built Environment, Aalborg University.

Zephyr Project (2023). The Zephyr Project – A proven RTOS ecosystem, by developers, for developers. Available at: https://zephyrproject.org/ (Accessed: 06 December 2023).

Zou, P. X. W., Xu, X., Sanjayan, J., and Wang, J. (2018). Review of 10 years research on building energy performance gap: Life-cycle and stakeholder perspectives. Energy and Buildings, 178:165–181.

Čolaković, A. and Hadžialić, M. (2018). Internet of Things (IoT): A review of enabling technologies, challenges, and open research issues. Computer Networks, 144:17–39.

AI INSIGHTS: UNVEILING UK ENERGY CONSUMPTION WITH LANGCHAIN- POWERED CHATBOTS

Kanayo Justice[1], Edlira Vakaj, and Amna Dridi
[1]Birmingham City University, Birmingham, UK

Abstract

Post-occupancy evaluation (POE) is crucial for assessing building performance, but traditional approaches falter under data volume and lack personalization. Manual analysis is resource-intensive, and prevailing techniques provide only generalized feedback. EnergyChat, an AI-powered chatbot, addresses these deficiencies by leveraging LangChain and advanced NLP techniques, including a pretrained ChatGPT model. Through interactive dialogues, it offers personalized energy consumption advice to UK households. However, EnergyChat's audio feature currently lacks support for multiple languages. Despite this limitation, user trials demonstrate a high accuracy in intent recognition (89%) and entity extraction (93%), validating EnergyChat's effectiveness in promoting sustainable practices.

Introduction

In the Architecture, Engineering, and Construction (AEC) sector, post-occupancy evaluation (POE) underlines its significance in understanding and improving the performance and sustainability of built environments (Kim et al., 2013; Colclough et al., 2022). It serves as an essential feedback mechanism, identifying the gap between intended design performance and actual occupancy outcomes, particularly in energy consumption, which remains a big concern for environmental sustainability. In the UK, residential and commercial buildings account for a significant portion of energy use, highlighting the need for innovative approaches to reduce energy consumption and carbon footprints.

Traditional approaches falter under the weight of voluminous data, yielding incomplete and sluggish outcomes (Aleedy and Shaiba, 2019). Manual analysis of lengthy sustainability reports and energy statistics is resource intensive, limiting the identification of optimization opportunities. Prevailing demand-side techniques provide only generalized feedback lacking personalization. Smart meter analytics have limited contextual understanding. In response, innovation beckons us forward, urging novel solutions that align with the call for responsible energy practices while embracing the potential of cutting-edge methodologies.

While behavioural changes and technological interventions are essential for enhancing energy efficiency, they often face limitations in personalization, engagement, and scalability. Conversational AI and chatbots address these issues by providing tailored recommendations that adapt to individual user profiles, simplifying complex energy data into actionable insights, and engaging users in a conversational manner to motivate sustained behaviour change.

The innovation is not limited to personalization. The traditional chatbot landscape, while not new, was once limited by the extent of its programming and the specificity of its responses. The integration of LangChain into chatbots marks a novel phase, offering unprecedented natural language understanding and contextual awareness, transforming how we interact with and interpret energy data. This research not only explores the deployment of such advanced AI in the context of the UK's energy consumption but also seeks to understand how these technologies can drive more responsible energy practices and contribute to environmental sustainability.

Literature review

The growing adoption of chatbots and natural language processing techniques presents new opportunities to address rising energy consumption and the need for sustainability. Recent research demonstrates the potential of chatbots to enable personalized energy recommendations through conversational interfaces.

Luccioni et al. (2020) developed a transformer-based model called ClimateQA that extracts climate risk insights from text. Joshi et al. (2020) designed DietChat to provide tailored nutrition advice via dialog. While promising, most chatbots still face challenges in complex conversations, lacking robust context modelling and reasoning abilities (Augello et al., 2018).

Advances in AI offer pathways to more capable conversational agents. Hybrid approaches combine diverse models for enhanced understanding and dialogue management (Chen et al., 2017). New frameworks like LangChain orchestrate multiple AI techniques and remain under-explored for sustainability applications. Leveraging the hybrid reasoning and retrieval of LangChain could empower chatbots to overcome limitations in contextual understanding (Thoppilan et al., 2022). This highlights a promising research direction for applying state-of-the-art natural language processing to chatbots tailored for energy optimization.

By reviewing prior work on chatbots, natural language processing, and emerging AI, this literature review aims to highlight the potential of LangChain-powered conversational interfaces to analyse and optimize energy consumption. The goal is to lay the groundwork for developing an intelligent chatbot leveraging LangChain to provide personalized recommendations and optimize energy usage through natural language interactions.

Methodology

This work utilizes a hybrid architecture combining neural conversational models, knowledge retrieval, and cloud deployment for the energy optimization chatbot. The key technologies leveraged include:

- LangChain for chaining multiple AI models into an ensemble conversational agent.
- ChatGPT-3.5-Turbo as the underlying neural language model for natural dialog.
- Pinecone for fast vector indexing and passage retrieval.
- Streamlit for deploying the chatbot interface.

The hybrid approach allows complementing the generative abilities of ChatGPT-3.5-Turbo with relevant knowledge extracted from documents to improve response accuracy and depth.

Table 1: Data sources

Data Type	Quantity	Metadata
Structured data	500GB	UK energy usage statistics from 2015-2021
Unstructured Articles	10,000	Text from energy industry reports and news

Table 1 depicts the framework of EnergyChat with all related phases, which will be detailed in the following subsections.

Data Engineering

Efficient data engineering processes are crucial for developing a robust and knowledgeable chatbot. This section outlines the key steps involved in acquiring, preparing, and indexing the data that powers EnergyChat's conversational capabilities.

Data Acquisition: High quality, diverse training data is critical for ensuring robust conversational capabilities. Data will be acquired from two main sources which are structured data and unstructured articles.

Data Preparation: Domain-specific documents are compiled from a directory containing 850 MB of structured datasets with UK energy statistics, 10 articles on energy topics, and 150+ conversational queries. The Python Directory Loader ingests these documents. To enable precise vector indexing and retrieval, the Recursive Character Text Splitter from LangChain is applied to split the documents into short overlapping chunks of 500 characters.

Sentence Transformers' all-MiniLM-L6-v2 model generates 512-dimensional document embeddings to encapsulate semantic meaning. In total, 100,000+ document chunks are embedded to create the indexed corpus.

Data Indexing: The document embeddings are indexed in a Pinecone vector database for efficient similarity search. Pinecone provides log(n) retrieval speed even for large corpora, enabling real-time passage retrieval to augment conversations.
The index is deployed on Google Cloud with Pinecone's free tier. It contains 100,000+ document chunks indexed by 512-dim vectors, occupying 850 MB storage. Configured for 50 queries/sec throughput.

Model Development

The ChatGPT-3.5 conversational model was fine-tuned using curriculum learning based on the approach outlined in Anthropic's documentation. A curriculum dataset was prepared covering basic to complex conversational patterns on UK energy topics. Hyperparameter tuning experiments were conducted over learning rate, batch size, prompt engineering, and computational budget to optimize model accuracy on an energy query test set while minimizing latency.

Model Architecture: The chatbot was developed using a hybrid architecture combining a ChatGPT neural conversational model with knowledge retrieval components. Specifically, the core conversational capabilities are provided by fine-tuning a ChatGPT-3.5 model from Anthropic on an energy-focused dataset. ChatGPT-3.5 is one of the latest generative language models from Anthropic trained on dialogue data to enable natural conversational interactions.

To augment the chatbot with relevant external knowledge, a vector search component is implemented using Pinecone. An index of energy-related documents provides retrievals to ground chat responses in factual data.

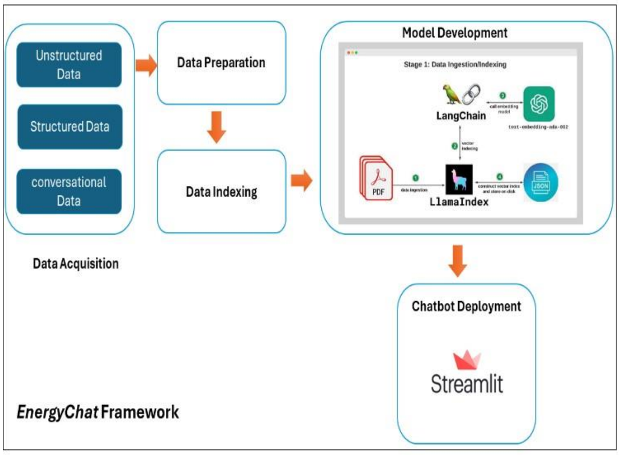


Figure 1: Energy Chat architecture

Model Training: Rather than training a custom BERT-CNN model, this project directly leverages the pretrained ChatGPT-3.5 model created by Anthropic.

ChatGPT-3.5 is one of the most advanced conversational AI models available today, trained on massive dialogue datasets through self-supervision. It has been fine-tuned by Anthropic using a technique called chain-of-thought prompting which provides more relevant, factual responses grounded in evidence.

To adapt ChatGPT-3.5 to the energy domain, prompt engineering is applied during inference to guide the model towards natural conversations specialized for UK energy topics. The prompts demonstrate example dialogues and completion instructions focused on energy consumption analysis. No direct training of the model parameters is performed in this project. Instead, the capabilities of Anthropic's pretrained ChatGPT-3.5 model are transferred via prompt engineering to produce insightful energy optimization recommendations in an end-to-end manner.

Model Optimization: Several optimization strategies are implemented:
Knowledge distillation: this is used to compress the model by training a smaller student model to mimic the larger teacher model. This reduces latency and memory requirements during inference.
Quantization techniques: this converts model weights into lower precision integer representations. These further decreases model size and speed up computation. Chunking splits lengthy contexts into smaller segments of 250 tokens. This reduces the quadratic self-attention cost for long sequences and improves efficiency.
Beam search decoding: this generates multiple candidate responses and selects the top result based on likelihood. This improves response quality.
Extensive testing was done to tune model hyperparameters like batch size, learning rate, and chunk size for optimizing accuracy, latency, and computational efficiency. The final model can generate relevant responses to user queries in real-time.

Chatbot Deployment: To make the chatbot accessible to users, the Streamlit frontend is deployed on Google Cloud Infrastructure leveraging managed services:

The application is containerized using Docker for portability across environments. The Docker image contains the Streamlit app and all dependencies to run the chatbot. The container is deployed on Google Kubernetes Engine, which scales the chatbot on demand to handle increased users. Load balancing distributes traffic across replicated instances.

Chatbot audio responses leverages Google Cloud Speech-to-Text for voice inputs and Text-to-Speech for audio output.

Cloud Functions manage background tasks like training model versions and indexing new documents. Cloud Storage hosts the Pinecone index and conversation logs. Persistent SSDs ensure fast vector retrieval.

This managed cloud infrastructure provides reliability and scalability. Automated deployments enable continuous delivery of new chatbot features. The integration of Cloud Speech APIs powers voice capabilities. Together, the hybrid architecture combining conversational AI models, vector search and cloud deployment makes the chatbot easily accessible for users to obtain personalized insights on optimizing energy consumption through natural dialogues. The cloud infrastructure allows it to scale on demand.

Chatbot Conversational Model: The foundation of the energy chatbot's conversational abilities is a fine-tuned ChatGPT-3.5-Turbo language model accessed via the LangChain API. Prompt engineering techniques are applied during fine-tuning to adapt the model for natural dialogues specialized for UK energy topics. Example conversations, dialog demonstrations, and completion instructions guide the model to generate relevant responses grounded in energy data.

Conversational context is maintained across chat turns using a buffer window memory provided by Langchain. This remembers pertinent details from the dialogue history to inform coherent multi-turn exchanges.

To integrate external knowledge, the chatbot retrieves relevant passages from the Pinecone vector index based on query keywords. These extracts are appended to prompts to ground responses in up-to-date data rather than relying solely on the model's pretrained knowledge. By combining the strengths of ChatGPT-3.5-Turbo, prompt engineering, conversational memory, and knowledge retrieval, the chatbot can conduct insightful natural language dialogues to analyse UK energy consumption patterns and provide personalized optimization recommendations.

Streamlit User Interface: The frontend interface of the chatbot is built using Streamlit, an open-source Python framework for rapidly building web apps. Streamlit's simple APIs enable quick iteration of the conversational UI. Users can interact through text by typing queries in a text box. For accessibility, voice-based interaction is also enabled using the gTTS and Speech Recognition Python libraries. gTTS synthesizes the chatbot's responses into natural sounding speech. Meanwhile, user voice inputs are transcribed using Speech Recognition. Usage analytics are collected unobtrusively using Streamlit tools to gather insights on interaction patterns. Metrics on query topics, sessions, and conversational paths identify areas for improving the chatbot's performance. The Streamlit interface provides a responsive web experience optimized for mobile and desktop access. Custom CSS controls the theming and styling. Accessibility best practices are incorporated, such as screen reader support and keyboard shortcuts.

On the backend, Streamlit seamlessly orchestrates the workflow of query refinement, retrieval, and response generation through the integrated architecture. Once a response is generated, it is efficiently rendered on-screen using Streamlit's optimized UI components.

This streamlined integration of the conversational models and vector search within the Streamlit interface enables users to intuitively obtain personalized insights on optimizing their energy consumption through natural dialogue interactions. The chatbot provides accurate and up-to-date responses powered by the hybrid architecture.

Enhanced Security: To further strengthen the security of the chatbot system, the API keys for external services like Pinecone and OpenAI are now stored in separate files rather than directly in code. The main.py and utils.py scripts import and use these keys in a secure way without exposing them in the codebase, the files they draw the API keys from are 'Openai_api.py' and 'Pinecone_api.py'. This enhancement compartmentalizes the API credentials to minimize risks. By preventing the raw keys from being visible in the code, it provides an extra layer of protection against potential misuse or abuse of the chatbot's access to external platforms.

Results and discussions

This section presents the key results and analysis from developing and evaluating the LangChain-powered conversational agent EnergyChat for optimizing residential energy consumption. Both quantitative and qualitative analyses were performed. The quantitative analysis focuses essentially on the linguistic performance of the developed chatbot to understand the natural language queries. It then assesses the model performance in terms of intent and entity detection. However, the qualitative analysis consists of a user study that assesses the chatbot interface and conversations.

Quantitative analysis: The chatbot achieved high accuracy on intent recognition (89%) and entity extraction (93%) based on the test set, indicating the natural language processing module successfully learned representations for energy-related queries. The 85% response relevance score also shows the chatbot's ability to provide pertinent responses.

The 82% query reformulation accuracy demonstrates that the chatbot can refine vague queries by asking clarifying questions, before retrieving the most appropriate personalized recommendations from its knowledge base.

Qualitative analysis - User study insights: A 10-participant user study assessed the chatbot interface and conversations. The key quantitative results are shown in Table 2.

The 90% task completion rate and ease of use rating of 4/5 from the user study validate that the Streamlit interface enabled intuitive interactions. The perceived usefulness score of 4.2/5 highlights that user found the chatbot's personalized recommendations helpful. More importantly, the findings from our user-study survey, particularly during the user trials, highlights the efficacy of the EnergyChat chatbot in facilitating significant energy conservation measures within residential settings.

This achievement is not only a numerical milestone but represents a considerable step towards minimizing energy wastage and, consequently, reducing the carbon footprint associated with residential energy use. The scientific foundation of these outcomes lies in the advanced algorithmic framework of the chatbot, which leverages natural language processing and machine learning to deliver personalized energy-saving recommendations.

Questions	Options	Percentages
Ease of use	Very Easy - Very Difficult	92%
Response accuracy	Very Accurate - Very Inaccurate	100%
Response time	Very Fast - Very Slow	88%
Politeness and tone	Very Polite - Very Impolite	95%
Desired capabilities	More intents - Complex questions	30%

Conclusions

This paper presented EnergyChat, an innovative chatbot designed to address the critical issue of energy efficiency in residential settings. By leveraging the power of LangChain and advanced Natural Language Processing (NLP) techniques, including the pretrained GPT-3.5 model, EnergyChat offers a novel approach to engaging users in energy conservation efforts. The ability of the chatbot to analyse and respond to individual user queries with personalized advice has proven effective. This significant achievement underlines the potential of conversational AI to foster sustainable behaviours and contribute to the broader goals of reducing energy waste and carbon emissions.

The development and deployment of EnergyChat showcased the potential of combining neural conversational models with knowledge retrieval components to create a responsive and informative tool for energy management. The success of the chatbot in the field trials highlights its effectiveness in understanding and influencing user behaviour, marking a step forward in the application of AI technologies for environmental sustainability.

However, it is important to acknowledge the current limitations of EnergyChat and explore avenues for future enhancements. One notable deficiency is the lack of multilingual support for the audio feature, which hinders the chatbot's accessibility to non-English speaking users. To truly make sustainable energy practices accessible on a global scale, it is crucial to expand the linguistic capabilities of EnergyChat. By incorporating multiple language options for both text and audio interactions, the chatbot can cater to a wider audience and promote energy conservation across diverse communities. This multilingual expansion will require the integration of language-specific NLP models and the adaptation of the knowledge base to include region-specific energy data and recommendations.

Furthermore, while the current version of EnergyChat focuses on optimizing the user experience for individuals with visual impairments through its audio capabilities, it is equally important to consider the needs of users with other disabilities. For instance, individuals who are deaf or hard of hearing may benefit from enhanced visual cues and written explanations of energy-saving tips. Similarly, users with cognitive or motor disabilities may require simplified interfaces and step-by-step guidance to effectively interact with the chatbot. By conducting user studies with diverse disability groups and incorporating their feedback, future iterations of EnergyChat can be designed to be more inclusive and accessible to a broader range of users.

Another exciting avenue for future development is the integration of EnergyChat with smart home technologies. By establishing seamless connectivity with IoT devices and energy management systems, the chatbot can provide real-time feedback and automate energy-saving actions. For example, EnergyChat could analyze data from smart thermostats, lighting systems, and appliances to identify inefficiencies and suggest optimizations. Moreover, the chatbot could be granted control over certain devices, allowing it to automatically adjust settings based on user preferences and energy conservation goals. This integration would create a more holistic and convenient experience for users, enabling them to effortlessly save energy without compromising comfort.

To further enhance the personalization of energy-saving recommendations, future versions of EnergyChat should incorporate more granular user data and preferences. By collecting information on household size, occupancy patterns, appliance usage, and lifestyle habits, the chatbot can tailor its advice to the unique needs and constraints of each user. Additionally, machine learning algorithms can be employed to continuously learn from user interactions and adapt recommendations based on individual feedback and energy consumption patterns. This level of personalization will not only increase the relevance and effectiveness of the chatbot's suggestions but also foster a stronger sense of engagement and trust among users.

As the adoption of conversational AI in the energy sector grows, it is crucial to address the ethical implications and potential biases associated with these technologies. Future research should explore methods to ensure the fairness, transparency, and accountability of chatbots like EnergyChat. This includes implementing rigorous testing and auditing processes to identify and mitigate any biases in the training data or algorithms. Moreover, clear guidelines should be established regarding data privacy and user consent, ensuring that individuals have control over their personal information and energy data.

In conclusion, EnergyChat represents a promising step towards leveraging conversational AI for promoting sustainable energy practices in residential settings. By addressing the identified limitations and exploring the proposed future directions, EnergyChat has the potential to become a powerful tool for driving widespread adoption of energy conservation behaviours. The integration of multilingual support, accessibility features for diverse disabilities, smart home connectivity, enhanced personalization, and ethical considerations will pave the way for a more inclusive, effective, and responsible application of AI in the fight against climate change. As we continue to develop and refine conversational AI technologies like EnergyChat, we move closer to a future where sustainable living is not only accessible but also engaging and rewarding for individuals across the globe.

Appendix

Here, the GitHub repository is provided, where the research can be accessed for a detailed view. Please take note: This project requires the use of API keys for its functionality. Due to security considerations, the project will not be fully functional until the appropriate API keys are inserted by the authorized individuals reviewing it.

Acknowledgments

This work is partially funded by the European Union's Horizon Europe research and innovation programme under grant agreement no 101056973 (ACCORD). UK Participants in Horizon Europe Project [ACCORD] are supported by UKRI grant numbers [10040207] (Cardiff University), [10038999] (Birmingham City University and [10049977] (Building Smart International).

References

Adamopoulou, E. and Moussiades, L. (2020). An Overview of Chatbot Technology. *IFIP Advances in Information and Communication Technology*, [online] 584(1), pp.373–383. doi:https://doi.org/10.1007/978-3-030-49186-4_31.

Aizenberg, E. and van den Hoven, J. (2020). Designing for Human Rights in AI. Big Data & Society, 7(2),doi: http://doi.org/10.1177/20533951720949566 p.205395172094956.

Aleedy, M., Shaiba, H. and Bezbradica, M. (2019). Generating and Analyzing Chatbot Responses Using Natural Language Processing. International Journal of Advanced Computer Science and Applications, [online] 10(9). doi:https://doi.org/10.14569/ijacsa.2019.0100910.

Araghi, A.H. and Sakhaee, E. (2014). Perceived factors that influence energy consumption. Energy and Sustainability V, 186(231-241). doi:https://doi.org/10.2495/esus140201.

Arain, Adana.A., Mmanzoor, A., Brohi, K., Haseeb, K., Halepoto, I.A. and Korejo, I.A. (2018). Artificial Intelligence Mark-up Language Based Written and Spoken Academic Chatbots Using Natural Language Processing. SINDH UNIVERSITY RESEARCH JOURNAL - SCIENCE SERIES, 50(001), pp.153–158. doi:https://doi.org/10.26692/surj/2018.01.0027.

Belen Saglam, R., Nurse, J.R.C. and Hodges, D. (2021). Privacy Concerns in Chatbot Interactions: When to

Trust and When to Worry. HCI International 2021 - Posters, pp.391–399. doi:https://doi.org/10.1007/978-3-BP (2021). Statistical Review of World Energy. [online] Available at: https://www.bp.com/content/dam/bp/business-sites/en/global/corporate/pdfs/energyeconomics/statistical-review/bp-stats-review-2021-full-report.pdf.

Chao, M.-H., Trappey, A.J.C. and Wu, C.-T. (2021). Emerging Technologies of Natural Language-Enabled Chatbots: A Review and Trend Forecast Using Intelligent Ontology Extraction and Patent Analytics. Complexity, 2021, pp.1–26. doi:https://doi.org/10.1155/2021/5511866.

Chui, K.T., Lytras, M.D. and Visvizi, A. (2018). Energy Sustainability in Smart Cities: Artificial Intelligence, Smart Monitoring, and Optimization of Energy Consumption. Energies, [online] 11(11), p.2869. doi:https://doi.org/10.3390/en11112869.

Jennifer, Chubba, SondesS, Chubba, Missaouib, Shauna, Concannonc, Liam, Maloneyb, James, Alfred and Walker (2021). Interactive Storytelling for Children: A Case-study of Design and Development Considerations for Ethical Conversational AI. arXiv: Human-Computer Interaction. [online] Available at: https://typeset.io/papers/interactive-storytelling-for-children-acase-study-of-design-51v4el4bp1 [Accessed 15 Sep. 2023].

G. Lakshmi Srinivas, Mane, R. and Javed, A. (2019). A Review on Optimization Methods to Enhance Energy Efficiency of Robots. EasyChair preprint. doi:https://doi.org/10.29007/djnw.

Graa, A. and Benhamida, F. (2020). A Review on Optimization Management Management, doi:https://doi.org/10.5937/sjm15-22519.

H J Burckhart, F Caspers, M Nonis, Doré, J P Burnet, C Martel, D Tommasini and L Gatignon (2013). Optimization of the Energy Usage at CERN. typeset.io. [online] Available at: https://typeset.io/papers/optimization-of-the-energy-usage-at-cern-4d6d4n2xyy [Accessed 15 Sep. 2023].

Han, X., Zhou, M., Turner, M.J. and Yeh, T. (2021). Designing Effective Interview Chatbots: Automatic Chatbot Profiling and Design Suggestion Generation for Chatbot Debugging. Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. doi:https://doi.org/10.1145/3411764.3445569.

Hoque, N. (2013). Analysing Sustainable Consumption Patterns: a Literature Review. Development, [online] 56(3), pp.370–377. doi:https://doi.org/10.1057/dev.2014.13.

Ibrahim Abdullahi, Ismail Abubakar, Iyadunni Ndububa Olufunmilayo, Lawal Batagarawa Rabi'A, Adebola Ajibike Ma'Aruf and Tsoho Usman (2018). Consumption Patterns of Public Water Supply Using Linear Programming Techniques. typeset.io, [online] 17(33), pp.257–276. Available at: https://typeset.io/papers/consumption-patterns-of-public- water-supply-using-linear4psx7ztcm0 [Accessed 15 Sep. 2023].

IEA. (n.d.). World total energy supply by region, 1990-2018 – Charts – Data & Statistics. [online] Available at: https://www.iea.org/data-and-statistics/charts/world- total-energy-supplyby-region-1990-2018 [Accessed 15 Sep. 2023].

Ischen, C., Araujo, T., Voorveld, H., van Noort, G. and Smit, E. (2020). securitcy Concerns in Chatbot Interactions. Chatbot Research and Design, 11970, pp.34–48. doi:https://doi.org/10.1007/978-3-030-39540- 7_3.

Kimon, Kieslich, Birte, Keller, Christopher and Starke (2021a). AI-Ethics by Design. Evaluating Public Perception on the Importance of Ethical Design Principles of AI. arXiv: Computers and Society. [online] Available at: https://typeset.io/papers/ai-ethics-by- design-evaluating-publicperception-on-the-17975rxsrd [Accessed 15 Sep. 2023].

Kimon, Kieslich, Birte, Keller, Christopher and Starke (2021b). AI-Ethics by Design. Evaluating Public Perception on the Importance of Ethical Design Principles of AI. arXiv: Computers and Society. [online] Available at: https://typeset.io/papers/ai-ethics-by- design-evaluating-publicperception-on-the-17975rxsrd.

Luccioni, A., Baylor, E. and Duchene, N. (2020). Analyzing Sustainability Reports Using Natural Language Processing. arXiv: Computation and Language. [online] Available at: https://typeset.io/papers/analyzing-sustainability-reports- using-natural-language-3mjf2qv0q1 [Accessed 15 Sep. 2023].

M C Marques, Boris Sucic, and Tomaz Vuk (2016). Context-based Decision Support for Sustainable Optimization of Energy Consumption. InImpact: The Journal of Innovation Impact, [online] 7(2), p.899. Available at: https://typeset.io/papers/context- based-decision- support-forsustainable-optimization-1lclctj6ra [Accessed 15 Sep. 2023].

Ma, Z., Dou, Z., Zhu, Y., Zhong, H. and Wen, J.-R. (2021). One Chatbot Per Person: CreatingPersonalized Chatbots based on Implicit User Profiles. Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval, [online] pp.555–564. doi:https://doi.org/10.1145/3404835.3462828.

Matic, R., Kabiljo, M., Zivkovic, M. and Cabarkapa, M. (2021). Extensible Chatbot Architecture Using Metamodels of Natural Language Understanding. Electronics, 10(18), p.2300. doi:https://doi.org/10.3390/electronics10182300.

Mihai, Cristian, Pîrvu, Alexandra, Anghel., Ciprian, Borodescu, Alexandru and Constantin (2018). Predicting user intent from search queries using both CNNs and RNNs. arXiv: Computation and Language. [online] Available at: https://typeset.io/papers/predicting- userintent-from-search-queries-using-both-cnns-17lgfn6yys [Accessed 15 Sep. 2023].

Mirko Z. Stijepović (2013). Issues and Solutions for Energy Consumption Optimization. typeset.io. [online] Available at: https://typeset.io/papers/issues- and-solutions-for-energyconsumption-optimization-n02rycvqmg [Accessed 15 Sep. 2023].

Nabavi, S.A., Aslani, A., Zaidan, M.A., Zandi, M., Mohammadi, S. and Hossein Motlagh, N. (2020). Machine Learning Modeling for Energy Consumption of Residential and Commercial Sectors. Energies, 13(19), p.5171. doi:https://doi.org/10.3390/en13195171.

Nicole Radziwill and Morgan C. Benton (2017). Evaluating Quality of Chatbots and Intelligent Conversational Agents. arXiv: Computers and Society. [online] Available at: https://typeset.io/papers/evaluating- quality-of-chatbots-and-intelligent-42wzhnly8o [Accessed 15 Sep. 2023].

Noureddine, A., Rouvoy, R. and Seinturier, L. (2013). A Review of Energy Measurement Approaches. ACM SIGOPS Operating Systems Review, 47(3), pp.42–49. doi:https://doi.org/10.1145/2553070.2553077.

Patil Rekha, S. and Rao , R. (2021). A Natural Language Understanding knowledge-based Chatbot over Linked Web Data. Journal of emerging technologies and innovative research, [online] 8(8). Available at: https://typeset.io/papers/a-natural-language- understandingknowledge-based-chatbot-wdd31yx7h8 [Accessed 15 Sep. 2023].

Sun, H.-C. and Huang, Y.-C. (2012). Optimization of Power Scheduling for Energy Management in Smart Homes. Procedia Engineering, 38, pp.1822–1827. doi:https://doi.org/10.1016/j.proeng.2012.06.225.

Ted Briscoe (2014). Introduction to Linguistics for Natural Language Processing. typeset.io. [online] Available at: https://typeset.io/papers/introduction-to-linguistics-for-natural-languageprocessing-4liy6yhvje [Accessed 15 Sep. 2023].

DATA ACQUISITION WORKFLOW UTILIZING LOD3 GEOMETRY TEMPLATES AND POINT CLOUDS FOR URBAN BUILDING ENERGY MODELLING

Elisa Iliste[1], Ergo Pikas[1], Targo Kalamees[1], Joosep Viik[1]
[1]Tallinn University of Technology

Abstract

This case study addresses the data acquisition problem for urban energy modeling (UBEM). A novel workflow is advanced to gather and integrate data from different sources to enable building performance modeling and simulation at the neighborhood scale. The building typology-specific geometry templates and airborne point cloud data to model surrounding buildings and trees were used for the integrated workflow. The benefits of the developed workflow are demonstrated through a solar exposure study, demonstrating the significant impact of detailed building geometry, surrounding buildings, and trees on buildings' solar exposure, ranging from the 0% difference between LOD2 and LOD3 to 47%.

Introduction

The renewed Energy Performance of Buildings Directive (EPBD) expands the scope from single buildings to district and building stock level planning, development, and implementation of renovation measures. Traditional single-site/building approaches, processes, methods, and tools are not fit for achieving the 2050 decarbonization targets (Iliste, 2023; Parts *et al.*, 2023). Moreover, the current standards overlook the integration of building surroundings in the assessment process, creating a disparity between the real-world conditions and simulation results. New digital capabilities are needed to evaluate existing building stock or buildings in the neighborhood and plan, develop, and evaluate renovation strategies and solutions.

Urban building energy modeling (UBEM) supports assessing building performance and facilitates the creation of decarbonization strategies on an urban scale for building clusters (Reinhart and Cerezo Davila, 2016). UBEM is the application of physics-based building energy models, often relying on bottom-up methods (Ferrando *et al.*, 2020), to predict energy use and indoor and outdoor environmental conditions for groups of buildings (Chen and Hong, 2018). Many UBEM digital tools have utilized the EnergyPlus engine (Kamel, 2022) ("EnergyPlus", 2014), which as an open-source engine, is flexible and highly customizable (Kamel, 2022).

The minimum information needed for UBEM modeling and simulation includes building envelope and geometry data. Building envelope data is commonly assigned by using archetypes or reference buildings (Ali *et al.*, 2019). In contrast, simplified shoebox models or 3D models are used for geometry data, commonly at the level of detail 1 (LOD1) or 2 (LOD2). However, these models have important limitations (Parts *et al.*, 2023). For example, these models do not contain necessary information on windows and (recessed) balconies, which significantly impact solar heat gains, shading, or photovoltage panels (PV) placement on the façade (Wang *et al.*, 2022).

Many studies to generate geometry data for windows have been conducted (Dochev *et al.*, 2020; Orenga Panizza and Nik-Bakht, 2023). The common method to address this data gap is to use window-to-wall ratios (De Jaeger *et al.*, 2020), assigned to archetypes. Abolhassani *et al.* scaled all wall surfaces by the window-to-wall ratio to get window areas (Abolhassani *et al.*, 2022). This method is easy to implement but it frequently either underestimates or overestimates window areas (Dochev *et al.*, 2020) and fails to consider the actual positioning and orientations of windows. Furthermore, apartment buildings often do not have windows in all orientations, which is why this method is often too simplistic.

Johari *et al.* (2022) compared the shoebox, LOD1, and LOD2 geometry and the related effect of window placements and thermal zoning on heat gains. Additionally, they examined three main zoning configurations: 1 zone per building, 1 zone per floor, and 5 zones per floor. The results showed that during the cooling period of the year, the energy demands of LOD1 and LOD2 were significantly lower than those of LOD3 (18% and 13% respectively) (Johari *et al.*, 2022). For the heating period, the differences were dismissible.The effect of the orientation of windows was not thoroughly examined.

Generally, two distinct bottom-up physics-based UBEM study approaches can be distinguished (Kamel, 2022): (1) independent simulation (not considering surroundings) of multiple buildings and aggregation of results, and (2) studying the microclimate effects in a selected region. Most studies address either the first or the second area. The combined study is less common mainly due to the increased computational needs of simulating multiple buildings and surroundings together.

However, in an urban environment, solar radiation, air temperature, and wind speed loads are significantly affected by surrounding buildings, ground surface materials, and vegetation (Shareef, 2021). This is why some studies have focused on the combination of these approaches and have inspected the effect of urban microclimate in UBEM (Dougherty and Jain, 2023; Katal *et al.*, 2022; Liu *et al.*, 2023). These have focused on how weather influences buildings or how surrounding shadings affect buildings.

For example, one study created a dataset of wall center points of neighborhood buildings, with ground and roof heights stored as attributes (Faure *et al.*, 2022). The points were then utilized to regenerate shadowing objects. This allowed to regenerate objects within a 250m radius as shadows to the target building, lowering the computational cost. They found that at a district scale, differences below 2% could be achieved by including all shadowing building surfaces within a 50m radius from the building's centroid. (Faure *et al.*, 2022).

Another study developed a SketchUp plug-in (MOOSAS-FastSolar), where the shadowing effects were accounted for by two indicators: Surrounding Building Factor (SBF) and Impact Factor (IF). The SBF considers the height and length of the surrounding buildings and their distance from the target building, and IF describes the total shading effect on the energy demand of the target building (Wen *et al.*, 2022).

Due to the computational burden, shading effects of trees on building energy consumption have been assessed in fewer studies (Abdel-Aziz *et al.*, 2015; Zhu *et al.*, 2022, 2023). Simpson (2002) created lookup tables to provide a simple way of accounting for the effect of trees as shading objects. Still considering vegetation in addition to surrounding buildings should be examined further. With the climate pact (Directorate-General for Climate Action (European Commission), 2020) and the goal of incorporating more green areas in cities, the effect of the vegetation will likely increase.

Shading has a great impact on both building heat gains as well as PV production potential. Especially in higher latitudes where low sun angles result in more intense shadows (Formolli *et al.*, 2023). According to the new EPBD, solar energy installations will become the norm for existing residential buildings starting in 2033 if technically suitable and economically and functionally feasible. Currently, the assessment of PV potential often occurs without contextual considerations, even at the level of individual buildings (Formolli *et al.*, 2023). This is mainly due to the lack of requirements but also due to the lack of data and increased computational cost (Abdel-Aziz *et al.*, 2015; Zhu *et al.*, 2022, 2023). This simplification can result in disparities between real-world conditions and simulation models.

The research gaps presented above motivated us to plan and develop a workflow to tackle the data acquisition problem for UBEM. Specifically, this research focuses on two research gaps in the context of data acquisition: (1) LOD2 versus LOD3 geometry, and (2) the impact of surrounding buildings and vegetation. This case study aims to streamline and automate the integration of LOD3 geometry and building surroundings into building performance assessment. The workflow consists of gathering, preparing, and integrating data from various sources to enable modeling and simulation at the neighborhood level. The benefits of such workflow are demonstrated through a solar exposure study.

Research Methods

Workflow Development and Testing

The design research method (Hevner and Chatterjee, 2010) was utilized to develop the workflow in the context of the case study. The development of the workflow for neighborhood scale building performance assessment included the following phases:

- Information needs were specified based on the literature review.
- Data availability and sources were determined.
- Data was gathered, prepared, and integrated from different sources to enable UBEM modeling and simulation at the neighborhood scale.
- Data (specifically, point clouds for modeling trees and typology-specific geometry) templates were developed and integrated into a common workflow.
- The benefits are demonstrated in a case study through a solar exposure study.

Defining Information Needs

Sufficiently accurate information about the building geometry is one major challenge in energy performance assessment at the neighborhood level. Solar exposure is an important parameter influencing heat gains and on-site electricity generation. For this the following information is needed: target building geometry, surroundings (landscape, neighbor buildings, and trees), geographic information, the solar path across the sky (throughout the day and year), and climate and weather conditions (including direct and indirect solar exposure).

To assess solar exposure effectively, the surfaces need to be categorized based on building number, surface type (external wall, window, balcony railing, illustrated in Figure 1), surface orientation, and floor numbers. This identification allows to evaluate how the positioning and orientation of specific surfaces impact solar exposure, offering insights into variations across different surfaces at various heights. Additionally, this approach allows to exclude the north-facing facades, which lack direct solar exposure, from the overall average values. This can allow us to examine the surfaces that are most affected by solar exposure. Spaces with north-facing exteriors have a reduced risk of overheating compared to other orientations. In addition, they are usually not considered for PV installation.

The study considers the surrounding environment, incorporating both neighboring buildings and trees. The shapes of these elements serve as shading objects in the simulation. To accurately model solar exposure, the material properties of these shading objects must be considered. Particularly for trees, where the selected material accounts for light passing through the voxelated shapes.

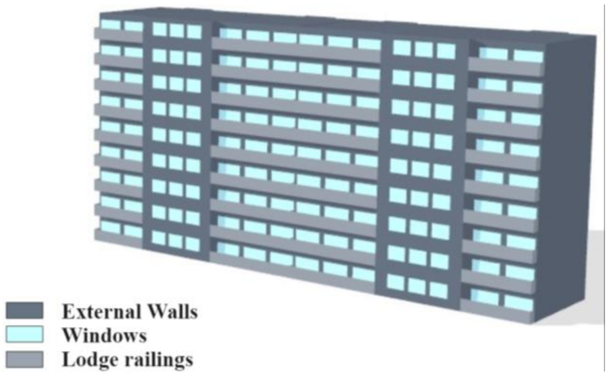


Figure 1: Vertical surfaces of the building.

Selected Pilot Site

An Estonian neighborhood (58°22' N Latitude, 26° 46' E Longitude) with 22 apartment buildings was chosen to develop and evaluate the workflow in the context of solar exposure study. The pilot area buildings represent one common Estonian archetype (Iliste, 2023): a typical not renovated apartment building with precast concrete external walls, built between the 1970s and 1990s. The main differences were in the number of floors and staircases, and orientation. It was also observed that all these buildings were constructed from different combinations of two staircase modules: (1) a side module, and (2) a middle module. This level of standardization makes them suitable for developing and applying geometry templates.

Table 1: Studied buildings' number of floors and staircases, and the neighborhood visualization from Estonian Landboard.

Building numbers	Number of floors	Number of staircases
12, 13	2	2
13	6	2
19	6	3
8	8	2
0, 1 ,2 ,3 ,4 ,5 ,6 ,7,9 ,10 ,11	9	2
20,21	9	3
15,16,17,18	9	4

Building Information and Geometry Template

General and technical information of buildings, LOD2 models, and design documents were acquired from the Estonian Building Register (EBR). It is a public database containing information for all Estonian buildings. It has an open data policy and web services to acquire data.

Geometry templates were developed to facilitate the generation of LOD3 models. For developing and constructing geometry templates, the dimensions of the buildings were acquired from floor plans and elevations of the buildings that had the design documentation available in EBR.

A geometry template is a 2D representation of a typical floor plan, consisting of a set of lines and points categorized according to their function or position (Figure 2). Two distinct templates were designed: (1) side module and (2) middle module. Each represents a staircase per floor and its apartments. Currently, the templates include one zone per floor and module. Additional details like internal walls could be added, which will be addressed within future research.

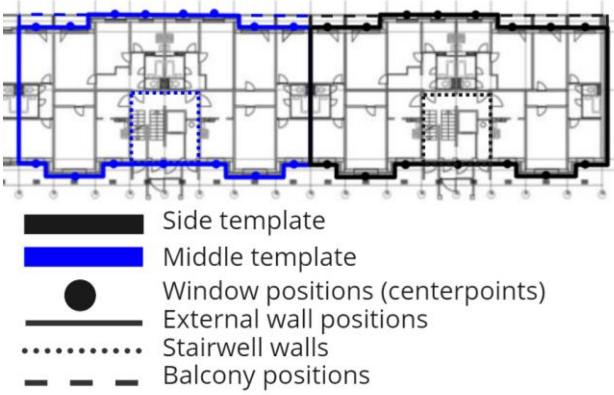


Figure 2: Extraction of geometry templates from design documents.

Point Clouds for Modelling Surrounding Trees

In addition to general and technical information of buildings, point cloud data were acquired from the Estonian Building Register (EBR). The point cloud data for the pilot neighborhood was acquired from EBR, processed, and segmented into landscape, trees, and buildings. The EBR LOD2 geometry, generated from LiDAR data, was used to position and orient the floor models to target buildings and to compare simulation results to current practices utilizing LOD1 or LOD2 models.

System Setup

An important issue in utilizing more detailed models and considering the surroundings is the computational cost. The simulations were run on a desktop setup: CPU: i7-13700 2.10 GHz, Graphic Card: Nvidia RTX A4500, Memory: 128GB. Rhinoceros software was chosen to develop the workflow. The built-in Grasshopper plug-in for visual programming was used to develop computational workflows. For the solar exposure study, the ClimateStudio plugin, utilizing the EnergyPlus engine, for Rhino was employed. For point cloud processing, trimming, segmentation, and conversion, CloudCompare was used.

Case Study Results

Workflow Development Criteria

The proposed workflow (illustrated in Figure 4) is divided into four phases: (1) Data Gathering and Aggregation; (2) Point cloud Processing, Segmentation, and Voxelization; (3) Generation of LOD3 Building Models; (4) Simulation and Visualization of Solar Exposure Study. The main criteria for developing the workflow included:

- **Integration of Existing Data Sources and Systems:** Ensuring seamless integration with existing data sources, systems, and tools to avoid redundancy, reduce manual data entry, and prototype the workflow.
- **Automation:** Automate gathering, preparing, and integrating data sources to reduce manual effort, minimize errors, and accelerate the overall workflow.
- **Flexibility and Scalability:** Ensuring the flexibility to adapt to changes in processes or requirements and accommodating growth.

In the following, the workflow phases and steps are described in detail.

EBR Data Gathering and Aggregation

The first phase involves data gathering and aggregation (see Figure 4). For the selected pilot site and its buildings, four types of data were gathered from the EBR and organized into a common database, including point clouds for the selected pilot site, design documentation, building LOD2 geometry, and EBR data on buildings. These data were acquired by directly downloading publicly available datasets or queried through EBR's open API services. The utilization of API services enables the automatization of data acquisition.

Point Cloud Processing, Segmentation and Voxelization

The second phase involved two steps for point cloud processing, segmentation, and voxelization to create the contextual information (See Figure 4 steps 2.1 and 2.2). For modeling and simulating shading effects on solar exposure, the airborne LiDAR (light detection and ranging) data available for the chosen neighborhood was acquired from the EBR. The point cloud was first separated into ground and non-ground measurements utilizing the cloth-based data filtering method (Zhang *et al.*, 2016), implemented in CloudCompare (Girardeau-Montaut, 2016). Second, off-ground points were categorized into buildings, trees, and rest through several iterations of applying statistical outlier filters, application of scalar fields (i.e., Planarity, Verticality, Sphericity), labeling of connected elements, and manual segmentation of clouds. The output was three separate sets of point clouds for ground, buildings, and trees in *.e57 file format.

Ground, buildings, and tree datasets were imported into Rhino, and the voxel method on trees' point cloud was applied to generate tree-like geometries. Different voxel sizes were tested and for this study, 1m voxel was used. The main criterion for this was the high computational cost of smaller-sized voxels. In the future, the impact of voxel size will be studied in more detail.

Generation of LOD3 Models

This phase involved two steps for generating LOD3 building models, including creating building archetype-specific geometry templates and floor models, positioning and orientating to the right location, and generating the entire LOD3 building model (Figure 4).

For creating a building geometry template, design documents for buildings in the selected pilot site were analyzed and two different basic modules were created: building side and middle modules. Depending on the number of staircases, these modules were aggregated into floor models. For example, a building with two staircases had only two mirrored side modules, building with three staircases had two mirrored side modules and one middle module between the two mirrored side modules (see Figure 3).

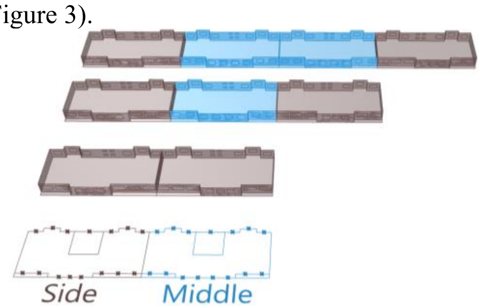


Figure 3: Geometry templates and aggregated floor models

After the creation of a floor model, it was positioned and orientated to the right location based on the LOD2 models from the EBR. After moving floor models to the right location, these were copied vertically as many times as the buildings had floors.

Simulation and Visualization of Solar Exposure Study

The fourth phase involved three steps for simulation and visualization of solar exposure results (see 4.1 – 4.3). In the first step, all previously generated data were integrated into a common simulation model.

In the second step, solar exposure simulations were performed for all 22 buildings of the selected pilot site. The study focused on assessing the impact of LOD3 geometry and the surrounding microclimate on solar exposure of buildings. Specifically, the focus was on vertical surfaces (external walls, windows, balcony railings see Figure 1) that are more likely to be affected by surroundings, especially when there is minimal variation in building heights.

Four solar exposure scenarios were established for simulations:

1. A LOD2 building in isolation (without surroundings).
2. A LOD3 building with detailed information on external walls, windows, and balconies in isolation (without surroundings).
3. A LOD3 building together with its surrounding buildings.
4. A LOD3 building surrounded by both other buildings and trees.

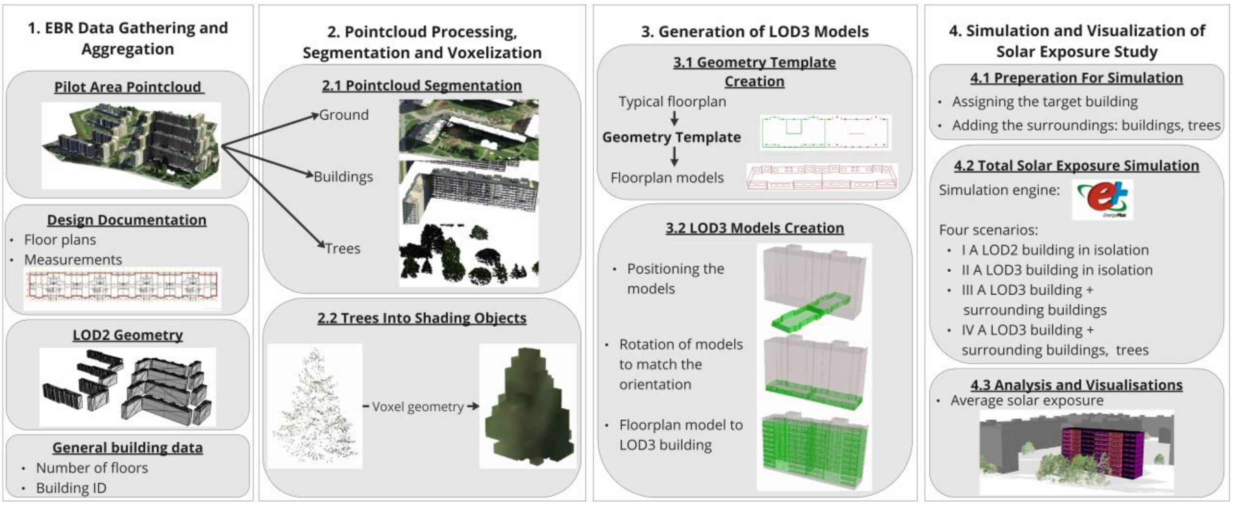


Figure 4: A workflow proposed and developed for neighborhood scale building performance assessment

In the first scenario, LOD2 geometry was employed. Sensor points were placed only on external walls since the LOD2 models lacked additional components such as windows and balconies. Surroundings were not considered. In the second scenario, solar exposure was simulated for walls, windows, and balconies, utilizing LOD3 building geometry. Similarly to the first scenario, the only shading object considered was the target building itself. In the third scenario, surrounding buildings were introduced as shading objects and used together with LOD3 building models. To optimize computational efficiency, LOD2 geometry was utilized for shading, while the target building retained LOD3 detail. In the fourth scenario, trees were incorporated as additional shading objects.

Demonstration: Solar Exposure Simulations

Solar exposure simulations were conducted for four scenarios and compared based on the total average solar exposure (kWh/(m²·yr)). North-facing surfaces were excluded, given their minimal exposure to solar. Climate Studio`s grasshopper plug-in was utilized for simulations.

Solar Exposure Comparison for LOD2 and LOD3 Models

In the first stage, a comparison between the average solar exposure of LOD2 (reference case) and LOD3 vertical surfaces was performed (see Figure 5). The LOD2 and LOD3 models for 22 buildings have similar side walls. Compared to the LOD3, LOD2 does not have different vertical wall, window, and balcony surfaces on the main facades, they are represented as a single surface. In this stage, the overall average values were compared: all vertical surfaces of the LOD2, and the average of all external walls and windows for LOD3.

The results revealed an average difference of 16% in solar exposure between the LOD2 and LOD3 geometry. The minimum difference between LOD2 and LOD3 buildings was 0% and the highest 47%. LOD2 overestimated solar exposure for buildings with the main facades facing East and West, while underestimating it for other buildings (Figure 5). For the East and West-facing buildings, the average solar exposure of LOD2 was 34% higher than LOD3. For North and South facing buildings, the average exposure of LOD2 was 1% smaller than LOD3.

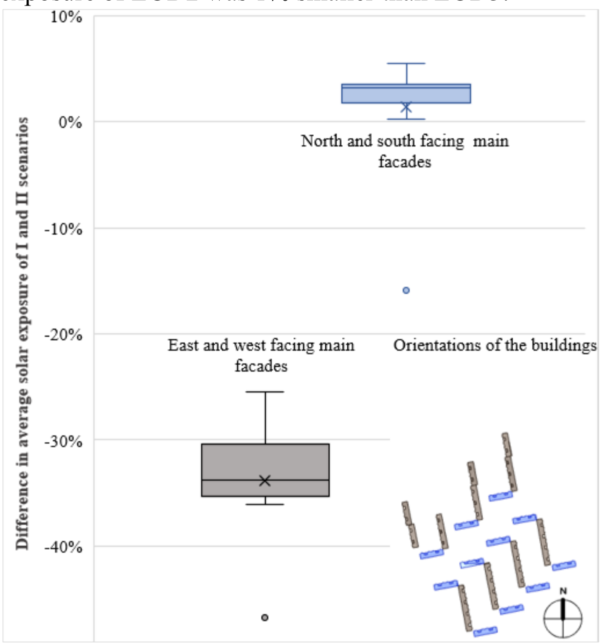


Figure 5: Difference in average solar exposure of scenario I (LOD2, reference case) and scenario II (LOD3).

These results demonstrate how the building's geometry level of detail impacts solar exposure. The main differences are in main facades. LOD2 simplifies main facades, excluding balconies (together with floors and railings), significantly influencing solar exposure on external walls and windows. This is why LOD2 overestimates the solar exposure for the East and West facing buildings.

For the North and South facing buildings, the South walls are side walls (similar for LOD2 and LOD3 geometry), receiving most of the sunlight. Also, these buildings have proportionally smaller main facades than East and West facing buildings have. This is why there are not that big differences in terms of solar exposure.

Solar Exposure Comparison Between II, III, and IV Scenarios

A comparative analysis was conducted for the second, third, and fourth scenario solar exposure results. Here, the second scenario was used as a reference case that the others were compared against.

Within these scenarios, three distinct surfaces were separately evaluated: (1) external walls, (2) windows, and (3) balcony railings. Additionally, the results were distinguished based on the floor level. Significant differences were observed on lower floors (see Figure 6). For the lower floors, the third and fourth scenarios exhibited a 17% and 27% reduction in solar exposure compared to the second scenario where only the target building without surroundings was considered.

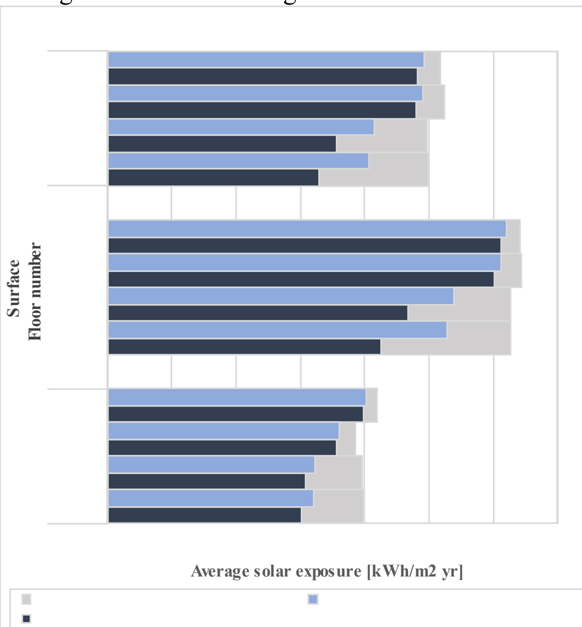


Figure 6: Variations in solar exposure across different surfaces and floors of lod3: scenario II (reference case) vs. scenarios III and IV.

When comparing the results of the three scenarios for higher floors, the disparities were minimal averaging only 4-6%. The average reduction in solar exposure over all the floors compared to scenario two was 12% for scenario three and 17% for scenario four.

The most substantial difference was observed in solar exposure on balcony railings with a notable 18% and 34% reduction on the first floor. This occurs because railings serve as shading objects for other surfaces (external walls and windows), while other components do not offer shading for the railings themselves.

Notably a small difference (3 % on average) in solar exposure between floors can be identified even when no surrounding objects are considered. This effect occurs due to the upper protruding parts of the façade being exposed to the sun before the rest of the façade (see Figure 7) but also when the protruding parts cast a shadow on the lower part of the façade leaving the upper parts exposed to the sun.

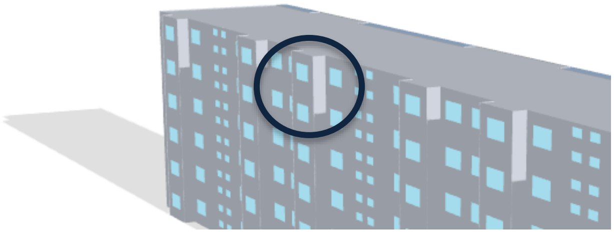


Figure 7: Early solar exposure on the upper protruding part of the facade.

Discussion

This case study approached the data acquisition problem in the field of UBEM by advancing data integration workflow. LOD3 geometry, surrounding buildings, and trees, as shading objects, were incorporated to facilitate neighborhood-level building performance evaluation. Solar exposure metric was used to assess the benefits of the developed workflow. Solar exposure is an important metric influencing many building performance aspects.

The utilization of geometry templates for highly standardized apartment buildings offers a simple solution to develop LOD3 models. Solar exposure analysis demonstrated a notable 16% difference between LOD2 and LOD3 geometry. The effect is mainly due to the incorporation of balconies that act as shading objects. This aligns with another study (Johari *et al.*, 2022) that demonstrated 18% and 13% lower energy demand for the LOD1 and LOD2 when compared to the LOD3.

Previous approaches to LOD3 creation have mainly relied on manual model creation, making it less cost-effective due to resource-intensive implementation. The main limitation of the geometry templates is their reliance on standardized solutions and designs. That is, geometry templates are most useful when there are multiple standardized buildings following the same design principle.

We also developed a workflow to simplify the incorporation of surroundings into building performance assessment and examined the benefits through a solar exposure study. We identified that the incorporation of surrounding buildings on average leads to a 17% reduction of solar exposure on lower floors, with an additional 10% reduction attributed to trees. The limitation of the solar exposure results is the exclusion of the north-facing facades. In future research, north-facing facades could be included to get a more comprehensive understanding.

Also, considering that the examined neighborhood was not intensely populated, these effects could be even more significant in more intensely populated areas. Some existing UBEM solutions have incorporated surrounding buildings but have disregarded trees due to the additional computational cost. This study demonstrated the need to also consider trees. However, the reliability of the developed approach and computational cost were not examined. This needs to be considered in more detail in future studies.

This case study primarily focused on workflow development and utilized solar exposure as an indicator to demonstrate its benefits. Future research could extend this work to include overheating modeling and simulation studies, additional shading requirements, heat gains, and their collective impact on building energy performance. This would allow us to further examine the benefits of accurate window placements and the impact of shading objects. Moreover, our developed workflow enables the incorporation of the ground surface. This could be also utilized for site analysis, for example, incorporating overflooding risk assessment.

Conclusions

This case study developed and demonstrated the utility of a data aggregation workflow for UBEM incorporating LOD3 models, surrounding buildings, and trees for building performance assessment at the neighborhood scale. Solar exposure was used to evaluate the workflow benefits. The proposed workflow was tested in a selected pilot site with 22 apartment buildings. The findings from our solar exposure study demonstrated the impact and need for LOD3 models of buildings with balconies and protruding façade elements. Additionally, it showed how the surrounding buildings and trees significantly influence solar exposure and consequently the reliability of UBEM workflows in general, ranging from the 0% difference between LOD2 and LOD3 to 47%. As a first step, this research contributes to the ongoing efforts towards renovation and digitalization as strategic measures for decarbonizing the European building stock by 2050. However, further studies regarding the reliability of typology-specific geometry templates and point cloud-based building trees are needed.

Acknowledgments

This work has been supported by the Estonian Research Council grant (PSG963); by the European Commission through LIFE IP BUILDEST (LIFE20 IPC/EE/000010), Horizon projects EBENTO (Grant Agreement no. 101079888), and oPEN Lab (VFP21038); and by the Estonian Centre of Excellence in Energy Efficiency, ENER (grant TK230) funded by the Estonian Ministry of Education and Research.

References

Abdel-Aziz, D., Shboul, A. and Al Kurdi, N. (2015), “Effects of Tree Shading on Building’s Energy Consumption -The Case of Residential Buildings in a Mediterranean Climate”, Vol. 2015, pp. 131–140, doi: 10.5923/j.ajee.20150505.01.

Abolhassani, S.S., Amayri, M., Bouguila, N. and Eicker, U. (2022), “A new workflow for detailed urban scale building energy modeling using spatial joining of attributes for archetype selection”, *Journal of Building Engineering*, Vol. 46, p. 103661, doi: 10.1016/j.jobe.2021.103661.

Ali, U., Shamsi, M.H., Hoare, C., Mangina, E. and , O’Donnell, J. (2019), “A data-driven approach for multi-scale building archetypes development”, Energy and Buildings, Vol. 202p. 109364, doi: 10.1016/j.enbuild.2019.109364.

Chen, Y. and Hong, T. (2018), “Impacts of building geometry modeling methods on the simulation results of urban building energy models”, *Applied Energy*, Vol. 215, pp. 717–735, doi: 10.1016/j.apenergy.2018.02.073.

De Jaeger, I., Reynders, G., Callebaut, C. and Saelens, D. (2020), “A building clustering approach for urban energy simulations”, *Energy and Buildings*, Vol. 208, p. 109671, doi: 10.1016/j.enbuild.2019.109671.

Directorate-General for Climate Action (European Commission). (2020), European Climate Pact, Publications Office of the European Union, LU.

Dochev, I., Gorzalka, P., Weiler, V., Estevam Schmiedt, J., Linkiewicz, M., Eicker, U., Hoffschmidt, B., *et al.* (2020), “Calculating urban heat demands: An analysis of two modelling approaches and remote sensing for input data and validation”, *Energy and Buildings*, Vol. 226, p. 110378, doi: 10.1016/ j.enbuild.2020.110378.

Dougherty, T.R. and Jain, R.K. (2023), “TOM.D: Taking advantage of microclimate data for urban building energy modeling”, *Advances in Applied Energy*, Vol. 10, p. 100138, doi: 10.1016/j.adapen.2023.100138.

“EnergyPlus”. (2014), *Energy.Gov*, 28 December, available at: https://www.energy.gov/eere/buildings/articles/ energyplus (accessed 14 December 2023).

Faure, X., Johansson, T. and Pasichnyi, O. (2022), “The Impact of Detail, Shadowing and Thermal Zoning Levels on Urban Building Energy Modelling (UBEM) on a District Scale”, *Energies*, Multidisciplinary Digital Publishing Institute, Vol. 15 No. 4, p. 1525, doi: 10.3390/en15041525.

Ferrando, M., Causone, F., Hong, T. and Chen, Y. (2020), “Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches”, *Sustainable Cities and Society*, Vol. 62, p. 102408, doi: 10.1016/ j.scs.2020.102408.

Formolli, M., Kleiven, T. and Lobaccaro, G. (2023), “Assessing solar energy accessibility at high latitudes: A systematic review of urban spatial domains, metrics, and parameters", *Renewable and Sustainable Energy Reviews*, Vol. 177, p. 113231, doi: 10.1016/j.rser.2023.113231.

Girardeau-Montaut, D. (2016), "CloudCompare", France: EDF R&D Telecom ParisTech, Vol. 11, p. 5.

Hevner, A. and Chatterjee, S. (2010), "Design science research in information systems", *Design Research in Information Systems*, Springer, pp. 9–22.

Iliste, E. (2023), *Creating a Typology for Soviet-Time Apartment Buildings Made of Stone Materials*, Master's Thesis, Tallinn University of Technology, Tallinn, Estonia.

Johari, F., Munkhammar, J., Shadram, F. and Widén, J. (2022), "Evaluation of simplified building energy models for urban-scale energy analysis of buildings", *Building and Environment*, Vol. 211, p. 108684, doi: 10.1016/j.buildenv.2021.108684.

Kamel, E. (2022), "A Systematic Literature Review of Physics-Based Urban Building Energy Modeling (UBEM) Tools, Data Sources, and Challenges for Energy Conservation", *Energies*, Multidisciplinary Digital Publishing Institute, Vol. 15 No. 22, p. 8649, doi: 10.3390/en15228649.

Katal, A., Mortezazadeh, M., Wang, L. (Leon) and Yu, H. (2022), "Urban building energy and microclimate modeling – From 3D city generation to dynamic simulations", *Energy*, Vol. 251, p. 123817, doi: 10.1016/j.energy.2022.123817.

Liu, S., Kwok, Y.T. and Ren, C. (2023), "Investigating the impact of urban microclimate on building thermal performance: A case study of dense urban areas in Hong Kong", *Sustainable Cities and Society*, Vol. 94, p. 104509, doi: 10.1016/j.scs.2023.104509.

Orenga Panizza, R. and Nik-Bakht, M. (2023), "Extraction of energy-influential parameters from building façade images through google street view", Vol. 4, presented at the EC3 Conference 2023, European Council on Computing in Construction, pp. 0–0, doi: 10.35490/EC3.2023.198.

Parts, E.-R., Pikas, E., Parts, T.M., Arumägi, E., Liiv, I. and Kalamees, T. (2023), "Quality and accuracy of digital twin models for the neighbourhood level building energy performance calculations", *E3S Web of Conferences*, EDP Sciences, Vol. 396, p. 04021, doi: 10.1051/e3sconf/202339604021.

Reinhart, C.F. and Cerezo Davila, C. (2016), "Urban building energy modeling – A review of a nascent field", *Building and Environment*, Vol. 97, pp. 196–202, doi: 10.1016/j.buildenv.2015.12.001.

Shareef, S. (2021), "The impact of urban morphology and building's height diversity on energy consumption at urban scale. The case study of Dubai", *Building and Environment*, Vol. 194, p. 107675, doi: 10.1016/j.buildenv.2021.107675.

Simpson, J.R. (2002), "Improved estimates of tree-shade effects on residential energy use", *Energy and Buildings*, Vol. 34 No. 10, pp. 1067–1076, doi: 10.1016/S0378-7788(02)00028-2.

Wang, C., Ferrando, M., Causone, F., Jin, X., Zhou, X. and Shi, X. (2022), "Data acquisition for urban building energy modeling: A review", *Building and Environment*, Vol. 217, p. 109056, doi: 10.1016/j.buildenv.2022.109056.

Wen, J., Yang, S., Xie, Y., Yu, J. and Lin, B. (2022), "A fast calculation tool for assessing the shading effect of surrounding buildings on window transmitted solar radiation energy", *Sustainable Cities and Society*, Vol. 81, p. 103834, doi: 10.1016/j.scs.2022.103834.

Zhang, W., Qi, J., Wan, P., Wang, H., Xie, D., Wang, X. and Yan, G. (2016), "An Easy-to-Use Airborne LiDAR Data Filtering Method Based on Cloth Simulation", *Remote Sensing*, Multidisciplinary Digital Publishing Institute, Vol. 8 No. 6, p. 501, doi: 10.3390/rs8060501.

Zhu, S., Causone, F., Gao, N., Ye, Y., Jin, X., Zhou, X. and Shi, X. (2023), "Numerical simulation to assess the impact of urban green infrastructure on building energy use: A review", *Building and Environment*, Vol. 228, p. 109832, doi: 10.1016/j.buildenv.2022.109832.

Zhu, S., Li, Y., Wei, S., Wang, C., Zhang, X., Jin, X., Zhou, X., *et al.* (2022), "The impact of urban vegetation morphology on urban building energy consumption during summer and winter seasons in Nanjing, China", *Landscape and Urban Planning*, Vol. 228, p. 104576, doi: 10.1016/j.landurbplan.2022.104576.

TRANSITION DYNAMICS IN RESIDENTIAL HEATING SYSTEMS: AN AGENT-BASED MODELING APPROACH

Xinyi Mu[1], Dujuan Yang, and Qi Han
[1]Eindhoven University of Technology, The Netherlands

Abstract

The urgent global need to reduce carbon emissions has intensified the focus on transitioning to sustainable energy solutions. Residential heating is a significant contributor to greenhouse gas emissions. As such, understanding the factors influencing this shift towards more sustainable heating solutions, such as gas-free heat pumps, is critical. This study employs an agent-based model (ABM) to examine the dynamics of adopting sustainable residential heating systems, with a special emphasis on the transition to a gas-free community. The model integrates individual household preferences, financial capabilities, environmental concerns, and social influences to understand the collective transition. The paper illustrated the methodology through a numerical simulation.

Keywords: Agent-based modeling, residential heating, energy efficiency.

Introduction

The imperative to curtail global carbon emissions has intensified the efforts to transition towards sustainable energy solutions. Among the significant contributors to greenhouse gas emissions, residential heating stands out, necessitating a profound shift towards environmentally friendly alternatives. For example, the Netherlands is ranked in the top 5 in the EU for per capita greenhouse gas emissions, largely due to residential heating, which contributes about 10% of national emissions (PBL, 2020). The predominant use of natural gas for home heating is a key factor (de Wildt et al., 2021).

The transition to a sustainable heating system faces challenges in collective decision-making. These challenges stem from varying priorities and constraints among homeowners, including economic considerations and personal preferences. The complexity increases with the need to align these diverse viewpoints for the community-wide adoption of sustainable heating solutions.

Traditional equation-based models (EBMs) have been widely used to predict the diffusion of energy technologies, however, their ability to capture the nuances of individual behavior and complex system dynamics is limited (Moglia et al., 2017). EBMs typically excel in forecasting aggregate behavior but often fall short in depicting interactions at the micro-level, which are crucial for understanding the intricacies of decision-making for energy-related investment.

Agent-based models (ABM) emerge as a robust alternative, addressing these complexities by facilitating a detailed and realistic representation of individual agents and their interactions within a socio-technical framework. ABMs offer an in-depth view of system dynamics, integrating agent heterogeneity, social behavior, and geospatial elements. They allow for an exploration of non-linear relationships and the impact of individual decisions, providing a comprehensive understanding of diffusion patterns at both individual and collective scales (Hansen et al., 2019; Hesselink & Chappin, 2019).

ABM's flexibility in simulating various scenarios helps in understanding the potential outcomes of different heating system implementations, considering the unique dynamics of the homeowner decision-making (Rahmandad & Sterman, 2008). This intricate decision-making process encompasses the delicate balance homeowners trike between economic incentives, environmental concerns, social pressures, and how individual preferences and constraints shape their decisions towards heating system investments. By simulating these varied decision-making processes, ABMs can capture the complexity and heterogeneity inherent in homeowner behaviors, providing insights that are often overlooked in traditional modeling approaches.

The motivation for this research stems from the urgent need to support the transition to sustainable heating systems in a way that is both environmentally beneficial and socially acceptable. By developing a model that simulates household decision-making processes and examines the influence of factors such as economic incentives and social influence, this study aims to contribute insights into the diffusion of sustainable heating technologies. Ultimately, the goal is to inform and guide effective planning strategies that can accelerate the adoption of sustainable heating solutions, aligning homeowner decisions with broader environmental objectives.

This paper begins with an overview of background and significance, then introduces the model in detail, including design, agent attributes, and decision-making processes. The subsequent section presents the simulation setup, execution, and findings, illuminating the model's insights into sustainable heating adoption. The next discussion section links these results to wider real-world implications, leading to a conclusion section that summarizes key findings and suggests avenues for further research.

Methodology

An agent-based model was constructed using the Mesa framework (Kazil et al., 2020) to examine the dynamics of adopting sustainable residential heating systems. This model is strategically designed to address the limitations commonly found in traditional equation-based models, namely their inadequate representation of individual behaviors and the static nature of social interactions

(Moglia et al., 2017; Natarajan et al., 2011; Rounsevell et al., 2012). This model brings agent heterogeneity into consideration, also adds spatial dynamics which traditional ones are hard to simulate. By integrating a detailed representation of household behavior within a socio-technical system, the approach aids in finding real-world emerging patterns to proposed heating solutions, thus enhancing the effectiveness of planning and implementation strategies.

Model Structure

Agents in the model are autonomous entities representing individual households. Each agent is characterized based on statistical distributions that reflect socio-demographic profiles and social clustering data, consistent with real-world demographics. The agents are programmed to evaluate and possibly adopt energy technologies at each simulation step, contingent upon their unique states and decision-making rules. These rules are algorithmic representations of how agents convert their internal states into actions.

The internal state of an agent encompasses a range of parameters such as income, savings, environmental concern, energy efficiency concern, and more. The state evolves as the agent interacts within its social network and makes decisions based on a utility function that integrates the Theory of Planned Behavior with the Net Present Value (NPV) of potential investments.

The Theory of Planned Behavior (TPB) (Ajzen, 1991, 2002; Schiera et al., 2019), which is a widely used human behavior theory, explains human behavior through three attributes: Attitude Toward the Behavior, Subjective Norm, and Perceived Behavioral Control. The intensity of the behavior and the individual's commitment to action are influenced by the weighted contributions of these attributes.

The utility function, which calculates the behavioral intention in this model, is weighted by the agent's attitudes toward each technology, perceived social pressure, and perceived behavioral control, alongside environmental and energy efficiency factors.

The environment provides a virtual space for agents to interact and includes resources such as energy technologies and market prices.

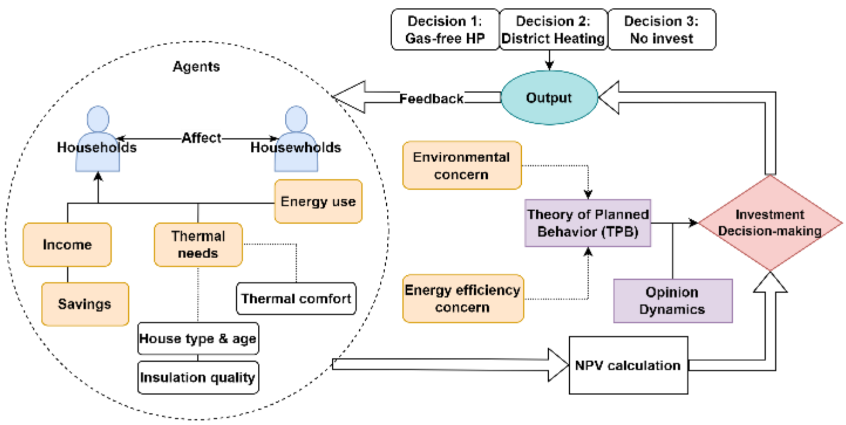


Figure 1. Conceptual ABM framework.

Figure 1 illustrates the conceptual ABM framework, where each household agent has some attributes: income, savings, energy use, and thermal needs. The thermal needs are influenced by house type, house age, the number and vulnerability of family members, and house insulation quality. The agents are faced with three primary decisions: to invest in a gas-free heat pump, district heating with gas, or to refrain from investing. These decisions are shaped by the agents' different concerns and the financial assessment. Feedback mechanisms are in place to reflect the evolving preferences and societal norms as agents interact and make investment decisions.

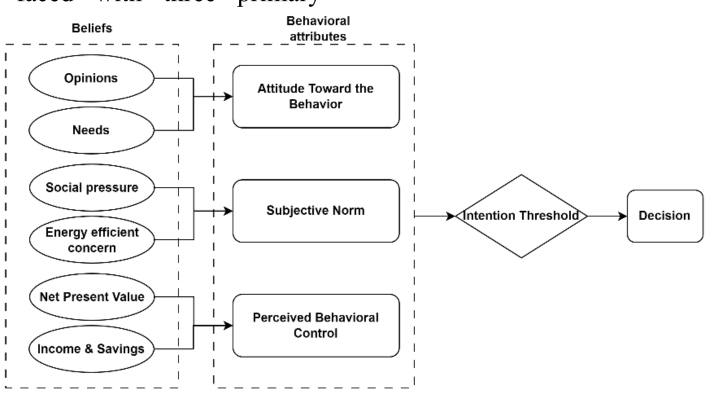


Figure 2. The behavioral sub model of decision-making process in the ABM model.

Figure 2 outlines the decision-making framework based on the Theory of Planned Behavior, which sets the decision-making rules for agents. The process begins with a range of input factors that influence an agent's beliefs, attitudes, and perceptions. 'Subjective Norm' is influenced by social aspects, such as the choices made by neighbors, reflecting the social pressure component. 'Perceived Behavioral Control' encompasses the agent's perception of the ease or difficulty of performing the behavior, which in this case could include financial constraints and thermal needs. The TPB weights are adopted from a survey conducted in the Netherlands (Niamir et al., 2018). Changing the weights would lead to different agents' behavior in the decision-making part.

These three constructs converge to form a 'Behavioral Intention' which is a critical determinant of whether the decision to invest will be made. If the value of this intention is higher than the intention threshold, the agent would decide to invest. Since there are two possible invest options in the model, only the one with higher value would be considered as household's decision. The decision is the final outcome of this process, indicating whether the agent will act on the intention to invest.

Decision-Making Process

To detail the model setting, the parameters adopted in the model is detailed in Tables 1 and 2. Table 1describes the difference between these investment options. Table 2 shows the parameters in the model, income and energy usage values are obtained from open-source datasets about the Netherlands, where N means normal distribution. These parameters collectively influence the final decision-making process.

Table 1. Comparison of Heating System Investment Options.

Attributes	Heat Pump*	District Heating	No invest/ keep original
Investment Cost	Very High	Medium	None
Energy Cost	Low to Medium	Medium	High
CO2 emission	Near Zero	Medium to High	Very High
Energy efficiency	High	Medium	Low
Lifespan	Around 15 years	Over 30 years	Around 15-25 years

***** Note: The Heat Pump category only includes gas-free options such as Air-source, Geothermal, or Solar Heat Pumps.

Table 2. Descriptions of Parameters/ Agent profiles in the model (euros). (Grimm et al., 2010; Müller et al., 2013)

Parameter	Value	Influence on Decision
	Household Characteristics	
Income (euros/m)	N(47120, 25210), Min: 3250 (*StatLine*, n.d.)	Determines investment capability
Environment Concern	N(0.35,0.1) (OECD, 2023)	Increases intention for environmentally friendly choices
Thermal Need	Calculate based on occupants and house situation	Affects the choices of heating systems
Occupants Number & Vulnerability	Uniform Integers [1,5]; See in Figure 3.	Influences thermal need
Insulation Quality	N(5,2)	Affects thermal need and perceived system efficiency
Energy Usage	N(2780, 1548), Positively clipped (Netherlands, 2023)	Affects energy costs and NPV calculations
	Investment Options	
Invest Costs	HP: 12000* DH: 4500	Upfront cost affects the decision
Net Present Value (NPV)	Calculated based on costs	Affects the long-term perceived benefits
Discount Rates	Calculated from interest rate	Used in NPV calculation
Energy Cost	Varies based on energy source	Affects running costs
Lifespan	HP: 15 years; DH: 30 years	Affects the discount rate and NPV calculation
CO2 Emission Factors	Varies by system type (Blum et al., 2010)	Impacts the environmental utility for each decisions

*Note: Initial costs, would change based on year used and energy cost.

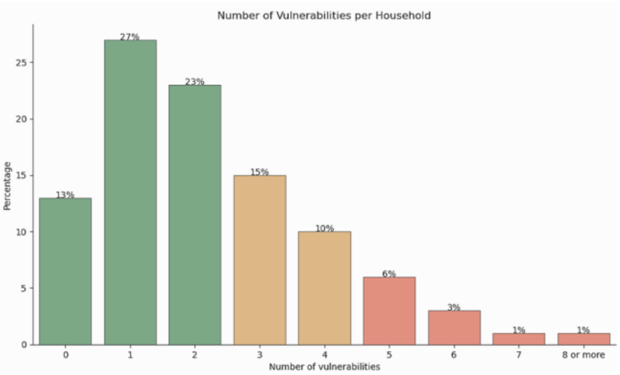


Figure 3. Number of vulnerabilities per household in the Netherlands. (*Many Dutch Are Doing Well, but for Some, Vulnerabilities Are Accumulating*, n.d.)

The thermal need is calculated by multiplying the number of occupants, the presence of vulnerable individuals, insulation quality with adjustment factors for house age, and heating system energy efficiency.

$$NPV = \sum_{0}^{t} (B_t - C_t)/(1+i)^t - C_0$$

The formula is for the calculation of Net Present Value. It calculates the present value of a series of benefits B_t minus costs C_t over a period t, discounted by the average annual bank interest rate i. C_0 is the initial cost.

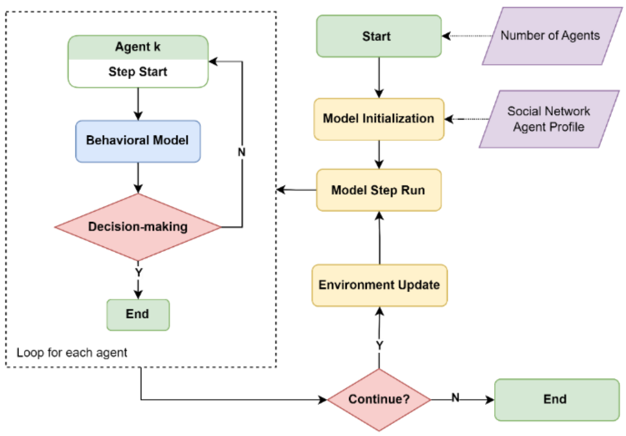


Figure 4. Decision-making process of the ABM framework.

The decision-making process is illustrated in Figure 4. The model is initialized with 1,000 households, each assigned a set of attributes, whose values are randomly generated based on the distributions, according to Table 2. The simulation runs over 15 steps in a loop, with each step representing a year. In each loop of the simulation, which represents one-time decision-making process, agents update their decisions based on various factors, including the recent choices made by their neighbors. This dynamic interaction introduces a layer of social influence, whereby the decisions made by an agent's neighbors in the current loop can affect the agent's own decision-making process in the same loop. Along with these social affects, households annually reassess and update their intentions and financial estimations, leading to potential changes in decision-making. Energy costs are also subject to an annual inflation rate, influencing the NPV calculations and thereby affecting investment decisions over time.

At each simulation step, results are collected on the number of households installing each heating system type and their corresponding investments. After running 15 steps, the agent decides on a heating system, which is the final of the simulation's set duration. This reflects the real-world scenario where investments in heating systems are long-term and rarely reversed. Post-simulation, the percentage of households opting for each heating system is calculated and analyzed to understand the prevailing trends and the impact of the considered factors.

Results

A numerical simulation has been performed in this study. The simulation demonstrates an increasing preference for gas-free heat pumps, indicating a collective inclination toward sustainable options. Heat pumps show higher adoption rates compared to gas district heating, as evidenced by the histogram distribution of final choices. The simulation results of 100 runs are presented below:

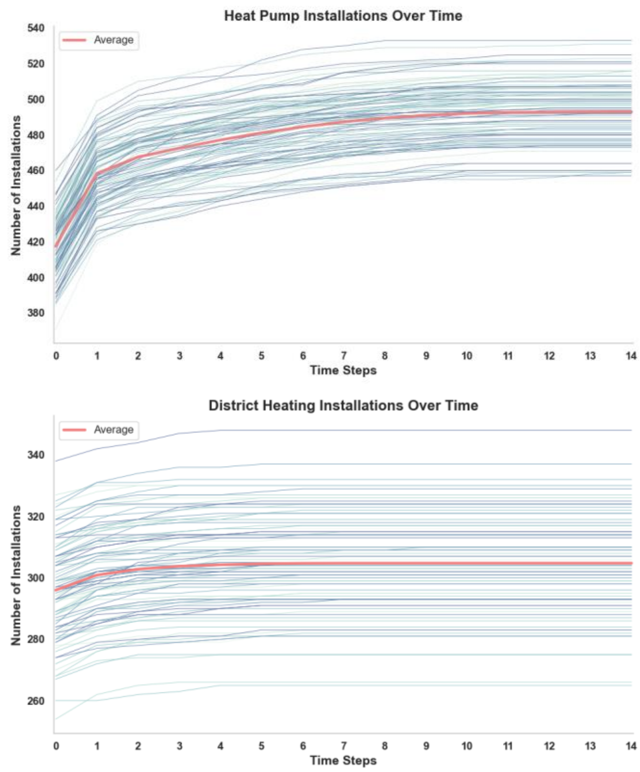


Figure 5. Installation numbers of each investment option across multiple simulation runs.

An upward trend for gas-free heat pump is evident, with a converging average suggesting a growing preference.

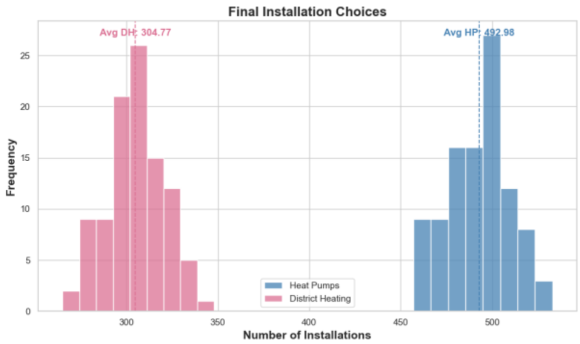


Figure 6. Final Installation Choices Histogram of the distribution of final choices.

Indicating a higher adoption rate for heat pumps over gas heating.

Table 3. Overall mean percentage of households' decisions.

Overall Mean Percentage in the Households		
Gas-free Heat Pump	District Heating	No Invest
49.30%	30.48%	20.23%

Discussion

The agent-based model's simulation process could reveal discernible patterns in household heating system decisions. The multi-objective nature of the model encapsulates the trade-offs between cost, efficiency, lifespan, and environmental impact.

There is a robust inclination towards gas-free heat pumps, despite their higher costs and reduced lifespan. This preference persists even when compared to gas district heating, which offers a longer lifespan and marginally improved efficiency than traditional separate heating systems but falls behind due to its CO2 emissions. The curve of district heating leveled off earlier, reflecting some residents' collective preference for cheaper, long-lasting district heating. The option to retain existing heating systems, which represent the least financial burden in terms of upfront costs but fail on efficiency and environmental grounds, is the least popular choice.

This outcome may signify an inherent bias within the simulation towards progressive environmental action. The bias is driven by the model's weighting of environmental concerns and a reflection of the agents' social pressure, where the adoption of greener technologies is perceived positively. This underscores the potential effectiveness of policy mechanisms, such as subsidies or carbon taxes, to steer communities towards greener alternatives, despite the associated costs and practicalities.

Moreover, the model offers a valuable tool for evaluating the influence of social norms and financial capabilities on the decision-making process. It highlights the critical role played by community behavior and collective environmental commitment in fostering the adoption of sustainable heating technologies. As such, the findings emphasize the interconnected nature of individual choices and broader community dynamics in the adoption of a more sustainable and environmentally conscious approach to residential heating.

Conclusion

This study applied an ABM to simulate the adoption dynamics of residential heating systems, focusing on gas-free heat pumps and district heating options. The model effectively captures the nuanced decision-making process of households in adopting heating systems. The model highlights the complex interplay between various factors that shape household decisions. It underscores the need for multi-faceted policy instruments that not only encourage the adoption of green technology but also address barriers related to cost and technology longevity.

Future research could focus on refining the model's parameters, building a more comprehensive real-world dataset to capture a broader spectrum of real-world behaviors, and integrating feedback loops that might emerge from widespread adoption of these technologies. Conducting a comprehensive sensitivity analysis is essential for validating the model. Furthermore, while the model captures individual decision-making within a physical neighborhood context—significantly shaped by social norms—it does not explicitly simulate collective decision-making scenarios. Such scenarios could include the influence of key decision-makers whose choices disproportionately impact the community's adoption patterns.

References

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Ajzen, I. (2002). Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior1. *Journal of Applied Social Psychology*, *32*(4), 665–683. https://doi.org/10.1111/j.1559- 1816.2002.tb00236.x

Blum, P., Campillo, G., Münch, W., & Kölbel, T. (2010). CO2 savings of ground source heat pump systems – A regional analysis. *Renewable Energy*, *35*(1), 122–127. https://doi.org/10.1016/j.renene.2009.03.034

de Wildt, T. E., Boijmans, A. R., Chappin, E. J. L., & Herder, P. M. (2021). An ex ante assessment of value conflicts and social acceptance of sustainable heating systems: An agent-based modelling approach. *Energy Policy*, *153*, 112265. https://doi.org/10.1016/j.enpol.2021.112265

Grimm, V., Berger, U., DeAngelis, D. L., Polhill, J. G., Giske, J., & Railsback, S. F. (2010). The ODD protocol: A review and first update. *Ecological Modelling*, *221*(23), 2760–2768. https://doi.org/10.1016/j.ecolmodel.2010.08.019

Hansen, P., Liu, X., & Morrison, G. M. (2019). Agent-based modelling and socio-technical energy transitions: A systematic literature review. *Energy Research & Social Science*, *49*, 41–52. https://doi.org/10.1016/j.erss.2018.10.021

Hesselink, L. X. W., & Chappin, E. J. L. (2019). Adoption of energy efficient technologies by households – Barriers, policies and agent-based modelling studies. *Renewable and Sustainable Energy Reviews*, *99*, 29–41. https://doi.org/10.1016/j.rser.2018.09.031

Kazil, J., Masad, D., & Crooks, A. (2020). Utilizing Python for Agent-Based Modeling: The Mesa Framework. In R. Thomson, H. Bisgin, C. Dancy, A. Hyder, & M. Hussain (Eds.), *Social, Cultural, and Behavioral Modeling* (pp. 308– 317). Springer International Publishing. https://doi.org/10.1007/978-3-030-61255-9_30

Many Dutch are doing well, but for some, vulnerabilities are accumulating. (n.d.). Retrieved January 12, 2024, from https://www.dnb.nl/en/general-news/dnbulletin-2023/many-dutch-are-doing- well-but-for-some-vulnerabilities-are-accumulating/

Moglia, M., Cook, S., & McGregor, J. (2017). A review of Agent-Based Modelling of technology diffusion with special reference to residential energy efficiency. *Sustainable Cities and Society*, *31*, 173–182. https://doi.org/10.1016/j.scs.2017.03.006

Müller, B., Bohn, F., Dreßler, G., Groeneveld, J., Klassert, C., Martin, R., Schlüter, M., Schulze, J., Weise, H., & Schwarz, N. (2013). Describing human decisions in agent-based models – ODD + D, an extension of the ODD protocol. *Environmental Modelling & Software*, *48*, 37– 48. https://doi.org/10.1016/j.envsoft.2013.06.003

Natarajan, S., Padget, J., & Elliott, L. (2011). Modelling UK domestic energy and carbon emissions: An agent-based approach. *Energy and Buildings*, *43*(10), 2602–2612. https://doi.org/10.1016/j.enbuild.2011.05.013

Netherlands, S. (2023, October 27). *Energy consumptionprivate dwellings; type of dwelling and regions* [Webpagina]. Statistics Netherlands. https://www.cbs.nl/engb/figures/detail/81528ENG

Niamir, L., Filatova, T., Voinov, A., & Bressers, H. (2018). Transition to low-carbon economy: Assessing cumulative impacts of individual behavioral changes. *Energy Policy*, *118*, 325– 345. https://doi.org/10.1016/j.enpol.2018.03.045

OECD. (2023). *How Green is Household Behaviour?* https://www.oecdilibrary.org/content/publication/2bbbb663-en PBL. (2020, October 29). *Klimaat- en Energieverkenning 2020* [Text]. PBL Planbureau voor de Leefomgeving. https://www.pbl.nl/publicaties/klimaat-en energieverkenning-2020

Rahmandad, H., & Sterman, J. (2008). Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models. *Management Science*, *54*(5), 998–1014. https://doi.org/10.1287/mnsc.1070.0787

Rounsevell, M. D. A., Robinson, D. T., & Murray-Rust, D. (2012). From actors to agents in socio-ecological systems models. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1586), 259–269. https://doi.org/10.1098/rstb.2011.0187

Schiera, D. S., Minuto, F. D., Bottaccioli, L., Borchiellini, R., & Lanzini, A. (2019). Analysis of Rooftop Photovoltaics Diffusion in Energy Community Buildings by a Novel GIS- and Agent-Based Modeling Co-Simulation Platform. *IEEE Access*, *7*, 93404–93432. https://doi.org/10.1109/ACCESS.2019.2927446 *StatLine—Inkomen van personen; inkomensklassen, persoonskenmerken*. (n.d.). Retrieved January 12, 2024, from https://opendata.cbs.nl/statline/?dl=D4D1#/CBS/nl/dataset/83931NED/table

AN EVOLUTIONARY ALGORITHM-BASED MODEL PREDICTIVE CONTROL FOR COMBINED ELECTRICAL AND THERMAL ENERGY SYSTEMS

Sarah NOYE, Asier URUEÑA ANDREU, Ruben MULERO MARTINEZ,
Ricardo ALONSO SEGADE and Angel CARRERA MARTIN
TECNALIA, Basque Research and Technology Alliance (BRTA), San Sebastian, Spain

Abstract

Energy storage is crucial to increase renewable energy adoption in construction. By optimizing their control strategies, operational costs decrease and return on investment improves. Model Predictive Controls (MPC) have been used to optimize the use of energy storage but are costly to implement. This paper presents an MPC with a generalized mathematical model for electrical and thermal storage. A methodology is introduced to account for physical restrictions. Three evolutionary algorithms were compared for the optimization and a Genetic Algorithm was found to best reduce the energy bill with average daily savings of 38.7 %.

Introduction

Driven by decarbonization targets, there is a strong push for increased electrification of building thermal demand coupled with the drive for increased use of renewable energy produced on site. One of the main constraints of Photovoltaic (PV) self-consumption in residential buildings is that a significant amount of the consumption takes place when there is no PV generation. A potential solution to overcome this is energy storage, electrical or thermal, allowing to use the PV excess generated during the central hours of daylight to cover the demand at night and other instants without PV generation. To date, electrical storage has been favoured, but thermal storage can be an additional source for storing surplus PV electricity in buildings with considerably less costs than electrochemical batteries, increasing the self-consumption rate and ultimately, usability of the local resource which yields better economics for the PV system (Psimopoulos et al., 2016).

To shorten the payback time of energy storage systems, it is important to have a sound control strategy. Most building systems use rule-based controls, which have the advantage to be easy to implement and to transfer between buildings (Noye et al., 2022). However, in the case of energy storage, the efficiency of a control strategy depends on the relation between future production and consumption. The advances in data-driven predictive models (Zhang et al., 2021) mean that it is possible to move towards predictive control. Optimal economic dispatch strategies for prosumers with energy storage have been widely investigated and reported in the literature (Liu et al., 2023). Model Predictive Controls (MPC) have shown to be effective at optimizing control prediction (Sultana et al., 2017). They consist in finding the optimal control action based on the constraints defined by a model over a finite receding horizon. MPC formulations over a short period of time, typically a day, to make battery charging/discharging decisions at each time step have been widely used to address different operational scheduling challenges, like the dispatch of energy storage in microgrids (Shang et al., 2020), operation control of multiple battery energy storage systems (Kim et al., 2018), market participation of smart home aggregators (Correa-Florez et al., 2018), or demand response for heat pump assisted solar water heater (Zhao et al., 2023).

Most of the studies address either battery or thermal storage. Combining both technologies lead to increased control complexity, as the storage charge and discharge efficiency, and self-discharge rate differ between technologies, and are time depended in the case of the thermal storage. A single study has been found to address the joint optimization of electrical and thermal storage (Iwafune et al., 2017), where the additional savings from combining both storage technologies are clearly stated. However, results are conditioned by the selected use case, where PV generation stands for around 5 times the averaged consumption, leading to a huge PV surplus that limits the margin for optimization of MPC management strategies.

The main challenge of MPCs in obtaining a model formulation, including physical constraints, that is compatible with optimization techniques, which often means reducing the order of the model and thus its accuracy (Noye et al., 2022). Evolutionary algorithms have been gaining interest for MPC, as they are able to solve non-linear optimization problems with non-differentiable cost functions (see for example Rodríguez del Nozal et al. (2019)).

The present paper builds on top of previous work from this research group, where a scalable and flexible optimization system based on production and load forecasting as a MPC for electrical storage scheduling (Lloréns et al., 2021). In the present work, the main contribution is that both, electrical and thermal storage, are simultaneously considered to optimize their jointly use. A mathematical formulation for an MPC to optimally control a system with renewable electricity production, electrical and thermal storage, and thermal and electrical consumption is presented, together with its optimization via evolutionary algorithms. The reparation method has been improved compared to previous work (Lloréns et al., 2021), so that it provides an interval of possible value, instead of truncating the values to their limits, reducing the exploration capability of the algorithm.

In the following section, the optimization problem is presented before describing the methodology used for the optimization. The methodology is divided between the mathematical formulation of the optimization problem and the methodology to solve it. The test case and how it was implemented is then described before presenting and discussing the results of the study.

Optimization problem definition

The considered system is composed of the following elements: electrical production, electrical and thermal storage, electrical and thermal consumption. A Heat Pump (HP) is used to charge and discharge the thermal storage. Figure 1 illustrates the energy flux between the different elements of the system. The aim of the MPC is to exploit the flexibility potential provided by the thermal storage and the electrical battery to maximize the use of local energy sources.

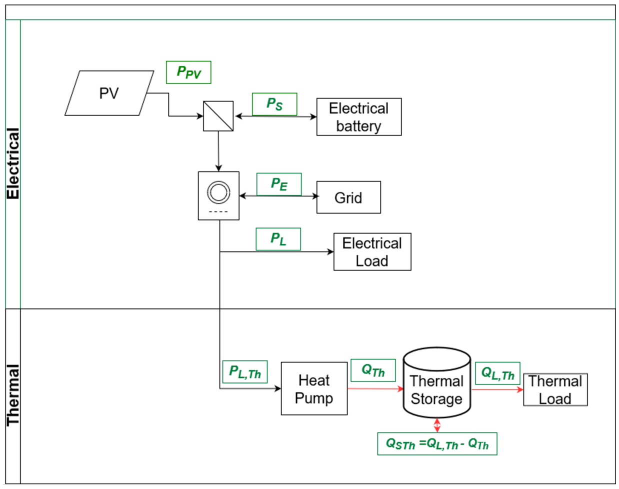


Figure 1: Main energy fluxes of the conceptual energy system to control

Methodology

Figure 2 illustrates the general schema of the MPC presented in this paper. First a general mathematical formulation of the problem is developed based on physical equations. It is divided in three sets of equations: 1) A cost function that represents the objective function to minimize; 2) A set of constraints to ensure the solutions of the optimization are viable; 3) A set of equations that link the different variables necessary for the calculation of the cost and the restrictions between time steps.
Evolutionary algorithms, a category of meta-heuristic optimizations, are then explored to find the optimal solution of the mathematical optimization problem. Three algorithms were compared: Estimation of Distribution Algorithm, Differential Evolution and Genetic Algorithm. The inputs of the optimization process are, on one hand, forecasting data (P_L the electrical load, P_{PV} the photovoltaic production, $Q_{L,Th}$ the thermal load and T_{amb} the outdoor ambient temperature) and, on the other hand, the current

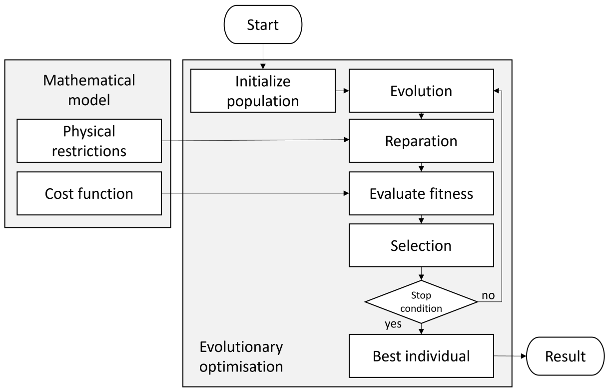


Figure 2: Overview of the MPC architecture

state of the system (charge of the battery SoC_{ini} and of the thermal accumulator $SoC_{ini,Th}$).
The optimization variables are the optimized values for charge and discharge of both the electrical battery (P_S) and the thermal storage (Q_S) for the N next time steps.
A baseline calculation, based on a rule-base reactive control sequence, is done in parallel to evaluate the performance of the evolutionary optimization process.

Mathematical model formulation

This section describes the mathematical formulation of the MPC control. It results in a cost function and a set of constraints used by the optimization algorithm.

Cost function

The mathematical definition of the optimization problem is based on the balance of electrical and thermal energies in the system. The convention used in this paper is that the positive numbers correspond to discharge actions and negative numbers to changer actions. For the electrical part, the energy balance can be expressed as follows:

$$P_{PV}^n - P_L^n - P_{L,Th}^n + P_S^n + P_E^n = 0 \qquad \forall n \in \{1,\ldots,N\} \quad (1)$$

Where P_{PV}^n is the PV production, $P_{L,Th}^n$ the electrical load of the heat pump, P_L^n the remaining electrical load, P_S^n the electricity charged (<0) or discharged (>0) from the battery, and P_E^n the electricity drawn from (>0) or injected to (<0) the electricity grid. n represents the time steps of the control system and N the optimization period.
The thermal energy balance is expressed as follows:

$$-Q_{Th}^n - Q_{L,Th}^n + Q_S^n = 0 \qquad \forall n \in \{1,\ldots,N\} \quad (2)$$

Where Q_{Th}^n is the thermal load of the system, $Q_{L,Th}^n$ the thermal energy produced by the heat pump (always <0), and Q_S^n the energy extracted (>0) or injected (<0) from the thermal storage.
The thermal and electrical equations are coupled by the heat pump performance, which is:

$$P_{L,Th}^n \cdot COP = -\,Q_{Th}^n \quad (3)$$

Where COP is the heat pump's Coefficient of Performance (COP), which is a function of both the storage temperature and the ambient temperature.
Combing Equations 1, 2 and 3, the equation of the conservation of energy is formulated as follows:

$$P_{PV}^n - P_L^n - \frac{1}{COP}(Q_{L,Th}^n - Q_S^n) + P_S^n + P_E^n = 0 \quad (4)$$

The objective of the MPC is to maximize the rate of self-consumption of the whole building, which is equivalent to minimizing the energy imported from the grid P_E. To maximize self-consumption, importing electricity need to be avoided, which corresponds to ($P_E > 0$). Thus the objective function to minimize becomes:

$$\sum_{n=1}^{N} \max\left\{0, (P_L^n + \frac{1}{COP}(Q_{L,Th}^n - Q_S^n) - P_S^n - P_{PV}^n)\Delta t\right\} \quad (5)$$

The use of the maximum condition between 0 and P_E avoids the exported electricity to be considered in the cost function and results in minimizing the imported electricity. One limitation of this cost function is that it does not account for the energy stored in the system within the optimization period, and would thus tend to discharge the battery. In addition, the cost function of Equation 5 does not differentiate between a solution with more energy stored in the system at the end of the period and one with an empty storage. To include this effect, the energy stored at the end of the period, weighted by a factor, is added to the cost function. The function to minimize becomes:

$$\sum_{n=1}^{N} \max\left\{0, (P_L^n + \frac{1}{COP}(Q_{L,Th}^n - Q_S^n) - P_S^n - P_{PV}^n)\Delta t\right\} - \alpha_1 \cdot Cap_{\max} \cdot \left[SoC^N\right] - \alpha_2 \cdot Cap_{\max,Th} \cdot \left[SoC_{Th}^N\right] \quad (6)$$

Where Cap_{max} is the net capacity of the electrical battery, SoC^N the State of Charge (SoC) of the battery at the end of the optimization period, $Cap_{\max,Th}$ the net capacity of the thermal storage, SoC_{Th}^N the SoC of the thermal storage at the end of the optimization period, α_1 a positive weighting factor for the electricity stored at the end of the period, and α_2 a positive weighting factor for the thermal energy stored at the end of the period. The drawback of this addition is the loss of physical meaning of the cost function.

Constraints

To ensure that the cost function leads to solutions that make physical sense, six constrains are introduced. The first two equations ensure that the charging and discharging of the battery meet the available battery capacity, one for the maximum capacity and the second for the minimum. The restriction on the discharge of the battery is expressed analytically as:

$$P_S^n \leq \psi_S \frac{cap_{max}(SoC_{ini} - SoC_{\min})}{\Delta t} - \psi_S \sum_{k=1}^{n-1} (min(0, P_S^k \psi_S) + max(0, P_S^k \frac{1}{\psi_S})) \quad (7)$$

And for the charge of the battery the restriction is:

$$-\frac{1}{\psi_S}(\frac{cap_{max}\Delta SoC}{\Delta t} - \frac{cap_{max}(SoC_{ini} - SoC_{min})}{\Delta t} + \sum_{k=1}^{n-1}(min(0, P_S^k \psi_S) + max(0, P_S^k \frac{1}{\psi_S}))) \leq P_S^n \quad (8)$$

Where ψ_S is the battery efficiency, SoC_{ini} the battery SoC at t=0, $SoC_{\min}$ the minimum allowable value for the battery SoC, and ΔSoC the allowed range for the battery SoC.
Similarly, two restrictions ensure that the charging and discharging of the thermal storage keep the tank operation in the allowed range. For charging, the equation is:

$$Cap_{max,Th} - Cap_{max,Th} SoC_{ini,Th} + \Delta t \sum_{k=1}^{n-1} (Q_S^k + G(T_{Tank}^k)) \geq \Delta t \cdot Q_S^n \quad (9)$$

And for discharging:

$$Cap_{\max,Th} SoC_{ini,Th} - \Delta t \cdot \sum_{k=1}^{n-1} (Q_S^k + G(T_{Tank}^k)) \geq \Delta t \cdot Q_{S,Th}^n \quad (10)$$

Where $SoC_{ini,Th}$ is the thermal storage's SoC at t=0, and $G^k(T_{Tank})$ a function of the tank temperature T_{Tank} that accounts for the thermal losses of the storage.
The fifth restriction limits the power available to charge and discharge the storage to the heat pump maximum capacity ($Q_{Th,max}$):

$$0 \leq Q_{L,Th}^n - Q_S^n \leq Q_{Th\_\max}^n \quad (11)$$

Finally, the sixth restriction limits the power requested from the grid to the maximum power (by contract), constraining the flow in or out of the battery by:

$$-P_L^n - P_{L,Th} + P_{PV}^n + P_S^n \geq -pow_{\max} \quad (12)$$

Where $pow_{\max}$ is the contracted electrical capacity.

Relations between time steps

The additional equations presented in this section are used to calculate how the two SoCs and thermal storage temperature are affected by charge and discharge of the energy system between time steps. For the SoC of the thermal storage, the temporal relationship is:

$$SoC_{Th}^{n+1} = SoC_{Th}^n - \left(\frac{Q_S^n - G\left(T_{Tank}^n\right) \cdot \Delta t}{Cap_{max,Th}}\right) \quad (13)$$

The tank temperature in the next time step is:

$$T_{Tank}^{n+1} = T_{Tank,min} + SoC_{Th}^{n+1} \cdot (T_{Tank,max} - T_{Tank,min}) \quad (14)$$

Where $T_{Tank,max}$ and $T_{Tank,min}$ are the maximum and minimum temperature allowed in the tank respectively. For the battery, the update of the SoC between time steps is:

$$SoC^{n+1} = SoC^{n} - \frac{\psi_s \cdot \min\left(0, P_S^n\right) + \left(\frac{1}{\psi_s}\right) \cdot \max\left(0, P_S^n\right)}{Cap_{max}} \quad (15)$$

Control optimization

Evolutionary algorithms

Evolutionary Algorithms are heuristic algorithms based on the Darwin evolution theory known to be efficient at solving complex optimization problems. They are easy to implement since the fitness function does not need to be differentiable. They are based on a population that is evolved stochastically to look for better solutions. Because they are good at keeping diversity, they perform well when it comes to escaping local optimum. Three evolutionary algorithms were tested for the MPC: Estimation of Distribution Algorithm (EDA), Differential Evolution (DE) and Genetic Algorithm (GA).

EDA it is an evolutionary algorithm that focuses on building and updating probabilistic models of promising solutions to guide the search for optimal solutions in optimization problems. DE is an algorithm that optimizes a population of candidate solutions by iteratively combining the differences between their parameter values to progressively converge toward the optimal solution. GA mimics the process of natural selection by iteratively evolving a population of candidate solutions by selecting, recombining, and mutating individuals to find optimal solutions in optimization problems.

Individuals definition

Common to all the algorithm implementation is the definition of the individuals. A solution is represented by a 2N array, where the N first values represent the battery control P_S^n and the N following values, the tank control Q_S^n:

$$\{P_s^0, P_s^1, \ldots, P_s^{N-1}, Q_{s,Th}^0, \ldots, Q_{s,Th}^{N-1}\} \quad (16)$$

Population initialization

An evolutionary algorithm starts by generating an initial population of individuals. Because of the restrictions, a random individual is likely not to be valid. Indeed, the possible values for P_S^n y Q_S^n are constrained by the previous values and how they have affected the SoC of the electrical and thermal storage respectively. A sequential sampling initialization process that incorporates the restrictions was implemented starting from the initial state $x_0 = x(t_0) = (SoC^0, SoC_{Th}^0)$ and where P_S^n y Q_S^n depend on P_S^{n-1} and Q_S^{n-1} respectively.

According to the restrictions on charge and discharge of the electrical and thermal storage (Eq. 7-10) the valid range for P_s^n y Q_s^n (the values at the next time step) is as follows:

$$P_s^n \in [-Cap_{max}(1 - SoC^{n-1}), Cap_{max}(SoC^{n-1} - SoC_{min})] \quad (17)$$

$$Q_s^n \in [-Cap_{max,Th}(1 - SoC_{Th}^{n-1}), Q_{L,Th}^n] \quad (18)$$

In physical terms, this means that for discharge, the valid values are those smaller than the energy available at the accumulator and battery, and for charging the valid values are limited by the capacity available in the storage. Besides, for the thermal accumulator the discharge cannot be larger than the thermal load.

Substituting the equation of restriction of maximum load of the thermal storage (11) in (18) gives:

$$Q_s^n \in [max(-Cap_{max,Th}(1 - SoC_{Th}^{n-1}), Q_{L,Th}^n - Q_{Th,max}), Q_{L,Th}^n] \quad (19)$$

And finally, substituting Equation 12 that restricts the maximum load of the electric battery in (17), gives:

$$P_s^n \in [max(-Cap_{max}(1 - SoC^{n-1}), P_L^n + P_{L,Th}^n - P_{PV}^n - pow_{max}), Capmax(SoC^{n-1} - SoC_{min})] \quad (20)$$

Equations 19 and 20 define the range of valid values. Because they are coupled by $Q_{L,Th}$ and $P_{L,Th}$, first a uniform random sampling is performed to determine the value of Q_S^n. The value of P_S^n is then uniformly sampled within the resulting range.

Evolution process

The evolution process is implemented according to the three algorithms tested. For the EDA, the EMNA algorithm proposed by Teytaud and Teytaud (2009) was implemented. It consists in re-weighting to remove bias, which is aimed at limiting premature convergence. The main hyperparameter is the λ which is the number of individuals retained at each generation.

For the DE, the original algorithm from Storn and Price (1997) has been implemented where vectors are recombined based on the trial vector that is a linear recombination of random vectors of the population. The main hyperparameter is the scaling factor β which determines the weight of the selected vectors compared to the individual that is being evolved.

The GA implements a Gaussian mutation of mean $\mu = 0$ and deviation σ and a uniform crossover which is a commonly used implementation according to the review of GA by Katoch et al. (2021).

As for the initial generation of the population, the evolution process might generate invalid individuals. Several approaches exist to address this problem: 1) Define a specialized operators so that the evolution process only produces valid solutions; 2) Include a penalty for invalid solutions in the fitness function. One limitation of this approach is that the algorithm then works on reducing the penalty and might not focus as much on the optimization objective; 3) Introduce a repair operator that transforms invalid solutions into valid solutions.

Approach (3) was implemented using Equations 19 and 20. When the value at n violates one of the constraints, the individual is repaired sequentially at all the time steps between n and N. This enables to keep diversity in the population despite the reparation process.
The fitness of the generated individuals is evaluated using the cost function defined in Equations 6. It is used to build a ranking of individuals and select among them which ones progress to the next generation.

Baseline definition

To assess the performance of the optimization strategy, a static control strategy was defined using a set of rules. This strategy consists in a basic sequence that, in case where there is a PV surplus, it is stored in the battery, and once this is full, the system starts filling the thermal storage through the heat pump. Once both storage are full, it sends the surplus to the electricity grid. In the case there is no PV surplus, the system uses, if possible, the energy available both at the battery and the storage, and if they are empty, it will use electricity from the grid. This control strategy is evaluated for each individual in the optimization process, to compare to the evolutionary algorithms' results.

Implementation

This section describes the implementation of the MPC descried in the previous section and the test case that was used to test it. The mathematical model was implemented in Python, and the DEAP python library was used to implement the evolutionary algorithms.

Test case

The test case is based on a real multi-family building located in Pasaia, Spain. The energy system to be controlled has two central air-to-water heat pumps with a total heating power of 34 kW at nominal conditions with a thermal storage of 2.000 litres located in the building garage. From this centralized system, a low temperature (30º C) distribution loop runs through the building to the individual water-to-water heat pumps located at each of the dwellings to provide heating and Domestic Hot Water (DHW). Besides, there is a 6 kWp PV system on the roof and an electrical battery for storage.
The controlled system does not include the storage for the DHW linked with the water-to-water heat pumps at each of the dwellings. They offer limited flexibility potential (less than 2 kWh compared to close to 70 kWh for the central storage) while including them would significantly add to the computational complexity. This means an increased computational effort, poorer performance, more expenses in monitoring and control and in general, a less robust implementation.

Heat pump and water tank models

The *COP* and the thermal losses of the tank use equations specific to the use case. Those are characterized by a correlation that is a function of the tank temperature and the ambient temperature. For the COP, the following function was used:

$$COP = 10.168 - 2.716{\cdot}10^{-1} \cdot T_{Tank} + 4.331{\cdot}10^{-2} \cdot T_{amb} + 2.372{\cdot}10^{-3} \cdot T_{Tank}^2 - 6.623{\cdot}10^{-4} \cdot T_{amb}^2 \quad (21)$$

This correlation was obtained from data generated by a simulation of the system in TRNSYS, where the heat pump was modeled by a performance map obtained from the manufacturer data. This correlation could be updated by monitoring data from the system if available.
$G(T_{Tank})$ is a function that characterizes the thermal losses of the water tank. To obtain it, a thermal storage with cylindrical geometry and insulated according to the regulatory specifications was modeled in TRNSYS. The function was obtained preforming a correlation of the results of simulations of this model:

$$G(T_{Tank}) = (3.6 + 1.9 \cdot Vol) \cdot (T_{Tank} - 15) \cdot 10^{-3} \quad (22)$$

Where Vol is the tank volume in m³ and T_{Tank} is the tank temperature in °C.

Test case parameters

The optimization horizon N was selected to be 24 hours with time steps Δt of 1 hour.
In the cost function, the value of α_1 was set to $0.9 \cdot \psi$, and the value of α_2 to 0.9/3.5, where 3.5 is a proxy of the heat pump average COP. The motivation for the 0.9 value was to give a slightly higher value to the energy already consumed than the one stored for the next day.
The values of the physical parameters of the system are summarized in Table 1. The value of the heat storage capacity was calculated from the sensible heat capacity of water as:

$$cap_{max,Th} = \frac{4.18 \cdot Vol \cdot (T_{Tank,max} - T_{Tank,min})}{3.6} [kWh] \quad (23)$$

Optimization inputs

To test the validity of the developed MPC, one year of synthetic input data were generated. The hourly production were generated with PVGis for the location of the building. The TMY weather file generated by PVGis was used as well for generating the ambient temperature, to ensure consistency between radiation and temperature data. The hourly thermal load data were generated based on the historical natural gas invoices of the occupants. From these invoices, the summer consumption was accounted as representative of the DHW load, and the rest of the consumption was accounted as space heating. To generate hourly values of DHW load, the software DHWCalc 2.0 was used, imposing that the yearly load fits the value estimated from the invoices. For the space heating, the generation of hourly loads from annual values was done by allocating the operation of the thermostat in a minute basis based mainly on ambient temperature, hour of the day, a setpoint temperature and radiator capacity. The profile is subject to fit the amount of energy derived from the natural gas invoices.

Table 1: System parameters of the test case

Parameter	Value
Battery capacity (cap_{max})	6 kWh
Valid SoC range for the battery (ΔSoC)	0.8
Battery minimum SoC (SoC_{min})	0.2
Battery charge/discharge efficiency (Ψ)	0.95
Maximum storage temperature ($T_{Tank,max}$)	55 ºC
Minimum storage temperature ($T_{Tank,min}$)	25 ºC
Accumulator volume (Vol)	2 m³
Storage capacity ($cap_{max,Th}$)	70 kWh
Heat pump maximum capacity ($Q_{th,max}$)	18 kW
Maximum electricity consumption (pow_{max})	18 kW

Table 2: Days selected for the evaluation of the optimizer

Quantile	Date	PV/Load ratio	Total load (kWh)
0	16/01/2022	0.04	46.12
0.1	07/03/2022	0.32	24.85
0.2	15/11/2021	0.59	15.97
0.3	09/12/2021	0.77	36.22
0.4	25/04/2022	1.06	18.19
0.5	06/05/2022	1.74	9.03
0.6	01/07/2022	2.03	8.68
0.7	01/08/2022	3.22	8.12
0.8	07/05/2022	2.59	12.99
0.9	06/09/2021	4.29	6.41
1	06/07/2022	4.45	7.83

The initial daily SoC were set to 0.6 for SoC_{ini} and 0.5 for $SoC_{ini,Th}$, so as to be in the middle of the range between minimum and maximum charge.

Test days

To select the best algorithm, a set of days representative of the different boundary conditions were selected. Since the optimization potential is strongly affected by the PV production and the thermal and electrical loads, the ratio between the PV production and the system load was used. After evaluating the metric for all the days in the test data, the days were ordered and days on the quantiles were sampled in increments of 0.1 (Table 2).
The days with a low PV/Load ratio have a high demand compared to the PV production, when the days with a high PV/Load ratio have a lot of production compared to the demand. The days with the most optimization potential are the ones where the PV/Load ratio is more balanced. The optimization results of the best algorithm were then evaluated on the full year.

Discussion and result analysis

Evolutionary algorithm selection

Several configuration were tested on the 10 reference days for the different algorithms. For the EDA, several values of λ between 300 and 1200 were tested. They were found to have a very small influence on the final results. Although the EDA gave acceptable results some days, there were days when it did not manage to improve on the baseline, mostly in days with a low PV/Load ratio. For the DE, different values of the scaling factor β were tested. Lower values of β performed better, but the DE got stuck in a local minimum and complexity would need to be added to the algorithm to increase its exploration rate, in order to have satisfactory results. For this reason and although the EDA and the DE could be improved to palliate their initial limitations, the GA was selected as the best algorithm as a basic implementation would give better results than the baseline.
For the GA, the main hyperparamters tested were the selection method and the σ value. The most important factor for the GA to find better solutions than the baseline is the selection of individuals that pass to the next generation. With the roulette method the GA converges in a local minimum and fails to improve the baseline, whether the tournament selection method enables to improve the baseline. Finally, σ was found to have a small influence on the results, and only affect the optimization outcome on the days with more solar production than demand. The final value selected is 0.5.
Once the GA algorithm was selected, some experiments have also been carried out to determine the best population size and optimal number of generations. It was found that 200 was generally a good number for both. Table 3 shows that the best individual is generally identified between generations 157 and 200. Even in the cases where the best individual was generated in the generation 200, it was found that running the optimization longer would only marginally improve the result.

Optimization results

This section summaries the optimization results obtained with the MPC with GA optimization. Table 3 shows the

Table 3: Results of the genetic algorithm for the test case

Date	PV/Load ratio	Total load (kWh)	Cost GA	Cost baseline	Absolute savings	Relative savings %	Best individual's generation
16/01/2022	0.04	46.12	30.47	30.12	-0.35	-1.17 %	179
07/03/2022	0.32	24.85	4.27	4.66	0.39	8.39 %	165
15/11/2021	0.59	15.97	-5.19	-3.56	1.64	46.00 %	191
09/12/2021	0.77	36.22	0.78	3.14	2.36	75.14 %	199
25/04/2022	1.06	18.19	-8.21	-8.10	0.11	1.31 %	198
06/05/2022	1.74	9.03	-15.59	-13.76	1.83	13.33 %	178
01/07/2022	2.03	8.68	-17.59	-17.39	0.21	1.20 %	187
01/08/2022	3.22	8.12	-21.69	-19.88	1.81	9.10 %	198
07/05/2022	2.59	12.99	-23.04	-18.99	4.06	21.36 %	200
06/09/2021	4.29	6.41	-23.39	-19.57	3.82	19.52 %	157
06/07/2022	4.45	7.83	-20.90	-21.21	-0.31	-1.48 %	199

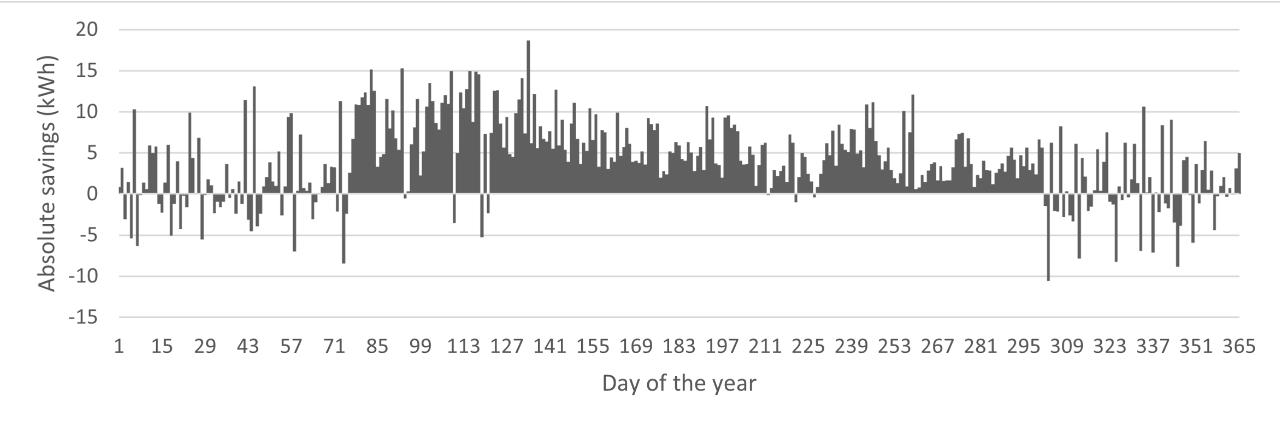


Figure 3: Daily absolute saving over one year

savings achieved for the 10 representative days. The absolute savings are difficult to interpret due to the additional term added to the cost function which makes it lose its physical meaning. Relative savings were thus calculated as the percentage of absolute saving compared to the cost of the baseline.

It can be seen that for two of the selected reference days, the MPC yields a performance under the baseline definition. Those days correspond to the limit cases. The 16/01, there is almost not PV production while the load is high. The 06/07, the demand is limited while the PV production is high. In both cases, there is no potential for optimization and the baseline provide the optimal solution. In those cases the GA failed to converge to the baseline. This can be solved in practice by implementing a control rule that does only run the MPC when there is optimization potential.

Two of the representative days have savings that are lower than 2 %. The 25/04 the demand during the hours with PV production is very close to the production, meaning the optimal solution is to consume the PV energy as it is produced leaving little room for optimization. The 01/07, the lack of optimization reserve comes from the fact that the pick of demand is before the period of PV production and the rest of the days then corresponds to a case with high solar production and no demand.

For the rest of the days (7 out of 11), the MPC is able to leverage the combined electrical and thermal storage much better than the rule base system.

Figure 3 shows the daily saving (without the additional term) that the MPC achieves compared to the rule base control. It can be seen that the MPC yealds less interesting saving in the winter month (beginning and end of the year) where there is statistically less solar production, but provides good results the rest of the year. Over the year 1454.23 kWh less is imported from the grid, which represents an average of 38.7 % of daily saving compared to the baseline. This could be further improved by using the baseline when it yields better results.

The MPC has been developed with PV in mind because of the test case at hand, but the mathematical formulation remains valid for any other electrical production (e.g. micro wind generation). The optimization results would need to be validated as the pattern between production and consumption would be different. Similarly, the formulation stays valid for a case where there is only one of the two storage types. In this case, the values of the system that is not present can simply be put to be zero.

Conclusions

In this paper, an evolutionary algorithm-based MPC optimizing the operation of both electrical and thermal renewable energy storage was proposed. A parametrisable mathematical formulation of the optimization problem was presented. Three evolutionary algorithms were tested to solve the optimization problem and a methodology to address physical restrictions was proposed. A genetic algorithm with tournament selection was shown to give the best results. The average daily saving of the MPC compared to a traditional rule-base control was found to be 38.7 %.

As future work, it would be interesting to test the results with variable energy prices, which enable to take even more advantage of the thermal and electrical storage.

Acknowledgments

This work was developed as part of the HAPPENING Project, which has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No. 957007.

References

Correa-Florez, C. A., Michiorri, A., and Kariniotakis, G. (2018). Robust optimization for day-ahead market participation of smart-home aggregators. Applied Energy, 229:433–445.

Iwafune, Y., Kanamori, J., and Sakakibara, H. (2017). A comparison of the effects of energy management using heat pump water heaters and batteries in photovoltaic - installed houses. Energy Conversion and Management, 148:146–160.

Katoch, S., Chauhan, S. S., and Kumar, V. (2021). A review on genetic algorithm: Past, present, and future. Multimedia Tools and Applications, 80(5):8091–8126.

Kim, S.-K., Kim, J.-Y., Cho, K.-H., and Byeon, G. (2018). Optimal Operation Control for Multiple BESSs of a Large-Scale Customer Under Time-Based Pricing. IEEE Transactions on Power Systems, 33(1):803–816.

Liu, J., Dai, Z., Bo, R., Meng, F., and Ou, M. (2023). Optimal economic dispatch policy for prosumer with energy storage considering self-consumption demand. Computers & Industrial Engineering, 176:108853.

Lloréns, I., Alonso, R., Gil-López, S., Riaño, S., and Del Ser, J. (2021). A Novel Formulation for the Energy Storage Scheduling Problem in Solar Self-consumption Systems. In 15th International Conference on Soft Computing Models in Industrial and Environmental Applications. Springer International Publishing.

Noye, S., Mulero Martinez, R., Carnieletto, L., De Carli, M., and Castelruiz Aguirre, A. (2022). A review of advanced ground source heat pump control: Artificial intelligence for autonomous and adaptive control. Renewable and Sustainable Energy Reviews, 153:111685.

Psimopoulos, E., Leppin, L., Luthander, R., and Bales, C. (2016). Control algorithms for PV and Heat Pump system using thermal and electrical storage. In EuroSun 2016. International Solar Energy Society.

Rodríguez del Nozal, A., Gutiérrez Reina, D., Alvarado-Barrios, L., Tapia, A., and Escaño, J. M. (2019). A MPC Strategy for the Optimal Management of Microgrids Based on Evolutionary Optimization. Electronics, 8(11):1371.

Shang, Y., Wu, W., Guo, J., Ma, Z., Sheng, W., Lv, Z., and Fu, C. (2020). Stochastic dispatch of energy storage in microgrids: An augmented reinforcement learning approach. Applied Energy, 261:114423.

Storn, R. and Price, K. (1997). Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces. Journal of Global Optimization, 11(4):341–359.

Sultana, W. R., Sahoo, S. K., Sukchai, S., Yamuna, S., and Venkatesh, D. (2017). A review on state of art development of model predictive control for renewable energy applications. Renewable and Sustainable Energy Reviews, 76:391–406.

Teytaud, F. and Teytaud, O. (2009). Why one must use reweighting in estimation of distribution algorithms. In Proceedings of the 11th Annual Conference on Genetic and Evolutionary Computation. Association for Computing Machinery.

Zhang, L., Wen, J., Li, Y., Chen, J., Ye, Y., Fu, Y., and Livingood, W. (2021). A review of machine learning in building load prediction. Applied Energy, 285:116452.

Zhao, Z., Wang, C., and Wang, B. (2023). Adaptive model predictive control of a heat pump-assisted solar water heating system. Energy and Buildings, 300:113682.

A SIMPLIFIED BAYESIAN APPROACH FOR THE CALIBRATION OF DISTRICT-BUILDING ENERGY MODELS

Said Bolluk, Özyeğin University, Istanbul [1]
Senem Seyis, Özyeğin University, Istanbul [1]
Reyhan Aydoğan, Özyeğin University, Istanbul [1]
Ece Kalaycıoğlu Özdemir, Özyeğin University, Istanbul [1]

Abstract

Bayesian optimization with surrogate modeling is widely used to calibrate building energy models. However, complexities arise in surrogate modeling due to the variability in building morphology at the urban scale. Thus, maintaining dynamic simulation accuracy is crucial. This study presents a novel optimization framework for calibrating district-building energy models using Bayesian decision theory. Once tested on a case study district, the approach reduces monthly calibration error by approximately 45%. Future works could be employing more robust classifiers and handling imbalanced target variables. The proposed approach can minimize computational demands for optimizing dynamic models while ensuring reliability.

Introduction

Building energy use contributed to 60 million tons of increase in greenhouse gas (GHG) emissions globally in 2022, comprising over 19% of the total rise in GHG emissions from the world's overall energy consumption (OECD, 2023). The construction industry should be investigated thoroughly to understand and reduce such a vast energy use. The occupant-related uncertainties (Santin et al., 2009) and the diversity in the characteristics of the urban building stocks (Oraiopoulos and Howard, 2022) complicate the analysis of the building energy demand with reliable assumptions. Similarly, the construction industry's partial failure in digitization and the resulting lack of comprehensive data hinder studies examining the energy consumption of urban building stocks (Reinhart and Davila, 2016) (Ferrando et al., 2020). However, well-equipped building energy models can provide sustainable solutions to monitor and control urban building energy use (Wang et al., 2018) (Buckley et al., 2021). Dynamic models in building energy modeling (BEM) and urban building energy modeling (UBEM) involve creating computer models that simulate the energy performance of buildings. During simulation, these models evaluate many factors, such as building materials, HVAC systems, occupancy patterns, and weather conditions. Unlike statistical models that analyze the correlation between the building energy-related parameters and the resulting energy consumption and provide predictions usually at lower temporal resolutions (e.g., annual predictions), dynamic models track how energy usage fluctuates throughout the day, week, month, or even year by considering the variations in the building operational schedules and the climate conditions (Pan et al., 2023).

The parameters of dynamic models often need calibration due to insufficient data to characterize buildings accurately, especially at urban level analysis, where diverse building stocks complicate calibration against metered consumption data (Chong et al., 2021). However, calibration is an iterative process that hinders the parametric optimization of dynamic models, which are computationally expensive due to exhaustive simulations covering the thermal interaction between building components. In this sense, Bayesian Optimization stands out as a powerful technique for calibrating simulation parameters (Hou et al., 2021). Traditionally, Bayesian Optimization employs surrogate models, such as regression models or Gaussian processes, to approximate the complex relationships within the dynamic simulation framework (Hou et al., 2021). These surrogate models replace the actual dynamic models and significantly reduce the computational time. They facilitate assessing various scenarios and obtaining optimal parameter combinations efficiently.

For instance, Markov chain Monte Carlo (MCMC) algorithms are prevalent optimization methods that employ surrogate modeling to estimate the posterior distribution of unknown parameters in Bayesian calibration. MCMC algorithms iteratively determine optimization directions by assessing parameter combination errors. In this regard, it resembles the online learning concept adopted in Artificial Neural Network (ANN) models (Rumelhart et al., 1986). Dynamic models are black box models whose simulation output depends on various complex assumptions and factors. Therefore, surrogate modeling approaches are valuable in BEM or UBEM when the dynamic models offer no analytical solution that refers to a direct mathematical method for determining the likelihood of parameter values and, consequently, the posterior probabilities (Hou et al., 2021). Surrogate models are effective when there is little or no knowledge about the possible values of the parameters and their prior probabilities.

However, replacing the dynamic models with surrogate models might yield a loss of accuracy in modeling complex patterns in building morphology at the district and urban scales. When the objective function is spoiled due to the high uncertainty raised by urban building stocks, the optimization process can be faulty. Therefore, the precision of models defining the optimization function becomes crucial even though utilizing an actual model instead of a surrogate model in the likelihood creation process can be exhaustive. Hence, this study proposes a simplified Bayesian optimization method to overcome the limitations arising from the computational requirements and reliability of surrogate modeling. Here, the primary aim is to filter and reduce the size of the parameter space. The proposed approach defines a simplified version of the Bayesian likelihood by assessing the probability of success in scenarios. This approach can be utilized for optimization problems that involve complex computational models (dynamic models) where the actual objective function is not easily accessible. It can be precious when uncertainty regarding the possible value ranges for simulation parameters can be significantly reduced through knowledge acquired from buildings' energy-related records or thorough research. Therefore, it can potentially enhance the optimization process without relying on surrogate modeling techniques.

Methodology

The methodology consists of five sequential steps. First, a dynamic model of the buildings in a selected district is constructed. Next, energy simulations are conducted for various scenarios, each comprising different combinations of simulation parameters, through the dynamic model. These scenarios are then compared with the measured data and labeled as desired and undesirable according to their monthly simulation errors. Subsequently, a Gaussian Naive Bayes Classifier is developed to classify scenarios as either desired or undesired. This classifier enabled the optimal parameter combination to be derived from the posterior distributions of the simulation parameters. Ultimately, a Bayesian Model is created using the optimal parameter combination, and a final energy simulation is performed to evaluate the proposed calibration of the dynamic model. The methodology workflow is demonstrated in Figure 1.

1. Building A Dynamic Model

The proposed calibration approach begins by creating a dynamic energy model for a sample district. Choosing a modeling tool with a user-friendly interface is crucial for easy adjustments to input data to optimize the dynamic model. Furthermore, the performance of the selected modeling tool must be validated, particularly in urban-level building energy assessment, to ensure its robustness and reliability. Therefore, the Urban Modeling Interface (UMI), a widely used tool in urban planning studies worldwide (Ang et al., 2022), is selected to assess the building energy consumption of the sample district. UMI is developed by MIT Sustainable Design Lab (Reinhart et al., 2013). UMI is integrated into a computer-aided design (CAD) program called Rhinoceros 3D (Rhino) that provides a user-friendly environment for the modeling and simulation steps (McNeel and Associates, 2008). UMI requires three main types of input data to create a 3D dynamic model. These are building templates, building geometries, and weather data. A building template is text-based data that includes various categories with the essential characteristics of the buildings, such as the building envelope properties, zone conditioning details, HVAC details, domestic hot water (DHW) details, and building operational schedules. Each category requires a detailed description of the mechanical systems, building parts, or operations. A building template should adequately represent the selected buildings in a district to ensure the accuracy of the energy simulation. Buildings are represented within archetypes of building energy models since modeling each building with its specific characteristics is time-consuming at the district and urban scales. The following input of the dynamic model is the building geometries. Building geometries are attained from existing building footprints or newly created ones in GEOJSON format. These footprints, along with building heights and glazing ratio of building surfaces, are processed in the UBEM.io tool (Ang et al., 2022) to create the building geometries. The final input of the dynamic model is weather data in EnergyPlus Weather (EPW) format. Typical EPW weather data includes comprehensive meteorological information such as temperature, humidity, wind speed and direction, solar radiation, and precipitation recorded at regular intervals. Once these three components of the dynamic model are gathered, the 3D UMI model is ready for the energy simulations.

2. Performing Energy Simulations

The dynamic model consists of various inputs (simulation parameters) that represent the energy consumption characteristics of buildings. These parameters need optimization to minimize simulation errors. Once the simulation parameters are identified for optimization, scenarios are crafted by combining them, and these scenarios are integrated into the 3D UMI Model as building templates. Monthly energy simulations are subsequently conducted for each scenario.

3. Labeling Scenarios

Monthly simulation errors of the scenarios against metered data are calculated using the Coefficient of Variation of the Root Mean Square Error (CV-RMSE) (Equation 1).

$$CV-RMSE = \frac{1}{\bar{r}}\sqrt{\frac{\sum_{t=1}^{N}(r^t-y^t)^2}{N}} \tag{1}$$

In Equation 1, $\bar{r}$ is the average of the monthly target consumption obtained from metered data, r^t represents the monthly cumulative target consumption, y^t is the monthly cumulative simulation result, and N is the total number of monthly simulation outputs, which is twelve. As seen from Equation 1, CV-RMSE decreases as the monthly simulation errors get smaller and more consistent. Hence, it simultaneously evaluates the stability and accuracy of the dynamic model. Additionally, an error metric named MAPE, which calculates the average percentage difference between simulation results and metered data, is utilized to measure the scenario errors (Equation 2).

$$MAPE = \frac{1}{N}\sum_{t=1}^{N}\left|\frac{r^t-y^t}{r^t}\right| \tag{2}$$

The main idea of the proposed approach is to label scenar-

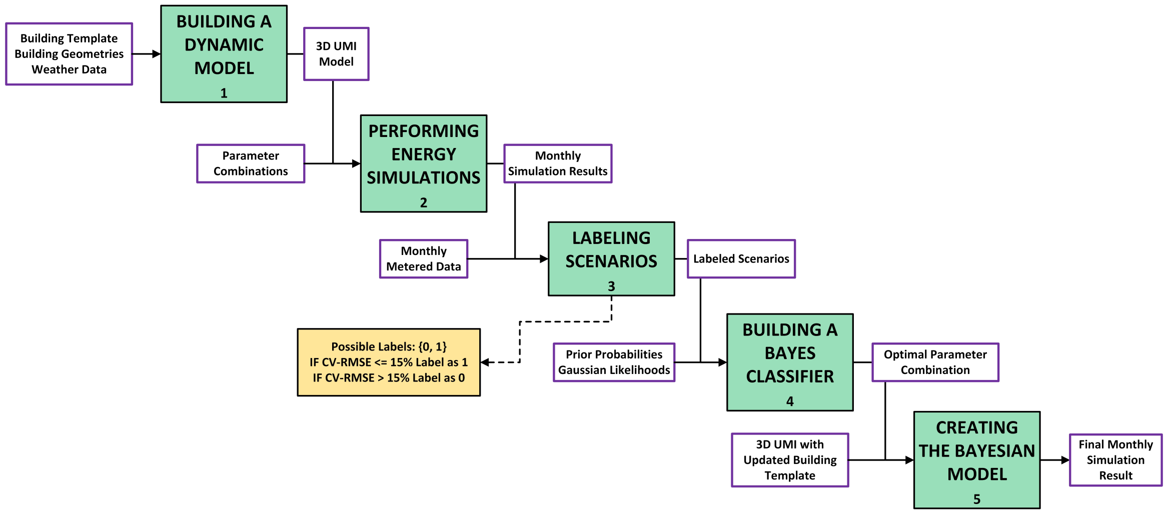


Figure 1: Methodology workflow

ios according to their simulation errors and subsequently perform classification to derive the optimal parameter values. To achieve this, a maximum allowable error limit for CV-RMSE is defined. ASHRAE suggests that the simulation (or prediction) error should not exceed 15% at the monthly resolution based on the evaluation metric CV-RMSE (ASHRAE, 2002). Thus, the maximum allowable error is set to 15% of CV-RMSE, and each scenario is labeled according to the total operational energy use intensity. For example, if a scenario's error is less than or equal to the maximum allowable error, the scenario is labeled one, and zero otherwise. The scenarios with label one are considered as desired scenarios.

4. Building A Bayes Classifier

The optimization becomes a binary classification task after labeling scenarios. Here, the goal is to derive the posterior probability distributions for the scenario labels (zero and one) using the Bayes formula in Equation 3:

$$P(L_i \mid X) = \frac{P(X \mid L_i)P(L_i)}{P(X)} \tag{3}$$

In Equation 3, X denotes the parameter combination, and L_i represents the scenario label. There are two possible labels for each scenario: $L_i = \{0, 1\}$. The posterior distributions for the labels zero and one are determined using the likelihood function $P(x \mid L_i)$ and prior probabilities $P(L_i)$ since the denominator $P(X)$ is the same for both labels. The term $P(X)$ in the denominator is called evidence, which is the probability of observing a certain parameter regardless of a condition. The prior probability $P(L_i)$ is the ratio of the number of scenarios with label i (zero or one) to the total number of scenarios. The likelihood function is the joint probability distribution of the simulation parameter combinations for each label. Here, the simulation parameters are assumed to be conditionally independent for simplicity of the optimization (Equation 4. This transforms the model into a Naive Bayes classifier.

$$P(X \mid L_i) = \prod_{m=1}^{n} P(x_m \mid L_i) \tag{4}$$

The likelihood function is the multiplication of the conditional probability of each simulation parameter over a given label. In Equation Equation 4, m is five since there are five parameters in scenarios. Each conditional probability $P(x_m \mid L_i)$ is assumed to have a Gaussian (normal) distribution with the parameters μ_m (mean) and σ_m (standard deviation) as shown in Equation 5.

$$P(x_m \mid L_i) = \frac{1}{\sigma_m\sqrt{2\pi}} \exp\left(-\frac{(x_m - \mu_m)^2}{2\sigma_m^2}\right) \tag{5}$$

Here, the objective is to find the optimal normal distribution parameters for each conditional probability $P(x_m \mid L_i)$. This is achieved by using the maximum likelihood estimation, wherein the gradient of the likelihood function is taken with respect to the distribution parameters (mean and standard deviation) to derive the optimal distribution parameters. These means and standard deviations are chosen to maximize the likelihood of the observed data given each label. Consequently, there are two posterior distributions with optimal mean and standard deviation corresponding to the labels zero and one. The expected value of a simulation parameter x_m, representing the average outcome expected to occur over many repetitions of the random event, can be calculated from its posterior probability distribution conditioned on a specific label L_i. Therefore, the expected values of the simulation parameters are the means of their posterior distributions based on labels one and zero, Here, it is essential to differentiate between simulation and distribution parameters to avoid confusion. Once the classification over the labeled scenarios using a Gaussian Naive Bayes classifier is implemented, each simulation parameter gets posterior distributions for the labels zero and one. The expected values of the simulation parameters are then derived from the means of these posterior distributions. As the optimal values of the simulation parameters should fall within the desired label distributions, the optimal parameter combination is obtained using the expected values of the distributions conditioned over label one.

5. Creating The Bayesian Model

The optimal values of the simulation parameters are integrated into the 3D UMI Model as an updated building template. This updated model forms the Bayesian model. Subsequently, the Bayesian model is simulated, and the monthly simulation results are compared to metered data to evaluate the optimization performance.

Case Study

The dormitory area of Özyeğin University campus in Istanbul was selected as the sample district. There are six dormitory buildings in this area Figure 2. Two energy audit reports from 2014 and 2020, which contain essential information about the energy performance of the dormitory buildings, were used to identify and characterize buildings in the building template. The dormitory buildings have similar building properties and energy consumption patterns. Therefore, a single archetype, named Dormitory, was created to represent all the buildings in the dynamic model. To create building geometries, polygons representing the building footprints in the district on the map were initially drawn in Yandex Map Constructor. These footprints were then saved as a map layer and exported as a GEOJSON file. Subsequently, this file was processed in the UBEM.io tool, in which a GIS interface ensured that the buildings were located in valid coordinates and that the footprints were not overlapping. Building heights were incorporated into UBEM.io to finalize the geometric model. There are multiple weather stations near campus. The hourly weather data between September 2018 and September 2023 were obtained from the Turkish State Meteorological Service (TSMS) from the nearest station to the campus. This data includes temperature, relative humidity, dew point temperature, wind direction and speed, sun exposure intensity, global solar radiation, and total precipitation at the hourly resolution within the given years. After obtaining three major components of UMI, a dynamic model of the district was established.

The metered data obtained from the campus energy management department includes natural gas and electricity consumption. This data forms the study's validation data. The dynamic model created via UMI considers several end-uses when determining the operational energy consumption of the dormitory buildings. Among these end-uses, cooling, equipment use, and lighting energy are set to be provided with electricity, whereas heating and DHW are supplied by natural gas. The electricity and natural gas consumption are recorded for each building at the monthly resolution. Even if small air conditioners cool the offices of the authorized dormitory personnel, it is assumed that the dormitory buildings do not have a cooling system. The validation data was derived from the average monthly metered consumption data between September 2018 and September 2023. Notably, the time intervals overlapping with the COVID-19 pandemic significantly impacted the energy consumption patterns of the dormitory buildings due to the absence of occupancy. Thus, the period spanning March 2020 to October 2021 was excluded from the weather and validation data.

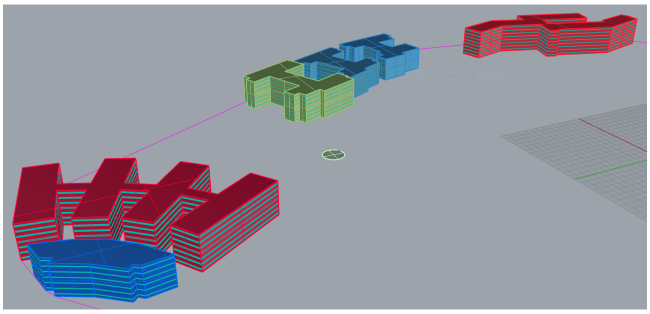


Figure 2: 3D dormitory model in Rhino

The building properties in the energy audit reports were processed, and the average values of the available parameters were used to characterize the archetype in the dynamic model. Table 1 shows the details of the essential parameters of the dynamic model. The dormitory buildings on the campus have similar characteristics. Y1, Y2, Y3, and Y4/5 buildings were built in 2012, whereas Y6 and GH were built in 2019. Hence, there is a small variation in the envelope properties of the dormitory buildings. Some parameters could not be obtained from the energy audit reports, such as infiltration rate, equipment power density (EPD), lighting power density (LPD), and DHW flow rate. There needs to be prior knowledge about the possible value ranges for these parameters. Thus, these parameters were assumed to hold the most uncertainty among the parameter space of the dynamic model and were selected for optimization. In addition, the heating set point was manually selected for optimization since this parameter can drastically affect the intensity of the zone conditioning and the resulting energy consumption.

The unit of the infiltration rate is the air change per hour (ACH), which denotes the ratio of the inside air volume replaced by the outside air due to the natural leakages on the building surface. The air leakage from the building surfaces is expected to be low in the campus buildings since dormitory buildings are constructed after 2012. Therefore, the upper limit of the infiltration rate was set to 0.5, where the minimum was 0.1. Similarly, an upper and lower limits were determined for EPD (Deru et al., 2011) (Mahajan et al., 2017) (Chang and Crawley, 2018), LPD (Deru et al., 2011) (ASHRAE., 2020), and DHW flow rate (Deru et al., 2011) (Murakawa et al., 2005) (Pérez-Lombard et al., 2008) using the literature. Overall, five parameters were chosen for optimization. The average values detected from the literature and the energy audit reports formed the input of the Baseline Model. The possible value ranges for these simulation parameters are outlined in Table 2. Five distinct values were generated for each

Table 1: Simulation parameters of the dynamic model

Parameter Name	Unit	Prior Knowledge
Wall U-Value	$W/m^2/K$	Available
Roof U-Value	$W/m^2/K$	Available
Ground U-Value	$W/m^2/K$	Available
Window U-Value	$W/m^2/K$	Available
Occupant Density	$people/m^2$	Available
Heating and Cooling Set Points	Degree Celsius	Available
Heating COP	%	Available
Infiltration	ACH	Not Available
EPD	W/m^2	Not Available
LPD	W/m^2	Not Available
DHW Flow Rate	$m^3/h/m^2$	Not Available

simulation parameter within the minimum and maximum bounds from Table 2 with equal intervals. This resulted in a total of 3125 parameter combinations. These diverse combinations were simulated in UMI. Monthly simulation errors against the metered data were then calculated using CV-RMSE (Equation 1) and MAPE (Equation 2) metrics. The metered data for monthly natural gas and electricity consumption are available for the dormitory buildings. However, for the sake of simplicity in the calibration process, the total operational energy (TOE) use intensities in kWh/m^2, representing the TOE over the gross floor area of the dormitory buildings, were utilized to compute scenario errors against the metered data. These scenarios were categorized according to their error, with those exhibiting an error smaller than 15% of CV-RMSE labeled as one, and zero otherwise.

Lastly, a simplified Bayesian optimization approach classified these scenarios by employing a Gaussian Naive Bayes classifier. The classifier determined posterior probability distributions for scenario labels based on a likelihood function derived from the simulation parameter combinations. Maximum likelihood estimation determined optimal distribution parameters (mean and standard deviation) for each conditional probability distribution. The expected values of simulation parameters were then obtained from the mean of the posterior distributions. The combination of these simulation parameters formed the Bayesian Model. Subsequently, UMI simulations were performed once more for the Bayesian Model.

Table 2: Optimization ranges for the selected parameters

Parameter	Baseline Value	Min. Value	Max. Value
Infiltration	0.30	0.10	0.50
Heating Set Point	21.50	20.00	23.00
EPD	7.50	5.00	10.00
LPD	6.00	3.00	9.00
DHW Flow Rate	2.55E-04	7.09E-05	7.09E-04

Results and Discussion

The optimal values for the simulation parameter were determined as the means of the posterior distributions for label one (Table 3). The new scenario with the optimal simulation parameters formed the Bayesian Model. A fundamental way to detect the optimal scenario combination is to select the scenario with the least CV-RMSE manually. The manually selected parameter combination with the least CV-RMSE was named the Deterministic Model. According to Table 4, the Deterministic Model decreased the simulation CV-RMSE of the Baseline Model by nearly 60%, whereas the Bayesian Model decreased the baseline error by around 45%. Even though the deterministic calibration seems to provide the optimal parameter combination with the least simulation error, this can be deceptive.

Table 3: Parameter combinations of different models

Parameter	Baseline	Deterministic	Bayesian
DHW Flow Rate	3.90E-04	5.49E-04	6.11E-04
EPD	7.50	5.00	7.41
Heating Set Point	21.50	21.50	20.93
Infiltration	0.30	0.40	0.39
LPD	6.00	3.00	5.77

Let us recall that the simulation parameters were calibrated based on a single target end-use, TOE use intensity. TOE provides an aggregated energy use as it is the summation of electricity and natural gas consumption values. Table 5 and Table 6 demonstrate detailed error comparisons of the calibrated model based on the monthly CV-RMSE and MAPE values. According to Table 5 and Table 6, it is evident that the Bayesian Model achieved a more balanced error distribution between electricity and natural gas. Specifically, the combined average CV-RMSE for electricity and natural gas is 0.267 for the Bayesian Model, compared to 0.318 for the Deterministic Model. This superior performance of the Bayesian Model can be attributed to its detailed parameter combination. As illustrated in Table 3, the simulation parameters of the Bayesian model exhibit distinct values, unlike the predetermined values within optimization ranges provided in Table 2. The likelihoods in the Gaussian Naive Bayes classifier enabled the determination of more precise parameter values. Consequently, these refined values contributed to a more stable and accurate electricity and natural gas consumption simulation at the monthly resolution. Therefore, the proposed Bayesian calibration is more robust than the deterministic calibration.

Table 4: Monthly simulation errors of different models

Model	CV-RMSE	MAPE
Baseline	0.196	0.149
Deterministic	0.080	0.069
Bayesian	0.107	0.101

The efficiency of the proposed Bayesian calibration can be further enhanced when considering the limitations of the Gaussian Naive Bayes classifier. This classifier relies on the naive assumption that simulation parameters are conditionally independent, which may not be entirely true in all cases. The correlation between these parameters can significantly impact the resultant energy consumption. Therefore, modeling the Bayesian likelihood in Equation 4 as a multivariate probability distribution, where the simulation parameters are considered conditionally dependent, could yield more reliable calibration results.

Table 5: Monthly CV-RMSE comparison of the calibrated models based on different end-uses

Model	TOE	Electricity	Natural Gas
Deterministic	0.080	0.478	0.159
Bayesian	0.107	0.277	0.257

The Gaussian Naive Bayes classifier also assumes that simulation parameters follow a Gaussian distribution with a single mean and variance. However, in reality, nonlinear relationships may exist between between the simulation parameters and the resultant energy consumption. Utilizing non-parametric models would enable the detection of these nonlinear patterns in the actual probability distributions of the simulation parameters.

Table 6: Monthly MAPE comparison of the calibrated models based on different end-uses

Model	TOE	Electricity	Natural Gas
Deterministic	0.069	0.466	0.260
Bayesian	0.101	0.156	0.295

Most parameter combinations were considered undesirable scenarios based on their simulation results. Figure 4 shows that almost 90% of the scenarios have more than 15% of CV-RMSE. This caused an imbalanced distribution in the target variable with 2774 zeros and only 351 ones. When the model was trained using these observations, the learning process predominantly focused on the majority class (zeros) rather than the minority class (ones). Therefore, the model tended to classify scenarios as belonging to the majority class. According to Figure 3, the Gaussian Naive Bayes classifier accurately predicted 94% of the zero-labeled, while it only correctly predicted 71% of the one-labeled observations. The imbalanced distribution of the zeros and ones in the target variable induced the classifier to struggle to identify which simulation parameter influences the scenario result and to what extent. Handling such imbalanced data can drastically improve the classification performance and the calibration results.

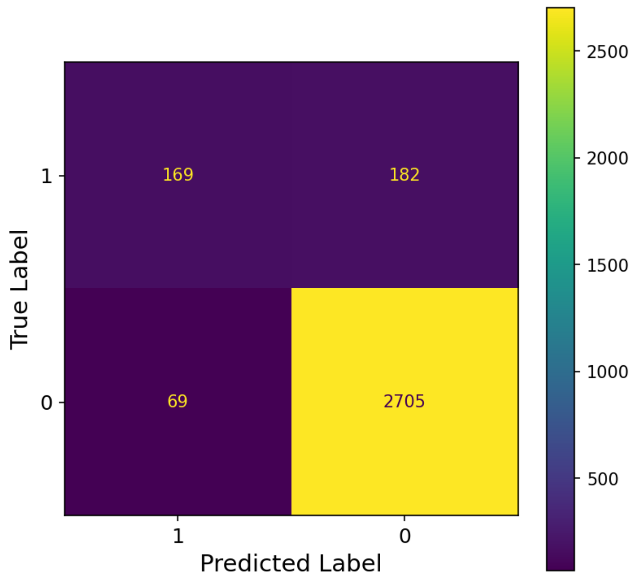


Figure 3: Confusion Matrix for Gaussian Naive Bayes Classifier

Conclusions

Bayesian optimization is widely used to calibrate building energy models. It involves surrogate modeling to reduce computational complexity since the dynamic models

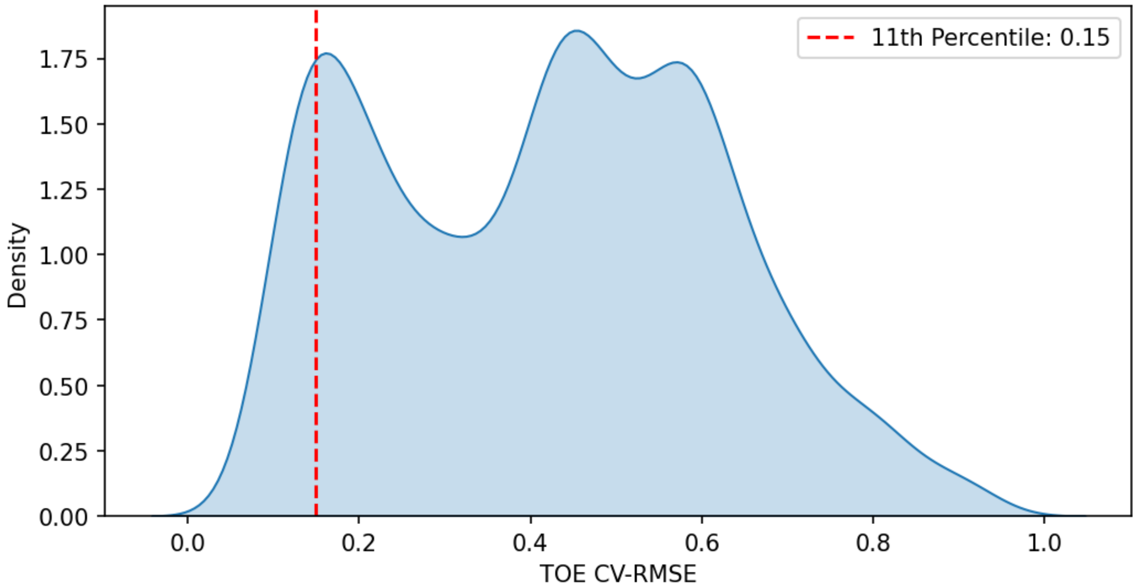


Figure 4: Probability Density Function for the CV-RMSE of the TOE

require extensive simulation time. However, the variability in building morphology and energy consumption patterns complicates surrogate modeling at the urban scale. This study proposes a simplified Bayesian optimization approach to calibrating building energy models at the district and urban scales while retaining dynamic modeling. The proposed approach introduces a novel framework for the parametric optimization of district-building energy models by filtering and reducing the parameter space to enhance the efficiency of the optimization process.

The proposed approach minimizes computational demands for optimizing dynamic models while ensuring reliability by leveraging Bayesian decision theory. Based on the monthly simulation errors, the methodology starts by labeling scenarios with varying combinations of the simulation parameters as either desired or undesired. Subsequently, a binary classifier is employed to categorize these scenarios. To achieve this, the Bayesian likelihoods are modeled as Gaussian probability distribution, with the simulation parameters being conditionally independent. Once the optimal distribution parameters (mean and standard deviation) are determined using the maximum likelihood estimation based on the scenario errors, the expected values of the simulation parameters obtained from the probability distributions constitute the Bayesian Model. This allows for acquiring more detailed and reasonable values of the simulation parameters without the need for exhaustive dynamic simulations. A case study implemented over a dormitory area on a university campus demonstrated the proposed approach's effectiveness. Here, the Bayesian Model significantly reduced the monthly simulation CV-RMSE of TOE for the Baseline Model by around 45% while also delivering balanced and accurate monthly simulation results for the dormitory buildings' electricity and natural gas consumption.

The approach's ability to derive optimal parameter combinations from labeled scenarios offers a practical solution for optimizing dynamic energy models. Urban planners can utilize it to enhance the accuracy of the district building energy models and reduce the energy consumption of the building stocks after testing various energy-efficiency scenarios. A possible future work of this study is to focus on integrating more robust classification models that involve correlations between the simulation parameters and the non-linear patterns in their probability distributions. Another future work can be overcoming the imbalanced label distribution in the target variable to improve the classification performance and the probabilistic calibration. Moreover, classifying scenarios based on natural gas and electricity consumption rather than TOE can help detect seasonal variations in building energy consumption patterns and identify essential parameters affecting seasonal building energy demand.

Acknowledgments

This study was supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under the Grant Number ”222M488”. The authors thank TUBITAK for their support.

References

Ang, Y. Q., Berzolla, Z., Letellier-Duchesne, S., Jusiega, V., and Reinhart, C. (2022). UBEM.io: A web-based framework to rapidly generate urban building energy models for carbon reduction technology pathways. Sustainable Cities and Society, 77:103534.

ASHRAE. (2020). ANSI/ASHRAE/IES Standard 90.1-2019: Energy Standard for Buildings Except Low-Rise Residential Buildings. ASHRAE.

ASHRAE, A. (2002). Ashrae guideline 14: measurement of energy and demand savings. American Society of Heating, Refrigerating and Air-Conditioning Engineers, 35:41–63.

Buckley, N., Mills, G., Reinhart, C., and Berzolla, Z. (2021). Using urban building energy modelling (UBEM) to support the new European Union's Green Deal: Case study of Dublin Ireland. Energy and Buildings, 247:111115.

Chang, R. and Crawley, D. (2018). Influence of plug and process loads and occupancy on ultimate energy savings â€“ a new approach. In 2018 Building Performance Analysis Conference and SimBuild co-organized by ASHRAE and IBPSA-USA, volume 8 of SimBuild Conference, pages 713–720, Chicago, USA. ASHRAE/IBPSA-USA.

Chong, A., Gu, Y., and Jia, H. (2021). Calibrating building energy simulation models: A review of the basics to guide future work. Energy and Buildings, 253:111533.

Deru, M., Field, K., Studer, D., Benne, K., Griffith, B., Torcellini, P., Liu, B., Halverson, M. A., Winiarski, D., Rosenberg, M., Yazdanian, M., Huang, J., and Crawley, D. B. (2011). U.S. Department of Energy Commercial Reference Building Models of the National Building Stock. Technical report.

Ferrando, M., Causone, F., Hong, T., and Chen, Y. (2020). Urban building energy modeling (UBEM) tools: A state-of-the-art review of bottom-up physics-based approaches. Sustainable Cities and Society, 62:102408.

Hou, D., Hassan, I., and Wang, L. (2021). Review on building energy model calibration by Bayesian inference. Renewable Sustainable Energy Reviews, 143:110930.

Mahajan, V., Srinivasan, R. S., Chini, A. R., and Ries, R. (2017). Space-Level Plug-Load densities of educational buildings on university campuses. Journal of Energy Engineering-asce, 143(2).

McNeel, R. and Associates (2008). Rhinoceros: Nurbs Modeling for Windows. Robert McNeel and Associates.

Murakawa, S., Nishina, D., Takata, H., and Tanaka, A. (2005). An analysis on the loads of hot water consumption in the restaurants.

OECD (2023). CO2 Emissions in 2022.

Oraiopoulos, A. and Howard, B. (2022). On the accuracy of Urban Building Energy Modelling. Renewable Sustainable Energy Reviews, 158:111976.

Pan, Y., Zhu, M., Liu, Y., Yang, Y., Liang, Y., Yin, R., Yang, Y., Jia, X., Wang, X., Zeng, F., Huang, S., Hou, D., Xu, L., Yin, R., and Yuan, X. (2023). Building energy simulation and its application for building performance optimization: A review of methods, tools, and case studies. Advances in Applied Energy, 10:100135.

Pérez-Lombard, L., Ortiz, J., and Pout, C. (2008). A review on buildings energy consumption information. Energy and Buildings, 40(3):394–398.

Reinhart, C. and Davila, C. C. (2016). Urban building energy modeling – A review of a nascent field. Building and Environment, 97:196–202.

Reinhart, C., Dogan, T., Jakubiec, A., Rakha, T., and Sang, A. (2013). UMI – an urban simulation environment for building energy use, daylighting and walkability. Building Simulation Conference proceedings.

Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning representations by back-propagating errors. Nature, 323(6088):533–536.

Santin, O. G., Itard, L., and Visscher, H. (2009). The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock. Energy and Buildings, 41(11):1223–1232.

Wang, D., Landolt, J., Mavromatidis, G., Orehounig, K., and Carmeliet, J. (2018). CESAR: A bottom-up building stock modelling tool for Switzerland to address sustainable energy transformation strategies. Energy and Buildings, 169:9–26.

PASSIVHAUS POST OCCUPANCY EVALUATION AND PROPOSED BIM-BASED INTELLIGENT MODEL

Adam Novogrodskis[1], Zaid Alwan[1], and Bahriye Ilhan Jones[1]
[1] Northumbria University, Newcastle upon Tyne, United Kingdom

Abstract

Passivhaus has been gradually adopted worldwide as the main voluntary building performance standard. While numerous governmental policies and environmental initiatives have promoted the use of low-carbon housing in the UK, the value of Passivhaus has not been fully realised. This paper focuses on UK Passivhaus homes through the perceptions of end-users living in large-scale developments. Case studies are used to obtain feedback on the buildings' performance in the use phase via questionnaire. All surveyed Passivhaus properties were perceived to be very comfortable during the winter, autumn and spring seasons; however, residents were less satisfied with thermal comfort in the summer. This suggests that the standard needs to continuously evolve to address occupant needs. Digital tools can play a role to achieve this. The study concludes a BIM-based process for post occupancy evaluation to overcome heating and thermal discomfort issues.

Introduction

Being responsible for 19% of total greenhouse gas emissions and over 25% of total energy consumption in the UK, decarbonisation of the building sector, especially housing, is crucial in meeting UK's overarching climate targets (UK Housing, 2019). To achieve this, the UK government over past 10 years announced a transition towards zero-carbon buildings and launched the Code for Sustainable Homes policy to assess, certify and promote zero-carbon sustainable design and construction in new buildings (McLeod et al., 2012). However, with the subsequent 'watering-down' of the zero-carbon legislation and eventual withdrawal of the Code for Sustainable Homes in 2015, the UK construction industry has started to debate the validity of the zero-carbon concept as a realistic target (Lynch, 2014). There has been an ongoing debate amongst Architecture Engineering and Construction (AEC) professionals as to how best to address significant carbon reductions, rather than zero- carbon in construction, and whether to deal with materials, fabric, or occupant behaviour. Unlike the net zero-carbon housing policy which primarily focuses on carbon emissions and offsetting by using intensive renewable energy, the Passivhaus standard concentrates on user comfort and conserving energy through a thermally-efficient fabric, passive solar design, airtightness, thermal-bridge-free construction and use of Mechanical Ventilation and Heat Recovery (MVHR).

Several studies have indicated that the Passivhaus standard performs considerably better compared to both conventional and low-carbon housing in the UK (8, 9, 10) and its delivery and adoption in the industry have been limited to approximately 124 units a year, which is 0.07% of 170,000 new-build homes annually constructed in the UK (Ministry of Housing, 2020). Factors suggest that slow Passivhaus standard uptake could be a result of the general private housing sector's reluctance to innovate and adopt energy-efficient home standards (Goodchild and Walshaw, 2011). Concerns with the Passivhaus standard such as its complexity and higher build costs as well as issues regarding excessive insulation, overheating, MVHR and necessary lifestyle changes as a hindrance to its delivery at scale are also factors affecting the adoption (Forde et al., 2020), (Zhao and Carter, 2020), (Pitts, 2017). Hence while there is a public aptitude for low-carbon housing, the construction sector has been not responsive to such needs.

This paper aims to investigate the contemporary drivers of and barriers to delivering the Passivhaus standard in the UK from the user perspective, covering both primary and secondary data. This is achieved through (a) analysing three distinct large-scale residential Passivhaus developments and comparing their occupant experiences, and (b) examining Passivhaus occupants' behavioural patterns and identifying habits/factors that influenced perceived thermal comfort through targeted post occupancy evaluations (POE). The study also proposes a BIM-based intelligent post occupancy evaluation model as a future direction.

Research method

The study adopts quantitative research by conducting a questionnaire through case studies. A total of 174 properties across three different passive house developments are included. Quantitative research in the form of POE was employed as it allowed to gathering of firm, factual data on how the completed Passivhaus buildings were being operated and provided a 'snapshot' of its users' experience and satisfaction at this particular time (Fellows and Liu, 2021). This is vital as the research is trying to identify trends and themes from a variety of cross-comparable accounts (Lucas and Lucas, 2016). POEs have become the preferred industry method to monitor in-use building performance and assess energy performance and air quality as well as collect feedback on thermal comfort, space use, and overall user satisfaction (Meir et al., 2009). The questionnaire was prepared following guidelines from the Post Occupancy Evaluation Toolkit (HEFCE, 2006) as well as specific POE examples (Mlecnik, 2013a). The POE was divided into seven sections including (1) background, (2) occupancy, building type, and use, (3) thermal comfort, (4) energy use, (5) building services and ventilation, (6) air quality and behavioural patterns, and (7) overall impressions. The questionnaire was comprised of 38 questions in total. Most questions were either a checkbox or multiple-choice; however, a 5-point Likert rating scale was also utilized in questions relating to user satisfaction, perceived thermal comfort, air quality, and energy-savings. POEs were collected between September and December 2019 through door-to-door canvassing, postal surveys as well as invited online questionnaires. A total number of 49 POEs were cumulatively collected across the three case studies. Within the scope of this study due to the space limitation, the focus will be on the thermal comfort and behavioural change, living standard improvement and overall satisfaction.

Collected POE questionnaires were then imported into statistical software, Statistical Package for the Social Sciences (SPSS), for further quantitative analysis. Each POE question and the collected answer was correspondingly coded into either a nominal, ordinal or scale variables and analysed to find frequencies, associations and correlations between the answers. Firstly, the analysis was run on nominal demographic variables such as age, gender, tenure, house type, etc. to identify reoccurring occupant and household types, which were then compared to determine if they presented differences regarding comfort, energy savings and satisfaction. The analysis was then conducted on occupants' motivations to live in a Passivhaus property. Lastly, collected data were examined to determine perceived behavioural change, living standard improvement and overall satisfaction with the properties across the case studies. As the majority of collected data was nominal, not normally distributed, and the sample size was fairly limited, non-parametric statistical methods had to be used (George and Mallery, 2019). Therefore, depending on the level of measurement, the collected data were analysed using either a crosstabulation test to determine the association; Mann-Whitney U-test to compare differences between independent groups; or Spearman rank-order correlation to measure the relationship between different variables (George and Mallery, 2019). Particular attention was given to ordinal Likert scale questions asking occupants to rate their satisfaction with a particular aspect of their Passivhaus dwellings such as thermal comfort.

Case studies

Three case studies of the main large-scale residential Passivhaus developments in the UK were conducted to analyse their inception, delivery, and real-life occupancy. The case study methodology was selected for the research as it enabled practical analysis of a broad phenomenon such as Passivhaus standard integration in the UK, which otherwise would have been difficult to examine (Feagin et al., 2016). Investigations by Robert K. Yin highlighted case studies can 'shed empirical light on some theoretical concepts or principles' (Yin, 2009). Just as a laboratory experiment, instead of trying to extrapolate empirical probabilities (statistical generalisation), it allows us to investigate, expand and generalise existing theories (analytical generalisation), e.g., living in a Passivhaus affects occupants' behaviour and perceived comfort (Yin, 2009).

Passivhaus case studies were chosen based on their scale, time of construction, location as well as type of tenure, thus aiming to acquire currently lacking typological data identified by the literature review. The Forgebank project in Lancaster was chosen as a case study because it is currently the only certified Passivhaus cohousing project in the UK, hence providing a unique insight into energy use and behavioural patterns in a highly interactive and social co-housing environment. Racecourse Estate project was chosen because of its specialised design for occupants with mobility needs as well as for having the best co-heating test results among all monitored Passivhaus dwellings in the UK. Finally, the Goldsmith Street project in Norwich was chosen because it is the biggest Passivhaus development in the UK to date as well as the only Passivhaus and social-housing project ever to win the Royal Institute Building Architects (RIBA) Stirling Prize 2019 for excellence in architecture. Besides, all three case studies were chosen so they would differ from each other in terms of their location, size, completion date, and tenure, therefore allowing us to examine the Passivhaus standard under varying conditions.

Case study 1: Forgebank Co-housing

Forgebank Co-housing is a 41-dwelling affordable community housing project in Halton, Lancaster built in 2013. The project comprises a range of house types ranging from flats to two and three-bedroom terraces. 35 of these dwellings are within the co-housing scheme with shared community facilities. All dwellings in the project were designed to Passivhaus, Lifetime Homes and Code for Sustainable Homes level 6 standards (Passivhaus Trust, 2020). At the time of its construction, Forgebank was the second-largest Passivhaus development in the UK and it still is the largest and only Passivhaus cohousing project in the UK. The development went beyond the fabric performance requirements of the Passivhaus standard. It utilised a biomass district heating system, solar thermal network as well a community micro-grid powered by the hydro-electric scheme and photovoltaic panel array mounted on most dwellings' roofs (Lancaster Cohousing, 2020). A review of the building performance evaluation (BPE) revealed that energy-performance aspirations and the Passivhaus methodology were adopted early in the project and Passive House Planning Package (PHPP) software was used to model and test different design options while an airtightness champion was appointed by the contractor to overlook air-barrier installation and air pressure tests (Innovate UK, 2016). Overall, the dwellings were built to a very high standard and performed exceptionally well in co-heating tests, thus practically eliminating the energy performance gap (Lancaster Cohousing Project, 2013).

Case study 2: Racecourse Estate

Racecourse Estate is a 28-bungalow housing development for elderly residents with mobility needs built in 2011 as part of a wider Racecourse Estate regeneration masterplan aiming to replace old housing stock with 4,000 sustainable new dwellings - comprised of 25 terraced (Passivhaus certified and CfSH level 4), and 3 detached bungalows (CfSH level 5) designed specifically for mobility-impaired tenants (Technology Strategy Board, 2014). At the time of its completion in 2011, Racecourse Estate was the largest Passivhaus development in the UK and currently is still the largest Passivhaus development in the North East of England (Passivhaus Trust, 2013). Each house was designed to have an open-plan living and kitchen area, two bedrooms and a mezzanine floor which functioned both as plant roof and loft space, with a total floor area of 66 m2. External walls and roofs were constructed using prefabricated timber cassettes filled with high levels of insulation, while the ground floor was built using a traditional reinforced concrete slab with 300mm insulation and a screed above, with several good areas of practice during construction including careful detailing, fixing damaged areas and use of appropriate gaskets and putty to seal any gaps in service penetrations (Technology Strategy Board, 2014). Moreover, very close collaboration between the project architect and the contractor was observed throughout the project with numerous training events and workshops to ensure on-site construction met raised airtightness and quality targets.

Case study 3: Goldsmith Street

Goldsmith Street is a large social housing scheme built in Norwich city as part of a city council's wider corporate plan to address increasing social housing needs and tackle fuel poverty (Priest, 2019). The development is comprised of 105 units: 45 two-bed terraces and 60 one-bed flats, and arranged in seven terrace blocks laid out east-to-west emulating the Victorian street layout of the adjacent 'Golden Triangle' district (Waite, 2019). In 2019 the Goldsmith Street development was nominated and won the RIBA's Stirling Prize for excellence in architecture becoming both the first Passivhaus and social housing project ever to win such an accolade (Passivhaus Trust, 2019). Moreover, with 105 housing units, Goldsmith Street became the largest Passivhaus scheme in the UK to date (Figure 1).

Figure 1: Goldsmith Passivhaus Development, image by http://www.mikhailriches.com/

Unlike many projects of such scale, the project was delivered through a traditional procurement route which in hindsight allowed the architects to have better control of the build quality, value engineering and retaining the project's initial design aspirations (Waite, 2019). Throughout the project, the design of the buildings took priority: special attention and care were taken to ensure the Passivhaus standard was working for the design rather than the design becoming subservient to make the Passivhaus work (Greengauge, 2020).

Results of Post Occupancy Evaluations

Post-occupancy evaluations were conducted in three different Passivhaus developments varying in their scale, location, typology and ownership (Table 1).

Table 1: House types identified across the case studies

Case	House Type	Construction Type
Forgebank, Halton [Co-housing]	Terrace, End-of-terrace, Flat	Masonry cavity
Racecourse Estate, Houghton-le-Spring [Affordable housing]	Terrace, End-of-terrace, Semi-detached	Timber frame
Goldsmith Street, Norwich [Social housing]	Terrace, End-of-terrace, Flat	Timber frame

Thermal comfort and heating control behaviour

Collected POE data indicated that all surveyed Passivhaus properties were perceived to be very comfortable during the winter, autumn and spring seasons; however, residents were less satisfied with thermal comfort in the summer months. Overall 22% identified they were either less or not pleased with indoor temperatures during the summer season, referring to overheating as the main issue. Concerning heating, SPSS analysis revealed there were some notable differences between the case studies in terms of heating control behaviour and devices used to control indoor temperature.

A comparison of winter heating patterns (Figure 2) highlighted that Goldsmith Street had the most habitual residents. They were likely to have the heating on at a predetermined time of day, i.e. mornings and evenings. In contrast, Forgebank occupants could be considered the most 'reactive'- typically heating their properties only when felt needed. Racecourse Estate residents, on the other hand, were found to be the most energy-conserving, over half of which indicated never needing to have the heating on in their properties.

In terms of heating control and ventilation devices (Fig-

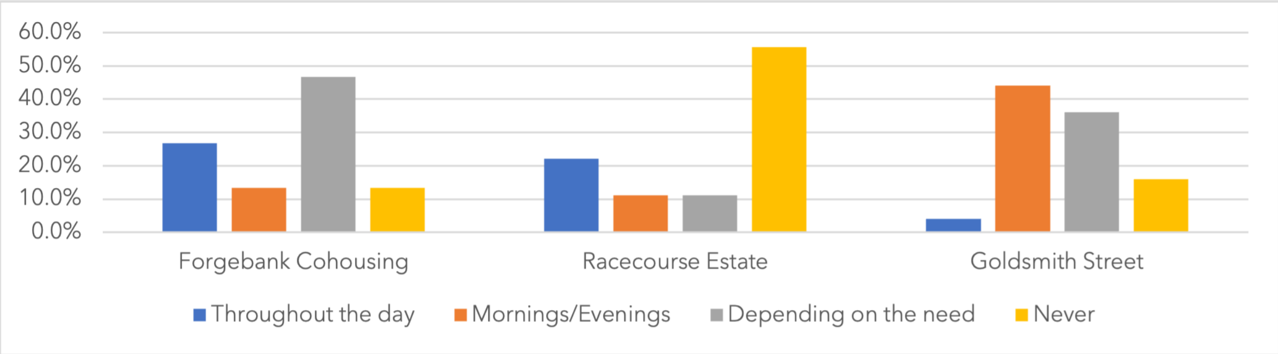


Figure 2: Comparison of winter heating patterns by users across the different case studies

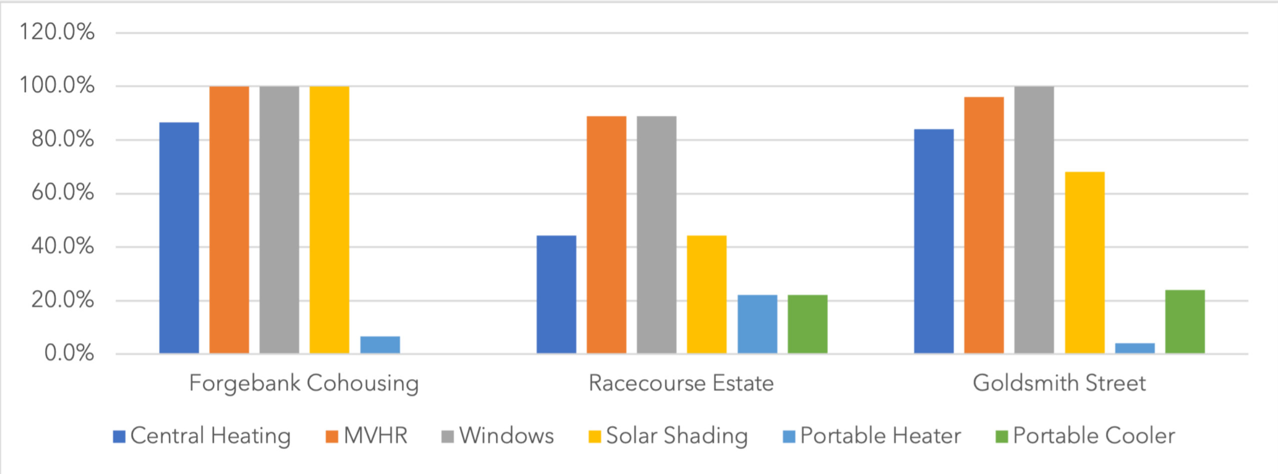


Figure 3: Comparison of winter heating patterns by users across the different case studies

ure 3), collected data indicated that Forgebank and Goldsmith St. residents were still heavily reliant on a conventional central heating system (used by >80% of respondents), in contrast to Racecourse Estate were only 44.4% of end-users reported needing the central heating. That said, Racecourse Estate inhabitants were also found to be the most likely to need a portable cooler to prevent summer overheating, as opposed to Forgebank and Goldsmith St. residents which indicated a much lower need for cooling. All case studies indicated a high use of MVHR and operable windows. Yet, notably, fewer Racecourse residents reported using solar shading, which might explain the summer overheating and need for additional cooling.

Living preferences, lifestyle and overall satisfaction

The last section of the POE research focused on occupants' general perceptions, overall satisfaction and changes to daily lifestyle. Firstly, research participants were asked to identify three main reasons why they chose to live in their Passivhaus property. Radar charts (Figure 4) showed that 'sustainability' was ranked to be the top priority for choosing to live in a Passivhaus across all three case studies (71%). Other top reasons were found to be 'house type'(47%) and 'house price' (43%) while 'community' was ranked as a top priority in the Forgebank development.

Research participants were then asked if living in a Passivhaus has caused a change towards a more sustainable way of life. Collected data revealed that the majority of occupants (41 out of 49) believed living in a Passivhaus has influenced them to adopt a more sustainable lifestyle with 33% of participants reporting a significant change, 49% reporting moderate change and 18% reporting a slight but noticeable change. In terms of overall satisfaction with Passivhaus buildings, collected POE data highlighted that surveyed residents were exceptionally pleased with their homes across all three case studies. Approximately 94% of all participants said they were satisfied or very satisfied with their homes, 6% were neither satisfied nor dissatisfied, and there were no occupants who were dissatisfied with their homes. Furthermore, 65% of residents reported that living in a Passivhaus has significantly improved their living standard, 33% identified a moderate improvement, and 2% found that living in a Passivhaus has not affected their standard of living. In retrospect, 98% of surveyed end-users reported they would recommend it to their friends and family as well as choose to live in a Passivhaus home again if given the opportunity.

Summary of the results

Post-occupancy evaluation data and subsequent SPSS analysis revealed there were significant differences between the case studies in terms of their occupant demo-

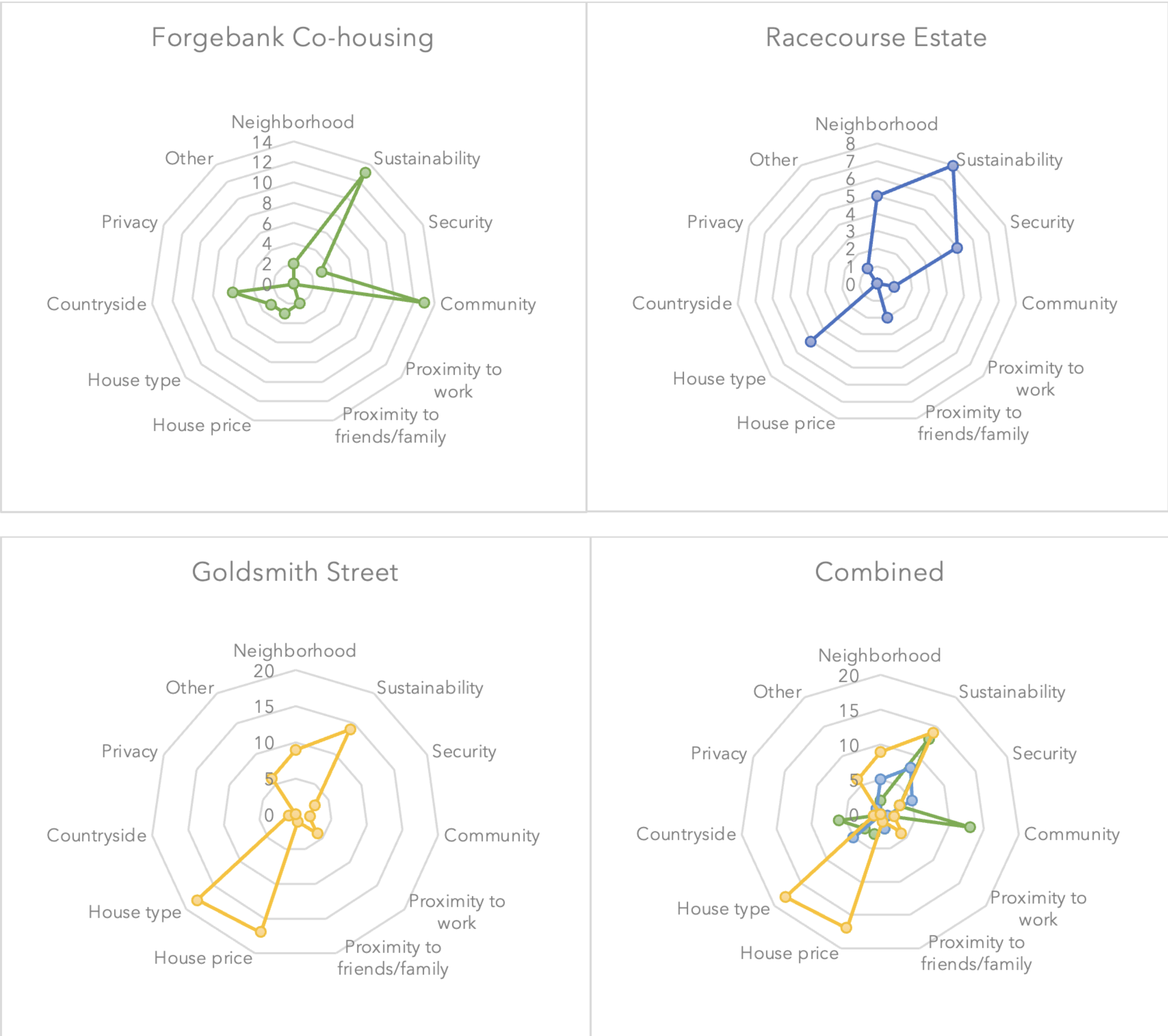


Figure 4: Residents' preferences for Passivhaus elements from the three different case studies

graphics, household types, occupancy and heating patterns as well as heating and cooling strategies. Nonetheless, the SPSS analysis also indicated many commonalities and reoccurring patterns between the case studies regarding thermal comfort and overall satisfaction.

Examined POE data revealed that Passivhaus dwellings across all three case studies were found to be exceptionally comfortable during the winter, autumn and spring seasons. However, all three case studies were also found to have consistently lower thermal satisfaction in the summer months due to reported overheating. Out of the case studies, Racecourse Estate was found to have the lowest summer comfort satisfaction with a mean value of ($\mu = 3.33$). Such results confirm previous study findings on the Racecourse Estate, where substantial overheating was recorded (Siddall et al., 2014), (Fletcher et al., 2017). In terms of overall satisfaction, collected POEs indicated that very high user satisfaction was reported across all three case studies despite significant differences between the Passivhaus properties. The vast majority of residents identified being either satisfied or very satisfied with Passivhaus thermal comfort, energy savings, building services, air quality as well as the overall effect it had on their standard of living.

Such high levels of reported occupant satisfaction confirm previous Passivhaus research findings from both Continental Europe (Cutland, 2012), (Mlecnik, 2013b) as well as the UK (Siddall et al., 2014) where Passivhaus properties were found to be exceptionally comfortable and energy-efficient.

Proposed POE model

Questionnaire results reveal that a more structured POE process is needed to monitor the Passivhaus performance during the occupancy. POE is not structured and not linked to the characteristics of BIM in the design phase. Materials used and the impact of building materials such as conductivity and location etc. are not evaluated during the use

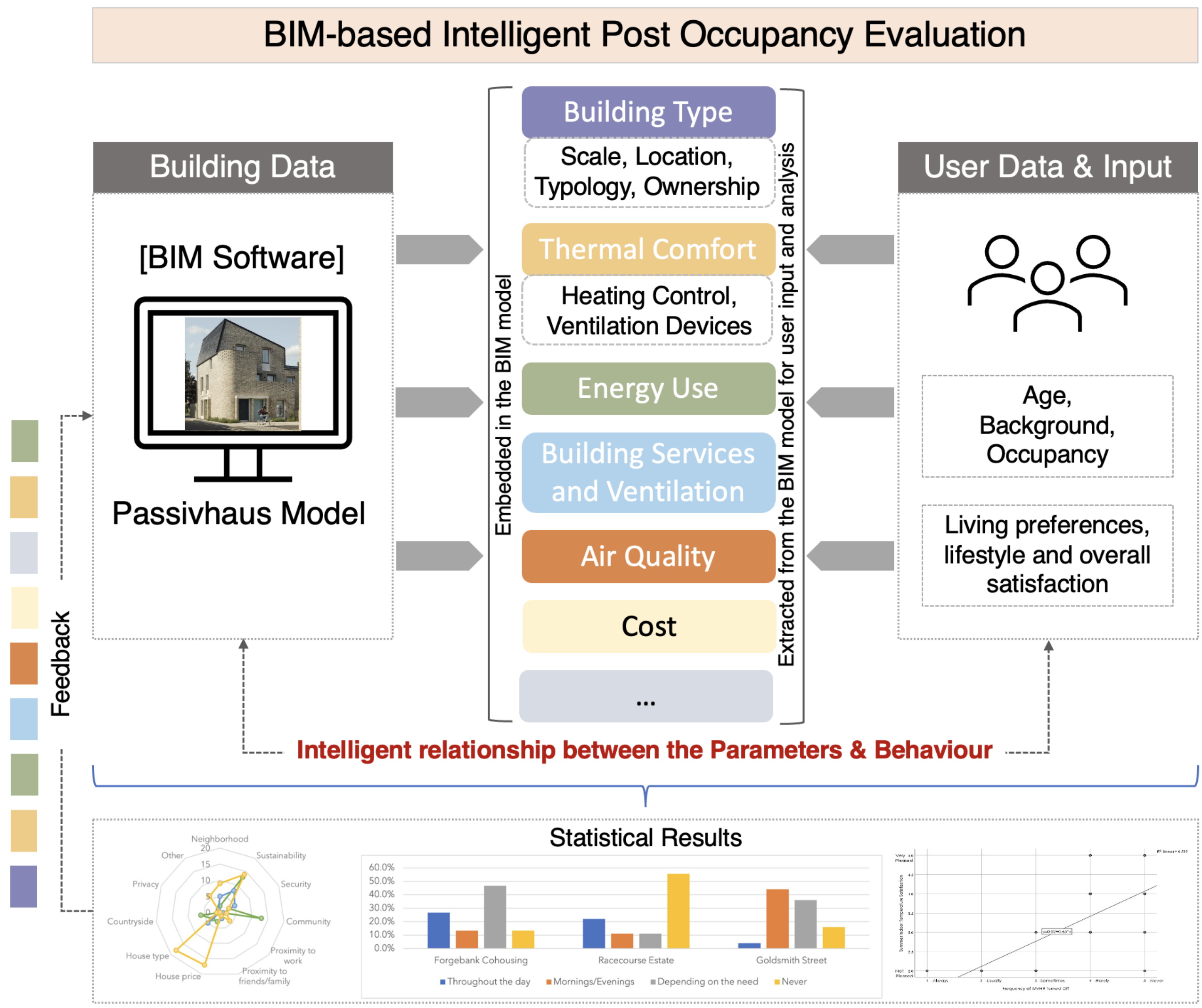


Figure 5: BIM-based intelligent post occupancy evaluation model

phase of the building life cycle.

BIM and other intelligent technologies can help adjust the optimum indoor comfort and energy use as well as identify the related problems based on the actual data for the occupants. Passive houses are very specific and have special features and details, the data should be reflected in the BIM models.

Within the scope of this study, a conceptual attempt was made due to the lack of passive house and BIM data. Enrichment of the BIM model and passive house is possible using the parameters such as building type, thermal comfort, energy use and cost etc. Based on the available data within the BIM model, for example, scale and location can be mapped easily while air quality and energy use cannot be easily integrated until data is available after several years of building occupation. As a baseline, the proposed model is expected to be used enriched for future research and provide solid results. Figure 5 presents a BIM-based intelligent post occupancy evaluation model.

Towards a more digitalised built environment, to facilitate the adoption of Passivehaus, and improvement of design decisions and consequently user satisfaction, automated and smart solutions should be utilised. A prototype sensor visualisation platform can be designed to connect in-use performance data to BIM context data to provide actionable advice for landlords and tenants for minimising repair and maintenance activities. This is mainly because assets do not perform as well as they should when built by comparison to the design phase (Lewry, 2015). Such variations have increased the appetite for carrying out post occupancy evaluation as mentioned in this paper for normally built domestic and commercial assets, however, this is not necessarily the case in the passive house due to the strict air tightness regulations and conditions. This study has revealed that there are still issues associated with important aspects such as the 6 factors mentioned in the model, which are not always integrated with a user-driven feedback process for better building operations for passive buildings. Application programme interfaces (API) could be linked into the system to provide further insight into the conditions that affect housing assets. The measurement of light levels requires further investigation to acknowledge tenant behaviour patterns such as only requiring light at certain times for certain activities (Rogage et al., 2020). In this case, APIs could record simple temperatures, and since the data could be used over time within passive house properties to influence thermal comfort levels, this can prevent issues such as overheating. The data centres could be connected to thermal satisfaction interfaces which could be modified every day depending on the rating given by the occupant on a sliding scale. Thus, over time an optimised model can be developed by feeding this data back into the BIM model specifically designed for asset management of passive houses. Furthermore, the development of artificial intelligence (AI) is a huge potential for programming such processes to take into account occupancy feedback within the building to reduce consumption even more.

Conclusion

This study aimed to analyse the Passivhaus standard in the UK in its present condition through current occupant experiences and discuss the future direction via the proposed POE model. Forty-nine post-occupancy evaluations conducted in three separate case studies across the UK revealed that occupants were generally very pleased with their Passivhaus dwellings. Residents reported high levels of satisfaction with both internal thermal comfort and perceived impact on health. In general, research results were consistent with previous literature findings and the hypothesis that the Passivhaus methodology can deliver highly comfortable homes and substantial energy and carbon savings regardless of the building's location, occupancy or construction. Participants believed that the Passivhaus standard had the potential to evolve to form a foundation for a more robust sustainable housing policy and improvement in understanding occupant behaviour currently lacking in the industry. The findings also highlight there has been a significant lack of longitudinal research and POE of Passivhaus properties over a long period to investigate the longevity of its thermal properties as well as the evolvement of end-user experience. Therefore, up-to-date research investigating the current construction industry's perceptions of the standard as well as revisiting occupant experience in large-scale Passivhaus properties to gain deeper understanding is needed over the long term as we move towards ever more stringent low-carbon targets, through more automated and digital solutions.

References

Cutland (2012). Lessons from germany's passivhaus experience, nhbc foundation. https://www.nhbc.co.uk/binaries/content/assets/nhbc/foundation/lessons-from-germanys-passivhaus-experience.pdf. Accessed: 2024-02-11.

Feagin, J. R., Orum, A. M., and Sjoberg, G. (2016). A case for the case study. UNC Press Books.

Fellows, R. F. and Liu, A. M. (2021). Research methods for construction. John Wiley & Sons.

Fletcher, M., Johnston, D., Glew, D., and Parker, J. (2017). An empirical evaluation of temporal overheating in an assisted living passivhaus dwelling in the uk. Building and Environment, 121:106–118.

Forde, J., Hopfe, C. J., McLeod, R. S., and Evins, R. (2020). Temporal optimization for affordable and resilient passivhaus dwellings in the social housing sector. Applied Energy, 261:114383.

George, D. and Mallery, P. (2019). IBM SPSS statistics 26 step by step: A simple guide and reference. Routledge.

Goodchild, B. and Walshaw, A. (2011). Towards zero carbon homes in england? from inception to partial implementation. Housing Studies, 26(6):933–949.

Greengauge (2020). Winner: Riba stirling prize 2019. https://ggbec.co.uk/portfolio/ambitious-me-design-for-uks-largest-passivhaus-social-housing-scheme/. Accessed: 2022-05-30.

HEFCE (2006). Post occupancy evaluation toolkit, space management group. http://www.smg.ac.uk/supp_occupancy.html. Accessed: 2023-05-15.

Innovate UK (2016). Building performance evaluation. https://www.ukri.org/councils/innovate-uk/. Accessed: 2023-04-30.

Lancaster Cohousing (2020). Lancaster cohousing. https://www.lancastercohousing.org.uk/About. Accessed: 2024-01-20.

Lancaster Cohousing Project (2013). A certified passivhaus / code for sustainable homes (level 6) and life time homes, affordable community housing project. https://ecoarc.co.uk/wp-content/uploads/2016/05/Lancaster-Green-Building-Magazine-Articles.pdf. Accessed: 2023-01-20.

Lewry, A. (2015). Bridging the performance gap: Understanding predicted and actual building operational energy. Journal of Building Survey, Appraisal and Valuation, 3:360–365.

Lucas, R. and Lucas, R. (2016). Research methods for architecture. Hachette UK.

Lynch (2014). Passivhaus in the uk: the challenges of an emerging market. https://discovery.ucl.ac.uk/id/eprint/1418470/. Accessed: 2019-12-20.

McLeod, R. S., Hopfe, C. J., and Rezgui, Y. (2012). An investigation into recent proposals for a revised definition of zero carbon homes in the uk. Energy Policy, 46:25–35.

Meir, I. A., Garb, Y., Jiao, D., and Cicelsky, A. (2009). Post-occupancy evaluation: An inevitable step toward sustainability. Advances in building energy research, 3(1):189–219.

Ministry of Housing (2020). Communities and local government. https://www.gov.uk/government/news/number-of-new-homes-built-soars-to-an-11-year-high. Accessed: 2020-01-23.

Mlecnik, E. (2013a). Improving passive house certification: recommendations based on end-user experiences. Architectural Engineering and Design Management, 9(4):250–264.

Mlecnik, E. (2013b). Innovation development for highly energy-efficient housing: Opportunities and challenges related to the adoption of passive houses.

Passivhaus Trust (2013). Uk passivhaus trust awards 2013. https://www.passivhaustrust.org.uk/uk_passivhaus_awards/2013/. Accessed: 2022-02-11.

Passivhaus Trust (2019). Goldsmith street crowned prize-winning passivhaus. https://www.passivhaustrust.org.uk/news/detail/?nId=859. Accessed: 2022-05-28.

Passivhaus Trust (2020). Lancaster cohousing project. https://www.passivhaustrust.org.uk/projects/detail/?cId=35. Accessed: 2020-04-30.

Pitts (2017). Passive house and low energy buildings: Barriers and opportunities for future development within uk practice. http://dx.doi.org/10.3390/su9020272. Accessed: 2020-05-29.

Priest (2019). The morality tale of norwich's social housing, riba journal. https://www.ribaj.com/buildings/mikhail-riches-goldsmith-street-social-housing-norwich-passivhaus-stirling-prize. Accessed: 2022-05-17.

Rogage, K., Clear, A., Alwan, Z., Lawrence, T., and Kelly, G. (2020). Assessing building performance in residential buildings using bim and sensor data. International Journal of Building Pathology and Adaptation, 38:176–191.

Siddall, M., Johnston, D., and Fletcher, M. (2014). Occupant satisfaction in uk passivhaus dwellings. In 18th International Passive House Conference 2014, pages 491–496.

Technology Strategy Board (2014). Case study, racecourse, building performance evaluation. https://goodhomes.org.uk/wp-content/uploads/2017/05/gha-case-study-racecourse-full.pdf. Accessed: 2020-02-11.

UK Housing (2019). Uk housing: Fit for future? committee on climate change. https://www.theccc.org.uk/publication/uk-housing-fit-for-the-future/. Accessed: 2020-12-06.

Waite (2019). Riba stirling prize 2019 goes to goldsmith street council housing scheme, architects' journal. https://www.architectsjournal.co.uk/news/riba-stirling-prize-2019-goes-to-goldsmith-street-council-housing-scheme#:~:text=Jury%20chair%20Julia%20Barfield%2C%20said,a%20development%20of%20this%20scale. Accessed: 2022-05-28.

Yin, R. K. (2009). Case study research: Design and methods, volume 5. Sage.

Zhao, J. and Carter, K. (2020). Do passive houses need passive people? evaluating the active occupancy of passivhaus homes in the united kingdom. Energy Research & Social Science, 64:101448.

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

A GRAPH BASED FRAMWORK TO SUPPORT DATA-DRIVEN URBAN BUILDING ENERGY SIMULATIONS

Usman Ali[1], Sobia Bano[1], Cathal Hoare[1], Muhammad Haris Shamsi[2], Divyanshu Sood[1] and James O'Donnell[1]

[1]School of Mechanical and Materials Engineering and UCD Energy Institute, UCD, Dublin, Ireland
[2]VITO, Hasselt, Belgium

Abstract

Urban planners and energy policymakers increasingly focus on sustainable urban development and the challenges of analyzing complex urban energy systems. Current models often lack the integration of diverse urban datasets and do not adequately address the dynamic nature of urban energy demands. This study proposes a data-driven framework that involves data collection and preprocessing, building archetypes, machine learning modeling, and parametric simulation. The novel contribution of this research lies in defining the scope, processes, information, data, and relationships for the ontology of urban building energy modeling, employing a graph-based approach for complex data integration. The proposed methodology is tested in residential buildings in Dublin City to examine and compare the modeling results. The study concludes that the proposed model offers a more comprehensive and adaptable approach to urban energy analysis compared to traditional methods. Furthermore, the study helps stakeholders by providing a scalable and flexible modeling framework for urban energy analysis.

Introduction

In the era of rapid urbanization, sustainable development in cities has become a crucial global challenge. The European Union has established a robust legislative framework to enhance sustainable planning and improve the energy performance of buildings. This framework is underpinned by two key directives, the Energy Performance of Buildings Directive EU/2010/31 and the Energy Efficiency Directive EU/2023/1791, both revised in 2023. These directives lead EU member states to adopt policies that aim to achieve a highly energy-efficient and decarbonized building stock by 2050 (EU-Energy, 2023).

Efficient and sustainable urban environments are critical for identifying scalable energy conservation strategies. A promising approach involves analyzing building energy performance data at an urban scale using a data-driven approach. However, the available urban data are often sparse, inconsistent, and lacking in diversity and heterogeneity. Despite these challenges, the last few decades have seen significant advancements in data-driven modeling, particularly in using sparse data to predict and estimate building energy usage. Nevertheless, a gap remains in these studies, primarily due to their focus on prediction and estimation without adequately integrating complex, multidimensional data sets. Therefore, a more comprehensive approach is essential for integrating complex urban building energy dynamics data to improve urban-scale modeling.

Urban planners and energy policymakers increasingly focus on the complexities of urban energy systems, driven by the crucial need to balance environmental sustainability with the growing energy demands of expanding urban populations. However, there is a lack of comprehensive models that effectively integrate diverse urban data sets to analyze urban energy systems. Current approaches often fail to address the dynamic and complex nature of urban energy demands. There is a growing recognition of the gap in the literature, where traditional models are insufficient to capture multifaceted interactions within urban environments. This gap necessitates a novel approach that can holistically encapsulate the intricate energy dynamics of cities.

This study introduces an innovative data-driven framework for urban building energy modeling, taking advantage of ontology and graph-based approaches. The ontology provides a structured representation of domain knowledge, allowing a clearer definition of concepts, attributes, and relationships (Poveda-Villalón et al., 2022; Curry et al., 2013; Hoare et al., 2022). Combining the ontology with a graph database enables creating of a comprehensive knowledge representation system that can efficiently handle diverse and interconnected data, promoting a more holistic understanding of urban energy systems.

Ontology and graph-based techniques are widely used in the field of Building Information Modeling (BIM), linked building data, urban planning, and energy management (Pritoni et al., 2021; Terkaj et al., 2017; Costin and Pauwels, 2022; Li et al., 2019). Kapsalis et al. (2022) use a graph analysis approach to efficiently query and analyze energy efficiency certificates. Wu et al. (2021) propose an ontology modeling solution for managing decentralized data for household energy systems. Van Dam and Keirstead (2010) initialize a model of an urban energy system built on an ontology using energy transformation. Wu et al. (2022) propose an ontology-based framework that can integrate data to build energy simulations from different data sources in the Operation Phase. Baumgärtel et al. (2014) used an ontology framework to assess the application of building performance regulations in design and operation. Daneshfar et al. (2022) proposed an ontology to represent geospatial data to support building renovation to collect data in IFC and CityGML format. Zadeh et al. (2019) developed a hybrid information infrastructure by integrating building design data in ifcXML format and 3D neighborhood models in CityGML format. Shi et al.

(2023) proposed a methodology based on ontology to create a digital twin city information model by integrating BIM, GIS, and IoT technologies. However, existing studies focus mainly on urban building energy management and building information modeling, but more research is needed in the field of urban building energy modeling. Furthermore, available urban building energy models typically do not emphasize data integration, flexibility, scalability, and performance, which are critical to adapting and managing the complexities of urban energy systems.

The novelty of this research lies in defining the scope, processes, information, data, and relationships for an urban building energy modeling ontology. The application of such an ontology uses a graph-based approach to assimilate and analyze complex urban energy data, enabling a more holistic understanding of urban energy dynamics. Furthermore, the framework enriches the predicted machine learning-based data of complex energy modelling scenarios, an aspect rarely explored in existing studies. This integration allows integration of diverse datasets, ranging from building stock to energy consumption patterns, and fosters a deeper understanding of the interdependent factors that influence urban energy systems. This study aims to offer a more comprehensive and adaptable approach to urban energy analysis, surpassing the limitations of traditional methods.

The article is organized as follows: Section 2 provides a detailed discussion of the methodology devised for urban energy modeling of residential buildings. Section 3 discusses the case study of Irish building stock. Finally, conclusions are discussed in Section 4.

Methodology

Modeling the energy performance of buildings on a large urban scale presents a formidable challenge for urban planners and policymakers. Accurate prediction of energy consumption and identification of energy efficiency opportunities are essential to promote the sustainable development of cities.

Therefore, this study proposes a framework for modeling the energy performance of buildings on an urban scale, which begins with the development of an ontology and its implementation in a graph database. This lays the foundational framework for representing and analyzing the complex interrelations of urban energy performance (Figure 1). The process involves comprehensive data collection and preprocessing to gather and refine building-related data. Subsequently, building archetypes are developed to define representative models of buildings, capturing the diversity of the urban built environment. These archetypes serve as the basis for parametric simulations, exploring the energy performance under various scenarios and conditions and generating a synthetic dataset for analysis. Machine learning models are then employed to predict energy performance across the urban scale, leveraging the insights gained from synthetic and real-world data to identify opportunities for energy efficiency improvements. Finally, this structured approach facilitates querying and data analysis, providing a detailed understanding of urban building energy performance. This holistic approach integrates sophisticated data analysis and modeling techniques to offer actionable insights for promoting sustainable urban development.

Urban Building Energy Modeling Processes

The proposed ontology for urban building energy modeling includes building archetypes and parametric simulation data, which are used to generate synthetic data (Figure 3). Ontology also incorporates machine learning modeling results, focusing on predicting the energy performance of buildings based on their respective archetypes. The proposed ontology includes entities such as *BuildingArchetypes, BuildingParameters, ConstructionTemplates, EnergyUsage, and MLPredictedEnergyPerformance*. As a result, this interconnected data enrichment creates a robust platform for urban modeling. This improves understanding of the energy dynamics of urban buildings and enables the development of targeted strategies to improve energy efficiency. The proposed ontology mapping on desired scenarios requires data collection and preprocessing, archetype development, parametric simulation, and machine learning modeling processes. These processes are designed to map complex urban data and scenarios from existing buildings onto ontologies and predict building energy performance.

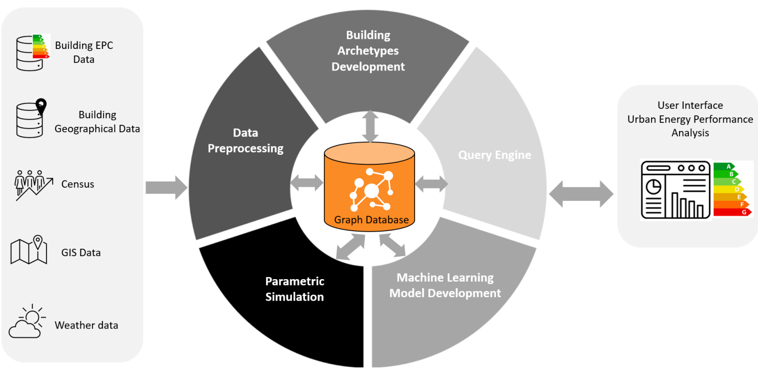


Figure 1: Methodology for data-driven graph-based urban building energy modeling

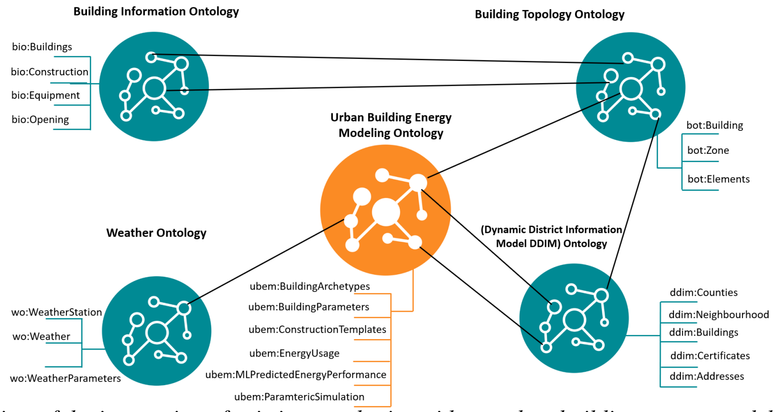


Figure 2: A conceptual view of the integration of existing ontologies with an urban building energy modeling ontology with required entities

Graph Database Development

A graph database is developed to efficiently manage and store complex interconnected data. This database enables the representation of relationships between various data points, such as the link between building types and energy usage patterns, facilitating more detailed analysis and insights (Curry et al., 2013; Zhu et al., 2022). The commonly used graph databases are Neo4j, Amazon Neptune, and OrientDB. A graphical representation of the complex graph model can be presented using the ontology (Hoare et al., 2022).

Ontology refers to a structured representation of knowledge that defines concepts, relationships, and properties within a specific domain, such as urban energy systems (Costin and Pauwels, 2022). These ontologies are represented using the Web Ontology Language (OWL). The proposed ontology for urban building energy modeling has been developed by integrating existing ontologies into the building domain, such as the Building Information Ontology, Building Topology Ontology, Weather Ontology, and Dynamic District Information Model Ontology (Figure 2). Building Information Ontology (BIO) provides a range of defined classes, axioms, and data types for reuse, including Building, BuildingElement, and BuildingParameter (TUWien, 2024a). Building Topology Ontology (BOT) is a minimal ontology to describe the core topological concepts of a building (Rasmussen et al., 2021). To represent weather-related information for a location, the specific Weather Ontology (WO) provides reusable patterns and terms (TUWien, 2024b). The Dynamic District Information Model (DDIM) offers a national-scale digital twin for domestic building stock, complete with geographic information (Hoare et al., 2022).

Data Collection

The ontology creation step starts with collecting available raw homogeneous data sources, including building energy performance certificate (EPC) data, geographic data, census information, GIS, and weather data. These data are gathered from various sources, including national databases and urban planning departments. The data are carefully selected to ensure relevance and accuracy for urban energy modeling.

Data Preprocessing

Once collected, the data undergo a rigorous pre-processing phase. This process involves cleaning the data and handling missing values to ensure consistency. Pre-processing also includes categorizing buildings based on characteristics such as age, region, and type of usage, which is crucial for developing accurate building archetypes. The preprocessed data are mapped to existing BIO, BOT, DDIM, and WO, further supporting the urban building energy ontology.

Building Archetypes Development

Building archetypes are developed to represent different categories of buildings with similar characteristics within the urban scale. Each building archetype serves as a core entity (*ubem:BuildingArchetypes*) for graph-based integration of urban buildings and requires specified features, such as geometry, layout, construction materials, insulation levels, and typical energy use patterns. These data can be extracted from existing ontologies, namely the BOT and the BIO. Similarly, the DDIM ontology provides all the geographical information for the building archetype. These archetypes serve as foundational elements for a parametric simulation framework that models energy performance in various types of buildings. In addition, these archetypes are essential for urban modeling because they encapsulate common building characteristics that can be used to generalize energy performance assessments.

Parameters Simulation

Parametric simulation generates synthetic data to simulate various energy consumption and performance scenarios to find the optimal solution, especially when a sparse data set is available for energy modeling (Ali et al., 2024). Parametric simulation uses developed archetypes and also sources data from the weather ontology for simulation. It stores all input and output results in a graph-based structure using building parameters (*ubem:BuildingParameters*), Construction (*ubem:ConstructionTem-*

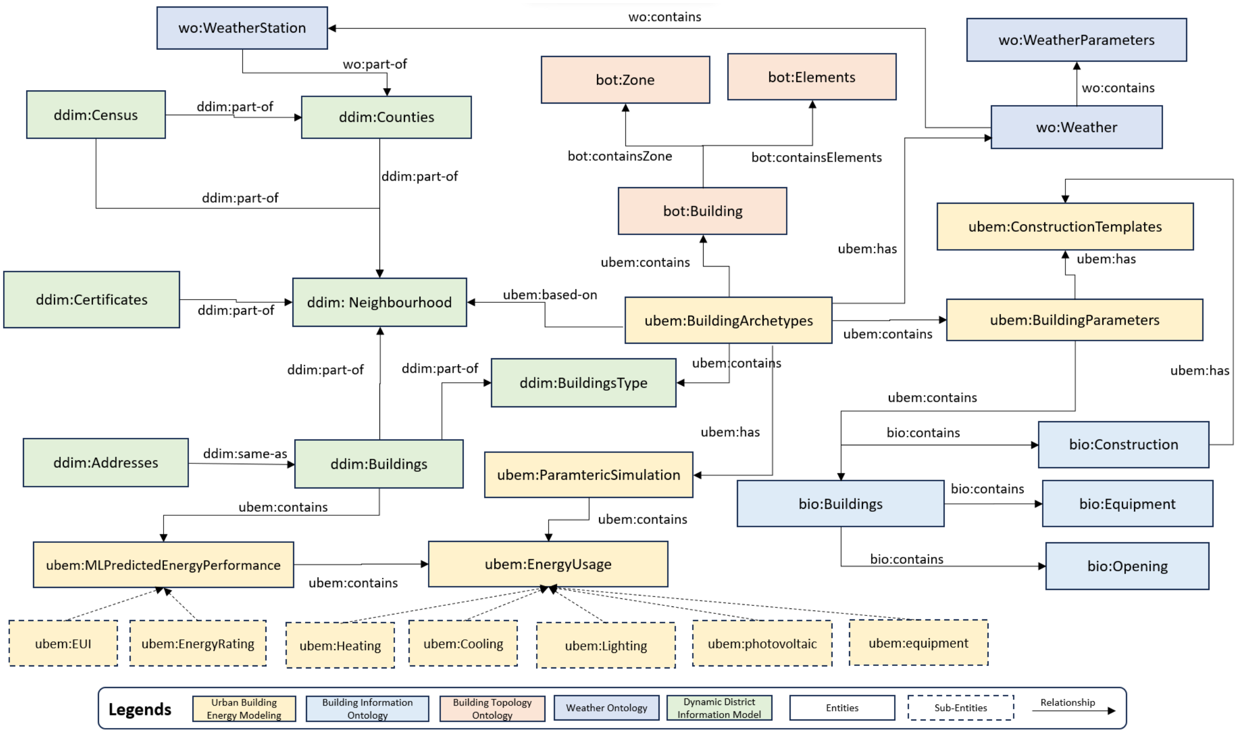


Figure 3: Graph structure for urban building energy modeling for residential building stock

plates), and the energy usage (*ubem:EnergyUsage*) entity. (*ubem:ParamtericSimulation*) entity estimates are determined using thermal modeling to assess heating, lighting, and water demands, and employing an energy simulation engine to evaluate the overall energy performance of the building under different conditions. A parametric tool performs numerous simulations using a simulation model (Ali et al., 2024) to perform complex parametric simulations involving multiple parameters. In this paper, JEPlus serves as a parametric tool for energy simulations. Additionally, JEPlus uses EnergyPlus for simulations, integrating various parameter values together with weather data and construction templates. However, due to the complex nature of the numerous parameters involved, generating simulated data for all parameters becomes nearly impossible. Therefore, synthetic data are generated using sampling methods such as Simple Random Sampling (SRS) and Latin Hypercube Sampling (LHS). These methods help generate the desired sample data that includes combinations of all parameters. Parametrically simulated data produce synthetic data sets stored back in the graph database and then used as input for developing machine learning models.

Machine Learning Modeling

Data-driven Machine-Learning (ML) models have been designed to enrich the predicted results for urban energy use scenarios in graph databases. The workflow involves formulating regression ML models to store predicted results in the energy performance entity of the building, specifically in terms of Energy Use Intensity (EUI) and energy rating. The model is trained using the data stored in the graph database. Generally, the models developed include the process of splitting data into training and testing sets, followed by the application of regression algorithms based on performance indices (Ali et al., 2024). These models are continuously refined to improve accuracy and adaptability to evolving urban dynamics. Furthermore, the models predict the intricate characteristics of buildings on an urban scale and store these predicted data in a graph database using *ubem:MLPredictedEnergyPerformance* entity for further analysis. The enrichment of ML-based building energy performance in graph-based databases offers dynamic, context-aware insights, significantly enhancing the precision and relevance of urban building energy modeling.

Query Engine

Developing a query engine enables efficient retrieval and data analysis from the graph database developed based on the proposed ontology. This engine supports complex queries, allowing users to extract specific insights, such as determining the influence of particular building characteristics on energy consumption or identifying potential areas for energy efficiency improvements. The most common graph database query languages are Cypher, GraphQL, and SPARQL (Kapsalis et al., 2022; Zhu et al., 2022; Hoare et al., 2022). Cypher is one of the most widely used query languages for graph databases, especially in the context of Neo4j. GraphQL is a query language primarily designed for APIs, but it can also be employed to query data from graph databases. On the other hand, SPARQL is a query language specifically used for querying RDF (Resource Description Framework) data. The proposed ontology, coupled with graph-based integration for urban

Table 1: Data requirements and associated data sources for Dublin graph-based database

Ontology	Prefix	Irish Data Source	Publisher
Building Information Ontology	bio	Irish EPC (BER) Database	SEAI
Weather Ontology	wo	Dublin EPW File	EnergyPlus
Building Topology Ontology	bot	Construction Templates, GeoIntel	An Post/ Ordnance Survey, DesignBuilder
Dynamic District Information Model	ddim	GeoDirectory, Irish Census database	An Post/ Ordnance Survey Ireland, Central Statistics Office
Urban Building Energy Modeling	ubem	Irish GIS database	Central Statistics Office

building energy modeling, can be further examined using a query engine based on data integration, flexibility, scalability, and performance. This study examines the query engine with and without machine learning-based data enrichment for estimating urban building energy performance.

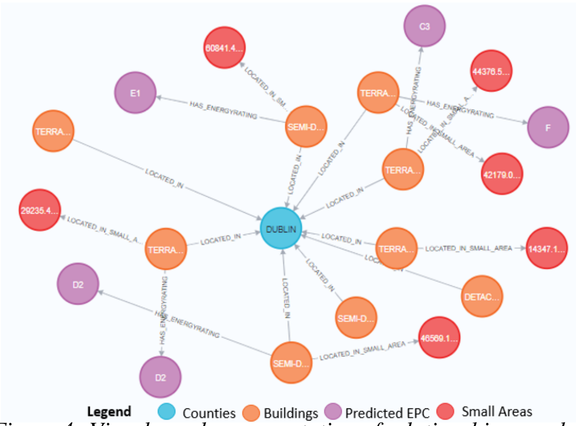


Figure 4: Visual graph representation of relationship complex urban building energy modeling for Dublin city

Urban Energy Performance Analysis

Finally, the methodology culminates in a comprehensive analysis of urban energy performance. This involves using the models and tools developed to evaluate and interpret buildings' energy efficiency and the overall energy dynamics of the urban area. Graphical or GIS-based modeling results can be integrated into the analysis using data stored in the DDIM ontology for geographical information. The analysis provides valuable information on current performance levels and identifies opportunities for optimization and improvement.

The study aims to provide a robust and comprehensive framework for urban building energy modeling, providing a deeper understanding of urban energy dynamics and paving the way for more sustainable urban development.

Case Study

The primary objective of this case study is to test the proposed methodology on the residential building stock of Dublin, Ireland. This study demonstrates the effectiveness of an ontology-driven approach in the management and analysis of complex urban building data. The application of this data-driven framework to buildings in Dublin has revealed significant insights into the dynamics of urban energy. Adopting an ontology-driven methodology supported by Neo4j graph database technology transforms the management and analysis of Dublin's residential building stock. Furthermore, this study helps in detail urban energy insights through a more structured, interconnected, and efficient data framework than conventional approaches.

Irish Graph Database Development

This study uses a Neo4j graph database to implement the proposed urban building energy modeling ontology. Neo4j efficiently stores nodes and relationships, significantly streamlining complex data management. Using Neo4j's capabilities, the study effectively represents and interlinks various urban data. The interconnected data enables comprehensive analysis of urban energy, facilitating the identification of patterns and insights that are difficult to identify with traditional methods due to complex relationships and interdependencies at an urban scale. The graph-based structure of Neo4j also allows high-performance querying and data retrieval, making it an ideal platform for handling complex queries essential for comprehensive urban energy modeling and analysis.

Gathering data on an urban scale for a building stock is a challenging task, as individual building information is often limited and sparse. The data collection process for this study involved acquiring raw building data from multiple sources to map them to the proposed or existing ontology. The case study includes data such as building energy performance certificates, building geographical data, census information, GIS, and weather data (Table 1).

In Ireland, the data set for EPC (also known as the Building Energy Rating (BER) certificate) of Irish residential stock represents a comprehensive measurement of the building stock, which includes more than 200 characteristics of the buildings. These characteristics include the building fabric, heating systems, estimated end use, CO_2 emissions, and estimated delivered and primary energy consumption. The Irish EPC dataset contained approximately 1.1 million residential buildings, with a substantial number of building ratings that fall within the C1 to D2 range (SEAI, 2023). EPC data assists in retrieving building-related data for archetype development, based on building topology and building information ontology requirements.

The Irish census provides spatial data on various scales and the number of buildings in each geographic area, such as small areas (neighborhoods), and counties (CSO, 2022). Furthermore, the GeoDirectory database, updated by An

Post (Irish Postal Service) and Ordnance Survey Ireland, offers statistical and geographical information on the entire Irish building stock. The Q2 2023 GeoDirectory report includes geocoded addresses of 2.1 million residential buildings (GeoDirectory, 2023; OSI, 2023). This study focuses on Dublin City, and the EPC data is available only for about 54% of the residential building stock in Dublin City. GeoDirectory is a useful resource for urban energy modeling and is required for the DDIM Ontology to analyze geographical aspects of modeling results. Furthermore, Dublin weather data is used to evaluate the impact of weather conditions on energy performance based on weather ontology requirements. This study integrates sparse and heterogeneous urban building data using a graph-based approach. Then, machine learning algorithms are applied to this integrated data set to predict the energy rating of the remaining 46% building stock.

After the initial data collection process, all data undergo a rigorous preprocessing phase. This phase involves cleaning the data and addressing missing values to ensure consistency. The preprocessed data mapped to the ontology are then used to create a Neo4j graph-based database. This step is crucial as it involves removing all irrelevant data and including only relevant data, which is essential to improve the performance of the query engine and the execution of complex queries for urban-scale modeling. For example, the Irish EPC database contains more than 200 features, and this process selects important features for building energy modeling based on existing studies (Ali et al., 2024).

This study uses parametrically simulated data comprising 1 million entries, generated using four residential archetypes such as semi-detached, detached, terraced, and bungalow. These archetypes were developed using available Irish building data based on data requirements of BIO, BOT, and DIMM ontologies. These parametric simulated data are created using DesignBuilder construction templates for building parameter mapping. Parametric data are stored in Neo4j and are further used for data-driven machine-learning modeling to predict the energy performance of urban buildings. Neo4j's graph database structure allows for the seamless integration of diverse Dublin data sources, providing a more comprehensive view of urban energy systems compared to traditional data approaches. Neo4j can help improve the generation of models for energy performance for urban buildings due to the richness of data and relationships available in the graph. The data are then partitioned into two subsets to create training and testing datasets, employing a cross-validation algorithm. These data sets were trained and evaluated using regression models (eXtreme Gradient Boosting) proposed in the existing study (Ali et al., 2024). Finally, the machine learning-based enrichment results are compared with those of traditional simple queries for building analysis.

```
MATCH (b:Buildings)-[:part-of]->(sa:SmallAreas),
    (b)-[:part-of]->(c:Counties {COUNTY: 'DUBLIN'})
WHERE b.BER_ENERGY_MEDIAN > 300
RETURN sa, collect(b) as Buildings
```

Figure 5: A complex query that finds buildings in small areas within Dublin County with an energy performance greater than 300 kWh/m²/yr EUI.

Dublin Energy Performance Analysis

The results show that the Neo4j graph-based database efficiently manages and stores complex data using the proposed ontology for Dublin, demonstrating its interoperability in combining various formats, including spatial and building energy data. The database stores data, including parametric simulated pre-processed data and predicted energy performance data for buildings based on their respective archetypes in Dublin (Figure 4). These interconnected data sets establish a robust foundation for urban modeling, deepening our understanding of the energy dynamics within urban buildings. Furthermore, the graph-based database developed for Dublin buildings exhibits sufficient scalability and flexibility, making it suitable for handling large and complex datasets. This accommodates the ever-evolving nature of urban energy systems and supports various urban planning scenarios.

The Cypher query offers a valuable tool for urban studies, specifically in the context of energy consumption within building infrastructures. For instance, the Cypher query helps to analyze energy efficiency within Dublin County's building infrastructure by identifying buildings with high energy use intensity, specifically those with a median energy performance above 300 kWh/m²/yr EUI (Figure 5). Similarly, the Cypher query can help with urban energy modeling in Neo4j, which is a vital tool input for the simulation of building energy performance within an urban context (Figure 6). The query gets the building parameters from BuildingParameter nodes and inputs them to Simulation nodes. With these parameters, the query sets the stage for parametric simulations that calculate energy performance metrics such as Energy Use Intensity (EUI), heating, lighting, and water usage. This process enables the detailed examination of building performances and facilitates a broader analysis of urban energy consumption patterns. The ability to dynamically adjust and simulate various building parameters provides a data-driven approach for stakeholders.

The developed graph can also help with the breakdown of buildings according to their Energy Rating from "A1" to "G." (Figure 7) The results showed a diverse range of energy efficiencies among the buildings, with the count of buildings for each energy rating varying significantly. The highest number of buildings fell into the "C2" rating, with 24,539 buildings indicating moderate energy efficiency. In contrast, the "A1" rating has the fewest buildings, with only 1,380, signifying the highest energy efficiency. Other notable findings include many buildings in the "D1" and "C3" ratings, with 24,307 and 23,252 buildings, respectively. This distribution highlights the poor energy efficiency in buildings across Dublin, providing insights into the current state of energy performance and potential areas for improvement in building energy standards.

```
MATCH (bp:BuildingParameter), (s:Simulation {id: bp.id})
SET s += {
  floorUValue: bp.`floor-uvalue`,
 doorUValue: bp.`door-uvalue`,
 roofUValue: bp.`roof-uvalue`,
 windowUValue: bp.`window-uvalue`,
 wallUValue: bp.`wall-uvalue`,
...
}
WITH s
RETURN s.id AS SimulationID,
    s.EUI AS EUI,
    s.EnergyRating AS EnergyRating,
    s.HeatingUsage AS HeatingUsage,
    s.LightingUsage AS LightingUsage,
    s.WaterUsage AS WaterUsage
```

Figure 6: A complex query that get building parameters for parametric simulation for energy performance calculation

```
MATCH (b:Buildings)-[:HAS_ENERGYRATING]->(p:BuildingsPredicted)
WHERE toUpper(b.COUNTY_NAME) = "DUBLIN"
RETURN p.BER_pred AS BER_pred, COUNT(p) AS Count
ORDER BY BER_pred
```

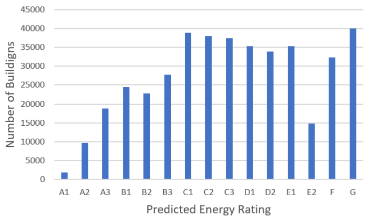


Figure 7: A cyhper query show the distribution of predicted building energy ratings in Dublin

Finally, the developed graph database allows stakeholders to formulate targeted strategies to improve energy efficiency in Dublin, including existing and predicted data, for use in future analysis. Cypher graph queries are used to analyze data within the database, enabling complex analysis. The results can be visualized on a GIS map, highlighting buildings with poor energy performance in a specific Dublin county area, based on Cypher graph queries (Figure 8). This visualization helps decision-makers identify key areas for energy efficiency improvements and informs policy decisions to promote sustainability and reduce energy consumption in the region. Moreover, the flexibility of the database in handling various data formats, including spatial and building energy data, enhances its utility for comprehensive urban planning and energy management initiatives.

Conclusions

This study underscores the growing importance of sustainable urban development and the need for an advanced framework to analyze complex spare urban energy data. The motivation behind this research was to address the

Figure 8: Visual map of buildings with poor energy performance in a specific Dublin county area based on cypher graph query

limitations of current models, which often need help integrating diverse urban data sets and addressing the dynamic nature of urban energy demands.

The key contributions of this research lie in developing a novel data-driven framework. This framework, which encompasses data collection, preprocessing, building archetypes, parametric simulation, and machine learning modeling, leads to the creation of a graph database with ontology and query engine. This approach enables a more holistic understanding of urban energy dynamics while accommodating diverse datasets. An ontology for urban building energy modeling facilitates the integration of complex urban energy data, ultimately offering a more comprehensive and adaptable approach to urban energy analysis compared to traditional methods. This study empowers various stakeholders to analyze and predict complex energy scenarios, thus supporting the creation of more sustainable and energy-efficient urban environments.

The application of the proposed framework could be expanded to include commercial buildings and energy suppliers to assess its applicability and effectiveness in diverse contexts. In general, this research lays the foundation for future advancements in sustainable urban development and energy policy, offering a promising direction for further exploration and innovation in this critical area.

Acknowledgments

This publication has emanated from the US-Ireland R&D Partnership supported by the Science Foundation Ireland through 20/US/3695, the U.S. National Science Foundation through Award Number 2217410, and the Department for the Economy in Northern Ireland through USI 167. We would also acknowledge NexSys project supported by the Science Foundation Ireland through Award Number SFI/21/SPP/3756. The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the funding agencies.

References

Ali, U., Bano, S., Shamsi, M. H., Sood, D., Hoare, C., Zuo, W., Hewitt, N., and O'Donnell, J. (2024). Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach. Energy and Buildings, 303:113768.

Baumgärtel, K., Kadolsky, M., and Scherer, R. J. (2014). An ontology framework for improving building energy performance by utilizing energy saving regulations. In Proceedings of European conference on product and process modelling, pages 519–526.

Costin, A. and Pauwels, P. (2022). Building information modeling and ontologies: Overview of shared representations. Research Companion to Building Information Modeling, 12:12–34.

CSO (2022). Census of Population 2022 - Profile 1 Housing in Ireland by Central Statistics Office. https://www.cso.ie.

Curry, E., O'Donnell, J., Corry, E., Hasan, S., Keane, M., and O'Riain, S. (2013). Linking building data in the cloud: Integrating cross-domain building data using linked data. Advanced Engineering Informatics, 27(2):206–219.

Daneshfar, M., Hartmann, T., and Rabe, J. (2022). An ontology to represent geospatial data to support building renovation. Advanced Engineering Informatics, 52:101591.

EU-Energy (2023). Energy for europe by european commission. https://energy.ec.europa.eu/index_en. [Online; accessed 08-Feb-2024].

GeoDirectory (2023). Geodirectory technical guide, an post and ordnance survey ireland.

Hoare, C., Aghamolaei, R., Lynch, M., Gaur, A., and O'Donnell, J. (2022). A linked data approach to multi-scale energy modelling. Advanced Engineering Informatics, 54:101719.

Kapsalis, P., Kormpakis, G., Alexakis, K., and Askounis, D. (2022). Leveraging graph analytics for energy efficiency certificates. Energies, 15(4):1500.

Li, Y., García-Castro, R., Mihindukulasooriya, N., O'Donnell, J., and Vega-Sánchez, S. (2019). Enhancing energy management at district and building levels via an em-kpi ontology. Automation in Construction, 99:152–167.

OSI (2023). Ordnance Survey Ireland. https://www.osi.ie.

Poveda-Villalón, M., Fernández-Izquierdo, A., Fernández-López, M., and García-Castro, R. (2022). Lot: An industrial oriented ontology engineering framework. Engineering Applications of Artificial Intelligence, 111:104755.

Pritoni, M., Paine, D., Fierro, G., Mosiman, C., Poplawski, M., Saha, A., Bender, J., and Granderson, J. (2021). Metadata schemas and ontologies for building energy applications: A critical review and use case analysis. Energies, 14(7):2024.

Rasmussen, M. H., Lefrançois, M., Schneider, G. F., and Pauwels, P. (2021). Bot: The building topology ontology of the w3c linked building data group. Semantic Web, 12(1):143–161.

SEAI (2023). Building Energy Rating Certificate Database by SEAI. https://ndber.seai.ie.

Shi, J., Pan, Z., Jiang, L., and Zhai, X. (2023). An ontology-based methodology to establish city information model of digital twin city by merging bim, gis and iot. Advanced Engineering Informatics, 57:102114.

Terkaj, W., Schneider, G. F., and Pauwels, P. (2017). Reusing domain ontologies in linked building data: the case of building automation and control. In 8th International workshop on formal ontologies meet industry, volume 2050.

TUWien (2024a). Building information ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/building/1_10/gbBuildingOntology.owl. [Online; accessed 01-Feb-2024].

TUWien (2024b). Weather ontology. https://www.auto.tuwien.ac.at/downloads/thinkhome/ontology/WeatherOntology.owl. [Online; accessed 01-Feb-2024].

Van Dam, K. H. and Keirstead, J. (2010). Re-use of an ontology for modelling urban energy systems. In Next generation infrastructure systems for eco-cities, pages 1–6. IEEE.

Wu, J., Orlandi, F., AlSkaif, T., O'Sullivan, D., and Dev, S. (2021). Ontology modeling for decentralized household energy systems. In 2021 SEST, pages 1–6. IEEE.

Wu, Z., Cheng, J. C., and Wang, Z. (2022). An ontology-based framework for building energy simulation in the operation phase. In International Conference on Computing in Civil and Building Engineering, pages 351–366. Springer.

Zadeh, P. A., Wei, L., Dee, A., Pottinger, R., and Staub-French, S. (2019). Bim-citygml data integration for modern urban challenges. Journal of Information Technology in Construction, 24.

Zhu, J., Chong, H.-Y., Zhao, H., Wu, J., Tan, Y., and Xu, H. (2022). The application of graph in bim/gis integration. Buildings, 12(12):2162.

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 14-17, 2024

AN INTEROPERABILITY APPROACH FOR SOLAR CADASTER AT HIGH LATITUDES

Martina Giorio[1], Mattia Manni[1], Gilles Desthieux [2], Nesrin Irmak Köker[1], and Gabriele Lobaccaro[1]
[1]Department of Civil and Environmental Engineering, Faculty of Engineering, Norwegian University of Science and Technology (NTNU), Trondheim, Norway
[2]Haute école du paysage d'ingénierie et d'architecture de Genève, (HEPIA), University of Applied Sciences and Arts Western Switzerland (HES-SO), Geneva, Switzerland

Abstract

The climate change represents an opportunity to revolutionize the urban planning paradigms, turning the role of the cities from massive producers of carbon emissions and pollution into a more sustainable built environment through the exploitation of renewable energy sources. In this transition solar energy plays a key role. In this regard, the novel contribution of this work is twofold: i) providing an interoperability approach for a preliminary version of the solar cadaster for the municipality of Trondheim (Norway), located at high latitude; ii) exploiting the use of the solar cadaster to identify the most usable rooftop areas for solar systems' installations and as supportive instrument for public and private stakeholders to facilitate decision-making for urban energy planning.

Introduction

The revolution of the urban planning paradigms emerges as one of the challenges posed by the climate change. Cities, recognized as significant contributors to carbon emissions and pollutants, should evolve into more sustainable environments. In the energy field, this leads to active engagement by policymakers in exploring the urban potential for energy generation from renewable energy sources (RES). Reducing dependency on fossil fuels by optimizing the utilization of RES enables municipalities to mitigate the carbon footprint associated with their energy consumption (Chatzigeorgiou & Martinopoulos, 2023). Advanced and innovative technology solutions for RES exploitation contribute to this aim. In that regard, comprehensive and interactive urban energy planning platforms can facilitate the implementation of similar strategies within the built environment (Desthieux & Thebault, 2024). In fact, such platforms are designed to empower private and public stakeholders with the necessary insights for informed and timely decision-making. Particularly, instruments for urban energy planning which focus on solar energy potential (i.e., solar cadasters) can boost the adoption of active solar systems, either applied or integrated into the building envelope (i.e., roofs and facades) (Manni et al., 2023). Solar cadasters facilitate users in quantifying the solar irradiation impinging on urban surfaces (i.e., roofs, facades, and ground), enabling them to identify the most suitable areas for the installation of solar systems. However, integrating such instruments into the territorial database requires addressing issues related to interoperability of software performing specific tasks such as geometry reconstruction, solar modelling, outcomes post-processing, and data visualization (Manni et al., 2023).

The hereby study investigates the preliminary application of the solar cadaster developed by Desthieux et al. (Desthieux et al., 2018) to high-latitude locations, with a focus on the municipality of Trondheim, Norway. The Web platform developed in Geneva *(https://apps.sitg-lab.ch/solaire/)* showcases the solar cadaster to a wide public audience and related energy and economic indicators by buildings. Although locations at latitude greater than 60°N have traditionally been assumed to receive low solar irradiation, recent studies have pointed out that these areas show only slight differences from Continental Europe in terms of annual solar energy potential (Formolli et al., 2021, 2023).

This study aims to investigate interoperability between software for handling CityGML files (i.e., open-source 3D city modeling format that enables the representation, storage, and exchange of virtual city model) and those for territorial data processing, such as ArcGIS Pro, for managing 2D/3D vector and raster data. By implementing a platform for urban solar energy planning in Trondheim municipality, this work emphasizes the significance of incorporating solar energy potential data into territorial regulatory instruments.

The literature mentions several examples of solar maps (Bieda & Cienciała, 2021; Fish & Calvert, 2016; Kanters et al., 2014). Some of them are also published on web platforms such as the Oslo Solkart *(https://od2.pbe.oslo.kommune.no/solkart/)* and the Norwegian Solkart *(https://solkart.no/search)*. These two examples represent the most recent web platforms, developed at high latitudes, in Norway. The reviews reveal the following gaps. For example, the solar cadasters implemented by Saretta et al. (Saretta et al., 2020) and Jurasz et al. (Jurasz et al., 2020) for the Swiss territory and Poland, respectively, are unable to visualize the usable area on roofs for PV installation. This consists of the portion of the roofs' areas more suitable for PV installations, excluding other service plants (e.g., chimney, storage/service equipment, swimming pool) and architectural elements (e.g., windows, overhanging parts) as well as poorly irradiated areas (i.e., shaded parts of the roof, unsuitable roofs' surfaces exposures and orientation). Current solar cadasters usually consider the entire roof area as equally useful (i.e., having the same solar energy potential level). However, to increase the reliability of the estimations, it is necessary to distinguish between area of the roof and usable area, as in London (Steadman et al., 2020) and Geneva (Desthieux et al., 2018) solar cadasters. There are very few studies and platforms providing estimations of the usable area, and none of them is developed for high-latitude applications. By examining the diverse web platforms, different levels of detail can be identified concerning the visualization of solar energy potential on the roofs. Among those, the most common two levels of visualization (LOV) are bidimensional (2D) exploiting either a solid (LOV1) or a gradient pattern (LOV2). The LOV1, which is the most common, corresponds to the lowest level of detail, and it associates a unique solar potential value with each roof slope. Conversely, the LOV2 enables visualizing the spatial variation of the solar energy potential with a color gradient. The level of visualization of this pattern strongly depends on the accuracy of the geometrical model. Solar cadasters with such level of visualization can show shadow cast from surrounding buildings as well as architectural and service plants elements of the roof (Amaro Silva et al., 2022), picturing the usable areas for installation of active solar systems. The Web platform of the solar cadaster of Geneva offers the LOV2 of visualization detail (Desthieux & Thebault, 2024).

Following the highlighted research gaps, the present study aims to investigate limitations and opportunities to improve the calculation and the visualization of the solar energy in the urban context, with a specific emphasis on high latitudes (Formolli et al., 2023). Identifying the available area for calculating the solar potential for a high latitude case study can be very valuable for advancing the research as there are still no studies with this level of detail.

Methods and materials

Case study selection

Three neighborhoods representative of low-, medium-, and high-density built areas in the city of Trondheim (N 63° 25' 41"), are selected as case studies to test the solar cadaster approach as introduced in Geneva. Their main characteristics are presented below in the Table 1 with the extension (m^2), the constructed volume (m^3), the built density, derived from the two previous values, and the normalized value of the density.

Table 1: List of the selected neighborhoods and information on their area and density

	Neighborhood area [m^2]	Buildings volume [m^3]	Density [m^2/m^3]	Normalized values
a	13662.12	5967.94	2.29	0.19
b	31098.13	4361.58	7.13	0.60
c	17665.09	2170.61	8.14	0.68

Besides the density, these neighborhoods differ in terms of their morphological characteristics and functions (Figure 1). The neighborhood (a) consists of a residential area with mostly two-story single-family detached houses, each with its private garden (i.e., high inter-building distances). Neighborhood (b) is also residential, but its houses are larger and taller compared to those in neighborhood (a). Gardens and courtyards are characterized by limited dimensions, thus increasing the risk of mutual shading between adjacent buildings. Finally, the neighborhood (c) is part of the downtown, and its buildings show a compact layout with small inner courtyards. In this area, stores open onto the ground level, while residential apartments are situated on upper floors.

Figure 1: Selected neighborhood (source: Google Earth). a) Pappenheim (low density), b) Rosenborg (medium density), c) Sentrum (high density).

Workflow

The workflow for the preliminary application of the solar cadaster platform developed by Desthieux et al. (Desthieux et al., 2018), at high-latitude locations, is outlined in Figure 2.

The solar model, encoded in JAVA, serves as the simulation engine, and constitutes the core of the workflow. Running it requires meteorological and geometrical input data. The meteorological input data are retrieved through the Meteonorm software version 7.3 *(https://meteonorm.com/)* provides statistical weather data representative of the period from 1990 to 2010 for a given location. Hourly data are averaged by month to shorten the dataset from 8,760 to 288 values. The geometrical input data consist of (i) digital surface model (DSM) of the analyzed neighborhoods for shadow casting, and the calculation slope and aspect on the area, (ii) the Shapefile containing 2D representation of buildings that provide an identifier and geometric data for each building, and (iii) the 3D model of the neighborhoods in CityGML format generated through the automated reconstruction technique from Kong and Fan (Kong & Fan, 2024). These are processed in ArcGIS Pro environment. Once the preparation of the input file is finalized, the solar simulation engine is started. The solar model is based on algorithms encoded in JAVA for the calculation of hourly, monthly, and yearly solar irradiation impinging on the studied area. The results are stored in Tag Image File (.TIF) format and can be visualized in ArcGIS Pro, utilizing various color gradients to represent the corresponding solar energy potential.

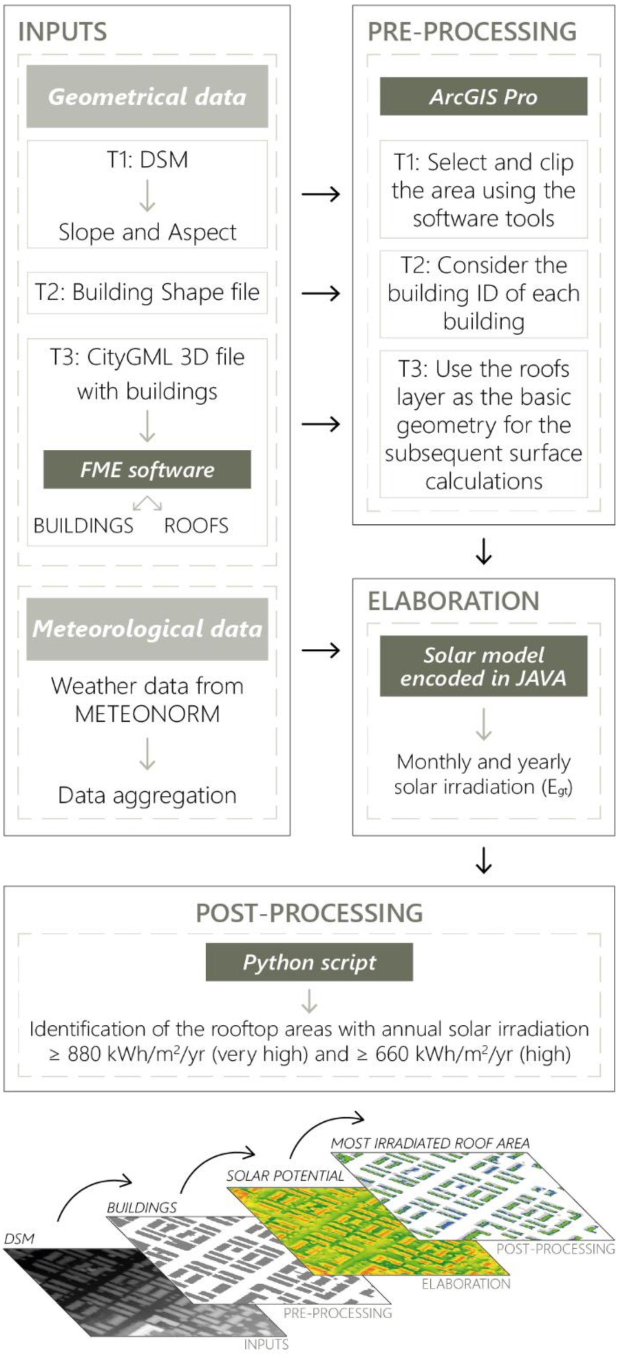


Figure 2: Workflow of the methodology developed in this study.

Following this, outputs from the solar model are post-processed to extract irradiation results on the roofs and identify the most suitable area on each roof for the installation of solar panels. Such an area consists of optimally exposed portions of the roof where there are no obstructions such as chimneys, roof windows, services, and machinery, hindering the installation of solar panels. Optimal exposition to solar irradiation is defined according to the threshold values identified for high latitudes in Lobaccaro et al. (Lobaccaro et al., 2019). Here in particular, the solar energy potential of any surface is labelled as 'very high' when the global tilted irradiation is greater than 880 kWh/m^2/yr; while it is classified as 'high' for global tilted irradiance amounts higher than 660 kWh/m^2/yr and lower than 880 kWh/m^2/yr. The range of low, medium, high, and very high solar irradiation is calculated based on the annual solar irradiation values for each surface, which are imported into 3D modelling software. The annual solar irradiation values are divided into five categories, for high latitudes: very low (0-220 kWh/m²/year), low (220-440 kWh/m²/year), medium (440-660 kWh/m²/year), high (660-880 kWh/m²/year), and very high (880-1.100 kWh/m²/year). These categories are determined based on the average values of solar irradiation in the middle point of each building's roof surface.

This part of the process (i.e., identification and classification of the most suitable area of the roof for solar panel installation), is conducting using a Python script. It represents a propaedeutic step for estimating the productivity of a potential PV system installed in the optimally irradiated portion of the roof.

The workflow is articulated and fragmented due to the utilization of different software and codifications, and it requires a high level of interoperability to integrate various type of information (i.e., meteorological, geometrical, and territorial vector and raster data) and data from different sources (e.g., HOYDEDATA.NO *https://hoydedata.no/LaserInnsyn2/*). This highlights the interdisciplinary and multi software domain of this research.

For this study, neighborhoods with an area ranging from 0.014 km^2 to 0.018 km^2 were selected to reduce the computational time for the solar analyses, compared to the whole city territory (340 km^2). Here the required time is represented by the time that the operator needs to manually perform the pre-processing tasks (T1-T2-T3). For each neighborhood, this can be estimated in 20 to 30 minutes. Considering the whole city of Trondheim, after the pre-processing time, needed for the selection of the different tiles, a total of around 50 hours are needed for the calculation of the solar potential results.

Pre-processing: preparation of input data for the solar modelling

Weather data are retrieved through Meteonorm software by choosing the nearest available weather station to the center of the city of Trondheim. The data are collected by the meteorological station at Værnes airport, located 35 km outside of the Trondheim municipality. The weather data file includes information about month of the year, hour of the day, global horizontal irradiation (Gh), diffuse horizontal irradiation (Dh), beam normal irradiation (Bn), and extraterrestrial irradiation (I_0). As mentioned in the workflow section, the solar analysis is conducted considering the average hourly values by month, rather than analyzing all days of the year (288 datapoints instead of 8,760). This permits to reduce both computational time and Central Processing Unit (CPU) requirements. The weather data is stored in a CVS format file. The other input data for initializing the solar model is related to the neighborhood geometry.

DSM, Slope, and Aspect. The DSM data of the studied area is provided by The Norwegian Mapping Authority *HOYDEDATA.NO* in .TIF format with one-meter spatial resolution. The three case study neighborhoods are extracted from the Trondheim DSM file. Edge dimension of the three clipping masks range from 220 m to 420 m. To create the new DSM (i.e., the clipped DSM file of the investigated neighborhood), slope, and aspect files ArcGIS tools are used. The raster pre-processing outputs are stored in .TIF format: the DSM file has the pixel type of 32-bit float, and both slope and aspect files have a pixel type of 16-bit signed.

Building Shapefile. This is a two-dimensional file containing the building layouts of the city of Trondheim. It is retrieved from the national territorial database *GEONORGE* (FKB-Buildings Dataset, *https://kartkatalog.geonorge.no/metadata/fkb-bygning/8b4304ea-4fb0-479c-a24d-fa225e2c6e97*), and it contains information about the identifier for each building (i.e., LOKALID). The building shapefile of Trondheim is also clipped to isolate the three neighborhood case studies.

CityGML file. A Level of Detail 2 (LoD2) building model for the urban area of Trondheim is created through two main steps: (1) roof plane segmentation from building point clouds conducted by LIDAR airborne plane survey and (2) roof line topology-based building model creation in LoD2. A deep learning-based method (Zhang & Fan, 2022) was applied to automatically segment the building point clouds into roof planes. The obtained roof plane information served as the basis for reconstructing 3D roof structures and subsequently creating building models. In the second step, the LoD2 building models were created using roof line topologies (Kong & Fan, 2024). This process brought to the automatic reconstruction of building models in LoD2 for the Trondheim urban area. A file in CityGML format was exported from this model, containing three-dimensional information of the buildings. Using the FME software, it is divided into two separate outputs: one for buildings and another for roofs. In this study, only the second file (concerning the roofs) is considered. This was used as the geometric basis of the subsequent solar analyses.

Other studies developed in Europe (Adjiski et al., 2023) and in Asia (Dahal et al., 2021) use similar approaches for creating the geometrical models from LiDAR data as well as for identifying the best exposed portions of the roof through Geographic Information System (GIS) tools.

Elaboration: solar radiation modelling on the studied area

The JAVA-based solar model is initialized with the files described in the previous section. Additionally, information concerning the geolocation of the case studies (i.e., longitude, area of the studied zone, DSM resolution) is manually set, allowing for a more accurate estimation of the global tilted irradiation. It processes hourly solar radiation for each component: direct, diffuse and reflection, as well as shadow casting impacting direct (hourly-based) and diffuse, based on Sky View Factor (SVF), components. The solar modelling approach is introduced in Desthieux et al. (Desthieux et al., 2018).

Outcomes from the elaboration stage consist of 14 files in .TIF format, including a solar irradiation map for each month (12), the annual solar irradiation map (1), and the SVF map (1). These files are in a gradient form and permit to visualize the solar energy potential of roof and ground surfaces.

Post-processing: calculations and visualization in GIS

The annual solar irradiation map is post-processed to identify the most suitable areas which are available for PV installation. The thresholds defined in the workflow section are considered to extract the highly (global tilted irradiance within the range 660-880 kWh/m^2/yr) and very highly (global tilted irradiance within the range 660-880 kWh/m^2/yr) irradiated portions of the roofs. Therefore, the potential for energy generation through PV panels is investigated for these areas. (Lobaccaro et al., 2019)

For extracting the roof areas with high and very high solar potential, the annual solar irradiation map is multiplied by a clipping mask based on the roof footprints. The highly irradiated areas are highlighted by selecting raster pixels corresponding to irradiation amounts greater than 660 kWh/m^2/yr. Then, the same procedure is reiterated considering the threshold of 880 kWh/m^2/yr for very highly irradiated roofs.

Additionally, the raster pixels representing roof elements which can obstacle the installation of PV panels (i.e. ventilation systems, mechanical bodies, chimneys) are filtered out. A new polygon surface is created from these pixels and subtracted from the roof footprints, considering an inner buffer of 0.5 m, as the solar panels could not be installed in proximity of the building edges. The identification of the most adequate areas for PV installation is completed by removing from the analysis all the polygons with an area of less than 5 m^2. These surfaces are deemed too small for installation of PV systems. Finally, the solar PV production is calculated considering the available and optimally exposed areas of the roofs, as well as the characteristics of the installed PV panels. In this study, the monocrystalline PV is considered, with a peak power of 220 Wp/m^2 based on the most powerful models available on the market. Although flat roof configuration is not present in the investigated neighborhoods, it is worth mentioning that the potential PV energy production of flat roofs is quantified by assuming that PV panels are installed with a default inclination ranging from 10° to 15°. A transposition factor (FT) accounts for this aspect. Indeed, the use of flat roof configuration is becoming increasingly common at high latitudes, owing to the development of effective techniques for mitigating the accumulation of snow.

Results and Discussion

Solar irradiation analysis

Results about the estimated annual solar irradiation for the three case study neighborhoods resulting on the annual solar irradiation values range between 0 kWh/m^2/yr and 1,170 kWh/m^2/yr (Figure 3).

In neighborhood ‘a’ (low-density), most of the surfaces are optimally exposed to solar irradiation. In fact, the houses are small and far, while streets are wide, allowing for a better solar accessibility. In neighborhood ‘b’ (medium-density), the areas subject to high solar irradiation are fewer and limited to south-exposed roof slopes. In the neighborhood ‘c’ (high-density), the rooftop areas suitable for PV installation are minimal due to mutual shading among buildings.

The results of this approach demonstrate the importance of conducting detailed solar analyses within urban environments, utilizing interoperable software.

Determination of suitable areas for PV installation

Solar irradiation outputs are post-processed to exclude parts of the roof that are either unavailable or characterized by global tilted irradiance lower than 660 kWh/m^2/yr. These surfaces are shown in differentiating between highly and very highly irradiated surfaces (Figure 4) (see the Methods and Materials section for the applied thresholds). The numerical results concerning the surface area are reported in Table 2.

Table 2: Calculation of total roof areas and actual usable parts after the analysis

	Total roof area (m^2)	Available area with E_{gt} > 660 kWh/m^2/yr (m^2)	Available area with E_{gt} > 880 kWh/m^2/yr (m^2)
a	2069.45	1023.44	400.20
b	2546.77	1121.32	545.37
c	1666.46	688.67	114.01

Results highlight the differences that exist among the investigated neighborhoods in terms of solar accessibility and solar energy potential. When considering the area of highly irradiated surfaces, this ranges from 40% (neighborhood ‘c’) to 50% (neighborhood ‘a’) of the total neighborhood area in each case study. Conversely, the portions of roofs with a very high solar potential varies between 6% (neighborhood ‘c’) to 20% (neighborhood ‘b’) of the total neighborhood area (Table 3).

Table 3: Percentages of usable area after changing the minimum irradiance threshold

	% of usable area with E_{gt} > 660 kWh/m^2/yr	% of usable area with E_{gt} > 880 kWh/m^2/yr
a	49.45 %	19.34 %
b	44.03 %	21.41 %
c	41.33 %	6.64 %

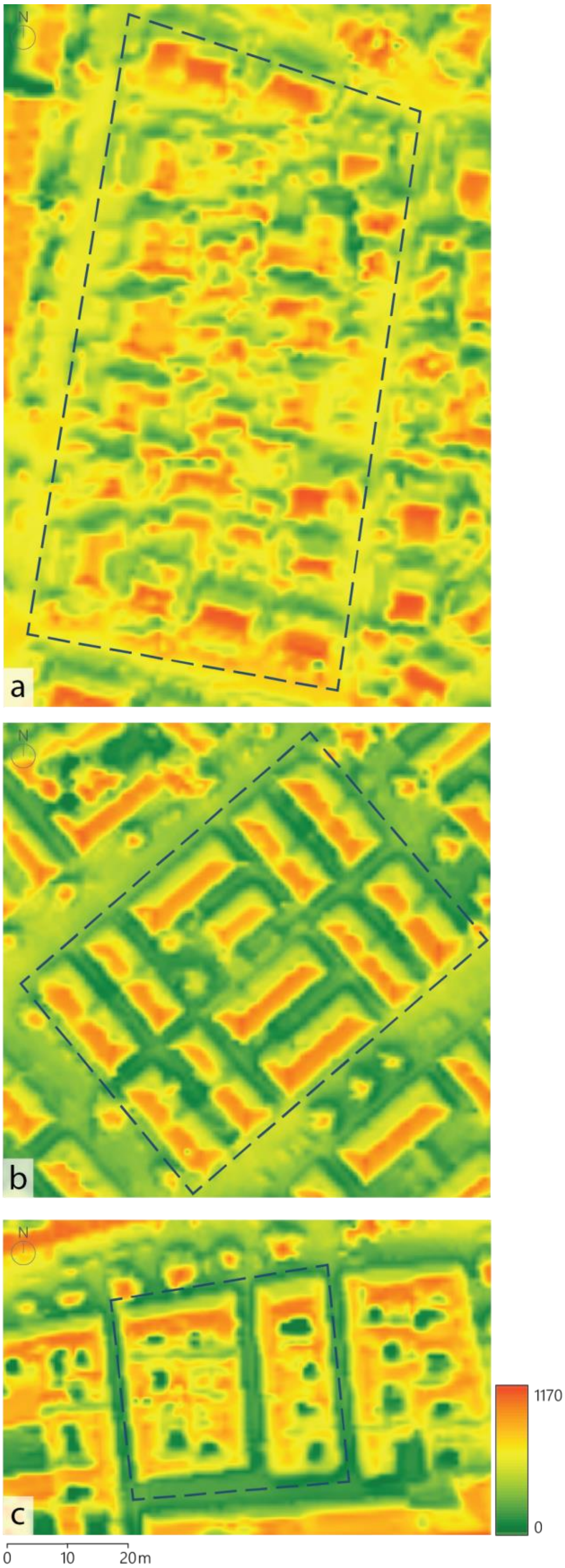


Figure 3: Annual solar irradiation (kWh/m^2/yr) calculated in the three neighborhoods.

Figure 4: Available roof area resulted considering E_{gt} higher than 660 kWh/m2/yr and 880 kWh/m2/yr.

The limited number of neighborhood case studies (3) prevents conducting robust correlation analysis between neighborhood density and the total area of available surfaces for PV installation. Nonetheless, it is worth highlighting that, among the percentages reported in Table 3, the lowest value is always associated to the high-density neighborhood.

This is due to (i) the proximity of the buildings, which facilitates mutual shading, and (ii) the lower ratio of the roof area to the total area of the building's envelope, resulting in a greater presence of elements installed on a rooftop compared to neighborhoods ‘a’ and ‘b’.

Comparison with existing solar modeling tools

ArcGIS geoprocessing tools such as the Raster Solar Radiation tool *(https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/raster-solar-radiation.htm)* allow for conducting solar analyses with shorter computational time. However, they require the utilization of other models for calibration (Choi et al., 2019; Kausika & van Sark, 2021). Furthermore, those methods are characterized by a low level of users’ customization (i.e., forced data inputs) of the simulation parameters, particularly, the temporal resolution that is fixed to one year. The use of more interoperable software permits to perform sub-yearly analysis, allowing the customization of input parameters such as weather data sources and time intervals.

Two online platforms for the visualization of the solar potential of roof surfaces are already available in Norway. These are characterized by different spatial scales: the Solkart spans across the whole country, while the Oslo Solkart covers the Oslo municipality. A visual comparison of the national Solkart and the solar cadaster approach applied in this study is shown in Figure 5.

There is a difference in the spatial resolution of the output data. The Solkart shows the solar potential of any roof slope as uniformly distributed, providing a single solar irradiance value for each surface. The Oslo Solkart has gradient values, but both existing Solkart neglect the identification of the rooftop areas suitable for PV installation considering the presence of obstructions or lower level of irradiation. The presented approach, on the other hand, provides a solar irradiance value for each pixel of the neighborhood’s raster image, presenting a heterogenous spatial distribution of the solar irradiation on any rooftop slope. The calculation of the irradiation value on each part of the roof made it possible to identify the areas that are exposed to irradiation greater than 660 kWh/m^2/yr and 880 kWh/m^2/yr, highlighting the area available for new installations.

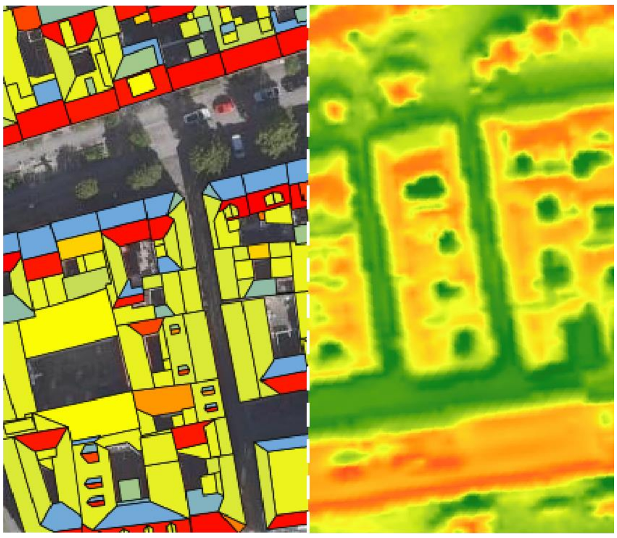


Figure 5: Visual comparison between the Solkart (left) and the solar cadaster approach from (Desthieux et al., 2018) *(right), applied to neighborhood 'b' (medium density).*

Finally, while the solar analysis performed with the Solkart does not account for the presence of trees, the method developer in the solar cadaster from (Desthieux et al., 2018) demonstrates the capability to process shadows casted from urban greenery included in the DSM.

Limitations of the study

This research study has three main limitations. Firstly, the weather data retrieved from Meteonorm v7.3 are collected by the meteorological station at Værnes airport, located 35 km outside of the Trondheim municipality. Indeed, this is the closest weather station to Trondheim which is available in Meteonorm database. Meteonorm was chosen over local weather stations to ensure a generic and easily replicable workflow for other case studies. Having more local data would allow for a more accurate analysis. Secondly, developing the process by implementing multiple software and tools allowed the results to be detailed, but at the same time having all the different working environments and code programming languages led to a fragmented process. The process is not automated, and the file recognition steps need to be filled in manually when switching from a software to another. Finally, only three case studies are considered for the calculation of the available roof area for PV installation. This limits the capability of identifying strong correlations between the model outputs and neighborhood density. Increasing the number of neighborhood case studies and expanding the analyses to the whole city of Trondheim can reinforce the conclusions drawn.

Processing the whole city of Trondheim will require to use the GPU machine installed at HEPIA (Stendardo et al., 2020), as for the solar cadaster in Geneva, which in particular allows to speed up the calculation of shadows implemented in the CUDA language.

Processing the solar analysis of a tile with a surface area of 9 km^2 takes approximately from one 1 to 1.5 hours. Therefore, analyzing the whole city of Trondheim, which has an area of 340 km^2, would take from 50 to 55 hours, which is acceptable for such an area. Computational time could be improved through parallelization.

Conclusions

The solar cadaster approach developed by Desthieux et al. (Desthieux et al., 2018) is applied to three neighborhoods in Trondheim (i.e., low-, medium-, and high-density neighborhoods). The proposed method is at the forefront of interoperability between software for managing urban geometry models and those for processing territorial data. The main findings of the study concern the identification of the best irradiated portion of the roofs suitable for installing PV systems. Results demonstrate the need for more refined analyses compared to existing online platform in Norway. Achieving this level of precision currently requires a combination of various software and types of information (i.e., meteorological, geometrical, and territorial vector and raster data). From the social perspective, the proposed solar cadaster approach aligns with the ongoing trend toward smart and sustainable cities. Informing citizens and municipalities about the solar potential of the building stock reflects a commitment to harnessing affordable and user-friendly technology for creating more efficient, sustainable, and livable built environments (Xue et al., 2021). Future advancements in this field may concern (i) enhancing the workflow through automation, parallelization, and reduction of the fragmentation to reduce the computational time especially for the whole city territory; (ii) exploiting the use of the local data and validating the results against real measurements; (iii) extending the analysis to the entire city territory of Trondheim; and (iv) creating an accessible online platform (i.e., solar cadaster) at the cityscape level to engage public and private stakeholders and the interaction with the individual users. Additionally, (v) extending the analysis to vertical surfaces (i.e., the facades) by including the optical properties of materials and the detection of architectural elements (i.e., windows, doors, balconies, overhanging parts) in the calculation will be important to increase the level of information on the most irradiated parts of the facades, especially at high latitudes where the vertical surfaces received the annual highest irradiance (Formolli et al., 2023). Therefore, it will be necessary to progress the research about material detection and image recognition techniques.

Acknowledgments

The authors gratefully acknowledge the support from the Norwegian Research Council (research project FRIPRO-FRINATEK no. 324243 HELIOS - eHancing optimal ExpLoitation of Solar energy in Nordic cities through the digitalization of built environment).

References

Adjiski, V., Kaplan, G., & Mijalkovski, S. (2023). Assessment of the solar energy potential of rooftops using LiDAR datasets and GIS based approach. *International Journal of Engineering and Geosciences*, *8*(2), 188–199. https://doi.org/10.26833/ijeg.1112274

Amaro Silva, R., Blanc, P., & Blanc OIE, P. (2022). Estimating Global Horizontal Irradiance at the Urban Level: A Sensitivity Analysis Using Different Digital Surface Models. *8th World Conference on Photovoltaic Energy Conversion*. https://www.researchgate.net/publication/363663602

Bieda, A., & Cienciała, A. (2021). Towards a Renewable Energy Source Cadastre - A Review of Examples from around the World. *Energies*, *14(23):8095*. https://doi.org/https://doi.org/10.3390/en14238095

Chatzigeorgiou, E., & Martinopoulos, G. (2023). Solar cadastre for assessment of near-zero energy districts. *IOP Conference Series: Earth and Environmental Science*, *1196*(1). https://doi.org/10.1088/1755-1315/1196/1/012003

Choi, Y., Suh, J., & Kim, S. M. (2019). GIS-based solar radiation mapping, site evaluation, and potential assessment: A review. *Applied Sciences (Switzerland)*, *9*(9). https://doi.org/10.3390/app9091960

Dahal, A., Chhetri, B., Sharma, K. R., & Neupane, M. (2021). Assessment of Solar Photovoltaic Potential of Building Rooftops Using Photogrammetry and GIS. *The Geographic Base*, *8*. https://doi.org/10.3126/tgbv8i01.43467

Desthieux, G., Carneiro, C., Camponovo, R., Ineichen, P., Morello, E., Boulmier, A., Abdennadher, N., Dervey, S., & Ellert, C. (2018). Solar energy potential assessment on rooftops and facades in large built environments based on lidar data, image processing, and cloud computing. Methodological background, application, and validation in Geneva (solar cadaster). *Frontiers in Built Environment*, *4*. https://doi.org/10.3389/fbuil.2018.00014

Desthieux, G., & Thebault, M. (2024). Solar governance for the transborder agglomeration of the Greater Geneva based on the solar cadaster development. *Frontiers in Built Environment*, *10*. https://doi.org/10.3389/fbuil.2024.1347056

Fish, C. S., & Calvert, K. (2016). Analysis of interactive solar energy web maps for urban energy sustainability. *Cartographic Perspectives*, *2016*(85), 5–22. https://doi.org/10.14714/CP85.1372

Formolli, M., Kleiven, T., & Lobaccaro, G. (2023). Assessing solar energy accessibility at high latitudes: A systematic review of urban spatial domains, metrics, and parameters. *Renewable and Sustainable Energy Reviews*, *177*. https://doi.org/10.1016/j.rser.2023.113231

Formolli, M., Lobaccaro, G., & Kanters, J. (2021). Solar energy in the Nordic built environment: Challenges, opportunities and barriers. *Energies*, *14*(24). https://doi.org/10.3390/en14248410

Jurasz, J. K., Dąbek, P. B., & Campana, P. E. (2020). Can a city reach energy self-sufficiency by means of rooftop photovoltaics? Case study from Poland. *Journal of Cleaner Production*, *245*. https://doi.org/10.1016/j.jclepro.2019.118813

Kanters, J., Wall, M., & Kjellsson, E. (2014). The solar map as a knowledge base for solar energy use. *Energy Procedia*, *48*, 1597–1606. https://doi.org/10.1016/j.egypro.2014.02.180

Kausika, B. B., & van Sark, W. G. J. H. M. (2021). Calibration and validation of ArcGIS solar radiation tool for photovoltaic potential determination in the Netherlands. *Energies*, *14*(7). https://doi.org/10.3390/en14071865

Kong, G., & Fan, H. (2024). Generating 3D Roof Models from ALS Point Clouds using Roof Line Topologies. Recent Advances in 3D Geoinformation Science. In Springer Cham (Ed.), *Proceedings of the 18th 3D GeoInfo Conference*.

Lobaccaro, G., Lisowska, M. M., Saretta, E., Bonomo, P., & Frontini, F. (2019). A methodological analysis approach to assess solar energy potential at the neighborhood scale. *Energies*, *12*(18). https://doi.org/10.3390/en12183554

Manni, M., Formolli, M., Boccalatte, A., Croce, S., Desthieux, G., Hachem-Vermette, C., Kanters, J., Ménézo, C., Snow, M., Thebault, M., Wall, M., & Lobaccaro, G. (2023). Ten questions concerning planning and design strategies for solar neighborhoods. *Building and Environment*, *246*. https://doi.org/10.1016/j.buildenv.2023.110946

Saretta, E., Bonomo, P., & Frontini, F. (2020). A calculation method for the BIPV potential of Swiss façades at LOD2.5 in urban areas: A case from Ticino region. *Solar Energy*, *195*, 150–165. https://doi.org/10.1016/j.solener.2019.11.062

Steadman, P., Evans, S., Liddiard, R., Godoy-Shimizu, D., Ruyssevelt, P., & Humphrey, D. (2020). Building stock energy modelling in the UK: the 3DStock method and the London Building Stock Model. *Buildings and Cities*, *1*(1), 100–119. https://doi.org/10.5334/bc.52

Stendardo, N., Desthieux, G., Abdennadher, N., & Gallinelli, P. (2020). GPU-enabled shadow casting for solar potential estimation in large urban areas. Application to the solar cadaster of Greater Geneva. *Applied Sciences (Switzerland)*, *10*(15). https://doi.org/10.3390/APP10155361

Xue, Y., Lindkvist, C. M., & Temeljotov-Salaj, A. (2021). Barriers and potential solutions to the diffusion of solar photovoltaics from the public-private-people partnership perspective – Case study of Norway. *Renewable and Sustainable Energy Reviews*, *137*. https://doi.org/10.1016/j.rser.2020.110636

Zhang, C., & Fan, H. (2022). An Improved Multi-Task Pointwise Network for Segmentation of Building Roofs in Airborne Laser Scanning Point Clouds. *Photogrammetric Record*, *37*(179), 260–284. https://doi.org/10.1111/phor.12420

INTELLIGENT RETROFITS IN RESIDENTIAL BUILDINGS: A KNOWLEDGE -BASED APPROACH

Sobia Bano, Usman Ali, Divyanshu Sood and James O'Donnell
School of Mechanical and Materials Engineering and UCD Energy Institute, UCD, Dublin, Ireland

Abstract

Building Life Cycle Assessment (LCA) is crucial to mitigate the environmental impact of residential retrofits. However, a significant gap exists in evaluating the complexity of building life cycles in the context of residential renovations. The proposed intelligent Knowledge Base System methodology integrates existing building materials, stock data, typologies, and retrofit scenarios, employing a hybrid approach that combines rule-based, case-based, and frame-based systems for comprehensive life cycle assessment. The case study focuses on retrofitting the semi-detached housing typology, transitioning from a boiler to a heat pump, and results in a reduction of total carbon emissions to 184 tCO_2e and 90 tCO_2e over 60 years. This study aids energy policymakers, urban planners, and researchers by offering an intelligent framework to promote eco-friendly renovations and advance sustainable building practices.

Introduction

Buildings are a major source of emissions, accounting for 39% of global energy-related carbon emissions, split between operational emissions (28%) and emissions from materials and construction (11%) (WGBC, 2023). The entire life cycle of a building, from material extraction to demolition, contributes to its carbon footprint. Emissions are categorized into operational (from daily operations) and embodied (from materials, construction, transportation, and demolition) (Fenner et al., 2020). Addressing these emissions is crucial for achieving near-zero carbon targets, requiring effective methods and strategies to assess and reduce carbon emissions throughout a building life cycle assessment (LCA).

A big challenge in applying LCA during initial design or retrofitting is the scarcity or complexity of data (Kadrić et al., 2023; Roberts et al., 2020). This complexity hinders stakeholders from developing and implementing effective retrofit measures (Zuo et al., 2017; Najjar et al., 2019). Conducting LCAs to upgrade residential building stock is crucial, as it allows for assessing long-term environmental and cost impacts throughout the buildings' lifespan, as discussed previously, and all subsequent phases. Despite the widespread adoption and standardization of LCA methodologies and the growing availability of data through tools like Environmental Product Declarations (EPD), a significant research gap still needs to be addressed. This gap mainly lies in integrating and interpreting complex data for specific scenarios. The complexity increases when dealing with residential renovations, where factors such as renovation type, materials used, and the building's operational phase add complexity (Roberts et al., 2020).

Furthermore, despite the benefits of sustainably retrofitted buildings, key stakeholders often lack a system to generate new knowledge while incorporating the entire lifecycle of a building. This includes selecting sustainable materials and transferring related information between stages. However, integrating complex data across different LCA phases and needing a flexible and scalable solution represents a significant research gap in Building LCA (Hollberg et al., 2021).

The importance of comparing different existing LCA analyses of buildings has been widely acknowledged. For instance, Röck et al. (2020) analyzed 238 buildings from 54 scientific articles, focusing on embodied and operational emissions. They emphasized the need for comparative LCA studies due to varying energy performance standards. Previous research has concentrated on implementing LCA frameworks in buildings to assess environmental impacts. Waldman et al. (2020) proposed a methodology for assessing construction materials during the specification and procurement stages of LCA. Balouktsi et al. (2020) explored the level of designers' awareness of environmental performance assessments and LCA in buildings. A study by Sharma et al. (2011) demonstrated large discrepancies in LCA results for 13 buildings, highlighting that commercial buildings have a greater environmental impact than residential ones. Similarly, Säynäjoki et al. (2017) reviewed 116 buildings and compared their GHG emissions during pre-use life cycle stages, finding significant inconsistencies in LCA results, ranging from 0.03 to 2.00 tons of CO_2 eq. per m^2. Data reliability and geographic relevance are crucial for accurate LCA methodologies across different buildings. Another challenge is the lack of collaboration among LCA practitioners, including designers, architects, building consultants, and engineers. Bridging this gap and promoting collaboration is a key objective of this study.

A survey by Jusselme et al. (2020) highlighted the time-consuming nature of LCA in the design stage, ranging from 18 to 33 hours from data collection to result interpretation. This adds to the time and cost of the process. Identifying strategies for efficient LCA analysis is therefore crucial. Various LCA databases, such as One Click LCA, Eco Invent, SimaPro, and the ULSCI database, offer different strengths and weaknesses. Each database, like Eco Invent v3 with its 10,000 interlinked datasets, or SimaPro, suitable for cradle-to-grave assessments, has unique features

(Wernet et al., 2016). Mohammadpourkarbasi et al. (2023) performed a comparative study of retrofit approaches for life cycle carbon assessment of decarbonizing UK's hard-to-treat homes. However, integrating data from multiple databases, establishing relationships, and reducing redundancies and discrepancies can greatly benefit decision-makers and stakeholders. Furthermore, studies rely on predefined retrofit scenarios and require intelligent solutions to recommend environmentally friendly retrofit scenarios for buildings based on their requirements.

The aim of this study is to propose a methodology for an intelligent Knowledge Base System (iKBS) for residential building LCA. The knowledge-base system integrates and analyzes complex, various data sources using knowledge acquisition, knowledge pre-processing (transforming unstructured data into structured data), and knowledge management. A novel feature of this system is knowledge inference, specifically tailored to generate new pieces of information from existing data. Furthermore, this research compares various knowledge-based approaches, such as rule-based, case-based, and frame-based systems for building LCA. This comparative analysis led to the development of a recommendation system for stakeholders that provides deep insights into the embodied energy aspects of residential renovations, a critical factor in understanding and improving the environmental performance of the building stock. The semi-detached housing typology case study shows that combining various knowledge leads to low-energy, cost-effective retrofit material recommendations.

Traditional LCA methods often rely on static, linear approaches to assess the environmental impacts of building projects. These methods may not adequately capture the complexities and interdependencies inherent in building life cycles, particularly for residential renovations. They typically struggle with dynamically incorporating new data or adjusting to novel scenarios without substantial manual intervention that accounts for updates during the life cycle. This can lead to inaccuracies in modeling the environmental impacts of different materials, processes, and retrofit options. In contrast, iKBS offers a dynamic, adaptive, and comprehensive framework for LCA. iKBS is built upon a hybrid model that combines the strengths of rule-based, case-based, and frame-based reasoning to handle diverse data types and complex analysis scenarios. The iKBS is capable of automating the data acquisition, preprocessing, and analysis processes, significantly reducing the time and effort required while enhancing the accuracy and relevance of the assessments. This integration allows for the sophisticated handling of diverse data types and complex analysis scenarios, facilitating the assimilation and processing of vast amounts of building material, stock data, residential typologies, and retrofit scenarios. the proposed intelligent framework can adapt and learn from new information, providing a more nuanced understanding of environmental impacts and facilitating the generation of actionable recommendations tailored to specific renovation projects.

The paper follows a structured format, with Section 2 presenting the methodology employed in the research. Section 3 examines a comprehensive Irish case study, providing real-world context for the study's findings. Finally, Section 4 concludes the paper and outlines potential avenues for future research.

Methodology

This paper presents a comprehensive methodology for an intelligent Knowledge-based System to facilitate detailed Life Cycle Impact (LCI) Assessments for the building and construction sector, integrating environmental and economic performance metrics for sustainable decision-making (Figure 1). The proposed development for residential building stock, based on the ISO 14040/44 standards (Principles and Framework for Life Cycle Assessment), comprises four phases: goal and scope definition, life cycle inventory, life cycle impact assessment, and interpretation. Each phase is crucial and builds upon the previous one, ensuring a comprehensive and systematic approach to assessing the environmental impacts of residential building renovations.

The proposed methodology encompasses several key steps: 1. Goal and Scope Definition to outline the study objectives, 2. Development of a Knowledge-based System for LCI to gather relevant data or knowledge; 3. Integration of Life Cycle Impact Assessment via an Inference Engine, and 4. Creation of an LCA Interface for Recommendations, facilitating informed decision-making. This comprehensive approach ensures thorough analysis and informed choices throughout the lifecycle assessment process, emphasizing requirements, steps, and effective implementation strategies.

Goal and Scope Definition

The methodology begins with defining the goal and scope of the building life cycle assessment. The goals are to understand its initiation reasons, identify potential applications, and specify the target audience. The scope includes specifying the application, approach, data requirements, study period, functional unit, and system boundary. These parameters are essential for guiding the overall development process and ensuring the assessment aligns with the intended purpose. The goal and scope definition serves as the foundation for developing an iKBS to strategically identify areas for enhancement in the building's environmental performance. Furthermore, the methodology is beneficial in determining the best retrofit options, providing in-depth cost-benefit analysis, and conducting comprehensive sustainability assessments by defining proper goals and requirements.

Knowledge-based System for LCI Development

Life Cycle Inventory (LCI) is a core step of LCA and represents the quantification of inputs and outputs for a product throughout its life cycle (Wernet et al., 2016). In the

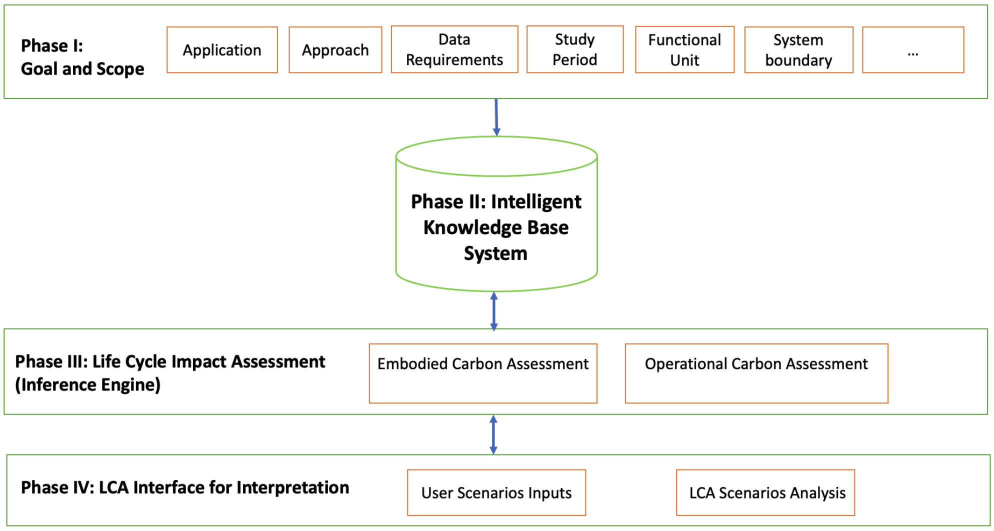


Figure 1: Methodology for intelligent Knowledge-based system for comprehensive life cycle assessment of residential buildings

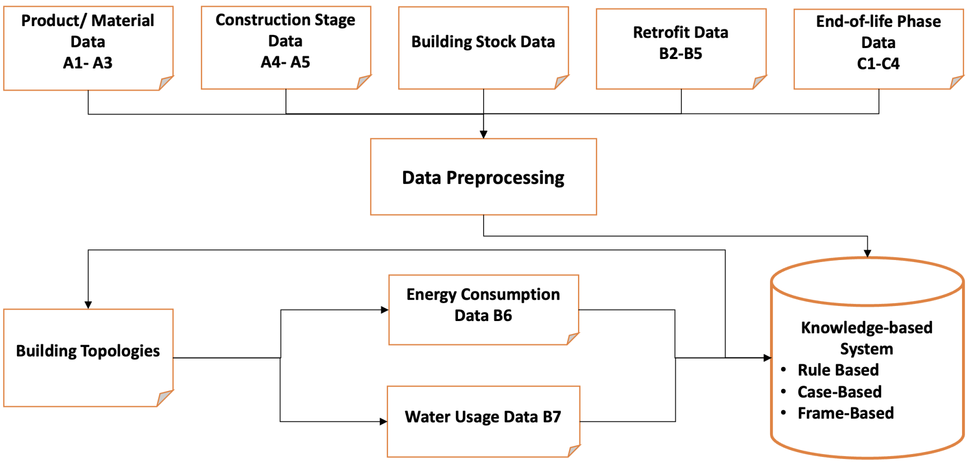


Figure 2: Sub-methodology for knowledge-based system for life cycle inventory development

context of buildings, LCI data include the energy, materials, and emissions associated with the construction, operation, maintenance, and eventual deconstruction of buildings. Several databases provide LCI data with embodied carbon values for various construction materials and processes, such as the Inventory of Carbon and Energy (ICE) database from the University of Bath (ICE, 2024), which can be used exclusively for calculating embodied energy from a prospective perspective, and Ecoinvent (Ecoinvent, 2024) considered to be the largest transparent unit-process. Similarly, various regions and organizations maintain databases of EPDs, which provide standardized LCA-based reports for specific products, including many construction products (Feng et al., 2023). Generally, Building LCA is categorized into different stages: A1-A3 for material extraction and transport, A4-A5 for construction and waste management, B1-B7 for use and renovation, C1-C4 for end-of-life disposal, and D for the environmental benefits of recovery. The process consists of multiple steps: Knowledge Acquisition and Preprocessing, Building Typologies, Knowledge-Based System Development, and Inference Engine for assessments (Figure 2).

Knowledge Acquisition and Preprocessing

The knowledge Acquisition phase for the iKBS involves gathering a wide range of information to create a comprehensive understanding of various aspects of building construction and maintenance. The knowledge acquisition in this phase is segmented into various categories:

- Product/Material Data (A1-A3): This involves gathering data on all materials and products used in the construction process, from extraction to processing.
- Construction Stage Data (A4-A5): Data related to the construction activities, including transportation and on-site processes.
- Building Stock Data: Information on the existing buildings, including their types, ages, and materials.
- Retrofit Data (B2-B5): Information on any renovation or updates made to existing buildings.

- End-of-life Phase Data (C1-C4): Data concerning the demolition, waste processing, and disposal or recycling of building materials or products.

The acquired knowledge undergoes a pre-processing stage to ensure it is consistent, accurate, and relevant for developing the iKBS. The methodology uses different data pre-processing methods such as data cleaning, data transformation, and data integration. The fuzzy-based method is used for identifying and merging records that refer to the same entities across different databases. This step is crucial to ensure the iKBS can effectively use the knowledge for intelligent decision-making in construction projects, sustainability assessments, and lifecycle management of buildings.

Building Typologies

Buildings often share common characteristics and can be classified into a building typology. This can also be referred to as a reference building, which is a representative structure capturing the typical characteristics and performance of a specific category or group of buildings within a larger building stock. This data can be sourced from established national building stock databases, such as TABULA or EPC databases (Ali et al., 2024). Building typologies data is further used for the B1 stage, which evaluates the building's performance in terms of energy and water usage. These data can also be generated using a thermal simulation engine such as EnergyPlus and IES for the generation of energy and water consumption data based on retrofit scenarios (Ali et al., 2024).

Intelligent Knowledge-based System Development

The proposed iKBS is an advanced framework to analyze the complex decision-making process. This process integrates pre-processed and simulated data from the Life Cycle Inventory Development, accommodating the vast array of data and variables in assessing a building's life cycle. The system comprises various intelligent knowledge representation and recommendation methods, including Rule-Based, Case-Based, and Frame-Based systems. Rule-based systems express knowledge concepts as a conjunction of conditions and conclusions (Santos et al., 2020). A rules-based system can be developed for LCA in residential buildings to evaluate environmental impacts based on specific criteria, such as material selection. For instance, a rule might state, *"If a building's construction year is old, then the materials used in the building are of low performance."* Case-based reasoning solves new problems by extracting and applying solutions previously adopted for similar problems (Zhao et al., 2019). In the context of LCA, past LCA studies or available data from similar residential buildings can be used to predict the environmental impact of a new building project. Frame-based systems use structured templates or frames to represent knowledge and infer new information based on existing structures and relationships within the frames (Nazaruks and Osis, 2021). These systems use detailed, available data in structured frames or templates. Each frame can represent different aspects of LCA, like materials, energy efficiency, or water usage, with slots for specific attributes and values.

Typically, KBS (Knowledge-Based Systems) stores all data in databases such as PostgreSQL and MySQL for structured data with well-defined schemas, and MongoDB, Cassandra, or Couchbase for unstructured or semi-structured data. In this paper, MongoDB and Python are used to implement a hybrid approach of Rule-Based, Case-Based, and Frame-Based systems that store different types of unstructured or semi-structured data.

This study comprehensively assesses each system based on accuracy, complexity, flexibility, and scalability. However, due to the complex nature of building lifecycle assessment and data requirements, this study proposes a hybrid approach for effective life cycle assessment. Furthermore, the hybrid approach offers a comprehensive, adaptable, and flexible method to handle unique scenarios, making it a robust solution for sustainable building practices.

Life Cycle Impact Assessment (Inference Engine)

An inference engine that evaluates the environmental implications of a building throughout its life cycle. Environmental performance assessment focuses on embodied energy (energy used in the production of the building materials) and operational energy (energy used during the building's operation).

LCA Interface for Recommendation

Stakeholders can input or query different scenarios to the inference engine to see how changes might affect the life cycle impacts. The inference engine gives recommendations based on input or query to analyze various scenarios to help decision-making. The system provides a comprehensive analysis that enables users to accurately interpret the life cycle impacts of buildings. It allows for assessing different scenarios and helps make informed decisions to improve environmental and economic performance.

The LCA interface empowers stakeholders to input or query diverse scenarios into developed iKBS using an intelligent inference engine. Furthermore, the interface provides a detailed and comprehensive analysis of environmentally sustainable and efficient materials. The IKBS is a pivotal tool in driving eco-friendly viable building and infrastructure development practices.

Case Study

This case study focuses on analyzing the impact of renovation on the life cycle of an Irish semi-detached building. This study demonstrates the application of an Intelligent Knowledge-based System (iKBS) for detailed residential building LCA (Figure 3).

Intelligent Knowledge-based System Development

This case study aims to evaluate the environmental impact of existing Irish semi-detached typology and recommend the best materials and components for retrofits. The scope encompasses assessing embodied energy, operational en-

Table 1: Data knowledge requirement for Irish case study knowledge-base development

knowledge Requirement	Irish Case Study
Product/ Material Data A1- A3	ICE, Ecoinvent, TM65, Irish National Material database
Construction Stage Data A4- A5	ICE, Ecoinvent, TM65,Irish National Material database
Building Stock Data	Ireland BER database
Retrofit Data B2-B5	Irish Building Regulation, Irish Retrofit Scenarios
End-of-life Phase Data C1-C4	ICE, Ecoinvent, TM65, Irish National Material database
Building Typologies	Irish TABULA, EnergyPlus

ergy, and potential retrofitting options for energy efficiency improvement. The study targets policymakers, contractors, and homeowners to provide insights into sustainable building practices. The scope of this case study includes specifying the application, approach, data requirements, study period, functional unit, and system boundary.

The knowledge-based development phase involves an extensive collection and analysis of knowledge for the life cycle of buildings, focusing on the context of Ireland. This process uses a variety of sources, including the Irish Energy Performance Certificate (EPC) database, the Irish database for embodied energy and materials, the International ICE database, and retrofit scenarios data relevant to Ireland (Table 1).

The knowledge of building materials for constructing Irish houses encompasses stages A1-A5 or C1-C5, ranging from raw material production to end of life. This information is gathered from the Irish National Material Database, the ICE, Ecoinvent, and TM65 databases. (ICE, 2024; Ecoinvent, 2024). These databases contained embodied energy and carbon emissions associated with different materials that contribute to global warming and climate change and information related to the transportation of materials and the actual on-site construction processes (Table 1). The building stock knowledge component offers detailed information about existing houses in Ireland, including their age, construction materials, and build type. This knowledge is collected from the Irish EPC database of 1 million buildings. Retrofit scenarios encompass information about renovations to enhance building energy efficiency collected from existing retrofit schemes in Ireland nationally (Figure 3)(CAP, 2024).

The study used typologies of existing residential buildings in Ireland from the TABULA project. These typologies aid in evaluating the buildings' energy and water usage, focusing on the usage data generation. TABULA provides typological data for the 34 Irish dwelling types across 10 age bands. The Irish TABULA project also provides baseline knowledge for building elements, such as exterior walls, roofs, windows, and floors. The baseline knowledge helps to identify recommended retrofit scenarios for target buildings. This study uses the EnergyPlus simulation engine for energy and water consumption, as well as B6 and B7 usage data generation based on retrofit scenarios.

The integration of acquired Irish knowledge within the iKBS begins with a critical pre-processing stage designed to ensure the accuracy and consistency of the data. This foundational step is crucial for preparing the information for sophisticated analysis and decision-making processes. The iKBS integrates pre-processed data and typologies knowledge using Rule-Based, Case-Based, and Frame-Based systems for intelligent decision-making. These knowledge-based approaches, commonly used based on existing studies, are considered optimal for Knowledge-Based Systems (Santos et al., 2020; Zhao et al., 2019; Nazaruks and Osis, 2021). In this study, MongoDB is used for storing knowledge because of its ability to handle unstructured or semi-structured data. This system enables the evaluation of environmental impacts based on criteria such as material selection and energy efficiency.

The Rule-Based system is pivotal in identifying construction materials used in typology, considering various attributes such as layers, thickness, conductivity, and cost. The Frame-Based system facilitates the retrieval of all data in frames, simplifying interaction with the inference engine for estimating total embodied energy and U-values. This study also compares the embodied carbon of materi-

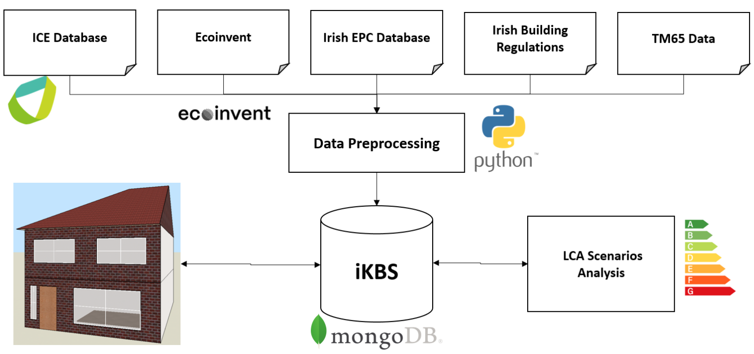


Figure 3: Irish case study knowledge-based development process diagram for LCA analysis

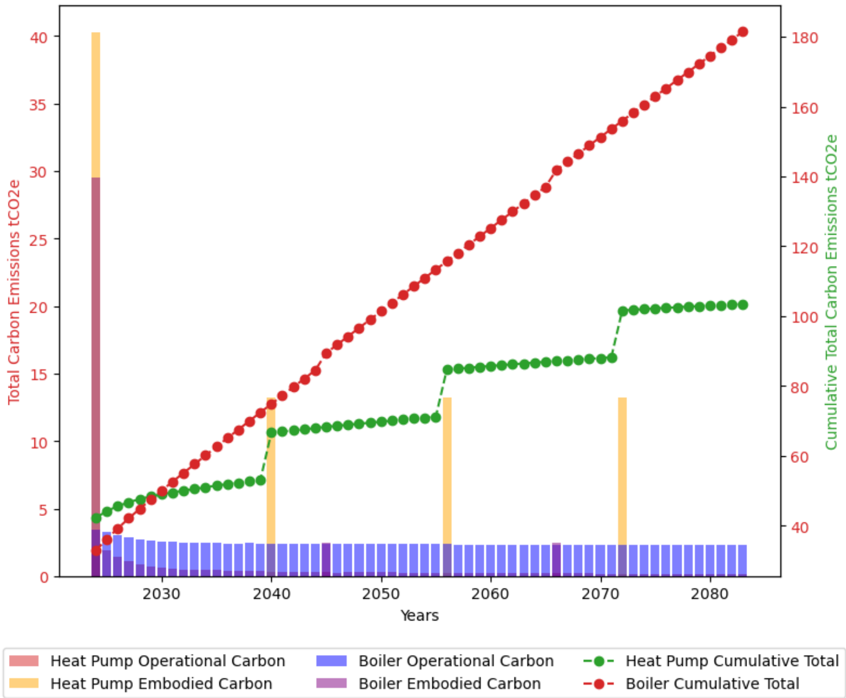


Figure 4: Long-Term impact of retrofitting measures on CO_2e emissions in semi-detached topology residential buildings: A 60-year lifecycle assessment

Table 2: Pre and post-retrofit scenarios for the semi-detached topology building for LCA analysis

	Pre-Retrofit	Post-Retrofit (Boiler)	Post-Retrofit (Heat Pump)
Wall U-value	0.6	0.21	0.21
Roof U-value	0.4	0.13	0.13
Floor U-value	0.64	0.21	0.21
Window U-value	3.1	1.6	1.6
Heating Efficiency/COP	0.7	0.9	3
EUI	290	148	78
Energy Rating	D2	B3	B1

als with those in the Irish material database and other available databases, such as ICE and Ecoinvent. The frame-based system organizes and stores data in a structured format known as frames, which facilitates the efficient retrieval and analysis of information related to construction materials.

Furthermore, the Case-Based system develops recommendations for stakeholders to find insulation materials with minimal embodied carbon. This system identifies similar cases using a fuzzy string-matching approach, helping to find materials by name or category. The fuzzy approach is the most robust in handling variations in input data and string matching, based on existing studies (Zhao et al., 2019). The case-based system uses historical knowledge of insulation materials to recommend options with minimal embodied carbon for new construction projects.

Overall, the hybrid approach is used in the inference engine to assess the environmental and economic impacts of the buildings. This integration offers a robust framework for intelligent decision-making, providing flexibility, adaptability, and comprehensive analysis. The system provides recommendations and a comprehensive analysis of existing and retrofit material, allowing users to explore various scenarios for improved environmental performance.

LCA for Semi-detached typology using iKBS

The case study of the semi-detached typology demonstrates how iKBS can lead to recommendations for materials with minimal embodied energy or operational energy for retrofitting. The typology of a semi-detached house built between 1983 and 1993 features cavity walls and pitched roofs with a floor area of 106.27 m^2. The pre-retrofit typology U-values are 0.6 $W/(m^2{\cdot}K)$ for walls, 0.4 $W/(m^2{\cdot}K)$ for roofs, and 0.64 $W/(m^2{\cdot}K)$ for floors, with double-glazed, PVC-framed windows at 3.1 $W/(m^2{\cdot}K)$, and a boiler efficiency of 70%, leading to an energy rating of D2 and a total Energy Use Intensity (EUI) of 290 $kWh/(m^2.yr)$. These figures indicate significant potential for improvement. The iKBS recommended various insulation materials based on their thermal conductivity and embodied energy values to enhance the U-values of the walls, roof, floor, and window (Table 2).

The iKBS recommends various insulation materials, including Extruded Polystyrene (XPS) and mineral wool, to improve the U-values of building structures such as walls, roofs, and floors. These recommendations based on Low thermal conductivity ensure that materials like XPS offer superior insulation performance, effectively reducing heat transfer through walls and floors. Furthermore, these materials enhance a building's energy efficiency and contribute to a more stable and comfortable indoor environment. On the other hand, mineral wool is highlighted for roof insulation due to excellent thermal resistance and ability to manage moisture, further optimizing the building's thermal performance. The iKBS aims to promote sustainable construction practices that minimize energy consumption and carbon footprint, aligning with broader environmental conservation efforts.

In this study, two heating system scenarios are considered for post-retrofit scenarios, such as employing a condensing boiler with 90% efficiency and an air source heat pump with a Coefficient of Performance (COP) of 3.0 (Table 2). The life cycle assessment of retrofitting measures in a semi-detached building, analyzed over a 60-year life span, highlights the significant environmental benefits of retrofitting despite initial increases in total CO_2 emissions due to embodied carbon from materials and heat pump installations (Figure 4). The analysis reveals a significant reduction in operational emissions post-retrofit when compared to both boiler and heat pump heating systems, validating the efficacy of retrofitting in reducing operational carbon.

Despite the initial environmental cost indicated by the surge in total emissions due to the heat pump (40 tCO_2e) and boiler installation (30 tCO_2e), the operational savings quickly compensate for these impacts, leading to a considerable reduction in the building's carbon footprint over the long term. The periodic accounting for the embodied emissions of heat pumps every 15 years and boilers every 20 years takes into account the replacement cycles of the equipment, underscoring the necessity of long-term sustainability planning.

Overall, the LCA demonstrates that retrofitting significantly contributes to environmental sustainability by lowering operational emissions, emphasizing the long-term benefits of such interventions in residential building management. Retrofitting with a boiler decreases the EUI from 290 to 149 kWh/(m^2.yr), enhancing the energy rating from D2 to B3, with a 184 tCO_2e emission. Similarly, Retrofitting with a heat pump reduces the Energy Use Intensity (EUI) from 290 to 78 kWh/(m^2.yr), upgrading the energy rating from D2 to B1 and resulting in a reduction of 90 tCO_2e emission.

This case study demonstrated the effective application of an iKBS in the context of Irish housing stock. The study provided valuable insights into sustainable building practices by integrating comprehensive data sources, highlighting areas for embodied energy efficiency improvements and potential retrofit options. The approach underscores the importance of detailed life cycle assessments in guiding sustainable decision-making in the building and construction sector.

The results show that iKBS enabled the quantification of energy and emissions savings with greater specificity, demonstrating significant improvements in energy use intensity and carbon emissions reductions through various retrofit scenarios. Traditional methods might not fully capture the nuanced impacts of different retrofit options due to complex interactions between various building materials, usage patterns, and lifecycle stages. This methodology offers a more dynamic and comprehensive tool for policymakers, urban planners, and researchers aiming to optimize residential renovations for sustainability and climate mitigation.

Conclusions

This study bridged a significant gap in Building LCA for residential renovations by integrating diverse data sources such as existing building stock data, residential typologies, retrofit scenarios, and building materials data into an intelligent Knowledge Base System (iKBS). This research has advanced the understanding of the environmental impacts of residential renovations. Knowledge-based system development represents a novel approach to assessing building life cycles, particularly in the national residential building stock context.

The comparative analysis of various KBS approaches, including rule-based, case-based, and frame-based systems, has provided valuable insights into operational and embodied energy aspects. This work is particularly noteworthy for contributing to eco-friendly renovation practices and advancing sustainable building methodologies. The findings of this study are a valuable resource for energy policymakers, urban planners, and researchers, offering a new and intelligent framework for promoting sustainable practices in the building sector.

Future work would expand the methodology's applicability by encompassing diverse geographic regions building typologies and retrofit scenarios thereby enhancing its utility and relevance. Furthermore, integrating emerging technologies such as machine learning and big data analytics could further refine the accuracy and efficiency of the LCA process. Collaborating with industry stakeholders and policymakers will be vital in implementing these advanced tools in practical scenarios. Finally, the ongoing evolution of this research can significantly contribute to the global effort to create sustainable and eco-friendly buildings.

Acknowledgments

This publication has emanated from research supported partly by a grant from the Sustainable Energy Authority of Ireland (SEAI) under the Research, Development and Demonstration Funding Programme Grant No. 22/RDD/822. The opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the SEAI.

References

Ali, U., Bano, S., Shamsi, M. H., Sood, D., Hoare, C., Zuo, W., Hewitt, N., and O'Donnell, J. (2024). Urban building energy performance prediction and retrofit analysis using data-driven machine learning approach. Energy and Buildings, 303:113768.

Balouktsi, M., Lützkendorf, T., Röck, M., Passer, A., Reisinger, T., and Frischknecht, R. (2020). Survey results on acceptance and use of life cycle assessment among designers in world regions: Iea ebc annex 72. In IOP Conference Series: Earth and Environmental Science, volume 588, page 032023. IOP Publishing.

CAP (2024). Climate Action Plan ireland.

Ecoinvent (2024). Ecoinvent database.

Feng, H., Kassem, M., Greenwood, D., and Doukari, O. (2023). Whole building life cycle assessment at the design stage: A bim-based framework using environmental product declaration. International Journal of Building Pathology and Adaptation, 41(1):109–142.

Fenner, A. E., Kibert, C. J., Li, J., Razkenari, M. A., Hakim, H., Lu, X., Kouhirostami, M., and Sam, M. (2020). Embodied, operation, and commuting emissions: A case study comparing the carbon hotspots of an educational building. volume 268, pages 122–081.

Hollberg, A., Kiss, B., Röck, M., Soust-Verdaguer, B., Wiberg, A. H., Lasvaux, S., Galimshina, A., and Habert, G. (2021). Review of visualising lca results in the design process of buildings. Building and Environment, 190:107530.

ICE (2024). Embodied carbon - the ice database.

Jusselme, T., Rey, E., and Andersen, M. (2020). Surveying the environmental life-cycle performance assessments: Practice and context at early building design stages. Sustainable Cities and Society, 52:101879.

Kadrić, D., Aganović, A., and Kadrić, E. (2023). Multi-objective optimization of energy-efficient retrofitting strategies for single-family residential homes: Minimizing energy consumption, co2 emissions and retrofit costs. Energy Reports, 10:1968–1981.

Mohammadpourkarbasi, H., Riddle, B., Liu, C., and Sharples, S. (2023). Life cycle carbon assessment of decarbonising uk's hard-to-treat homes: A comparative study of conventional retrofit vs enerphit, heat pump first vs fabric first and ecological vs petrochemical retrofit approaches. Energy and Buildings, 296:113353.

Najjar, M., Figueiredo, K., Hammad, A. W., and Haddad, A. (2019). Integrated optimization with building information modeling and life cycle assessment for generating energy efficient buildings. Applied Energy, 250:1366–1382.

Nazaruks, V. and Osis, J. (2021). An overview of knowledge representation with frames. Advancements in Model-Driven Architecture in Software Engineering, pages 46–63.

Roberts, M., Allen, S., and Coley, D. (2020). Life cycle assessment in the building design process–a systematic literature review. Building and Environment, 185:107274.

Röck, M., Saade, M. R. M., Balouktsi, M., Rasmussen, F. N., Birgisdottir, H., Frischknecht, R., Habert, G., Lützkendorf, T., and Passer, A. (2020). Embodied GHG emissions of buildings – The hidden challenge for effective climate change mitigation. volume 258, pages 114–107.

Santos, G., Vale, Z., Faria, P., and Gomes, L. (2020). Bricks: Building's reasoning for intelligent control knowledge-based system. Sustainable Cities and Society, 52:101832.

Säynäjoki, A., Heinonen, J., Junnila, S., and Horvath, A. (2017). Can life-cycle assessment produce reliable policy guidelines in the building sector? Environmental Research Letters, 12(1):013001.

Sharma, A., Saxena, A., Sethi, M., Shree, V., et al. (2011). Life cycle assessment of buildings: a review. Renewable and Sustainable Energy Reviews, 15(1):871–875.

Waldman, B., Huang, M., and Simonen, K. (2020). Embodied carbon in construction materials: a framework for quantifying data quality in epds. Buildings and Cities, 1(1).

Wernet, G., Bauer, C., Steubing, B., Reinhard, J., Moreno-Ruiz, E., and Weidema, B. (2016). The ecoinvent database version 3 (part i): overview and methodology. The International Journal of Life Cycle Assessment, 21:1218–1230.

WGBC (2023). World Green Building Council. `https://worldgbc.org/`. [Online; accessed 31-Aug-2023].

Zhao, X., Tan, Y., Shen, L., Zhang, G., and Wang, J. (2019). Case-based reasoning approach for supporting building green retrofit decisions. Building and Environment, 160:106210.

Zuo, J., Pullen, S., Rameezdeen, R., Bennetts, H., Wang, Y., Mao, G., Zhou, Z., Du, H., and Duan, H. (2017). Green building evaluation from a life-cycle perspective in australia: A critical review. Renewable and Sustainable Energy Reviews, 70:358–368.

Product and Process Modelling

AUTOMATED GEOMETRIC-SEMANTIC DIGITAL TWINNING OF STAIRCASES FROM DENSE LASER SCANNER POINT CLOUDS BY PARAMETRIC PROTOTYPE MODELS

Mansour Mehranfar[1], Alexander Braun[1], and André Borrmann[1]
[1]Technical University of Munich, Munich, Germany

Abstract

Digital Twins (DTs) have become as valuable tools for management, and operation of assets. A main challenge, however, persists in the automated creation of high-quality buildings DTs, providing both precise geometry and semantics. This paper introduces a novel hybrid bottom-up, top-down approach for the automated creation of DT models of staircase structures using laser scanner point cloud. The proposed workflow involves separating inclined staircase points, designing parametric DT models, and model fitting through an optimization process. The results demonstrate the effectiveness of the proposed with an average accuracy of about 4 cm in determining the dimensions of the model's elements.

Introduction

The automatic creation of DTs for the built environment has emerged as a prominent and demanding field within the AEC domain (Volk et al., 2014; Bosché et al., 2015; Drobnyi et al., 2023). A DT is a virtual replica of physical assets that facilitates real-time simulation and monitoring. DTs offer numerous possibilities for resource and facility management, enabling more intelligent analysis and interactions (Drobnyi et al., 2023). When it comes to creating DT models for exiting buildings with rich semantic content and coherent geometry, an important first step is to acquire relevant data, especially visual and spatial data. In this regard, laser scanning technology plays a crucial role, enabling the meticulous collection of point cloud data and the creation of virtual replicas for both indoor and outdoor environments (Noichl and Borrmann, 2022; Pan et al., 2023; Abdollahi et al., 2023; Martens and Blankenbach, 2023). This technology plays a pivotal role in advancing the domains of building performance and sustainability.

In the past decade, significant progress has been achieved in the field of point cloud processing techniques and 3D scene understanding methodologies, with a specific emphasis on automatically generating accurate geometric-semantic DT models from point cloud data. Nevertheless, developing a comprehensive and precise algorithm for this purpose has encountered numerous challenges, primarily arising from noise, clutter, obstructions, and the representation of volumetric-semantic models (Ochmann et al., 2016). Furthermore, most of the existing approaches have primarily focused on creating 3D models of structural elements such as ceilings, floors, and walls while neglecting the intricate modeling of other structural components, notably staircases, which play a pivotal role in the reconstruction of DTs for multi-story buildings (Nikoohemat et al., 2020). This oversight can be attributed to the inherent complexities associated with indoor scenes and the intricacies involved in reconstructing staircase structures (Schmittwilken et al., 2009). Staircase structures assume diverse shapes and configurations based on varying specifications, making the automatic reconstruction of 3D models a challenging task.

In the realm of urban planning and facility management, digital staircase models are vital for optimizing pedestrian flow and enhancing public safety. In this regard, the automatic creation of digital staircase models from point cloud data empowers facility managers and builders to conduct virtual walkthroughs, identify potential conflicts, and reduce construction and maintenance costs.

In the research presented in this paper, the primary objective is to introduce an automated workflow that combines both bottom-up and top-down approaches to extract staircase points within the indoor environment and generate parametric DT models with coherent geometry. The key contribution lies in utilizing domain knowledge to formulate parametric models for the building's staircase structures and subsequently fitting the designed rough models to point cloud data to achieve a close representation of reality.

Background and related work

Digital building twin creation

With the increase in demands for the creation of DT models for the built environment, indoor digital twinning has become an intensively researched topic in AEC domain (Borrmann et al., 2023). In this context, laser scanners and photogrammetry technologies are the most modern and efficient measurement tools, facilitating the acquisition of accurate geometric and semantic information, which are widely used in the realm of automatic creation of 3D digital models. However, despite all progress made, the automatic creation of digital building twin from point cloud data remains only partially resolved, and most developed methods are designed for specific types of buildings and restricted to reconstructing specific kinds of objects based on use cases.

In this regard, Xiong et al. (2013) proposed an automated 3D reconstruction framework utilizing a voxelized point cloud to identify patches, such as walls, ceilings, and floors, by adhering to boundary constraints. Monszpart et al. (2015) proposed the regular arrangements of planes (RAP) technique for reconstructing indoor 3D scenes using point cloud data. The proposed method uses local plane-based approximations and global inter-plane relations to simplify the arrangement of planes fitted to the points of the environment. Ochmann et al. (2019) proposed an innovative approach for reconstructing volumetric building DT models, encompassing floor, ceiling, and wall elements. They employed a linear optimization framework to disjoint distinct 3D spaces and determine the positions of common wall instances shared among them. Tran and Khoshelham (2020) employed a reversible jump Markov Chain Monte Carlo (rjMCMC) algorithm to facilitate the application of shape grammar rules in the procedural-based reconstruction of 3D indoor models from dense point cloud data. Abdollahi et al. (2023) introduced a progressive model-driven approach for the 3D modeling of indoor spaces employing watertight predefined models. This approach initially segmented spaces into rectangular and non-rectangular regions with an even number of sides. Subsequently, a point density occupancy map is used to enhance the level of detail in the intrusion and protrusion parts of Manhattan and non-Manhattan models.

Recent advancements in artificial intelligence (AI) and machine learning (ML) technology provide an effective solution for accurate point cloud processing, reducing the manual effort required in creating DTs. In this regard, Mehranfar et al. (2022) proposed a hybrid bottom-up, top-down method using AI semantic segmentation network to separate the 3D space in simple and complex indoor point cloud. Mehranfar et al. (2023) also combined domain knowledge in construction and architectural design with the capabilities of AI techniques in scene understanding to create highly parameterized building models with rich semantic and coherent geometry from dense point clouds. The proposed method employs the parametric modeling approach and model fitting through an optimization process to overcome common indoor point cloud challenges, including noise, gap, and clutter. Kellner et al. (2023) developed a multi-step data-driven algorithm enriched with AI techniques for the 3D reconstruction of building models at Level of Development (LoD) 400. The proposed approach utilizes an AI semantic segmentation network to separate the following classes: doors, door leaves, windows, walls, ceilings, floors, and clutter. Subsequently, a 2D projection combined with a neighborhood graph structure is employed to partition 3D space within indoor environments, followed by applying the RANSAC plane fitting algorithm to reconstruct 3D models of walls. Ultimately, a bounding box is fitted to the points pertaining to each individual instance of door and window elements.

Staircase points detection and modeling

Staircase detection within point clouds and the subsequent creation of staircase DT models is an area of research that has received less attention compared to other applications of point cloud processing. Staircases exhibit diverse shapes and dimensions, and their visual representation can be influenced by different factors, such as structural design and the quality of the point cloud data. This complexity introduces challenges in developing robust and broadly applicable staircase detection algorithms. Additionally, utilizing ML approaches for staircase detection heavily relies on huge labeled datasets for network training purposes. The scarcity of such datasets specifically for staircases has impeded progress in research within this domain. Among the little research conducted on staircase point detection and digital model reconstruction, Schmittwilken et al. (2009) introduced a low-level module based on the random sample consensus (RANSAC) algorithm to generate planar polygonal patches for building facades and the surrounding ground. They employed Conditional Random Fields (CRFs) to classify patches into facade, window, door, and staircase classes, considering local neighborhood information and incorporating attribute grammar, including object partonomy and observable geometric constraints.

In 2009, Schmittwilken and Plümer (2009) proposed a top-down approach for reconstructing symmetric and partly recursive objects, such as a triple-run staircase from point cloud data. The proposed method utilizes an attribute grammar formulation using geometric dependencies for the design of 3D models and subsequently employs the random sample consensus paradigm for model selection and extraction of the geometric parameter of each 3D object. Sinha et al. (2014) presented a data-driven approach that employs a minimal 3D map representation and calculates step-like local features using point neighborhoods to detect staircase points. Yang et al. (2019) employed a bottom-up hierarchical semantic classification method, incorporating various semantic definitions, such as planarity of wall, ceiling, and floor surfaces at the geometric primitive level, to establish relationships between the staircase connection space and indoor space. For the coarse segmentation of staircase points, the height histogram of points is used to identify the void region between the planar surfaces of solid slabs and the connection space with staircases. Subsequently, the connected component algorithm is applied to cluster distinct pieces, planes, and clusters of staircases. Finally, the α-shapes algorithm is employed to construct the surface model for each step of the staircase.

In conclusion, the automatic creation of digital building twins using point cloud data is a research field constantly facing challenges such as complex space layouts, clutter, obstructions, and the need to provide methods for representing volumetric semantic models. The challenges become even more complex for elements such as staircases, which can have varying appearances based on their design. In the research presented in this paper, we aim to develop a hybrid bottom-up, top-down method for the automatic creation of a DT model for staircase structures using point cloud data. The proposed method leverages domain knowledge in construction and design to create para-

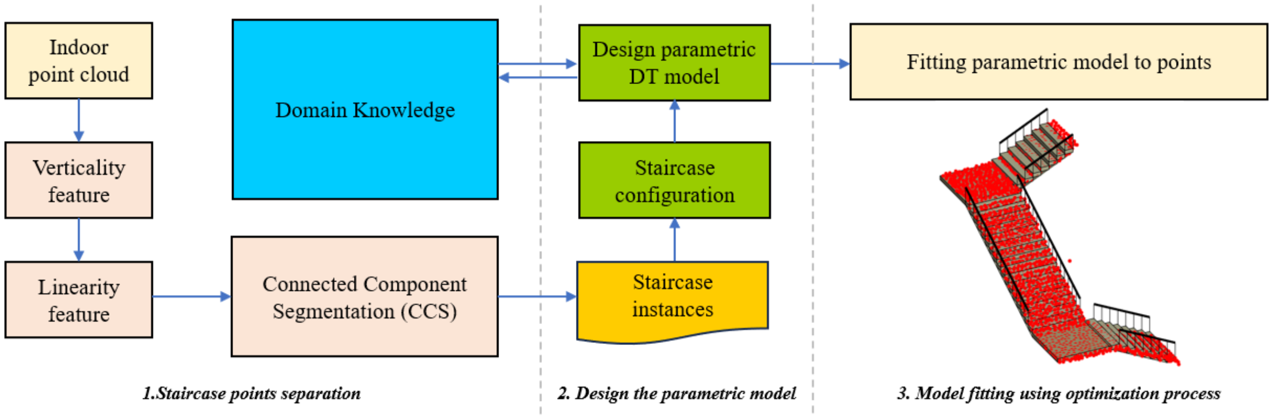


Figure 1: The proposed workflow for creating the DT model of the staircase.

metric models that can consider semantic relationships between components, allowing to overcome challenges such as noise and obstructions.

Methodology

As shown in Figure 1, the proposed workflow for the automatic creation of the DT model for staircase elements consists of three major steps, including 1) staircase points separation, 2) designing the parametric model of the staircase, and 3) model fitting using an optimization process. The details of the steps are given in the following subsections.

Staircase points separation

Staircase structures are typically inclined and constructed with a specific slope. Various factors, such as building regulations, architectural preferences, and the intended purpose of the staircases determine the inclination angle of a staircase. In residential and educational buildings, staircases are often designed with a relatively gentle incline to provide comfort and ease of use, such as for the convenience of (elderly) occupants. In the proposed method, the Verticality geometric feature is employed to distinguish inclined staircase points from other structural elements within the point cloud space, including vertically oriented walls and horizontally oriented ceilings and floors (Figure 2b).

To calculate the Verticality feature for each point in the point cloud space, the equation 1 is used with a neighborhood of 25 cm (Grilli et al., 2019). The selected neighborhood radius value ensure the presence of a minimum number of neighboring points for accurate estimation of Verticality features for the main structural elements, as well as small furniture objects.

$$Verticality = 1 - nz, \tag{1}$$

where nz is the normal vector value toward the Z axis, which is calculated by the covariance matrix and eigenvector values of the nearest neighboring points using equation 2 (Grilli et al., 2019):

$$c = \frac{1}{k}\sum_{i=1}^{n}(pi - \bar{p}).(pi - \bar{p})^{T}, \tag{2}$$

where k is the number of neighboring points, pi and p also refer to the 3D coordinates of the points being considered. Additionally, the eigenvalues and eigenvectors are determined by equation 3 where λ and $\vec{v}$ are eigenvalues and eigenvectors respectively (Grilli et al., 2019):

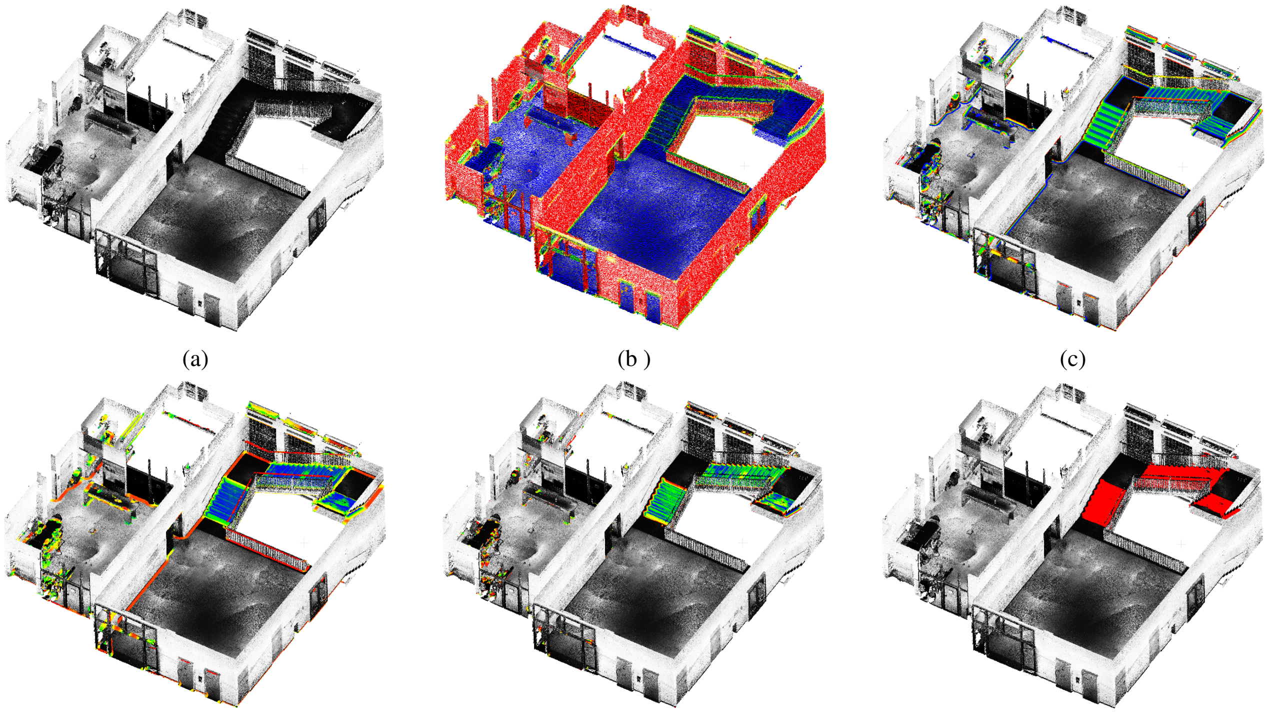
$$c.\vec{v}_j = \lambda_j.\vec{v}_j, \quad j \in \{0,1,2\} \tag{3}$$

Points associated with entirely horizontal elements are assigned a Verticality value of zero, while points pertaining to completely vertical elements are assigned a Verticality value of one. In this case, a confidence interval of 45%, characterized by approximate slopes ranging from 0.05% to 0.5%, is considered to extract the inclined points (Figure 2c).

The output of this process includes inclined staircase points, edge points of walls and ceilings, floor elements, noise points, and clutter. To extract points specific to each staircase instances, we adopt a two-step approach. Initially, the Linearity feature of the points is leveraged using equation 4 to eliminate the edge points of wall and ceiling elements (Figure 2e). Subsequently, the Connected Component Segmentation (CCS) algorithm is employed to segment points into distinct groups based on their connectivity, effectively isolating small segments that often appear as clutter within the point cloud space (Figure 2f) Trevor et al. (2013).

$$Linearity = \frac{\lambda_1 - \lambda_2}{\lambda_1} \tag{4}$$

As shown in Figure 3a, the extracted staircase lacks points corresponding to the horizontal planes of the staircase landing tread. This omission is due to the exclusion of horizontal points in the filtering process of the previous step. However, this information gap proves advantageous in subsequent steps for separating staircase components and their configuration analysis.



(a) (b) (c)

(d) (e) (f)

Figure 2: Staircase points separation: (a) original point cloud, (b) Verticality feature calculation, and removal of empty cells from the initial 2D bounding box model, (c) filtering the points with the Verticality feature between 0.05 to 0.5, (d) Linearity feature calculation to filter noise and unwanted furniture points, (e) staircase points after noise removal, (f) applying the CCS method to segment staircase points.

Design the parametric model of staircase

The structural configuration of interconnecting staircases between building levels frequently comprises multiple stair tread parts and incorporates two or more landing treads. The design and structural morphology of these staircases is inherently influenced by factors such as the flow of ingress and egress and the placement of landing treads. To quantify the number of landings and determine their direction, the central portion of the staircase points is examined using 3D point density analysis (Figure 3b). This analysis reveals that the extracted staircase element comprises three flows and two landing treads that run from left to right, top to down, and right to left, respectively (Figure 3c). This information serves as valuable input for subsequent steps, particularly within the parametric design of the staircase structure.

The staircase structure typically consists of distinct primary components such as step riser, step tread, railing etc. The differences between staircases primarily pertain to the number and dimensions of the primary components (e.g. number of steps, width, length, depth of steps, and landing treads) as well as the specific modes of flow rotation and positioning of the landing tread. Specifically, we consider four possible rotation types, denoted by numerical identifiers 1, 2, 3, and 4, corresponding to left-to-right, right-to-left, bottom-to-top, and top-to-bottom orientations, respectively (Figure 4a). For instance, the primary configuration of the extracted staircase is represented by the array ”142” (Figure 4b). This signifies that the staircase in question consists of three flows and two landing treads, each being successively positioned from left to right, top to bottom, and right to left. Consequently, a library of parametric models is systematically generated based on the number of landing treads and flow patterns.

Model fitting using optimization process

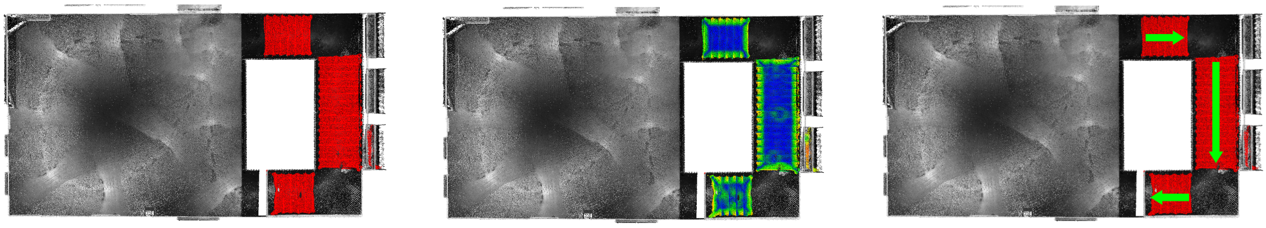
The designed parametric model has low geometric accuracy but a consistent semantic topology. To determine the optimal geometric values of the DT model's components, the optimization process is used to fit the designed parametric model to the extracted staircase points. The Nelder-Mead optimization algorithm fits the initial parametric model to points (Nelder and Mead, 1965). Within this context, the objective function employed during the model-to-point fitting process is defined as follows (5) :

$$Obj = min(G + \frac{F}{10}) \tag{5}$$

The term G refers to the distance between the points and the step planes, which is a critical factor for the vertical alignment of the parametric model toward the staircase points (6) (Figure 5):

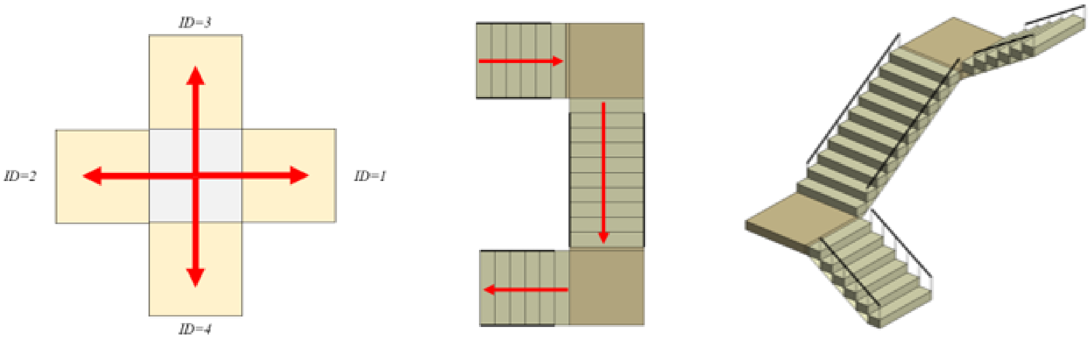
$$G = \sum_{i=1,j=1}^{n,k} |p_i - plane_j| \tag{6}$$

where the p_i is the staircase point i, $plane_j$ is the jth plane of parametric model, p_{stairs} is the number staircase points



(a) (b) (c)

Figure 3: Staircase configuration: (a) staircase points, (b) central part of staircase, (c) separated flows and their orientation.



(a) (b) (c)

Figure 4: Determining the type of staircase based on the number of flows and the number of landing treads: (a) possible rotation directions, (b) design the configuration of the parametric model for the staircase instance with two landing treads and flow's ID array [1 4 2], (c) highly parameterized DT model of the staircase.

and p_{in} is the number staircase points inside the step box of staircase model.

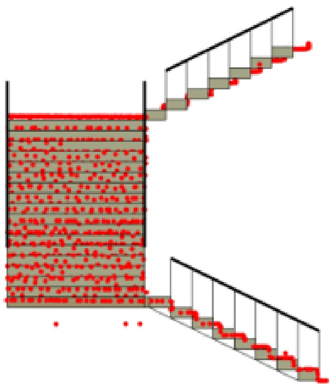


Figure 5: Fitting parametric model of staircase to points; Vertical alignment of the parametric model toward the staircase points (term G).

The term F is related to the number of steps present within each flow connecting the landing treads, as well as the placement of the parametric model planes on the staircase points on the X-Y plane (7) (Figure 6):

$$F = |p_{stairs} - \sum_{i=1}^{steps} p_{in}| \tag{7}$$

where the p_{stairs} is the number of staircase points and p_{in} is the number of staircase points inside the step box of the staircase model. To enhance the process of fitting a model to the points, in each iteration, the 3D points inside each box of the model are extracted and used in the optimization process. This includes the points on the landing treads that were omited during the staircase points separation step.

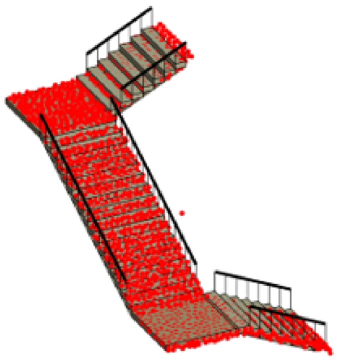


Figure 6: Fitting parametric model of staircase to points; placement of the parametric model planes on the staircase points (term F).

During the model fitting process, the points inside each box of the DT models are extracted and appended to the previously extracted staircase points to improve the model fitting, specifically in the parts of landing treads. In between, adding excessive steps may not change the overall distance value between the staircase points and the model planes. To address this challenge, a penalty factor is incorporated into the objective function. Specifically, if an additional step does not encompass any points within the 2D X-Y plane, a penalty value of 10000 is appended to the final value of the objective function (8):

$$if\, p_{in} == 0 \;\; then \;\; Obj = min(G + \frac{F}{10}) + 10000 \tag{8}$$

Result

To evaluate the proposed method, the dense point cloud dataset is used that has been captured by the authors at the City campus of the Technical University of Munich (TUM). The TUM building in question (Building 1) is a multi-story structure comprising four floors. The point cloud data was captured using NavVis laser scanner with a spatial resolution of 0.1 cm (https://www.navvis.com/). Figure 7 illustrates an overview of the data.

Figure 7: Overview of the TUM Building 1 data.

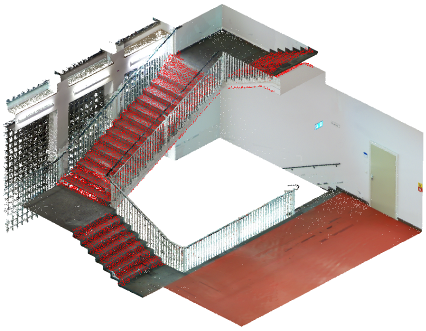
Experimental results on staircase points separation

To extract the staircase points within each floor's point cloud, first, the Verticality feature is calculated for the points in each floor by considering the neighborhood radius of 0.25 cm. Next, the inclined points of staircases are extracted from the point cloud space considering a confidence interval of 45%, which encompasses the points with the Verticality feature between 0.05 to 0.5. As mentioned in Section 2, the result of this step contains noise and clutter. In this regard, the linearity feature is calculated for each point considering the neighborhood radius of 0.50 cm to detect the noise and boundary points between the ceiling, floor, and wall elements. Finally, the CCS method is employed to segment the points belonging to each staircase instance. In this case, the distance threshold for connecting the points in each segment is considered as 25 cm and the minimum number of points to segment each staircase instance is specified equal to 5000. Figure 8 shows the result of staircase point separation for different floors of the TUM Building 1 dataset.

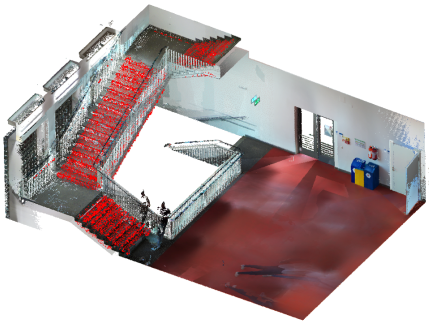
To assess the effectiveness of the staircase points separation step, the manually annotated point cloud and the extracted staircase points are compared, and the results are reported in Table 3 based on Equations (9-11). In this regard, an overall accuracy of about 95% for separating staircase points from other elements in the point cloud space highlights the performance of the proposed workflow for separating staircase points, which significantly influences the quality of the model fitting steps.

$$Precision = \frac{TP}{TP+FP} \tag{9}$$

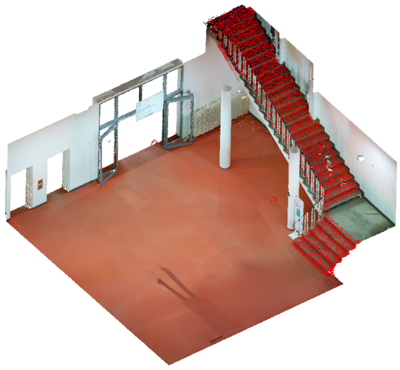
$$Recall = \frac{TP}{TP+FN} \tag{10}$$



(a)



(b)



(c)

Figure 8: The result of staircase points separation for the TUM Building 1 dataset: (a) floor-2, (b) floor-3, and (c) floor-4.

$$F\text{-}score = 2.\frac{Precision.Recall}{Precision+Recall} \tag{11}$$

Experimental results on creation parametric model of staircase

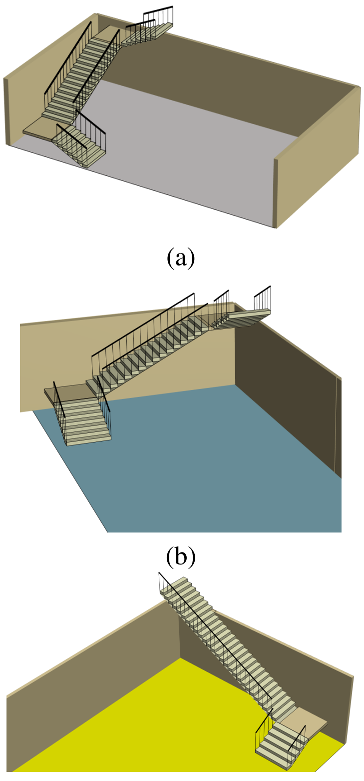
First, the central parts of the extracted staircase points are examined to design the parametric DT model for the extracted staircase instances. In this regard, the structure of staircase instances in the point cloud of the second to third floor is determined by array 142. This means the corresponding staircase structure consists of two landing treads and three flows formed from left to right, up to down, and right to left, respectively. Also, the staircase structure extracted in the point cloud floor fourth is determined by array 32, meaning that this staircase consists of one landing tread element and two flows, which are formed from bottom to top and from right to left.

Subsequently, parametric DT models are designed based

Table 1: Accuracy evaluation of staircase points separation.

Dataset	F (2)	F (3)	F (4)
Precision	0.75	0.60	0.70
Recall	0.80	0.73	0.85
F-Score	0.77	0.64	0.75
Overall accuracy = 0.95			

on the provided information for each staircase instance. These models incorporate various parameters and consider their relation, including the number of steps between landing treads and dimensional values of geometric properties such as width, depth, and height, which, in addition to maintaining the semantic relation between the staircase components, also promotes geometric integrity. Following this, the designed parametric models are fitted to the extracted staircase points to estimate optimal values for the parameters of the created DT models. As discussed earlier, the Nelder-Mead optimization method and the objective function are employed to fit the model to the points and extract optimal parameters for the designed DT models, ensuring that the models closely resemble real-world structures.



(a)

(b)

(c)

Figure 9: The result of staircase points separation for TUM Building 1 dataset: (a) floor-2, (b) floor-3, and (c) floor-4.

To evaluate the proposed method for the automatic creation of the staircase DT model, the differences between distances and dimensional values of the staircase components in the reference DT models and the generated models are calculated. For each dataset, the standard metric of mean error is presented in Table 4. Furthermore, the distance between the model and the extracted staircase points is measured for each dataset, indicating the closeness of the created model to the captured point cloud. The overall mean accuracy of about 0.04 m highlights the performance and effectiveness of the proposed method for the automatic creation of DT models for staircase structures in the built environment.

Table 2: Accuracy evaluation of staircase DT creation.

Dataset	F (2)	F (3)	F (4)
Flights:			
Width	0.04	0.05	0.04
Depth	0.03	0.04	0.04
Height	0.02	0.03	0.03
Number of Steps	100%	97%	93%
Landing treads:			
Width	0.04	0.05	0.04
Depth	0.03	0.04	0.05
Height	0.02	0.03	0.04
Overall accuracy	0.03	0.04	0.04
Points to Model	0.08	0.11	0.07

Conclusion

This research presents a novel hybrid bottom-up, top-down method to create the highly parameterized DT model of a staircase structure. The proposed method aligns existing knowledge in the design and construction of building elements with the parametric modeling framework, allowing us to consider different degrees of freedom to model a wide range of staircase models in the real world. By considering semantic relationships between staircase components, the presented approach provides solutions to overcome challenges such as point cloud obstruction, noise, and complexity in the representation of 3D geometric models. In this regard, an overall mean error of about 0.04 m in creating a highly parameterized DT model for a staircase structure with different shapes and designs promises significant progress in the field of "Scan-to-BIM" that ultimately provides high-quality DTs with rich semantics and coherent geometry. Despite diligent considerations, the proposed method has inherent limitations in creating digital models of round staircases with non-Manhattan designs, which will be addressed in future works.

Acknowledgments

The research presented has been performed in the frame of the project AI4TWINNING (Artificial Intelligence for the automated creation of multi-scale digital twins of the built world) funded by the TUM Georg Nemetschek Institute Artificial Intelligence for the Built World (GNI).

References

Abdollahi, A., Arefi, H., Malihi, S., and Maboudi, M. (2023). Progressive model-driven approach for 3d modeling of indoor spaces. Remote Sensing, 23(13).

Borrmann, A., Biswanath, M., Braun, A., Chen, Z., Cremers, D., Heeramaglore, M., Hoegner, L., Mehranfar, M., Kolbe, T., Petzold, F., Rueda, A., Solonets, S., and Zhu, X. (2023). Artificial intelligence for the automated creation of multi-scale digital twins of the built world – ai4twinning. In Proc. of the 18th 3D GeoInfo Conference.

Bosché, F., Ahmed, M., Turkan, Y., Haas, C., and Haas, R. (2015). Heritage building information modeling (h-bim) applied to a stone bridge. Automation in Construction, 49.

Drobnyi, V., Hu, Z., Fathy, Y., and Brilakis, I. (2023). Construction and maintenance of building geometric digital twins: State of the art review. Sensors, 23(9).

Grilli, E., Farella, E., Torresani, A., and Remondino, F. (2019). Geometric features analysis for the classification of cultural heritage point clouds. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W15.

Kellner, M., Stahl, B., and Reiterer, A. (2023). Reconstructing geometrical models of indoor environments based on point clouds. Remote Sensing, 15(18).

Martens, J. and Blankenbach, J. (2023). Vox2bim + - a fast and robust approach for automated indoor point cloud segmentation and building model generation. PFG – Journal of Photogrammetry Remote Sensing and Geoinformation Sciencen, 91(4):273–294.

Mehranfar, M., Braun, A., and Borrmann, A. (2022). A hybrid top-down, bottom-up approach for 3d space parsing using dense rgb point clouds. In Proc. of European Conference on Product and Process Modeling.

Mehranfar, M., Braun, A., and Borrmann, A. (2023). Automatic creation of digital building twins with rich semantics from dense rgb point clouds through semantic segmentation and model fitting. In Proc. of the 30th Int. Conference on Intelligent Computing in Engineering (EG-ICE).

Monszpart, A., Mellado, N., Brostow, G., and Mitra, N. (2015). Rapter: rebuilding man-made scenes with regular arrangements of planes. ACM Transactions on Graphics, 34(4):1–12.

Nelder, J. and Mead, R. (1965). A simplex method for function minimizatio. International Journal of Mathematics and Computers in Simulation, 7:308–313.

Nikoohemat, S., Diakité, A., Zlatanova, S., and Vosselman, G. (2020). Indoor 3d reconstruction from point clouds for optimal routing in complex buildings to support disaster management. Automation in Construction, 13:103109.

Noichl, F. and Borrmann, A. (2022). Automated deterministic model-based indoor scan planning. In Proc. of European Conference on Product and Process Modeling.

Ochmann, S., Vock, R., and Klein, R. (2019). Automatic reconstruction of fully volumetric 3d building models from oriented point clouds. ISPRS Journal of Photogrammetry and Remote Sensing, 151:251–262.

Ochmann, S., Vock, R., Wessel, R., and Klein, R. (2016). Automatic reconstruction of parametric building models from indoor point clouds. Automation in Construction, 38:94–103.

Pan, Y., Braun, A., and Borrmann, Andre Brilakis, I. (2023). 3d deep-learning-enhanced void-growing approach in creating geometric digital twins of buildings. Smart Infrastructure and Construction, 176(1):24–40.

Schmittwilken, J. and Plümer, L. (2009). Model selection for composite objects with attribute grammars. In 12th AGILE International Conference on Geographic Information Science. Leibniz Universität Hannover, Germany.

Schmittwilken, J., Ying, M., FörstnerLutz, W., and Plümer, L. (2009). Construction and maintenance of building geometric digital twins: State of the art review. Annals of GIS, 15(2):117–126.

Sinha, A., Papadakis, P., and Elar, M. (2014). A staircase detection method for 3d point clouds. In 13th International Conference on Control Automation Robotics & Vision (ICARCV).

Tran, H. and Khoshelham, K. (2020). Procedural reconstruction of 3d indoor models from lidar data using reversible jump markov chain monte carlo. Remote Sensing, 12(5).

Trevor, A., Gedikli, S., Rusu, R., and Christensen, H. (2013). Efficient organized point cloud segmentation with connected component. In Proceedings of Semantic Perception Mapping and Exploration, pages 1–6.

Volk, R., Stengel, J., and Schultmann, F. (2014). Building information modeling (bim) for existing buildings — literature review and future needs. Automation in Construction, 38:109–127.

Xiong, X., Adan, A., Akinci, B., and Huber, D. (2013). Automatic creation of semantically rich 3d building models from laser scanner data. Automation in Construction, 31:325–337.

Yang, F., Liang, Y., Li, D., Su, F., Zhu, H., Zuo, X., and Li, L. (2019). Detection of space connectivity from point cloud for stair reconstruction. Environmental Science.

AUTOMATING AS-BUILT-MODELLING BY USING DELIVERY DOCUMENTS

Niels Bartels[1], Oliver Eischet[2], and Henric Meinhardt[3]
[1]University of Applied Sciences Cologne, Cologne, Germany
[2]Specter Automation GmbH, Cologne, Germany
[3]comstruct ICT GmbH, Munich, Germany

Abstract

Building Information Modeling (BIM) promises advantages for Facility Management (FM) by integrating the lifecycle phases of a building. Especially the optimized start-up phase is mentioned as one key advantage of the BIM method. In order to implement a successful start-up, a holistic as-built model is necessary, in which all relevant information for FM is stored. Actually, for the generation of an as-built model a high manual effort is necessary. This paper aims to develop an automated as-built model by using delivery documents. Therefore, this paper presents a process using Optical Character Recognition and analyze the process in a proof-of-concept.

Introduction

Building Information Modeling (BIM) aims to integrate all stakeholders by using a digital building model throughout the lifecycle (Borrmann et al. 2021). The focus here is particularly on the consistent exchange of information using open data exchange formats (Baranova 2021). One important standard to exchange data throughout the lifecycle is the Standard of the Industry Foundation Classes (IFC) (ISO 16739-1). Especially IFC-based as-built models could support an efficient information exchange between BIM systems and FM systems, as Computer Aided Facility Management (Matarneh et al. 2019; Ensafi et al. 2022).

As-built models hereby describes the actual state of the building, as it is constructed and commissioned (Bartels 2020). It represents the update of the as-planned model, which describes the target status of the building and includes graphical and alphanumerical data (Pilling 2016). By integrating all FM-relevant data and the data of the service provisioning during the Operations & Maintenance (O&M) phase, an As-Operated model is generated (RealFM 2021). The maturity of the models is shown in Figure 1.

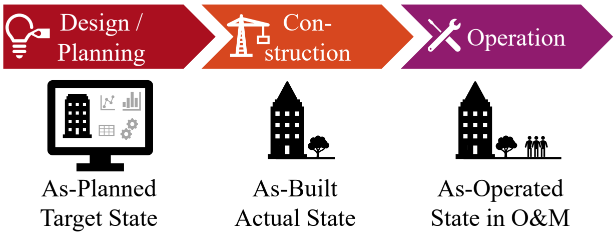


Figure 1: Maturity Levels of the BIM models throughout the lifecycle

While as-planned states the status of the design phase, the as-built modelling requires a finer level of detail, a more precise modelling as well as an accurate data import and maintenance (Pătrăucean et al. 2015). But these data is often poorly collected, processed and revised in practice, what especially applies to asset-specific engineering data (Asmar et al. 2022; Hellenborn et al. 2023). This is especially due to the manual data entry during the construction and commissioning phase and the need for manual verification of Facility Managers (Klein et al. 2012).

1. In order to define the necessary data in the as-built model, it is necessary to unambiguously and clearly define the owner requirements, specially through exchange information requirements (EIR) and Asset Information Requirements (AIR). Research studies show, that the client's requirements are often not clear and specified (Hellenborn et al. 2023). Furthermore, the requirements change from client to client (Di Filippo et al. 2021).
2. In addition to that, especially for Construction Managers the usage of BIM software tools is often not common. Actually, the Construction Managers are using specific tools, such as MS Project or ERP to manage their processes as well as MS Office products, such as Word or Excel. This particularly applies to medium-sized companies (Lauer 2021).
3. The needed information is stored in documents, which are handed over in paper form or as a digital file. In particular, smaller craft businesses that work as subcontractors hand over their documents in paper form, which leads to a high documentation effort (Hausmann et al. 2022).

Actually, there is no connection between the delivery documents and the revision documents with the digital building model, although all needed information is stored in these two documents. Various studies focused on the processing of geometry to develop an as-built model (e.g. (Klein et al. 2012)) or on the documentation of construction specifics, such as timetables (e.g. (Park und Cai 2017)). The needed information for the asset owner and Facility Management are mostly not taken into account. Especially the information is not automatically read out of documents and assigned to objects in the model.

Therefore, it is necessary to implement an IFC-based interface, that enables the construction managers and their companies to automatically fill in all relevant data on the objects in the digital building model. By developing such an interface, the modelling of as-built models could be highly improved.

Based on this, the Facility Management is able to transfer and use all data in their Computer Aided Facility Management (CAFM) systems. This would lead to a better understanding of the building and improvements of efficiency by the FM as well as an increasement of user satisfaction (Bartels 2020).

Literature Review

Existing Solutions

In order to evaluate the state-of-the-art of as-built models, existing solutions for data acquisition, data integration, and data visualization need to be analyzed.
In recent years, various solutions for data acquisition for as-built models have been published. The earliest presented solutions are based on 3D laser scanning for building or geotechnical sites in order to generate geometric as-built models (Su et al. 2006; Deruyter et al.). Based on that research, further solutions for the acquisition of geometric data use advanced technologies, such as Autonomous Unmanned Aerial Vehicles, Imaging technologies or Augmented Reality were published (Omar und Nehdi 2016; Freimuth und König 2019). Not only geometrical but also alphanumeric data was integrated in as-built models, such as construction dates or production data (Son et al. 2017; Son et al. 2015). The analysis of these case studies shows, that the processes are regularly done manually. One presented solution for automated as-built models are sensors (Moselhi et al. 2020).
For data integration the solutions can be divided in solutions that (1) use proprietary exchange formats and (2) that use the open exchange format of the Industry Foundation Classes (IFC). Moretti et al 2020 use IfcSharedFacilitiesElements for the integration of as-built data (Moretti et al. 2020). Also solutions for as-built-schedules are presented by using an integrated entity information model based on ID of attributes (Fagiar et al. 2023).
For data visualization especially site photography is mentioned. Therefore, workflows were developed to efficient processing of unordered photos (Jadidi et al. 2015). This case study as well as the majority of the other case studies emphasizes that humans need to be trained and that humans are a central part for generating as-built models successfully. Furthermore, the analysis shows, that especially the geometric aspect is in the foreground.

Interoperability and Data Exchange standardization

This case study focusses on open exchange standards. Therefore, it is necessary to analyze the Model View Definitions (MVD) that are defined in the IFC standard. Furthermore, the state of the art of the Information Delivery Specifications (IDS) need the be analyzed.
The MVD is used to define a subset of the IFC data and defines exchange requirements (buildingSMART international 2024b). An IDS is a document, that can be interpretated by a computer and defines how information are delivered in a project. It hereby combines IFC and other extensions, such as company specific properties or properties of the buildingSMART Data Dictionary (buildingSMART international 2024a). While the concept of MVD deals with the content that is relevant for software manufacturers (e.g. mapping the class hierarchy and the transfer of geometry) and is the basis for use-case based Software Certification, the concept of IDS defines the alphanumeric information content of models and is therefore the more relevant concept for users and as-built models, because it describes all data that is needed.
One important definition for the handover to FM and therefore for as-built modeling are the FM-Handover-Aquarium and the FM Basic Handover, that is defined as MVD in IFC 2x3. The MVD aims to define the handover of information from Design and Construction Phase to FM software, such as CAFM and CMMS (buildingSMART international). The MVD are based are based on the entities and Property Sets (PSet) in IFC.
IFC is a central data format for IDS related to FM. Although there is currently no separate FM domain in IFC 4.3 (this still existed in IFC 2x3), the entities, such as

- Shared Facilities Elements – IfcInventory, IfcFurniture, IfcAsset and IfcOccupant
- Shared Management Elements – IfcActionRequest, IfcOccupent, IfcOrderAction, IfcProjectOrder, IfcPermit, IfcCost, IfcCostSchedule, IfcWorkCalendar, IfcWorkPlan, IfcTask and IfcEvent.

and Property Sets, such as, PSet_ProjectOrderMaintenance-WorkOrder, PSet_Utility, PSet_Warranty or PSet_ActionRequest are integrated for FM in IFC. There are also other class-specific PSets, quantity sets and other attributes relevant to FM. The classes defined in IFC are particularly suitable for describing as-built data. These can be used to fundamentally describe the building and the associated technical systems. There are also individual classes that can map FM process data (Bartels 2020; Bartels und Wimmer 2024).

Methodology

In order to implement an interface that allows an automated filling of the missing attributes from the as-planned state to the as-built state by taking IFC and IDS into account, four steps were taken. These steps are shown in Figure 2.

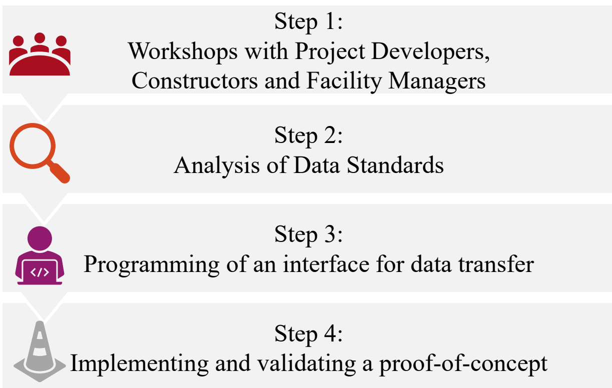


Figure 2: Methodology of this paper

First, workshops were conducted with Developers, Constructors, Planners and Facility Managers in order to get to know the needed attributes and the state-of-the-art. Secondly, data standards were analyzed to develop data specifications, with whom it is possible to filter the needed data out of documents and to transfer the needed data into an IFC-file by applying IDS. Thirdly, an interface was programmed and implemented to transfer the files.

Fourth, the programming and the process was implemented in a proof-of-concept to validate the programming, its acceptance and to analyze the results.

Data Analysis

Evaluation of the basis for information exchange

To implement an interface for transferring data and information from delivery and revisional documents to the digital building model by using IFC, the Exchange Information Requirements and Asset Information Requirements were evaluated with workshops.

Therefore, all in all 9 workshops were conducted, respectively 3 workshops with project developers and owners, 3 workshops with construction companies and 3 workshops with FM companies. The workshops took place from June 2023 to September 2023 and were conducted in person as well as online.

The workshops showed, that as-built models are more and more requested in building projects. In most of the cases, IFC is asked as exchange format for as-built models. But also proprietary formats (especially Revit) are requested.

Furthermore, the workshops showed, that there is a need for the automated reading and transferring of information out of delivery and revisional documents to the digital building model. Especially the project developers stated, that this would improve the information exchange across the whole lifecycle and would perfectly fit to their aim of having as-built data for all developed building. For the Construction companies, the linking of documents and the digital building model also offers opportunities and is viewed positively; in addition, the aspect of a possible tracking of deliveries and thus a should-is comparison is emphasized as an advantage.

Lastly, the workshops aimed to evaluate the needed information of the stakeholders. The project developers stated, that the EIR and AIR are still not finalized and change slightly in the projects. The same effects are stated by the construction and FM companies. This often causes incomplete or wrong information in the as-built models due to missing definitions of the EIR and AIR in contracts. Therefore, some of the construction and FM companies developed their own AIR and EIR. Although the AIR and EIR vary slightly, the experts stated, that the information requirements are usually based on the Construction-Operations Building information exchange (COBie) or the German standard CAFM-Connect (CC). Especially CC is used in most of the projects as basis for the Information Requirements. An analysis of these exchange formats is shown in the next section.

Relevant Exchange formats

Due to its complexity, IFC is not specific and fully suitable transferring and storing FM data throughout the lifecycle of a building (Yalcinkaya und Singh 2019). Therefore, formats have been developed to transfer FM data properly and to get as-built models. As shown in the next section, the IFC-based exchange formats CAFM-Connect and COBie are the most mentioned formats for the delivery of as-built models. The workshops showed, that the EIR and AIR are based on the attributes mentioned in these two standards.

COBie was developed in 2007 by Bill East form the United States Army Corps of Engineers (Lavy et al. 2019). The format uses spreadsheets to transfer BIM information from the Design and Construction process to FM (FM Handover). COBie defines three types of information delivery, one for information, that is created by designers, one that is created by contractors and one that is created by both of them. Hereby COBie defines information requirements for spatial and equipment assets (East 2013).

CAFM Connect is a German IFC-based initiative of various CAFM vendors. CC aims to capture the data collected in the Design and Planning phase and transfer it in a usable way into CAFM systems (Verband für die Digitalisierung im Immobilienbetrieb, CAFM RING e. V. 2019). Therefore, various lists were developed, to define all relevant data for the building objects, e. g. based on the German standard DIN 276 (Otto und Bartels 2018).

In the following, a door will be used as example for the requested data in CC. IFC defines a door with the entity IfcDoor, a subtype of IfcProduct. Various Property Sets assign to the entity IfcDoor, such as PSet_IfcDoorCommon, PSet_DoorLiningProperty. But also FM-data could be assigned by using PSet_Warranty, PSet_Manufacturer or PSet_MaintenanceStrategy (buildingSMART international 2022). These attributes are the basis for CAFM Connect and further defined in the lists. An example for the standardized attributes of a door (IfcElement IfcDoor) is shown in Table 1.

Table 1: Example of CAFM-Connect Attributes (ifcDoor)

IfcElement	DIN 276	Attribute	Type
IfcDoor	334.10	Door Stop	IfcText
IfcDoor	334.10	Number of wings	IfcText
IfcDoor	334.10	Automatic drive	Ifc Boolean
IfcDoor	334.10	Year of manufacture	IfcReal
IfcDoor	334.10	Description	IfcText
IfcDoor	334.10	DiBT approval number	IfcReal
IfcDoor	334.10	Hold-open system	IfcText
IfcDoor	334.10	Fire resistance class	IfcText
IfcDoor	334.10	Escape door	Ifc Boolean
IfcDoor	334.10	Manufacturer	IfcText
IfcDoor	334.10	Emergency exit lock	Ifc Boolean
IfcDoor	334.10	Overhead door closer	Ifc Boolean
IfcDoor	334.10	Opening type	IfcText
IfcDoor	334.10	Protection requirement	IfcText

Furthermore, CAFM Connect developed so called BIM profiles, in which relevant attribute for FM services are defined (Verband für die Digitalisierung im Immobilienbetrieb, CAFM RING e. V. 2023).

Development of an attribute list and process

Based on these data exchange formats, a consolidated list of attributes for building objects has been developed. The following Figure 3 shows the attributes of a door, all based on the global unique identifier (GUID), as example and in which phase data is actually included in the model. It furthermore shows, that various stakeholders need to enter data throughout the lifecycle. At the beginning of the project, the planners (e.g., Architects) enter various planning data. These data are needed for tendering and awarding.

Maturity Level	Source of Information	Attributes / Information
As-Planned	Digital Building Model / Door List (BIM-Software) *Planner (here: Architect)*	GUID Name Description Emergency Exit lock Protection Requirements Number of wings Opening Type Overhead door closer Fire resistance class Emergency Exit Fire Protection Door Stop Hold-open system
As-Built	Delivery Note *Construction Manager*	Manufacturer Year of Construction Serial number Automatic drive DiBT approval number
	Revision Documents *Construction Manager / Commissioning Manager*	Inspection intervall Maintenance intervall Commissioning date Qualification of the examiner
As-Operated	Contract *Facility Manager*	Contracter Fault Priority
	Service Provisioning Documentation (CAFM) *Facility Manager*	Inspection Report Result Description Responsible Person Date of Maintenance Next Maintenance Result *plus* *Building Automation System and IoT-Data*

Figure 3: Data entry throughout the lifecycle

Based on that data, the Constructor execute his services. Due to the incomplete data and changes in the construction phase, it is necessary to enter further data. During the construction phase, the Construction Manager need to enter various data to achieve an as-built model. Especially data about the manufacturer and the year of Construction as well as serial numbers need to be entered in the digital building model. Beside that data also data of the commissioning is needed to receive an as-built model. Therefore, after the Commissioning additional data need to be entered in the digital building model. These data are especially Commissioning Dates and data about the service provisioning for Facility Management.

The remaining data is entered during the O&M phase of the building by the Facility Management and includes data about the service provisioning. These data could be transferred out of the CAFM system by using IFC.

Especially the data, that need to be entered by the Constructors is a key success criterion. On the other hand, that data is often not fully entered in the digital building model. The workshops with the constructors showed, that the Construction Manager is not capable to enter all relevant data due to

1. a lack of knowledge regarding BIM, especially in entering data
2. an insufficient hard- and software equipment on the construction site,
3. especially a lack of time due to the task on the construction site
4. belayed information of the sub-contractors due to missing documents.

The numbers 1-3 could be solved by the suggested approach of this paper. Furthermore, number 4 could also be solved, if the delivery of the documents is ruled in the contracts and the delivery of documents could even better be controlled.

Development of an automated document-based as-built model

Based on the data standards and processes, an automated exchange process to receive an as-built model has been developed.

Extraction of information on delivery documents

Currently the technical capabilities of material suppliers vary substantially. Around 1 percent of suppliers can share delivery bills via an API with defined End Points that can be accessed from external parties. Approximately 5 percent of material suppliers offer customer portals, where they upload electronically generated PDFs . The majority of suppliers, around 74%, are able to send scans or electronically generated PDFs of the delivery bills to an agreed upon mailing address. Only 20% of suppliers are not able to electronically transfer the delivery receipt. The only way for them to transmit the information is to bring the physical document to the construction site or send it via post. To mitigate this problem and enable all delivery receipts to be pushed into the digital workflow a scanning application on the construction site is unavoidable.

After each document is at least available as a digital picture the next step is to retrieve the information of the scanned documents. This can already be done with common OCR applications (to transform the scan into a machine-readable string. (Memon et al. 2020; Chaudhuri et al. 2017) After all documents are in machine-readable format the varying formats need to be mapped on one central standard. To guarantee a very high data quality this mapping is currently still done manually but large language models already show to be a great way to automate this task. An extract of a delivery bill of the example project is shown in Figure 4.

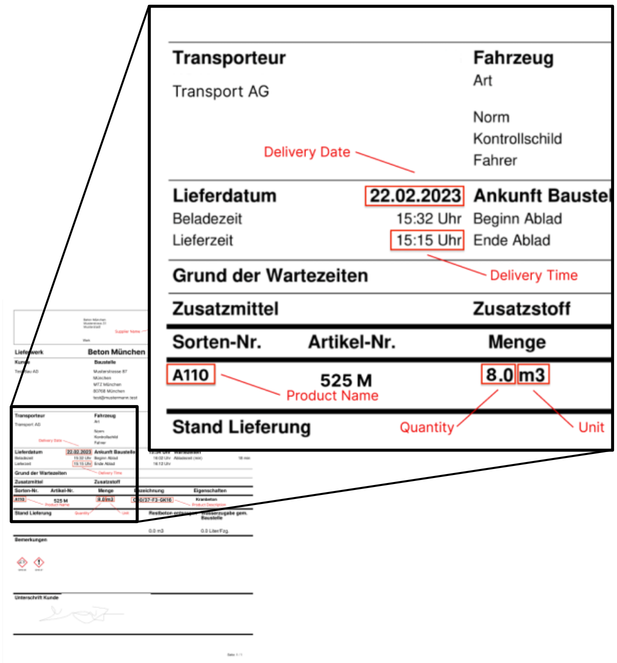


Figure 4: Extract of the exemplary delivery bill

IFC integration – Developing a process of automated as-built-documentation

The process of automatic as-built documentation on the IFC model is represented by a two-track, interlinked workflow between the construction site ("on-site") and the material dispatcher ("off-site") and is shown in Figure 5. As in the conventional workflow, the construction site is the starting point: The site manager or foreman determines the time of a material delivery. For this purpose, a task is created in a model-based construction management software, which consists of the following information: (1) time of delivery (date and time), (2) location of delivery (linked components or GUIDs), (3) product specifications (e.g., compressive strength of the concrete), and the (4) order quantity (e.g., m^3 of the delivery). While information (1) and (2) is entered individually by the ordering party, (3) can be read from the stored IFC file and thus pre-filled for the order. The construction site management would then only have to send the order information in the next step after checking the order. The order information, including the linked GUID(s), is then automatically sent via an interface (API) to the ERP system of the dispatcher, who checks the incoming order electronically and releases and picks the goods accordingly. After the final check, the physical delivery process can be started while the information flows back into the construction site management software via the updated delivery information in the ERP system. This means that the status of the delivery is visible at all times, as well as the expected time of arrival (ETA) and the underlying material information. When the delivery finally arrives and all materials have been delivered as ordered, the delivered and thus installed material information must be integrated back into the building model. There are two possible ways of doing this: using ML, as described in the section before, or sending delivery bills electronically. If the information contained in the delivery bill cannot be transferred from the dispatcher's ERP system to the construction site management software via the interface, the delivery bill would have to be scanned manually on site. The use of ML allows the automatic capture and evaluation of the information depicted on the delivery bill and thus the transfer into a digital format. By matching the previously placed digital order and uploading the delivery bill, the delivery bill can be subsequently linked to the order and thus to the underlying GUIDs. Ideally, however, the use of a paper-based delivery bill should be avoided anyway.

In this case, the delivery bill would be sent digitally via API when the physical materials are dispatched. The previously placed order is linked to the components, making it easy for the construction site team to compare them when goods are received for verification purposes.

Once the information on the delivery bills has been successfully processed electronically, the semantic interoperability of the collected information is permanently guaranteed. This is due to the storage of the information in a model-based and document-oriented NoSQL database management system. Although this database system is initially independent of the IFC standard, the export of the data from the software allows it to be converted into a wide variety of formats. Both API requests and the export as an IFC file itself access the database and convert it into the desired format with respective PSets.

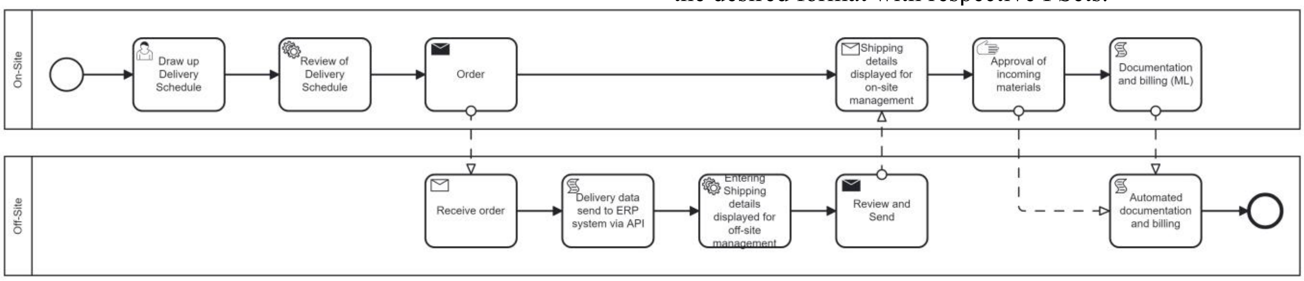


Figure 5: Process of integrating information in BIM

Critical reflection on surplus value

The process of automatic as-built documentation on the basis of model-based deliveries provides great added value, particularly for users on the construction site, but also for all other project stakeholders. The results show, that a majority of the demanded information (see previous chapters) could be filled in automatically by using OCR in combination with BIM.

The proof-of-concept also shows, that the added value arises primarily in the following four areas for the construction site: (1) order information, (2) communication, (3) documentation and (4) archiving.

Firstly, material information already contained in the model does not have to be entered again when filling in the order information. Instead, the quality and quantity can be pre-filled and only needs to be checked before the order is placed. In shell construction in particular, where measurements are still taken using analog, laborious and error-prone processes based on 2D plans, time savings of over 60% (equivalent to around 2 minutes per order) can be achieved for a typical order. Whereas previously, for example, the length, width and height of foundations were multiplied and then added together by number, the site team now obtains this information by simply clicking on the components.

Secondly, the information can subsequently be transmitted electronically via an interface to the dispatcher's ERP system. This process eliminates the need to prepare the order information again (e.g., as a PDF for sending by e-mail) or to call the material supplier. As a result, time savings of around 80% (equivalent to around 1 minutes per order) can be observed.

However, the greatest added value for the construction site comes from automated documentation. Today, manual diaries ("delivery lists"), for example in the form of an Excel spreadsheet, are created at great expense, particularly for concrete but also steel deliveries. Almost all the information typically contained there, such as the installation date, the delivery bill number, the quantity, the component or the location, could be automated using the process outlined above. What would currently require time-consuming updating could be generated and pre-filled at the touch of a button, leaving just one verification process. Time savings of over 90% (equivalent to approx. 3 minutes per order) seem plausible.

Finally, the filing of delivery bills in physical orders or the filing of scanned or digitally transmitted delivery bills in digital folders is also overdue. Delivery bills would be automatically linked to the IFC model and made available to everyone at any time via an interface to Common Data Environments or billing software, without the need for manual intervention. Time savings of almost 100% (equivalent to approx. 1 minutes per order) can therefore be realized.

For the construction site team alone, this results in a time saving of 7 minutes per order and documentation process. For a medium-sized construction project (roughly 5 Mio EUR construction volume) with approx. 300 delivery bills per months, resulting in roughly 10,000 orders, this would mean 1,170 hours in total. Additional time and cost savings can be expected from the reduction in human error when planning and rescheduling orders, as well as from automatic version management and the resulting traceability of changes. Finally, there are possibilities for data analysis and evaluation, e.g., with regard to target/actual materials or CO_2 values, which were previously unheard of or only possible with great effort, if a material database is connected accordingly.

Conclusion

This paper presented an automated approach to generate as-built models by using the BIM method in combination with OCR. It hereby combines an IFC-based model with information out of delivery documents. A process has been developed and implemented in a proof-of-concept.

By doing so, the efficiency of the information exchange and maintenance has been increased in the proof-of-concept. Furthermore, the automated exchange of information – based on IFC – lead to a more precise delivery of the information, that were evaluated in the workshops with project developers and facility managers.

This paper is a first approach for automated as-built models, that was evaluated in practice. the PoC shows that the process outlined in this paper can be used to automate data enrichment from BIM models to as-built models, further research is needed. The PoC shows in particular, that not all relevant information for FM could be extracted of the delivery documents. That means, that further documents – especially the revisional documents – need to be integrated in the process. In addition to that, the workshops with the experts showed, that not all companies are using digital delivery documents, but some still have delivery bills in paper form. This makes it necessary to scan or photograph the delivery documents. One solution is to use other internal systems (such as ERP).

All in all, this paper shows, that the manual integration of information could be reduced to a minimum by using OCR or related systems. Furthermore, it shows, that all relevant software systems, such as ERP-systems, need to be connected to the digital building model in order to fill in information automatically. Furthermore, PSets for IFC should be developed to generate as-built models out of delivery bills.

References

Asmar, Daniel; Daher, Rema; Hawari, Yasmine; Khoury, Hiam; Elhajj, Imad H. (2022): Automated building and evaluation of 2D as-built floor plans. In: *Machine Vision and Applications* 33 (3). DOI: 10.1007/s00138-022-01289-8.

Baranova, Olga (2021): Open data formats in building information modeling. In: *E3S Web Conf.* 263, S. 4062. DOI: 10.1051/e3sconf/202126304062.

Bartels, Niels (2020): Strukturmodell zum Datenaustausch im Facility Management. Wiesbaden: Springer Fachmedien.

Bartels, Niels; Wimmer, Reinhard (2024): BIM von A bis Z etablieren, inklusive „Tuning“-Maßnahmen. Teil 4: Schnittstelle zum Facility Management (4), S. 36–39.

Borrmann, André; König, Markus; Koch, Christian; Beetz, Jakob (2021): Building Information Modeling. Wiesbaden: Springer Fachmedien Wiesbaden. buildingSMART international (Hg.): Basic FM Handover View. Online verfügbar unter https://technical.buildingsmart.org/wp-content/uploads/2021/01/GSC-001%20Basic%20HandOver%20to%20Facility %20Ma-nagement.zip, zuletzt geprüft am 01.04.2024.

buildingSMART international (Hg.) (2022): IfcDoor. Online verfügbar unter https://ifc43-docs.standards.buildingsmart.org/IFC/RELEASE/IFC4x3/HTML/lexical/IfcDoor.htm, zuletzt geprüft am 11.10.2023.

buildingSMART international (Hg.) (2024a): Information Delivery Specification IDS. Online verfügbar unter https://technical.buildingsmart.org/projects/information-delivery-specification-ids/, zuletzt geprüft am 01.04.2024.

buildingSMART international (Hg.) (2024b): Model View Definition (MVD) - An Introduction. Online verfügbar unter https://technical.buildingsmart.org/standards/ifc/mvd/, zuletzt geprüft am 01.04.2024.

Chaudhuri, Arindam; Mandaviya, Krupa; Badelia, Pratixa; Ghosh, Soumya K. (2017): Optical Character Recognition Systems. In: Arindam Chaudhuri, Krupa Mandaviya, Pratixa Badelia und Soumya K Ghosh (Hg.): Optical Character Recognition Systems for Different Languages with Soft Computing, Bd. 352. Cham:
Springer International Publishing (Studies in Fuzziness and Soft Computing), S. 9–41.

Davenport, Thomas (1998): Putting the Enterprise into the Enterprise System. Analytics And Data Science. Hg. v. Harvard Business Review. Online verfügbar unter https://hbr.org/1998/07/putting-the-enterprise-into-the-enterprise-system, zuletzt geprüft am 27.11.2023.

Deruyter, Greta; Hennau, Marc; De Wolf, Vicky; Dewulf, Niek: Approach for comparing design and as build models based on data acquisition using a 3D terres-trial laser scanner, a case study. In: Proceedings of the 4th International Workshop on 3D Geo-Information, Ghent University, 4-5 November, 2009, Ghent, S. 2–14.

Di Filippo, Andrea; Cotella, Victoria Andrea; Guida, Caterina Gabriella; Molina, Victoria; Centarti, Lorena (2021): BIM Interoperability and Data Exchange Support for As-Built Facility Management. In: Osvaldo Gervasi, Beniamino Murgante, Sanjay Misra, Chiara Garau, Ivan Blečić, David Taniar et al. (Hg.): Computa-tional Science and Its Applications – ICCSA 2021,
Bd. 12950. Cham: Springer International Publishing (Lecture Notes in Computer Science), S. 702–711.

East, Bill (2013): Using COBie. In: Paul Teicholz (Hg.): BIM for Facility Managers: Wiley, S. 107–143.

Ensafi, Mahnaz; Harode, Ashit; Thabet, Walid (2022): Developing systems-centric as-built BIMs to support fa-cility emergency management: A case study approach. In: *Automation in Construction* 133, S. 104003. DOI: 10.1016/j.autcon.2021.104003.

Fagiar, Muaz; Mohamed, Yasser; AbouRizk, Simaan (2023): Simulation-Assisted Project Data Integration for Development and Analysis of As-Built Schedules. In: *Buildings* 13 (4), S. 974. DOI: 10.3390/buildings13040974.

Freimuth, Henk; König, Markus (2019): A toolchain for automated acquisition and processing of as-built data with autonomous UAVs. In: Proceedings of the 2019 European Conference on Computing in Construction. 2019 European Conference on Computing in Construc-tion, Jul. 10, 2019: University College Dublin (Compu-ting in Construction), S. 9–18.

Hausmann, Maik; Lücke, Michael; Lechtenberg, Simon; Guhlemann, Kerstin; Best, Christina (2022): Entwicklung eines Geschäfts- und Betreibermodells für digitale Plattformen im Handwerk am Beispiel von Athene. In: Andreas Ehlert und Hans Jörg Hennecke (Hg.): Perso-nenzentrierte Digitalisierung zur Zukunftssicherung des Handwerks. Ergebnisse des Forschungsprojekts „Athene 4.0“. Düsseldorf: Handwerk NRW w.V., S. 26–34.

Hellenborn, Benjamin; Eliasson, Oscar; Yitmen, Ibrahim; Sadri, Habib (2023): Asset information requirements for blockchain-based digital twins: a data-driven predictive analytics perspective. In: *SASBE*. DOI: 10.1108/SASBE-08-2022-0183.

ISO 16739-1, 2018: Industry Foundation Classes (IFC) for data sharing in the construction and facility management industries.

Jadidi, Hossein; Ravanshadnia, Mehdi; Hosseinalipour, Mujtaba; Rahmani, Fatemeh (2015): A step-by-step con-struction site photography procedure to enhance the effi-ciency of as-built data visualization: a case study. In: *Vis. in Eng.* 3 (1). DOI: 10.1186/s40327-014 -0016-9.

Khan, M. Arsalan (2023): Integrating BIM with ERP Systems Towards an Integrated Multi-user Interactive Database: Reverse-BIM Approach. In: M. S. Ranadive, Bibhuti Bhusan Das, Yusuf A. Mehta und Rishi Gupta (Hg.): Recent Trends in Construction Technology and Management, Bd. 260. Singapore: Springer Nature Sin-gapore (Lecture Notes in Civil Engineering), S. 209– 220.

Klein, Laura; Li, Nan; Becerik-Gerber, Burcin (2012): Imaged-based verification of as-built documentation of operational buildings. In: *Automation in Construction* 21, S. 161–171. DOI: 10.1016/j.autcon.2011.05.023.

Lauer, Thomas (2021): Wandel als Folge von Digitalisierung. In: Thomas Lauer (Hg.): Quick Guide Change Management für alle Fälle. Berlin, Heidelberg: Springer Berlin Heidelberg (Quick Guide), S. 105–123.

Lavy, Sarel; Saxena, Nishaant; Dixit, Manish (2019): Effects of BIM and COBie Database Facility Management on Work Order Processing Times: Case Study. In: *J. Perform. Constr. Facil.* 33 (6), Artikel 04019069. DOI: 10.1061/(ASCE)CF.1943-5509.0001333.

Matarneh, Sandra T.; Danso-Amoako, Mark; Al-Bizri, Salam; Gaterell, Mark; Matarneh, Rana T. (2019): BIM for FM. In: *F* 38 (5/6), S. 378–394. DOI: 10.1108/F-07-2018-0084.

Memon, Jamshed; Sami, Maira; Khan, Rizwan Ahmed; Uddin, Mueen (2020): Handwritten Optical Character Recognition (OCR): A Comprehensive Systematic Liter-ature Review (SLR). In: *IEEE Access* 8, S. 142642– 142668. DOI: 10.1109/ACCESS.2020.3012542.

Moretti, Nicola; Xie, Xiang; Merino, Jorge; Brazauskas, Justas; Parlikad, Ajith Kumar (2020): An openBIM Ap-proach to IoT Integration with Incomplete As-Built Data. In: *Applied Sciences* 10 (22), S. 8287. DOI: 10.3390/app10228287.

Moselhi, Osama; Bardareh, Hassan; Zhu, Zhenhua (2020): Automated Data Acquisition in Construction with Remote Sensing Technologies. In: *Applied Sciences* 10 (8), S. 2846. DOI: 10.3390/app10082846.

Omar, Tarek; Nehdi, Moncef L. (2016): Data acquisition technologies for construction progress tracking. In: *Automation in Construction* 70, S. 143–155. DOI: 10.1016/j.autcon.2016.06.016.

Otto, Jens; Bartels, Niels (2018): Integration von FM-Prozessdaten in ein digitales Gebäudemodell. In: *Bautechnik* 95 (12), S. 823–831. DOI: 10.1002/bate.201800044.

Park, Jaehyun; Cai, Hubo (2017): WBS-based dynamic multi-dimensional BIM database for total construction as-built documentation. In: *Automation in Construction* 77, S. 15–23. DOI: 10.1016/j.autcon.2017.01.021.

Pătrăucean, Viorica; Armeni, Iro; Nahangi, Mohammad; Yeung, Jamie; Brilakis, Ioannis; Haas, Carl (2015): State of research in automatic as-built modelling. In: *Advanced Engineering Informatics* 29 (2), S. 162–171. DOI: 10.1016/j.aei.2015.01.001.

Pilling, André (2016): BIM - Das digitale Miteinander. 1. Auflage. Berlin, Wien, Zürich: Beuth Verlag GmbH (DIN).

RealFM (2021): BIM2FM - Leitfaden zur Anwendung der BIM-Methodik in der Betriebs- und Nutzungsphase. Mit Handlungsempfehlungen für Betreiber und Nutzer zum Informationsfluss in allen Lebenszyklusphasen von Bauwerken. 1. Auflage. Berlin: RealFM.

Son, Hyojoo; Bosché, Frédéric; Kim, Changwan (2015): As-built data acquisition and its use in production moni-toring and automated layout of civil infrastructure: A survey. In: *Advanced Engineering Informatics* 29 (2), S. 172–183. DOI: 10.1016/j.aei.2015.01.009.

Son, Hyojoo; Kim, Changwan; Kwon Cho, Yong (2017): Automated Schedule Updates Using As-Built Data and a 4D Building Information Model. In: *J. Manage. Eng.* 33 (4), Artikel 04017012. DOI: 10.1061/(ASCE)ME.1943-5479.0000528.

Su, Y. Y.; Hashash, Y. M. A.; Liu, L. Y. (2006): Integration of Construction As-Built Data Via Laser Scanning with Geotechnical Monitoring of Urban Excavation. In: *J. Constr. Eng. Manage.* 132 (12), S. 1234–1241. DOI: 10.1061/(ASCE)0733-9364(2006)132:12(1234).

Verband für die Digitalisierung im Immobilienbetrieb, CAFM RING e. V. (Hg.) (2019): CAFM Connect. Nahtloser Übergang der Gebäudedaten von Planung und Bau in den Betrieb. Online verfügbar unter https://cafm-connect.org/, zuletzt geprüft am 09.10.2023.

Verband für die Digitalisierung im Immobilienbetrieb, CAFM RING e. V. (Hg.) (2023): BIM-Profile. Online verfügbar unter http://cafm-connect.org/bim-profile/, zuletzt geprüft am 09.10.2023.

Yalcinkaya, Mehmet; Singh, Vishal (2019): VisualCOBie for facilities management. In: *F* 37 (7/8), S. 502–524. DOI: 10.1108/F-01-2018-0011.

MODELING AND ANALYSIS OF JOINT WORK PACKAGE SIZING AND PROJECT SCHEDULING CONSIDERING RESOURCE CONSTRAINTS

Nianmin Zhang[1] and Xiao Li[1]
[1]Department of Civil Engineering, The University of Hong Kong SAR, China

Abstract

Effective work package sizing and project scheduling are essential problems in construction project management. In current studies, they are often treated as separate optimization problems. However, in practical situations where resources are limited, there is a growing recognition of the need to integrate these two processes to achieve more efficient and effective project management. This research aims to fill this gap by developing an integer non-linear programming model incorporating work package sizing and project scheduling while considering resource limitations. To evaluate the effectiveness of our model, we utilized Gurobi optimization software to solve nine project instances obtained from RanGen, each characterized by its unique complexities. Furthermore, we analyzed the impact of resource utilization on work package sizing and the project schedule, and we found that both excessive and inadequate resource availability can increase project costs.

Introduction

Effective project scheduling is a crucial problem in project management, as highlighted by several studies (Li et al., 2018; Tsz Wai et al., 2023). This problem becomes even more challenging in a resource-constrained environment, where project tasks exhibit interdependencies in two key aspects. Firstly, tasks compete for limited resources, and sufficient availability of resources is essential for initiating and executing construction tasks. Secondly, the project workflow imposes precedence constraints, whereby certain tasks must be performed based on the completion of specific scheduled tasks. These two interdependencies are encompassed within the resource-constrained project scheduling problem (RCPSP), which is widely recognized and formalized in the project scheduling literature (Hartmann and Briskorn, 2022).

From a perspective of scheduling scope and complexity, RCPSP can be categorized into two groups. The first group is the resource-constrained single project scheduling problem (RCSPSP), which aims to optimize the schedule and resource allocation for a single project to ensure timely project completion. In RCSPSP, the scheduling process is task-based, and each task to be scheduled has specified work content, such as processing duration and resource requirement, as well as an execution scope, such as precedence relationships among tasks. The second group is the resource-constrained multiple project scheduling problem (RCMPSP), which involves coordinating and prioritizing schedules across multiple projects that share common resources. The objective is to balance the workload across projects, ensure the timely completion of each individual project, and achieve the overall goals of the project portfolio (Issa and Tu, 2020).

With the implementation of the work breakdown structure (WBS) method in the construction industry, the RCPSP for construction projects has evolved into a work package-based scheduling problem, which can be considered as a complex extension of the RCPSP (Li et al., 2022). For example, the production of a modular room can be decomposed hierarchically into six work packages, including structure production, door/window production, wall production, print, electric service, and test, using WBS tools for planning or scheduling (Liu et al., 2023). Each work package typically consists of one or more elemental tasks that are not further decomposed (Li and Hall, 2019). Specifically, the work package for structure production may include elemental tasks such as board production, 2D assembly, and concrete curing. Therefore, in construction projects, not only should the resource and precedence constraints be considered in the RCPSP, but the size of the work package also has a significant impact on project performance. Specifically, smaller work packages tend to increase project complexity and workload while diminishing economies of scale. Conversely, larger work packages require a higher allocation of resources, potentially straining resource availability and leading to bottlenecks or delays. Furthermore, as demonstrated by Blazewicz et al. (1983), the RCPSP belongs to the category of strongly NP-hard problems, presenting a significant challenge in obtaining the optimal solution for RCPSP and acting as a bottleneck for further improvements in project performance. By considering a more comprehensive set of influential factors in the RCPSP, there is potential to enhance project management capabilities and improve the efficiency of resource utilization.

Work package sizing has garnered considerable attention, with studies demonstrating the significance of optimizing work package size to reduce project costs (Li and Hall, 2019; Li et al., 2021b). This optimization encompasses various factors, including minimizing costs associated with inaccurate estimation, monitoring, and control of work packages. Moreover, the sizing of work packages has implications for the project's economies of scale and discounted cash flow, further emphasizing the importance of this optimization in cost reduction endeavors (Liu et al., 2023).

While the optimization of work package sizing has enabled a trade-off among multiple objectives, there are still limitations in current studies that hinder the improvement of project performance. One limitation is that the work package sizing process often takes place during the initial planning stage of a construction project, without considering detailed task information such as task durations, resource requirements, and precedence relations among tasks (Li et al., 2023). However, resource constraints are common in construction projects and significantly impact the feasibility and effectiveness of work packages (Li et al., 2021a). This can lead to the creation of unsatisfactory project schedules during the project scheduling stage, ultimately resulting in poor project performance. Moreover, the work packages in RCPSP are predetermined and do not allow for splitting or recombination, limiting the flexibility of project managers to schedule work packages and effectively utilize resources in the project scheduling stage (Du et al., 2021).

To summarize, work package sizing and construction project scheduling are interdependent activities that should be considered simultaneously to mitigate project delays, cost overruns, and disruptions, particularly in resource-constrained and uncertain project scenarios. Work package sizing involves grouping tasks into cohesive work packages, which determine the resource requirements and processing time. These factors then influence scheduling decisions and the overall project cost. Project scheduling decisions, in turn, have a significant impact on resource availability, which affects the feasibility of work package splitting and recombination (Pellerin et al., 2020). By integrating work package sizing and project scheduling, project stakeholders can better manage resources, optimize project timelines, and mitigate potential bottlenecks or resource conflicts (Zaman et al., 2021).

However, most studies treat work package sizing and project scheduling as separate problems, neglecting the interrelationship between them in resource-constrained construction projects. To the best of our knowledge, no research to date provides an integrated resource-constrained project scheduling and work package sizing method while considering the interrelationship between them. Therefore, it is crucial to propose a new integrated method that combines work package sizing and project scheduling under resource and precedence constraints to reduce project costs while improving resource efficiency (Zhang et al., 2024).

To address the existing research gap, we present a joint model for project scheduling and work package sizing that takes into account both precedence and resource constraints. Our proposed approach makes the following contributions: 1) For the first time, we introduce an integrated model that considers the interplay between project scheduling and work package sizing. 2) To validate the effectiveness of our model, we conducted various experiments using Gurobi with nine project instances obtained from RanGen. 3) We analyze the impact of resources on project schedule and work package size by changing availability.

Methodology

Assumptions

- We assume that the project has been planned into a given set of tasks that are elemental. That is, no further breakdown of those tasks is possible for reasons that may be technical, logistical, financial, administrative, or cultural.
- Only the active tasks can be grouped into a work package due to the specific role in the construction project.
- Considering the mass production in real scenarios (e.g., reduction in set-up time, material preparation, mass workforce, and new production tech), lag time is considered among tasks that are grouped into the same package, while the package-to-package is still regarded to be hard precedence-constrained.

Problem Modeling

Consider an task-on-node (TON) project network described by $G(N,A,W)$, where N denotes a set of tasks in the project ($N = \{0,1,2,\ldots,\bar{n},\bar{n}+1,...,\bar{n}+m,\bar{n}+m+1\}$). Each activity has a given work duration $d_i \geq 0, i \in N$ and given work content $x_i \geq 0, i \in N$. We assume that there are K types of renewable resources, and each resource $k \in \{1,2,3,...,K\}$ has a limited capacity a_k available. When a task i is executed, its requirement for resource type $k \in \{1,...,K\}$ at each time period is a fixed integer $r_{i,k} \geq 0$. Tasks 0 and $\bar{n}+m+1$ are the dummy start and end of the project ($(Task_0, Task_{\bar{n}+m+1}) \in N_{Dummy}$), whose process duration, work content and required resource are both 0. More specifically, let $N_A = \{1,...,\bar{n}\}$ denotes the active tasks that can be grouped into work packages, while $N_I = \{\bar{n}+1,...,\bar{n}+m\}$ is the inactive task set, which can only be regarded as single work package due to several fac-tors, such as the responsibility assignment, internal cohe-sion or risk control (Li and Hall, 2019); thus we can get that $N = N_A \cup N_I \cup N_{Dummy}$. A represents a set of precedence relationships among tasks, ($A \subseteq N \times N$), if $(i, j) \in A$ that means $Task_j$ cannot start before $Task_i$ is finished. W is the set of weights on the arcs that denotes the minimum time lags between tasks grouped into a single work package, $W = \{\delta_{i,j} \geq 0 | (i, j) \in A\}$. When $s_i - s_j > \delta_{i,j}, (i, j) \in A$ is satisfied, the $Task_i$ can be conducted, where s_i and s_j are the start times of tasks i and j,respectively. More specifi-cally, (1) if $\delta_{i,j} = 0$, then tasks i and j can be completely overlapped, i.e. they can be executed in parallel; (2) if $d_i \geq \delta_{i,j}$, then tasks i and j can be partially overlapped, i.e. activity j can only begin after activity i has been in ex-ecution for at least $\delta_{i,j}$ time units; (3) if $\delta_{i,j} \geq d_i$, then tasks i and j are executed sequentially, i.e. activity j can only begin after activity i is completed. Figure 1 presents an example illustrating the above three cases of activity overlap, namely, complete overlap, partial overlap, and no overlap with respect to two tasks. These variables and additional notation are summarized in Table 1.

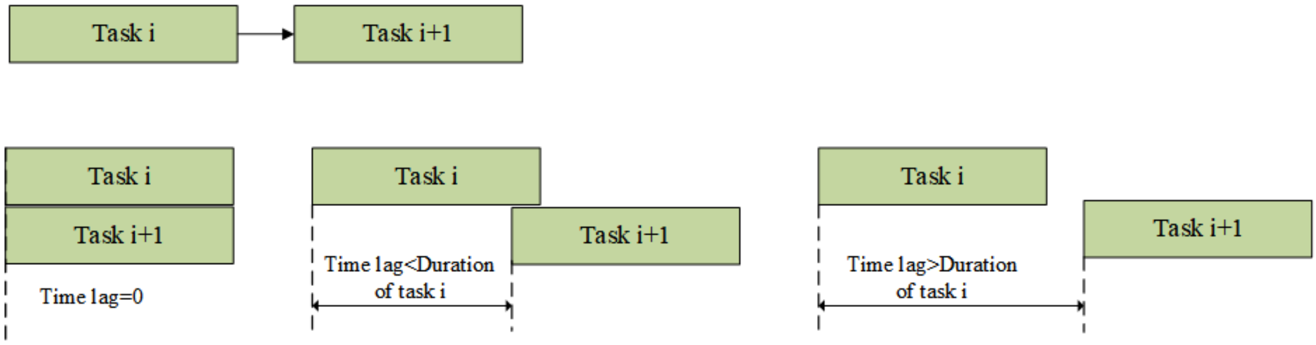


Figure 1: The lag time of the tasks.

The Objective Function

In this study, we adopted a combination of objectives (minimize the total project makespan and project cost) to optimize the work package sizing and work package scheduling in a resource-constrained environment (Li and Hall, 2019). Assuming that all the tasks are ready to be allocated to a total of $\bar{n}+m$ work packages, the objective function can be expressed as follows:

$$Minimize \qquad J = \lambda F_t + (1-\lambda)F_c, \tag{1}$$

$$F_c = \omega(p+m) + \sum_{j=1}^{\bar{n}+m} F(X_{w_j}) + \xi \sum_{j=1}^{\bar{n}+m} X_{w_j}(1-e^{-\alpha C_{w_j}}), \tag{2}$$

$$F_t = Max(C_{w_j}), \qquad \forall j \in \{1,2,...,\bar{n},...\bar{n}+m\} \tag{3}$$

$$F(X_{w_j}) = f(X_{w_j}) + g(X_{w_j}) + h(X_{w_j}), \quad \forall j \in \{1,2,...,\bar{n},...\bar{n}+m\} \tag{4}$$

where F_c and F_t are the project cost and makespan, and λ is the weight of the time part. For the most part, $p \in \{\bar{p}, \bar{p}+1,...,\bar{n}+m\}$ is the number of work packages that contain at least one task, where $\bar{p}$ is the lower bound for any feasible solutions that can be derived from the theory of critical path. ω is the fixed cost of the work package and X_{w_j} is the j_{th} work package's work content. C_{w_j} is the complete time of the j_{th} work package. We model the cost of inaccuracy in time estimation of a work package as function $f(X_{w_j})$, the cost for monitoring and controlling the progress of the work package as function $g(X_{w_j})$ and the economies of scale from repetition and similarity of tasks within a work package as function $h(X_{w_j})$, which are functions of the work packages' work content. Additionally, work package sizing poses impacts on cash flow, which is modeled as $\xi \sum_{j=1}^{\bar{n}+m} X_{w_j}(1-e^{-\alpha C_{w_j}})$, where ξ and α are parameters for cash flow cost function.

Work Package Sizing Constraints

work package sizing is conducted based on the structure of the given activity network and relies on the requirement that the network of work packages cannot contain cycles (Lambrechts et al., 2008). Additionally, the active tasks need to be allocated to the work package $P_A \in \{1,2,3,...,\bar{n}\}$, while the inactive tasks should be allocated to $P_I \in \{\bar{n}, \bar{n}+1,...,\bar{n}+m\}$. $b_{i,j}$ is the decision variable, which is 1 if activity $i \in N$ is allocated to work package $j \in P_A \cap P_I$, and 0 otherwise. The work package sizing constraints are as follows:

$$b_{i,j} + b_{i',j} \le 1, \qquad \forall i,i' \in N_A, \quad \forall i' \in \bigcup_{i'' \in Succ(i) \cap N_I} Succ(i''), \forall j \in \{1,2,3...,\bar{n}\} \tag{5}$$

$$b_{i,j} + b_{i',j} \le 1, \qquad \forall i,i' \in N_A, \quad \forall i' \in \bigcup_{i'' \in Pre(i) \cap N_I} Pre(i''), \quad \forall j \in \{1,2,3...,\bar{n}\} \tag{6}$$

$$b_{i',j} b_{i'',j} \le b_{i,j}, \qquad \forall i,i',i'' \in N_A, \quad \forall i' \in Succ(i), \quad \forall i'' \in Pre(i), \quad \forall j,j' \in \{1,...,\bar{n}\} \tag{7}$$

$$\sum_{j \in P_A} b_{i,j} = 1, \qquad \forall i \in N_A \tag{8}$$

$$\sum_{j \in P_I} b_{i,j} = 1, \qquad \forall i \in N_I \tag{9}$$

$$\sum_{i \in N_I} b_{i,j} = 1, \qquad \forall j \in P_I \tag{10}$$

$$z_j = max\{b_{i,j}\}, \qquad \forall j \in P_A, \quad \forall i \in N_A \tag{11}$$

$$p = \sum_{j \in P_A} z_j, \tag{12}$$

$$X_{w_j} = \sum_{i \in N} b_{i,j} x_i, \qquad \forall j \in \{1,...,\bar{n}+m\} \tag{13}$$

$$r_{w_j,k} = \sum_{i \in N} b_{i,j} r_{i,k}, \qquad \forall j \in \{1,...,\bar{n}+m\}, \quad \forall k \in K \tag{14}$$

$$b_{i,j} \in \{0,1\}, \qquad \forall i \in N, \quad \forall j \in \{1,...,\bar{n}+m\} \tag{15}$$

$$z_j \in \{0,1\}, \qquad \{\forall j \in 1,2,...,\bar{n}+m\} \tag{16}$$

where constraints (5)-(6) refer to that if an inactive task i'' is a predecessor/successor of activity i, then no predecessor/successor of task i'' should be grouped into the same work package with activity i. Constrain (7) shows that any pairs of a predecessor and a successor activity of active activity i are not allowed to be grouped into a single work package while without activity i. Otherwise, there will be a cycle in the work packages network. constraints (8)-(10) represent that an active/inactive activity can only be allocated to an active/inactive work package, and each inactive work package only contains a single inactive task. We il-

Table 1: Nomenclature

Notaion	Explanation
N	The set of project tasks.
N_A	The set of active tasks.
N_I	The set of inactive tasks.
N_{Dummy}	The set of dummy tasks.
K	The set of resource types.
T	The set of time.
P_A	The set of work packages for active tasks,$P_A = \{1,2,3,...,\bar{n}\}$.
P_I	The set of work packages for inactive packages, $P_I = \{\bar{n}+1, \bar{n}+2, \bar{n}+3, ..., \bar{n}+m\}$.
d_i	Duration of task $i \in N$.
$r_{i,k}$	The quantity of required resource $k \in K$ for task $i \in N$.
$\delta(i', i)$	Time lag between task $i \in N_A$ and task $i' \in N_A$.
a_k	The initial quantity of resource $k \in K$.
c_k	The cost of each unit of resource $k \in K$.
x_i	Work content of task $i \in N$.
$b_{i,j}$	1 if task $i \in N_A$ is allocated to active work package $j \in P_A$; 0 otherwise.
$q_{i,t}$	1 if task $i \in N$ starts at time t; 0 otherwise.
$a_{j,t}$	1 if work package $j \in P_A \cap P_I$ is processed at time t; 0 otherwise.
z_j	1 if active work package $j \in P_A$ is not empty; 0 otherwise.

lustrate the relationship between the active tasks and the active work packages in Constraint (11) and Constraint (16). If some active tasks are grouped into active work package j, z_j is 1, which means that the active work package is not empty, and 0 otherwise. Constraint (12)-(14) shows that we can get p effective active work packages that are not empty, and the work content and the resource requirement of these work packages are the sum of the tasks in the corresponding work package. Constrain (15) shows that $b_{i,j}$ is a binary decision variable.

Work Package Scheduling constraints

Assuming that the project is conducted within $T \in (T_{lower}, T_{upper})$ time units, where T_{lower} and T_{upper} are the lower bound and upper bound of a project. They can be estimated by the project's activity graph. Taking into account the presence of precedence and resource constraints both among work packages and tasks within the same work package, we outline the following work package scheduling constraints:

$$\sum_{t \in T} t q_{i,t} = s_i, \qquad \forall i \in N \tag{17}$$

$$\sum_{t \in T} q_{i,t} = 1, \qquad \forall i \in N \tag{18}$$

$$s_i \geq b_{i',j} C_{w_j} (1 - b_{i,j}), \qquad \forall i \in N, \quad \forall j \in P_A \cup P_I, \quad \forall i' \in Pre(i) \tag{19}$$

$$s_i \geq b_{i,j} b_{i',j} (s_{i'} + \delta(i', i)), \qquad \forall i \in N_A, \quad \forall j \in P_A, \quad \forall i' \in Pre(i) \cap N_A \tag{20}$$

$$c_i = d_i + s_i \qquad \forall i \in N, \tag{21}$$

$$S_{w_j} = min\{s_i b_{i,j}\}, \qquad \forall i \in N, \quad \forall j \in \{1, ..., \bar{n}+m\} \tag{22}$$

$$C_{w_j} = max\{c_i b_{i,j}\}, \qquad \forall i \in N, \quad \forall j \in \{1, ..., \bar{n}+m\} \tag{23}$$

$$a_{j,t} = \sum_{i \in N_I} \sum_{t'=t-d_i+1}^{t} b_{i,j} q_{i,t'}, \qquad \forall j \in P_I \tag{24}$$

$$\sum_{i \in N_A} \sum_{t'=t-d_i+1}^{t} b_{i,j} q_{i,t'} - M a_{j,t} \leq 0, \qquad \forall j \in P_A \tag{25}$$

$$a_k \geq \sum_{j=1}^{\bar{n}+m} r_{w_j,k} a_{j,t}. \qquad \forall t \in T \tag{26}$$

$$q_{i,t} \in \{0,1\}, \qquad \forall i \in N \tag{27}$$

$$a_{j,t} \in \{0,1\}, \qquad \forall j \in P_A \cap P_{I]} \quad \forall t \in T \tag{28}$$

where the variables s_i and c_i represent the starting time and completion time of activity i, while S_{w_j} and C_{w_j} represent the starting time and completion time of work package j, respectively. The decision variable $q_{i,t}$ in constrain (17) corresponds to the start processing time of the i_{th} activity, taking a value of 1 if activity i starts at time t, and 0 otherwise. Constraint (17) indicates that task execution is non-preemptive. Once a task starts at time t, it will continue without interruption until the task is complete. Constraints (19)-(20) enforce precedence constraints on the start time of task i. More specifically, constraint (19) ensures that task i can only begin after all the predecessor work packages have been completed. Constraint (20) specifies that the start time of activity i within work package j must not exceed the start time of the preceding activity within the same work package, and any required lag time must be satisfied prior to execution. Constraints (22)-(23) indicate that the starting and completion times of the work package are determined based on the earliest and latest starting and completion times of the tasks within that work package. Specifically, constraint (22) specifies that the starting time of the work package cannot be later than the earliest starting time among its tasks. On the other hand, constraint (23) ensures that the completion time of the work package is not earlier than the latest completion time among its tasks. After that, constraints (24)-(26) give the resource constraints for the work packages. $a_{j,t}$ is 1 if any tasks in work package w_j is processing at time t, 0 otherwise. Constrain (24) gives $a_{j,t}$ for inactive work packages while constrain (25) gives it for active work packages using the big-M method. Constraint (26) illustrates that the amount of resources $\sum_{j=1}^{\bar{n}+m} r_{w_j,k} a_{j,t}$ needed by all the executing work packages at any time t should not exceed the current amount of available resources a_k.

Experiments and Analysis

This section validates the effectiveness of the proposed work package sizing and scheduling model with nine projects from Rangen (Demeulemeester et al., 2003; Vanhoucke et al., 2008). These nine projects are categorized into three types, each consisting of a different number of tasks: 10, 20, and 30, respectively. The experiments are conducted using Gurobi 11.0 optimization software and executed on a PC platform running the Windows 11 64-bit operating system, Intel i9-13900KF processor.

Validation of Our Model

To obtain the nine projects, we first generate the task networks using RanGen2 with the following network measures:

- Network size indicator I_1: the number of the tasks including the active tasks and inactive tasks.
- Serial or parallel indicator I_2: $I_2 = (\eta - 1)/(I_1 - 1)$, where η is the depth of the network, and $0 \leq I_2 \leq 1$.
- Resource factor RF: an indicator of average resource availability

$$\mathrm{RF} = \frac{1}{nK} \sum_{i=1}^{n} \sum_{k=1}^{K} \begin{cases} 1, & \text{if } r_{ik} > 0 \\ 0, & \text{otherwise} \end{cases}$$

- Resource strength (RS): an indicator of resource availability at the peak time, where $r_k^{\min} = max\{r_{i,k}\}, \forall i \in N$ and $r^{\max}$ is the peak requirement of resource $k \in K$ in a schedule with the precedence constraints

$$\mathrm{RS}_k = \frac{a_k - r_k^{\min}}{r_k^{\max} - r_k^{\min}} \quad (29)$$

If the value of I_2 approaches 1, the generated network exhibits characteristics similar to a serial network, where tasks are sequentially executed. Conversely, if I_2 approaches 0, the network resembles a parallel network, where all tasks can be executed simultaneously. RF represents the typical distribution of resource types that are requested by each activity, while RS serves as a metric for assessing the level of complexity involved in scheduling a project that is constrained by limited resources. As for the remaining network measures of RanGen, we utilize their default settings to ensure consistency and comparability in our analysis. We have defined the following parameters for the three groups of project instances: I_i is set to $\{10, 20, 30\}$, I_2 is set to $\{0.8, 0.8, 0.8\}$, RF is set to $\{0.4, 0.4, 0.4\}$, RS is set to $\{0.2, 0.2, 0.2\}$ and $K = 4$. To evaluate the performance of our model, we have generated three test instances for each group, resulting in a total of nine instances. In accordance with the work by Li and Hall (2019), we have set the discount rate as $\alpha = 0.00025$, and used values of $\lambda = 0.5, \omega = 50$, $\xi = 50$, $f(x) = 3x^{0.8}$, $g(x) = h(x) = x^{1.2}$, to define our objective function. Although Gurobi is more proficient in solving linear models, our problem includes a nonlinear objective term. However, by introducing auxiliary variables, we can still utilize Gurobi to solve it.

The results of the nine projects from RanGen are presented in Table 2. In Group 1, each project instance consists of 10 tasks (nodes), with 8 active tasks and 2 inactive tasks. All experiments for this group were completed within 520 seconds, yielding objective values of 317.31, 325.29, and 304.84 for Project 1, Project 2, and Project 3, respectively. In Group 2, each instance comprises 20 tasks, with 14 active tasks and 6 inactive tasks. However, the running time for each experiment significantly increased, with Project 4, Project 5, and Project 6 taking 15,007, 24,428, and 23,674 seconds, respectively. The corresponding objective values for these projects are 516.16, 579.00, and 532.21. Projects in Group 3 consist of 21 active tasks and 9 inactive tasks, with a total of 30 tasks packaged into approximately 18 packages. Although we were able to achieve objective values of 815.70, 869.14, and 765.43 for Project 7, Project 8, and Project 10, respectively Balouka and Cohen (2021).

Overall, the results of these experiments provide compelling evidence to support the effectiveness of our modeling method that integrates work package sizing and scheduling in a resource-constrained environment. Additionally, our model takes into account the significance of striking a balance between timely completion, reasonable management cost, and maintaining a healthy cash flow in real-world construction projects. This is demonstrated in Table 2, where the makespan is included as part of the objective. By incorporating these relevant practical considerations, our model offers a more comprehensive and realistic solution for effectively managing construction projects in real-world scenarios.

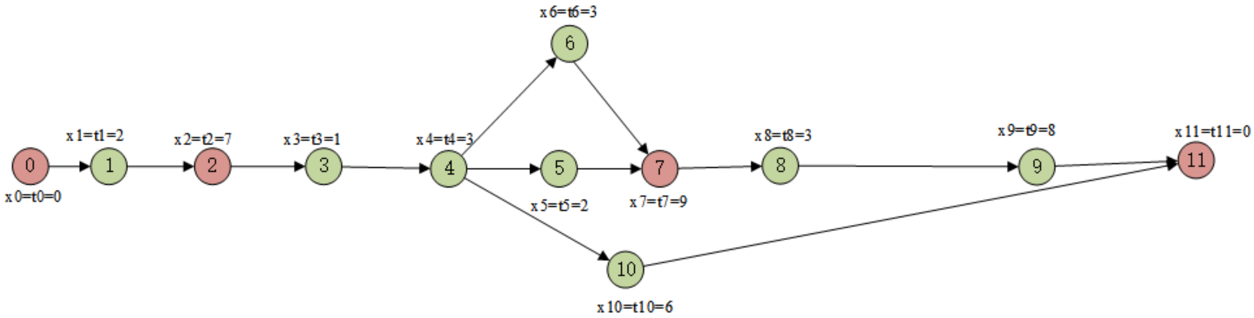


Figure 2: The initial structure of project 1 in group 1.

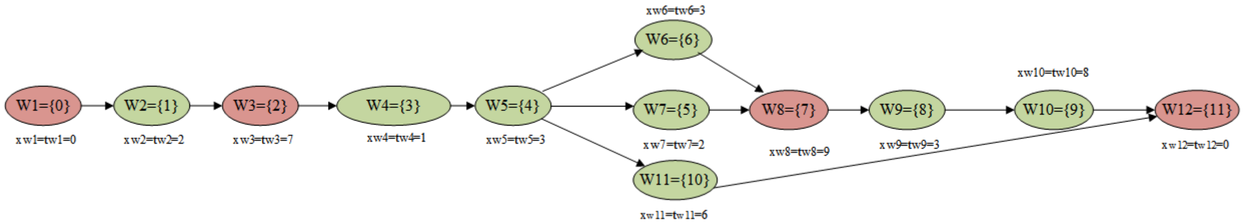


Figure 3: The work packages in the environment with 10 units of resource.

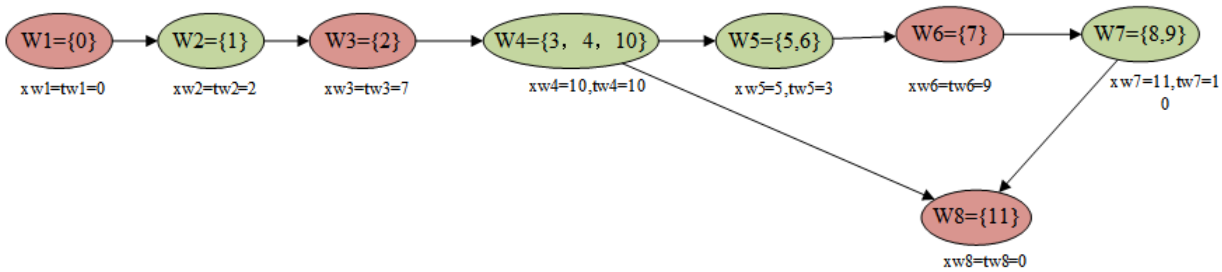


Figure 4: The work packages in the environment with 20 units of resource.

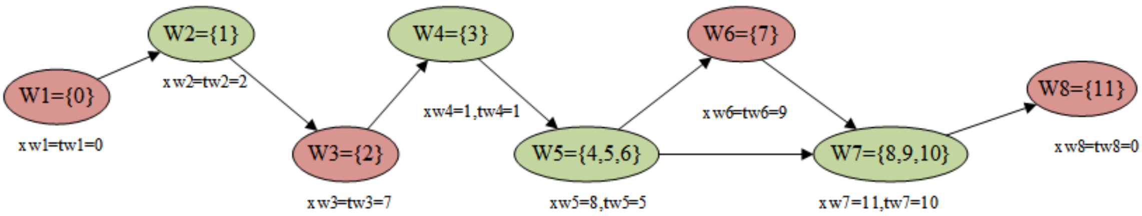


Figure 5: The work packages in the environment with 30 units of resource.

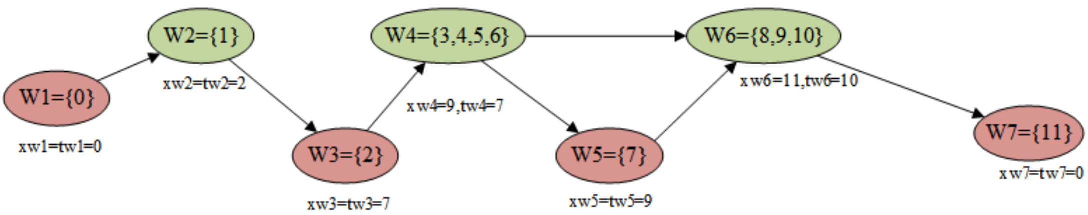


Figure 6: The work packages in the environment with 50 units of resource.

Table 2: Solutions of the nine projects via Gurobi

Indicator	Group 1			Group 2			Group 3		
	P 1	P 2	P 3	P 4	P 5	P 6	P 7	P 8	P 9
Node	10	10	10	20	20	20	30	30	30
$Node_A$	8	8	8	14	14	14	21	21	21
$Node_I$	2	2	2	6	6	6	9	9	9
Runtime	131	520	473	15007	24428	23674	98078	107645	188763
Packages	8	8	8	11	12	12	18	17	18
Makespan	41	39	30	75	88	56	105	89	98
Objective	317.31	325.29	304.84	516.16	579.00	532.21	815.70	869.14	765.43

Sensitivity Analysis

Given the significant impact of resource availability on the work package sizing and scheduling process in our problem, we conducted a sensitivity analysis to further investigate this relationship. During the sensitivity analysis, we systematically varied the resource availability levels with project 1 in group 1 (with resource {10,10,10,10}, {20,20,20,20}, {30,30,30,30} and {50,50,50,50}) and observed the corresponding changes in work package sizes. Figure 2 illustrates the initial structure of the tasks, while Figure 3 shows the work package structure under an environment with 10 units of each type of resource. Due to the limited resource availability, it is not feasible to group tasks into a single package. As a result, the work package structure remains the same as the initial one.

As the resource availability increases, more tasks can be grouped into packages. In Figure 4 and Figure 5, where there are 20 and 30 units of resources respectively, we observe the presence of 8 packages, with each package accommodating a maximum of 3 tasks. Figure 6 demonstrates a scenario where resource availability is no longer a constraint. This allows tasks to be packaged as long as the precedence constraints are satisfied. As a result, the number of packages is minimized. Overall, these figures highlight the impact of resource availability on the work package structure. With limited resources, fewer tasks can be grouped together, resulting in a larger number of packages. However, as resource availability increases, more tasks can be efficiently packaged, leading to a decrease in the number of packages.

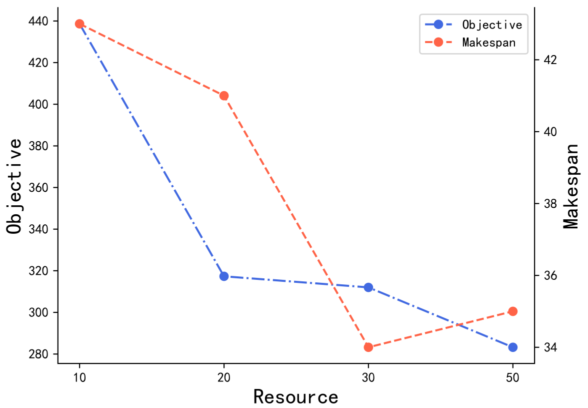


Figure 7: The comparisons of the makespan and objection.

Figure 7 presents a line graph showcasing the relationship between the objective function and makespan as the resource availability varies. As the resource availability increases from 10 to 30, the objective function shows a declining trend, decreasing from 440 to 290. This indicates that with more resources allocated, the total project cost decreases, reflecting improved efficiency and cost-effectiveness. However, when the resource availability reaches 50, the objective function starts to increase, reaching a value of 305. This suggests that excessive resource allocation beyond a certain point can lead to resource waste and increased project costs. Therefore, it is crucial to determine the appropriate level of resource availability that strikes a balance between cost reduction and resource utilization. In contrast, the makespan shows a declining trend as resource availability increases. This means that with more resources, the project duration or completion time decreases. This aligns with the intuition that additional resources enable tasks to be completed more quickly, resulting in a shorter overall project duration.

Conclusions

This paper presents an integrated model that addresses the challenges of work package sizing and scheduling in resource-constrained construction projects. To validate the effectiveness of our model, we utilized Gurobi, a powerful optimization solver, to search for optimal solutions, which demonstrates our model's effectiveness. Furthermore, we conducted an analysis to investigate the influence of resource availability on the work package sizing and scheduling process. By considering different resource levels, we find that proper resources are crucial for reducing the project's cost. However, solving our problem using Gurobi is time-consuming and not suitable for practical applications. In the future, we plan to develop algorithms that can efficiently solve our problem.

Acknowledgments

The work described in this paper is supported by grants from the Research Grants Council of the Hong Kong SAR of China, the National Natural Science Foundation of China, The University of Hong Kong (RGC Project No.15219422 G-HKU502/22, NSFC Project No. 72201228, HKU Project No. 2201100548 109000053)

References

Balouka, N. and Cohen, I. (2021). A robust optimiza-tion approach for the multi-mode resource-constrained project scheduling problem. European Journal of Oper-ational Research, 291(2):457–470.

Blazewicz, J., Lenstra, J. K., and Kan, A. R. (1983). Scheduling subject to resource constraints: classification and complexity. Discrete applied mathematics, 5(1):11–24.

Demeulemeester, E., Vanhoucke, M., and Herroelen, W. (2003). Rangen: A random network generator for activity-on-the-node networks. Journal of scheduling, 6:17–38.

Du, B., Tan, T., Guo, J., Li, Y., and Guo, S. (2021). Energy-cost-aware resource-constrained project scheduling for complex product system with activity splitting and recombining. Expert Systems with Applications, 173:114754.

Hartmann, S. and Briskorn, D. (2022). An updated survey of variants and extensions of the resource-constrained project scheduling problem. European Journal of operational research, 297(1):1–14.

Issa, S. and Tu, Y. (2020). A survey in the resource-constrained project and multi-project scheduling problems. Journal of Project Management, 5(2):117–138.

Lambrechts, O., Demeulemeester, E., and Herroelen, W. (2008). Proactive and reactive strategies for resource-constrained project scheduling with uncertain resource availabilities. Journal of scheduling, 11(2):121–136.

Li, C.-L. and Hall, N. G. (2019). Work package sizing and project performance. Operations Research, 67(1):123–142.

Li, C. Z., Xu, X., Shen, G. Q., Fan, C., Li, X., and Hong, J. (2018). A model for simulating schedule risks in prefabrication housing production: A case study of six-day cycle assembly activities in hong kong. Journal of cleaner production, 185:366–381.

Li, F., Xu, Z., and Li, H. (2021a). A multi-agent based cooperative approach to decentralized multi-project scheduling and resource allocation. Computers & Industrial Engineering, 151:106961.

Li, H., Zhang, X., Sun, J., and Dong, X. (2023). Dynamic resource levelling in projects under uncertainty. International Journal of Production Research, 61(1):198–218.

Li, X., Chi, H.-l., Lu, W., Xue, F., Zeng, J., and Li, C. Z. (2021b). Federated transfer learning enabled smart work packaging for preserving personal image information of construction worker. Automation in Construction, 128:103738.

Li, X., Wu, C., Xue, F., Yang, Z., Lou, J., and Lu, W. (2022). Ontology-based mapping approach for automatic work packaging in modular construction. Automation in Construction, 134:104083.

Liu, Z., Li, X., Wu, C., Ma, J., Yang, Z., and Guo, Y. (2023). Automatic work package sizing for cost-effective modular construction. Automation in Construction, 154:105003.

Pellerin, R., Perrier, N., and Berthaut, F. (2020). A survey of hybrid metaheuristics for the resource-constrained project scheduling problem. European Journal of Operational Research, 280(2):395–416.

Tsz Wai, C., Wai Yi, P., Ibrahim Olanrewaju, O., Abdelmageed, S., Hussein, M., Tariq, S., and Zayed, T. (2023). A critical analysis of benefits and challenges of implementing modular integrated construction. International Journal of Construction Management, 23(4):656–668.

Vanhoucke, M., Coelho, J., Debels, D., Maenhout, B., and Tavares, L. V. (2008). An evaluation of the adequacy of project network generators with systematically sampled networks. European Journal of Operational Research, 187(2):511–524.

Zaman, F., Elsayed, S., Sarker, R., Essam, D., and Coello, C. A. C. (2021). An evolutionary approach for resource constrained project scheduling with uncertain changes. Computers & Operations Research, 125:105104.

Zhang, Y., Li, X., Teng, Y., Shen, G. Q., and Bai, S. (2024). A heuristic rule adaptive selection approach for multi-work package project scheduling problem. Expert Systems with Applications, 238:122092.

DYNAMIC MATERIAL PASSPORTS FOR SUSTAINABLE MATERIAL MANAGEMENT: A CONCEPTUAL FRAMEWORK

Ioannis Markou[1], Derek Sinnott[1], and Ken Thomas[1]
[1]Department of Architecture & Built Environment, South East Technological University (SETU), Waterford City, Co. Waterford, Ireland

Abstract

The Construction sector faces a pressing need for whole life cycle sustainable material management. This paper presents the development of a theoretical framework for Dynamic Material Passports. Existing literature identified that Material Passports concept is new, lacking real-time updates and standardized methods with limited data and accuracy. Additionally, the current extensive manual input process discourages adoption. The framework, aligned with the ISO19650 standard, encompasses vital material information, environmental metrics such as Life-Cycle Assessment Data and Circularity Index, with real-time updates. This innovative framework enhances whole life material management, empowering industry stakeholders to make more informed Design and Procurement decisions.

Introduction

The construction industry stands as a substantial contributor to environmental harm, encompassing issues like natural resource depletion, waste generation, and greenhouse gas emissions (MacArthur Foundation, 2022; UN, 2016). In response to these pressing challenges, there has been a growing emphasis on the adoption of circular economy principles in recent times. The Circular Economy (CE), as an economic model, is aims to maximize the utilization of resources, minimize waste, and reduce adverse environmental impacts (MacArthur Foundation, 2022) In this context, Dynamic Material Passports (DMPs) emerge as a promising tool within the Architecture, Engineering, and Construction (AEC) sector.

In the Built Environment (BE), Material Passports (MPs) function as digital records, meticulously documenting the materials employed in construction, from their composition to their origin and location, throughout their lifecycle. This important information opens avenues for the efficient reuse, recycling, or repurposing of materials, significantly reducing waste, and promoting resource efficiency (Çetin, Straub, et al., 2022; Honic, Kovacic and Rechberger, 2019a, 2019b).

Nevertheless, considering its potential, the deployment of MPs faces certain challenges, particularly lack of standard methods of creation, compatibility, and currently have static form, without real-time updates (Honic et al., 2021)To address those obstacles, this research paper seeks to provide with a BIM-based framework for the development of DMPs, aligned with the information management processes described in EN ISO 19650, part 2 for a BIM level 2 project, that will promote CE practices in the BE. BIM is a form of digital technology that involves the creation and management of a digital model of a building, which include both its functional and physical properties (Becker et al., 2018). When it is used during design, construction, and operational stages, it creates a detailed and simple to access record of the materials used (M. Honic et al., 2019).

Hence, this research paper discusses the gaps based on the current state of MPs, the preliminary findings, and a proposed conceptual framework, which aims to fill the gaps identified from the literature review. The proposed framework involves four key stages: data collection, data management, data analysis, and data dissemination.

Literature Review

MPs and CE

Over the pass years, the construction industry is experiencing a significant transformation, and the emerging concept of MPs plays a crucial role to it (Çetin et al., 2022b). They have the ability to significantly improve resource management processes and the same time to simplify the transition of the BE to a CE perspective (Honic et al., 2021, 2019a, 2019b; Soman et al., 2022).

MPs, as comprehensive documentation systems, provide in-depth information about construction materials and products throughout their life cycle, covering details about their origin, characteristics, and environmental and sustainability attributes (Honic et al., 2019a; Kedir et al., 2021) Some example of data that may be included in MPs are manufacturing information, physical properties, chemical or biological properties, quality of material, design related information, location and transportation information, disassembly/ deconstruction manuals and sustainability information, such as embodied carbon and Circularity Index (Stella et al., 2023).

The Circularity Index or Indicator was developed by Ellen Macarthur Foundation (2016), as a tool for measuring the materials circularity, considering the amount of recycled material content the potential for reuse and recyclability of materials. It is calculated on the scale of 0% to 100%, with the higher percentage indicating the higher circularity. This indicator it can be used to assess the circularity of individual materials and products (Ellen Macarthur Foundation, 2016). Similarly, Madaster, a Dutch company that developed a digital platform for producing MPs, developed the Building Circularity Index based on MacArthur's Circularity Indicator, which measures the circularity of a building, taking into consideration the amount of recycled materials used in a building (Madaster, 2018).

Also, it was found that there are different levels of MPs (Çetin et al. 2023). At the product level, "Material-Level Passports" set the foundation, while at the “Building level”, specific design stages come into play (Çetin et al. 2023a; Stella et al. 2023). During “Design for construction stage” detailed material specifications are introduced while at construction stage with documentation of “as-built” information is complied, and “in used” stage focusing on the operational phase with regular updates to the MPs. For “End-of-life” stage marks the identification of information crucial for the considerations (Stella et al., 2023).

Historically, the BE has followed a linear economic model defined by resource consumption, use and disposal of building materials with the creation of significant amount of waste (MacArthur Foundation, 2022). Such an approach is both ecologically untenable as well as financially inefficient. Considering the current concerns about resource scarcity, deterioration of the environment, and global warming to move to a more circular and sustainable system for managing construction materials (Çetin et al., 2022a; Charef, 2022).

Therefore, as a fundamental concept, CE aspires to reduce waste and increase resource efficiency by encouraging recycling, reuse, and repurposing of materials across their life span (MacArthur Foundation, 2022). It encourages to “close the loops”, from a linear economy to a more circular, in which items and materials are continuously refurbished, reused, and remanufactured, lowering the demand for raw materials and reducing the ecological effects of waste disposal (Pomponi and Moncaster, 2017).

Also, the implementation of MPs in the BE is completely compatible with the CE principles (Çetin et al., 2023; Çetin, Straub, et al., 2022). They can assist in the tracking, control and optimization of the use of building materials (Atta et al., 2021; Honic, Kovacic and Rechberger, 2019a; Munaro and Tavares, 2021). Additionally, they can provide comprehensive information on each material used in a building, providing accurate and thorough documentation through the building’s entire life span. MPs can also assist stakeholders to make informed decisions about material procurement, selection and consumption with focus on sustainability and resource efficiency (Honic et al. 2019a; Munaro and Tavares 2021).

MPs and Building information Modelling (BIM)

BIM constitutes an advanced technological method that effectively combines planning, design, construction, and asset management in the BE. This digital tool, which assist to the representation both the functional and physical features of building assets allow successful cooperation among stakeholders (Becker et al., 2018).

Furthermore, BIM encourages transparency, collaboration, and decision-making based on data throughout the life circle of a building boosting sustainable practices and waste reduction (Pishdad-Bozorgi et al., 2018; Sacks et al., 2018).

The combined approach of BIM and MPs offer great potential for sustainable building practices, since it provides a digital basis for a detailed 3D model (Alshammari et al., 2021; Pehlken and Baumann, 2020; Shahzad et al., 2022). Hence, a BIM-based method of producing a MP can be employed to improve the building design by considering resource efficiency and detailed documentation of materials used. BIM is used as a resource of information regarding the material quantities, overall building geometry and location within the building, while linking secondary databases to obtain their physical properties and evaluate their environmental impact or recycling potential(Honic et al., 2019a).

Additionally, the combination of MPs with BIM contributes to the successful collection and storage critical information on building materials increasing transparency and efficiency (Atta et al., 2021). A good example of that is a case study in Egypt, where a parametric BIM model permitted the automatic development of sustainable metrics for a residential building. This demonstrated that BIM improved the sustainability of the project through the use of resources efficiently, decreasing waste and enabling material monitoring for reuse (Atta et al., 2021). The BIM-based method takes use of the technologies capacity to integrate environmental impact evaluations via Life Cycle Assessment (LCA) via, giving a holistic approach to developing of sustainable buildings (Atta et al., 2021; Honic et al., 2019a; Kedir et al., 2021). Additionally, this combination can foster CE concepts in the BE by increasing transparency, traceability and accountability during the asset’s life cycle (M. Honic et al., 2019).

LCA is a standardised method for assessing the environmental impact of a product, throughout its lifecycle and is based on ISO 14040, It includes greenhouse gas (GHG) emissions which is measured through Global Warming Potential (GWP). GWP is expressed in the equivalent amount carbon dioxide (CO_2) (McGrady, 2022). Other indicators that are used in LCA are Acidification Potential (AP) measuring the potential of emissions to cause acidification, and ISO 14040 describes that LCA is carried out in four steps: goal and scope definition, inventory analysis, impact assessment and interpretation (Muralikrishna and Manickam, 2017).

However, BIM is more than 3D visual representation, they contain essential data about the building’s structure and its elements (Volk et al., 2014). An effective exchange of information can be achieved through Industry Foundation Classes (IFC) (buildingSMART, 2023). IFC was created by BuildingSMART as an open standard for transferring building data amongst different BIM tools. It also assures the compatibility and vendor neutrality, allowing project teams to use their preferred software while ensuring data usability (Jiang et al., 2019) IFC classes are also specified in ISO 19650 series as am appropriate format of storing the BIM model, due to interoperability (Earley Michael et al., 2022). Moreover, Madaster platform, allows uploading the BIM model in IFC format, in order to produce MPs (Madaster, 2023).

MPs and Artificial Intelligence (AI)

The data compilation for creating MPs has been traditionally a time-consuming operation. As some researchers stated, manual procedures need precise data gathering, which is a laborious and inclined to errors (Honic et al., 2021, 2019a; Hunhevicz et al., 2023; Kovacic et al., 2020). Hence, they are an urgent need for standardized protocols and tools (Atta et al., 2021; Göswein et al., 2022) which created the ground work for Deep Learning (DL), a particular type of AI. DL specializes in optimizing and automating complex procedures in MPs development (Akanbi et al., 2020; Oluleye et al., 2023), providing immediate updates and precision across the building's lifespan (Çetin et al., 2022a).

DL, was influence by the functions of human brain, emerging this way, as a transforming force in CE activities. Its forecasting ability is aligned with the requirements of sustainable material management (Lu and Chen, 2022), enabling the creation of a more sustainable BE adaptable to changing demands. Contrary, Machine Learning (ML), is based on algorithms that analyses and forecast based on data (Noman et al., 2022). ML is extremely useful for CE as it can assess material consumption patterns and foreseeing trends for enhancing resource (Çetin et al., 2023). Through sophisticated algorithms, both DL and ML provide possibilities for developing a circular BE, outcoming from technological, economic environment and social aspects (Akanbi et al., 2020; Oluleye et al., 2023).

The selection of an AI approach for developing an MP framework requires a considerable deliberation. Current research describes three approaches for developing MPs: the top-down, the bottom up and hybrid methods (Honic et al., 2021), that demand an extensive understanding of the building materials. As a result, such methods may be easier to be examined by DL given that they have higher accuracy, flexibility and capacity to handle large datasets, making it the best choice, enabling stakeholders taking better decisions.

Current state of MPs

The present form of MPs in the BE includes technological, economic, environmental, and social factors. Current studies proposed that a standardized method for generating and managing MPs, pointing out the need of uniformity and accessibility (Atta et al., 2021; Honic et al., 2019a). Several studies restate the need for consistent guidelines and protocols across the sector (Göswein et al., 2022; Negendahl et al., 2022).

In respect to technological factors, current methods of creating MPs can be quite laborious. Their creation requires substantial time as it involves to extract material information from the BIM model to an excel file, collect information from material inventories and then manually input in the excel document, which will finally be linked back to the BIM model, using unique identifiers (Atta et al., 2021; Honic et al., 2019a; Kovacic et al., 2020).

The implementation of MPs is also considered as economically feasible with possible cost decreases through material reuse (Smeets et al., 2019). Furthermore, significant economic benefits can be achieved through the digitalization of MP (Çetin et al., 2022a), in conjunction with the best practices (Heinrich and Lang, 2019).

From an environmental perspective, MPs are touted as catalysts for sustainability, enabling high levels of material recycling (Honic et al., 2019b)s and stock flow modelling emphasize the potential to calculate circularity (Göswein et al., 2022; Negendahl et al., 2022).

On the social front, transparency, accountability, and stakeholder engagement take centre stage. Digital platforms integrating MPs with stakeholder engagement are recommended (Hunhevicz et al., 2023; Kovacic et al., 2020). The importance of stakeholder engagement, especially in the end-of-life stage of buildings, is highlighted (Honic et al., 2021). Additionally, local scenarios play significant role on the choice of sustainable construction materials, as evidenced by a case study suggesting that a building MP for wood construction system in Brazil, could significantly reduce the environmental impact of construction versus more conventional materials, such as a concrete and steel (Munaro et al., 2019).

Gaps and Challenges

MPs have great potential with their deployment in the BE although they are facing significant challenges. Their major disadvantage is their static form, that limit them to no frequent updates, that are necessary to accurately reflect real-time changes in a material's composition, condition, and lifecycle, enabling informed decision making for sustainable resource management and end-of life strategies (Atta et al., 2021; Kovacic et al., 2020). Their stationary state compromises MPs' adaptive perspective, in terms or real-time updates, reducing their level of accuracy and efficacy through sustainable material management. Due to that static nature and high level of manual input for updating material data also their inability of optimizing and automate processes reduce their long-term effectiveness (Atta et al., 2021; Honic et al., 2019a; Kovacic et al., 2020) .

Another challenge for the development of MPs is the lack of a standardized methods for compiling MPs. MPs appear in a variety of formats, including paper-based and digital, with varied degrees of details, causing in inaccuracies and limiting sharing of data among stakeholders (Heinrich and Lang, 2019; Honic et al., 2021). Developing a consistent approach for producing MPs is critical that contributes to the reduction of errors and ensuring that correct information is exchanged between the stakeholders such as AEC professionals, waste managing and material suppliers (Kovacic et al., 2020).

Another challenge that affects the implementation of MPs in the BE, is the insufficient incentives for their implementation. MPs adoption is not adequately motivated by the current economic and regulatory frameworks, but also the BE gives higher priority to the speed and cost reduction rather than sustainability (Göswein et al., 2022). In order to promote the broad use of MPs, policymakers must provide regulations and incentives, such as tax reliefs or financing green procurement for sustainable construction (Smeets et al., 2019).

A further obstacle involves the quality, accessibility and data administration. All building materials have different origin, making it difficult to collect and maintain their specific data (Atta et al., 2021; Honic et al., 2019a). To guarantee the accuracy and accuracy of MPs data, improved data management systems and quality control techniques are required. Additionally, the environmental implications of building materials are frequently neglected, when implementing MPs. There is no assessment of the materials' long-term environment impacts in the existing MPs frameworks (Heisel and Rau-Oberhuber, 2020). To provide a thorough understanding about their environmental effects, using standardized methods for environmental assessments into MPs frameworks.

MPs also face challenges related to the end-of-life state of buildings, such as disassembly or deconstruction instructions. Therefore, buildings are often thought to be demolished as the best solution, rather than carefully dismantled, leading to destruction of possible healthy reusable materials, creating this way construction waste. Hence, by integrating MP data into waste management strategies can help avoid unnecessary demolition and waste, preserving essential materials and resources (Honic et al., 2019b). Successful reuse and recycling necessitate improved coordination among stakeholders and waste management professionals.

Methodology

The present research provides a conceptual framework for Dynamic Material Passports (DPMs) using a strong approach based on in-depth literature review. The main emphasis is on analysing the present status of MPs in recent literature understanding the data needs for constructing MPs, defining how their features could be implemented in the DMPs in accordance with existing standards and finally identifying the gaps in the current state. The process entails a comprehensive literature review, followed by an inductive research approach to develop this conceptual framework. The framework is refined by a comparison of various MPs approaches, as described in "current state". The compilation of qualitative data from different methodologies adds to a practical conceptual framework, providing useful insights for its future application in the BE.

Proposed Conceptual Framework

The conceptual framework intends fill the gaps for the static data of current MPs approaches and to address the tracking of the material differenced in terms of future retrofitting and refurbishment, through a project's life circle (See Table 1). The emphasis will be initially on the creation pf DMPs at "Material level" with information that collected during the design stages. The DMPs will incorporate technologies such as BIM, GS1 barcoding standards and AI. This strategy addresses will address towards the improvement of data accessibility, accuracy and regular updating thought the building's life span.

Table 1: "Current State" Vs "Future State"

Feature	Current State (MPs)	Future State (DMPs)
Definition	Tracks and documents materials used in construction or manufacturing	Real-time, digital tracking throughout the lifecycle
Scope	One-time documentation at construction/manufacturing	Continuous monitoring and updates
Data Update	Static information; that cannot update	Dynamic, allows real-time updates
Tech. Integration	Relies on traditional methods with substantial manual input	BIM, AI, QR codes
Use	Snapshot for end-of-life recycling/reuse	Critical for asset management throughout life cycle

In the context of this study, a BIM model for a proposed new building is generated as the initial step (Step 1, Fig.1). This involves creating a digital representation of the physical structure using BIM tools like Revit and adhering to the ISO 19650 for "BIM Stage 2" modelling guidelines. Clause 5 of the ISO 19650-2 describes in detail the processes and processes that are necessary to compile the Project Information Model (PIM) during the design stages. Also, the information that will be required for DMPs will be included in the Project Information Requirements (PIR) and Exchange Information Requirements (EIR), both compiled during the process 5.1: Assessment for Need, as described in the relevant clause.

Following the PIM creation, the primary components for constructing the DMPs are identified. Recognizing the absence of a standard method for data requirements, the material data considered is refined in subsequent stages, drawing from recent literature. The data encompasses three categories: material information (type, physical properties), sustainability indicators (such as the Circularity Index), and component information (including manufacturing, installation, maintenance, disassembly instructions, and expected lifespan and performance data). Simultaneously, while setting up the initial BIM model, a material inventory is established based on information for each piece, using unique identification and names specified in ISO 19650 (Step 2, Fig.1). Also, is worth

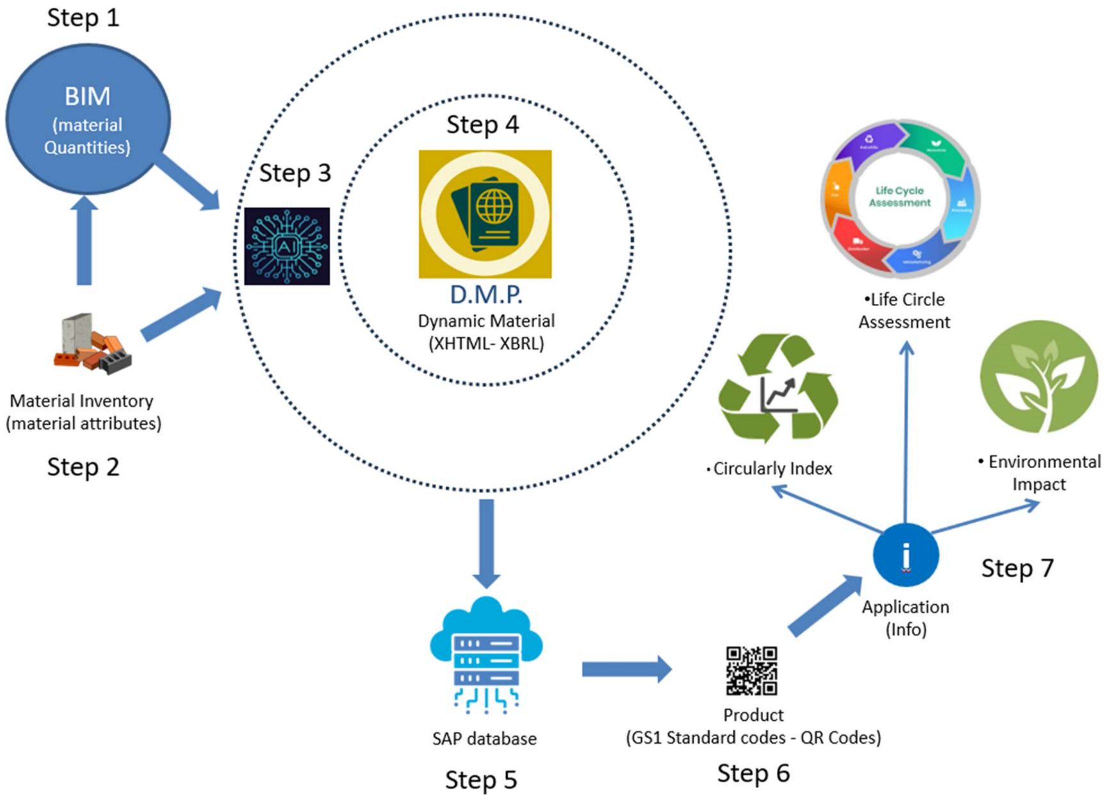


Figure 1: Conceptual framework for DMPs (own illustration)

mentioning, part of the dynamic nature of DMPs, it would require to be linked to 4D ontology. 4D ontology refers to the integration of time (the fourth dimension) into the digital representation of a building. This will enable the capturing and visualisation of how elements of the project change and interact over time, but also simulate and anticipate future changes throughout the lifecycle of the project.

AI techniques, including ML and DL, may be deployed to collect diverse attributes of the supplied content (Step 3). ML algorithms can analyse existing data sources, such as supplier documentation and manufacturing records, to identify patterns and relationships. Additionally, ML predictive models can assist in forecasting potential gaps and areas of uncertainty, prompting proactive measures. On the other hand, DL models, equipped with neural networks, are advance in processing complex and unstructured data. Furthermore, DL models, can handle sequential data, ensuring a more refined understanding of materials' dynamic attributes over time. These techniques contribute to the identification, categorization, and tracking of materials, assessing their origins, environmental impact, and even predicting future performance (Çetin et al. 2023b). DL algorithms automate material-related data acquisition from various sources like product databases, material suppliers, manufacturers, and construction documents. They can navigate websites to extract essential data from digital sources, eliminating the need for extensive manual labour.

DMPs may be created as an Extensible Hypertext Mark-up Language (XHTML) document adorned with eXtensible Business Reporting Language (XBRL) tags (Step 4, Fig.1). XHTML, as opposed to HTML, has more rigid syntax and formatting restrictions. It is also guaranteeing that web browsers and interpreter programs can process it seamlessly. Usually, it is standard practice to combine XHTML with XBL in order to ensure data integrity and interoperability across systems, particularly on the internet (Stewart, 2020).

In financial reporting, XHTML formats interconnected with XBRL tags are commonly employed, giving a dual-readable structure for both people and machines. XBRL tags are used to insert standardized and machine- readable data into an XHTML file, facilitating financial report processing and analysis. This unified technique reduces inconsistencies and simplifies the filing process, resulting in a strong resolution for financial reporting requirements (Siddle 2020; Stewart 2020). The use of XHTML with XBRL tags is an effective form for storing and evaluating material and sustainability data. Hence, the ability of XBRL to stipulate context, combined with XHTML is quite beneficial for DMPs.

The next step is proposed that store the files in Single Access Point (SAP) database, similar to European Single Access Point (ESAP) which is used for (Step 5, Fig.1). As in ESAP, SAP will provide access to publicly available information related to MPs data. Therefore, this will be consistent whit the EU' Digital Strategy and the targets of

European Green Deal, fostering transparency and accelerating decision-making for stakeholders (European Commission 2023).

SAP will constitute a regulated platform, will encompass various features, including a web portal, API for data retrieval, multilingual search function, information viewer, machine translation service, download service, and notification service for updates. This can be particularly crucial for Small and Medium Enterprises (SMEs), for future sustainable related decision-making (European Commission 2023).

The subsequent step involves exporting BIM model and data related to building elements, their layers, volumes, thicknesses, and individual components to be linked to a Quick Response code based on GS1 standards (Step 6, Fig1). Essential standards include the Global Trade Item Number (GTIN), GS1-128, Global Location Number (GLN), and Global Data Synchronization Network (GDSN) to ensure accurate and updated data accessibility across the supply chain (GS1 2023).

A method of barcoding involves Quick Response (QR) codes, with dynamic QR codes being considered for this research. Dynamic QR codes have the capability to employ an intermediate concise URL for redirecting, enabling modification of stored data within the same printed QR code. Recent research highlights the effectiveness of QR codes in tracking materials, providing essential information about each building element (Byers et al., 2022). These QR codes improve material tracking in the construction process, offering real-time updates to the MP, especially concerning location.

Real-time update is crucial will be enabled by scanning the QR code, especially during and construction phases, or refurbishments which allows access to the material's geo-location through GIS, by scanning by using a phone or tablets. Some of the building materials like concrete, steel, and brick exhibit relatively stable properties over time. Unlike dynamic assets that undergo frequent changes or degradation, the characteristics of these materials remain largely consistent throughout their lifespan. Therefore, the need for real-time updating to capture sudden variations or anomalies may be minimal, as gradual changes can be adequately monitored through periodic assessment. AI techniques, including computer vision, enable the scanning of QR codes and linking to relevant data, facilitating estimations such as embodied carbon due to transportation.

For example, a DMP for items such as electrical motors within ventilation systems which have relatively shorter lifecycle than the building structure and envelop, will allow instant access to current data on the origin, composition, and quality of its components. As the motor has been replaced, then with DMP will be easier to be tracked in order to obtain the embodied carbon. With the aid of AI, the DMP will be able to collect missing information, in respect to the physical or manufacturing characteristics, as well as LCA calculations, and updated information. The combination of the BIM 4D ontology will enable to track the motors' state over time, such operation and maintenance stages, capturing crucial details about the history and condition of motor. Also, the use of the QR codes will facilitate easier retrieval of the motor information using mobile devices, as well as tracking the geolocation. In this way, informed decision-marking is enabled, in respect to the mor' maintenance, repair, or end-of-life options and ultimately will assist to reduce the environmental impact of the motor's lifecycle. Additionally, the DMPs will include automated sustainability indicators such percentage of recyclable materials, waste, and environmental impacts, LCA calculations expressed by the use of algorithms, in terms of Global Warming Potential (GWP), Acidification Potential (AP), and Potential Energy Index (PEI) (Step 7, Fig.1). Also, the inclusion of Circularity Index created by McArthur Foundation to DMPs will allow to score the circularity of the building based on the percentage of recycled materials used. AI plays a significant role in enhancing the accuracy of these environmental impact assessments. ML algorithms process large datasets to identify patterns and correlations, providing insights into material performance and environmental impact.

Limitations and Challenges for Implementing the Conceptual Framework

DMPs implementations in the BE faces significant limitations. Those major challenges include industry's opposition, anchored in old procedures, necessitates careful positioning to explain the befits of DMPs while addressing concerns from moving traditional approaches. A further challenge is the financial limitations, particularly for SMEs, that need to deliberate cost reductions. Also, interoperability concerns between many stakeholders who use different BIM software, demand careful planning and possible plugin development. In addition to that, maintaining data accuracy among the supply chain is difficult, requiring coordinated collaborations common data dictionaries. In relation to data, concerns about privacy and security is another limitation.

Finally, the lack of standard and regulations for creating MPs and educating professionals in the use of new technologies such as DMPs could be really challenging. Considering these problems as possibilities for improvement, can lead to a more robust implementation of DMPs in the BE.

Conclusions and Future Work

The proposed conceptual framework of DMPs intends to fill the gaps of current state of MPs, such as its current static form, lack of regular real-time updates for the material information throughout a buildings life cycle which current requires a lot of manual input. This is may lead to errors and included information to be included in the MPs current state.

DMPs is proposing the use of BIM model, in conjunction with the BIM standards of ISO19650-2 "Delivery phase of an Asset", which give a detail the processes of creating a BIM model in addition to information management for a BIM level 2 project. The BIM model will provide the quantities and specific location of materials in the building. In conjunction to AI techniques of collecting and analysing complex data, and also scanning codes based on GS1 standards, will enable the DMPs to provide up-to-date information of about the construction materials. Furthermore, the incorporation of LCA into DMPs will add sustainability aspect to the framework, contribute to the reduction of the carbon footprint of the BE.

The future research will focus on the creation of the prototype DMP and further investigation of the use of appropriate AI techniques. Following the creation of the prototype, collaboration with industry will take place to run a Pilot Case study for the DMPs implementation in the BE that will allow the assessment of its effectiveness. This will allow feedback from industry that will be assessed and contributes to the continues improvement of the prototype and based on the lessons learned from this pilot study.

Acknowledgments

The Author would like to thank the co-writers of this paper for their assistance in completing this paper. Also, the author would like to express his sincere gratitude to Ms. Anastasia Stella, Sustainability Associate at Waterman's Group (UK) and co-founder of "CIRCULAND", for her support during this research. The research is funded by "Build Digital Project" from Department of Public Expenditure, Republic of Ireland.

References

Akanbi, L.A., Oyedele, A.O., Oyedele, L.O., Salami,8 R.O., 2020. Deep learning model for Demolition Waste Prediction in a circular economy. J Clean Prod 274, 122843.

Atta, I., Bakhoum, E.S., Marzouk, M.M., 2021. Digitizing material passport for sustainable construction projects using BIM. Journal of Building Engineering 43.

Becker, R., Falk, V., Hoenen, S., Loges, S., Stumm, S., Blankenbach, J., Brell-Cokcan, S., Hildebrandt, L., Vallée, D., 2018. Bim – towards the entire lifecycle. International Journal of Sustainable Development and Planning 13, 84–95.

buildingSMART, 2023. What is openBIM ? - buildingSMART International [WWW Document]. URL https://www.buildingsmart.org/about/what- is-openbim/ (accessed 10.22.22).

Byers, B., Cheriyamulla, S., Ewason, J., Hall, D., De Wolf, C., 2022. USING ENGRAVED QR CODES TO CONNECT BUILDING COMPONENTS TO MATERIALS PASSPORTS FOR CIRCULAR CONSTRUCTION.

Çetin, Gruis, Straub, 2022a. Digitalization for a circular economy in the building industry: Multiple-case study of Dutch social housing organizations. Resources, Conservation and Recycling Advances 15.

Çetin, Raghu, Honic, Straub, Gruis, 2023. Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. Sustain Prod Consum 40, 422–437.

Çetin, S., Raghu, D., Honic, M., Straub, A., Gruis, V., 2023. Data requirements and availabilities for material passports: A digitally enabled framework for improving the circularity of existing buildings. Sustain Prod Consum 40, 422–437.

Çetin, Straub, Gruis, 2022b. How Can Digital Technologies Support the Circular Transition of Social Housing Organizations? Empirical Evidence from Two Cases.

Charef, R., 2022. Supporting construction stakeholders with the circular economy: A trans-scaler framework to understand the holistic approach. Clean Eng Technol.

Earley Michael, Montague Ralph, Voortman Bernard, O'Connell David, De Palma Rich, Hinshelwood Michael, Bracken Daire, Slatter Pat, Prunty Eoin, 2022. RIAI BIM Pack 2 BIM Update with ISO 19650 Series.

Ellen Macarthur Foundation, 2016. Material Circularity Indicator [WWW Document]. URL https://www.ellenmacarthurfoundation.org/materia l-circularity-indicator (accessed 12.13.23).

European Commission, 2023. Council adopts regulation easing access to corporate information for investors - Consilium [WWW Document]. URL https://www.consilium.europa.eu/en/press/press-releases/2023/11/27/council-adopts-regulation-easing-access-to-corporate-information-for-investors/ (accessed 12.13.23).

Göswein, V., Carvalho, S., Cerqueira, C., Lorena, A., 2022. Circular material passports for buildings - Providing a robust methodology for promoting circular buildings. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics.

GS1, 2023. GS1 Standards [WWW Document]. URL https://www.gs1.org/standards (accessed 11.13.23).

Heinrich, M., Lang, W., 2019. MATERIALS PASSPORTS-BEST PRACTICE. Innovative Solutions for a Transition to a Circular Economy in the Built Environment - BAMB 1–74.

Heisel, F., Rau-Oberhuber, S., 2020. Calculation and evaluation of circularity indicators for the built environment using the case studies of UMAR and Madaster. J Clean Prod 243.

Honic, Kovacic, Aschenbrenner, Ragossnig, 2021. Material Passports for the end-of-life stage of buildings: Challenges and potentials. J Clean Prod 319, 128702.

Honic, Kovacic, Rechberger, 2019a. Concept for a BIM-based Material Passport for buildings. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics Publishing.

Honic, Kovacic, Rechberger, 2019b. Improving the recycling potential of buildings through Material Passports (MP): An Austrian case study. J Clean Prod 217, 787–797.

Honic, M., Kovacic, I., Sibenik, G., Rechberger, H., 2019. Data- and stakeholder management framework for the implementation of BIM-based Material Passports. Journal of Building Engineering 23, 341–350.

Hunhevicz, J.J., Bucher, D.F., Soman, R.K., Honic, M., Hall, D.M., De Wolf, C., 2023. WEB3-BASED ROLE AND TOKEN DATA ACCESS: THE CASE OF BUILDING MATERIAL PASSPORTS.

Jiang, S., Jiang, L., Han, Y., Wu, Z., Wang, N., 2019. OpenBIM: An Enabling Solution for Information Interoperability. Applied Sciences 2019, Vol. 9, Page 5358 9, 5358.

Kedir, F., Bucher, D.F., Hall, D.M., 2021. DEPARTURE Material passport as a tool for a circular economy.

Kovacic, Honic, Sreckovic, 2020. Digital Platform for Circular Economy in AEC Industry. Engineering Project Organization Journal 9.

Lu, W., Chen, J., 2022. Computer vision for solid waste sorting: A critical review of academic research. Waste Management 142, 29–43.

MacArthur Foundation, E., 2022. What is a circular economy? | Ellen MacArthur Foundation [WWW Document]. URL https://ellenmacarthurfoundation.org/topics/circula r-economy-introduction/overview (accessed 9.25.22).

Madaster, 2018. Madaster Circularity Indicator explained Audience: Author: Users Madaster.

Madaster, 2023. Madaster platform [WWW Document]. URL https://madaster.com/platform/ (accessed 2.22.23).

McGrady, C., 2022. Life Cycle Assessments (LCA) [WWW Document]. URL https://www.mcgradyclarke.com/services/carbon/li fe-cycle-assessment/ (accessed 10.20.22).

Munaro, M.R., Fischer, A.C., Azevedo, N.C., Tavares, S.F., 2019. Proposal of a building material passport and its application feasibility to the wood frame constructive system in Brazil. In: IOP Conference Series: Earth and Environmental Science. Institute of Physics Publishing.

Munaro, M.R., Tavares, S.F., 2021a. Materials passport's review: challenges and opportunities toward a circular economy building sector. Built Environment Project and Asset Management 11, 767–782.

Munaro, M.R., Tavares, S.F., 2021b. Materials passport's review: challenges and opportunities toward a circular economy building sector. Built Environment Project and Asset Management 11, 767–782.

Muralikrishna, I. v., Manickam, V., 2017. Life Cycle Assessment.

Negendahl, K., Barholm-Hansen, A., Andersen, R., 2022. Parametric Stock Flow Modelling of Historical Building Typologies. Buildings 12.

Noman, A.A., Akter, U.H., Pranto, T.H., Haque, A.K.M.B., 2022. Machine Learning and Artificial Intelligence in Circular Economy: A Bibliometric Analysis and Systematic Literature Review. Annals of Emerging Technologies in Computing.

Oluleye, B.I., Chan, D.W.M., Antwi-Afari, P., 2023. Adopting Artificial Intelligence for enhancing the implementation of systemic circularity in the construction industry: A critical review. Sustain Prod Consum 35, 509–524.

Pishdad-Bozorgi, P., Gao, X., Eastman, C., Self, A.P., 2018. Planning and developing facility management-enabled building information model (FM-enabled BIM). Autom Constr 87, 22–38.

Pomponi, F., Moncaster, A., 2017. Circular economy for the built environment: A research framework. J Clean Prod 143, 710–718.

Sacks, R., Eastman, C., Lee, G., Teicholz, P., 2018. BIM Handbook. BIM Handbook.

Siddle, J., 2020. ESEF - Filing in iXBRL or XHTML? - CoreFiling website [WWW Document]. URL https://www.corefiling.com/2020/05/04/esef-filing-in-ixbrl-or-xhtml/ (accessed 11.13.23).

Smeets, A., Wang, K., Drewniok, M.P., 2019. Can Material Passports lower financial barriers for structural steel re-use? IOP Conf Ser Earth Environ Sci 225.

Soman, R.K., Kedir, F., Hall, D.M., 2022. TOWARDS CIRCULAR CITIES: DIRECTIONS FOR A MATERIAL PASSPORT ONTOLOGY.

Stella, A., Terndrup, M., Pilcher, N., Nilsson, S., Buhagiar, J., 2023. Waterman Materials Passports Framework Introducing a Standardised Approach to Materials Passports in the Construction Industry. Waterman Group.

Stewart, A., 2020. iXBRL OR XHTML: AN ESEF TECHNICAL BRIEFING [WWW Document]. URL https://1stopxbrl.com/ixbrl-or-xhtml-an-esef-technical-briefing/ (accessed 11.13.23).

UN, 2016. United Nations Sustainable Development Action Plan [WWW Document]. URL https://www.un.org/sustainabledevelopment/cities/ (accessed 12.17.22).

Volk, R., Stengel, J., Schultmann, F., 2014. Building Information Modeling (BIM) for existing buildings — Literature review and future needs. Autom Constr 38, 109–127.

ADVANCING AEC PROJECT MANAGEMENT: A MODEL-BASED AND DATA-DRIVEN APPROACH FOR SUSTAINABLE PRACTICES

Gertraud Wolf[1], Martin Wogan[2,3], Nijanthan Mohan[2,4], Nazereh Nejat[5], Rolf Gross[4], Karsten Menzel[2]
[1]Institute of Construction Management, University of Bundeswehr Munich, Germany
[2]Institute of Construction Informatics, Technical University Dresden, Dresden, Germany
[3]B/M Consult GmbH, Braunschweig, Germany
[4]Institute of Smart Building Engineering, Aachen University of Applied Sciences, Aachen, Germany
[5]Infrakit, Munich, Germany

Abstract

The construction industry is transitioning from traditional linear methodologies to the Total Life Cycle Process (TLCP) based on DIN EN ISO 19650 principles. TLCP employs database-supported information models (dIMs) and Information Requirement Matrices (IRMs) to enhance project management, focusing on sustainability. This method improves communication, ensures timely access to critical information and optimizes resource use. By integrating the Level of Information Need (LOIN) concept and web-based tools, TLCP supports agile transitions and technological adaptation. This paper showcases TLCP's practical applications, promoting sustainable construction management and encouraging industry-wide discourse, ultimately contributing to long-term environmental and economic benefits.

Introduction

The construction industry has focused on conserving resources in recent years to align with the Sustainable Development Goals (SDGs) (United Nations, 2023) and the Green Deal (European Commission, 2023). The Green Deal promotes forward-thinking actions and responsible resource management through terms such as sustainable production and sustainable decisions. These terms serve as specialized language. The importance of forward-looking planning and conscientious resource utilization is emphasised by these terms. A comprehensive strategy for managing raw materials is presented, which emphasizes savings while also promoting reuse and reduction of excess, such as construction waste. Although foundational elements in the construction industry, the sustainability factor involved in action and decision-making has not received enough attention in public discourse, despite its critical role in achieving the aforementioned goals.

The state of disrepair in infrastructure (Norddeutscher Rundfunk, 2023) has long been a public concern. The list of reasons for recent failures highlights structural flaws in the industry that neglect the life cycles of infrastructure such as bridges, roads, and railways. It is crucial to understand the temporal significance of structures and their relationship with sustainability. Infrastructure should not be viewed as disposable, but rather as part of a recurring cycle. They should not be forgotten once built. Adherence to academic norms, objectivity, and formal language is essential. The object life cycle can be divided into five phases: initiation (0), planning (1), execution (2), operation (3), and optimized initiation renewal (4). However, as the physical existence of the object is limited to phases 2 and 3, the term project life cycle is used to emphasize its place within the broader process. Sustainability is an important consideration in project management, which involves the overall execution of projects. Project management provides a more comprehensive framework for considering sustainability compared to the project life cycle.

This study aims to establish links between temporal processes and modern digitisation techniques. Additionally, this paper proposes a novel approach to developing project-specific temporal processes. This will result in clearly defined tasks that enhance project management in the construction industry. The ultimate goal is to contribute to the long-term sustainability and future prospects of the German construction sector through the use of TLCP.

Sustainability integration in lifecycle processes

In contemporary discourse, sustainability is moving beyond its traditional definition to become a multifaceted concept, often divided into three domains: economic, environmental and social (von Hauff et al., 2005). Current research efforts focus on effectively integrating these domains within work systems to facilitate concerted action, taking into account both immediate and future impacts. This holistic approach to sustainability, often referred to as the sustainable polygon, requires a thorough understanding of the underlying mechanisms, which requires iterative development and analysis in real-world contexts for practical applicability (Flemisch et al., 2023).

Skilled planners play a crucial role in the efficient use of available resources to achieve cost effectiveness in building construction, while ensuring suitability throughout the life cycle. This involves careful consideration of materials, processes and responsible resource management over the life of the building. In Germany, the Fee Structure for Architects and Engineers (HOAI) further delineates these phases, resulting in up to ten phase levels, with phase (0) serving as an initialisation or preliminary feasibility study.

However, this traditional approach to project billing in Germany has its limitations, particularly in addressing gaps and information loss during project execution, given the iterative life cycle of products and technological advances such as model-based computer systems and

Building Information Modelling (BIM). Several factors contribute to these shortcomings, including compliance with regulatory requirements, prioritisation of social needs over technical optimisation, and challenges such as lack of digitisation, transparency and effective communication (Weber-Lewerenz, 2021).

Figure 1: Stages within the project Management

Integrating sustainability principles into project management practices is essential to promote project sustainability. According to the German National Standard (DIN) 69901-5 (DIN e.V., 2023), project management encompasses various tasks, techniques and means necessary for initiating, defining, planning, controlling, coordinating and completing projects. This includes overseeing phases (0)-(3) of the product lifecycle, as shown in Figure 1. However, project management responsibility typically only extends to the end of the warranty period of the contracted services in phase (3).

Efforts to embed sustainability in project management include the simultaneous consideration of economic efficiency, material selection and social aspects (Silvius, 2017). For example, during the planning phase, conducting a product life cycle analysis can help assess the impact of actions on the three pillars of sustainability. Project management plays a critical role in either analysing the impact of sustainability measures or coordinating their implementation (Project Management Institute, 2021).

Holzbauer et al. (2021) further categorize sustainability processes in project management into three areas: internal processes, management of projects, and project-specific processes. Emphasizing a broader perspective on projects, sustainability in project management entails coordinating the overall process, networking internal and external influences, and applying known sustainability principles (Stumpf et al., 2012).

Sustainability in project management practices

To effectively embed sustainability into project management practices, it is imperative to adopt a comprehensive approach that encompasses various aspects of project execution. This involves integrating sustainability considerations into internal processes, such as resource allocation and risk management, to ensure that economic, ecological, and social factors are adequately addressed throughout the project lifecycle. Moreover, managing projects with sustainability in mind requires proactive measures to minimize environmental impact, promote social equity, and optimize economic viability.

Project-specific sustainability processes entail tailoring sustainability initiatives to meet the unique requirements of each project. This may involve conducting thorough assessments of environmental impact, implementing green building practices, and fostering stakeholder engagement to ensure alignment with sustainability goals. By incorporating sustainability principles into project management practices, organizations can enhance their reputation, reduce costs, and contribute to long-term environmental and social well-being.

In conclusion, sustainability in project management is integral to achieving lasting success in today's dynamic business environment. By integrating sustainability considerations into project planning, execution, and evaluation, organizations can mitigate risks, capitalize on opportunities, and create value for stakeholders. Through collaborative efforts and strategic initiatives, project managers can drive positive change and foster sustainable development for future generations.

Managing Information Transfer Complexity in AEC Projects

Effective information transfer is crucial for the success of AEC projects, yet it often presents significant challenges due to the multifaceted nature of project requirements and stakeholder involvement. This subsection delves into the various facets of managing information transfer complexity within AEC projects, highlighting key considerations and strategies for addressing these challenges with a focus on the German construction industry.

Navigating the regulatory landscape, including frameworks such as the HOAI, introduces complexities in information transfer processes. Compliance with regulatory requirements adds layers of intricacy to data exchange, necessitating clear understanding and adherence to legal standards. Moreover, ambiguity in information requirements further complicates the transfer process, as stakeholders may struggle to discern the specific data needed for each project phase. This ambiguity can result in inefficiencies and delays, underscoring the importance of clarifying information needs and ensuring compliance with regulatory guidelines (Mohan et al., 2021).

Effective information transfer hinges on seamless collaboration among project stakeholders, including architects, engineers, contractors, and clients.

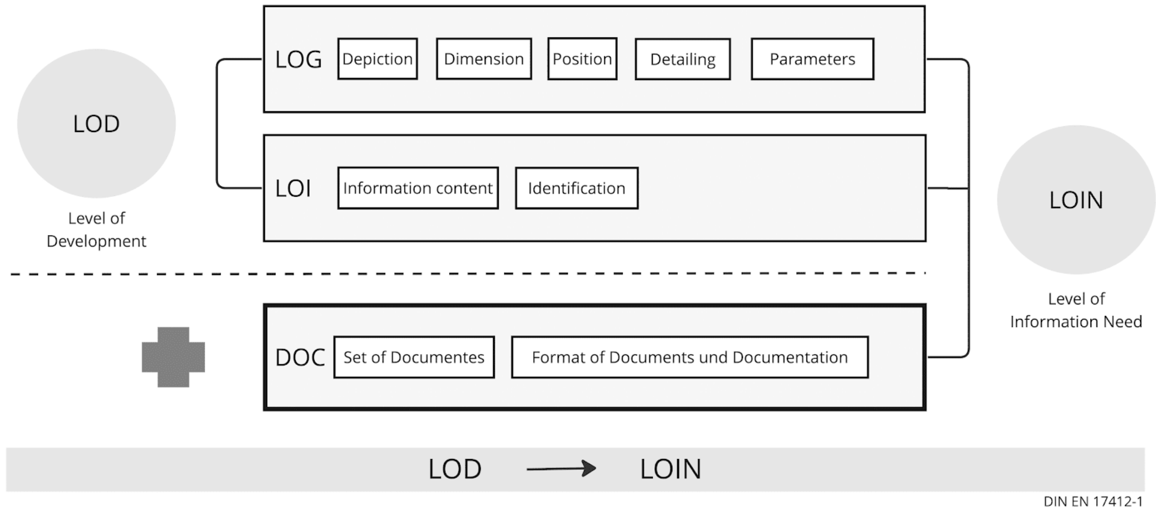


Figure 2: Level of Information Need (LOIN) (DIN e.V., 2023)

Coordinating diverse teams and aligning their efforts towards shared project goals requires robust communication channels and established information-sharing protocols. Clear delineation of roles and responsibilities, along with regular communication and collaboration platforms, fosters synergy among stakeholders, facilitating the smooth flow of information throughout the project lifecycle.

As all project participants have different starting points, it is essential to have a defined digital knowledge transfer to ensure accurate and transparent information transfer.

Table 1: Table 1: IRM concept based on LOIN use case with consideration of TLCP (Mohan et al., 2023)

Key Data	Content
Use Case	Why is the information required?
Milestone	When is the information required?
Objects	What / *Which Information* is required? (LOG/LOI/DOC)
Actors	Who required the information? ***Who delivers the information?***

Implementation of Digital Knowledge Transfer Solutions and the Effect of TLCP

In response to the challenges posed by information transfer complexity, digital knowledge transfer solutions emerge as indispensable tools for streamlining communication and data exchange. Centralized databases and information management systems provide a centralized repository for project-related information, enabling real-time updates and seamless accessibility for all stakeholders. Standardized information formats and protocols ensure consistency and accuracy in data transfer, reducing the risk of errors and discrepancies. Additionally, digital tools and technologies, such as Building Information Modelling (BIM) software, offer advanced capabilities for visualizing project data and facilitating collaborative decision-making, further enhancing the efficiency of information transfer processes (Vaatz et al., 2023).

Implementing the TLCP not only provides a comprehensive view of the product lifecycle in the AEC industry but also offers a strategic approach towards sustainable project management. Through a dIM, TLCP facilitates the structured capture, management, and exchange of project-relevant information, ensuring its availability and quality while fostering efficient communication, transparency, and collaboration among all stakeholders.

Real-time updates and easy accessibility of information in the dIM significantly enhance decision-making and planning accuracy, which are essential components in sustainable project management practices.

Incorporating sustainability principles into TLCP, Table 1 and Figure 2 outline the IRM and its components, such as Level of Geometry (LOG), Level of Information (LOI), and Documentation (DOC), forming the foundation for managing the dIM. Informed by insights gleaned from past and ongoing projects, TLCP development underscores the importance of standardized information requests and effective communication channels in sustainable project management efforts.

TLCP's capabilities, organizations can elevate their project management practices, leading to improved outcomes and greater sustainability across the industry.

Proposed methodology

Implementing TLCP in the current process relies on key factors, including adherence to national and international standards, digitalization through information models,

Table 2: IRM for Infrastructure se case with consideration of TLCP

Key Data	Information Requirements Matrix (Infrastructure)
Use Case	Position of road alignment
Milestone	Preliminary Design to Conceptual Design
Objects	LOG: Geology, hydrography data, environmental constraints, protected areas details, As-is information, information from previous phases (As-is, Feasibility, alternative variations) LOI: Weather data, socio-economic aspects related information, cost-time analysis, infrastructure and utilities, national standards and guidelines, material details, construction phase considerations DOC: Explanatory report, soil report, construction phase line report, safety report
Actors	From Preliminary Design (Who delivers): Owner (Public/Private), Legal& state bodies such as state-approved surveying and information department, specialist/consultant, Fire protection expert, energy suppliers In Conceptual Design (Who requires): Planners, Owners, state authority, civic bodies, energy suppliers

By defining appropriate LOIN use cases adhering to DIN EN 17412-1 (DIN e.V., 2023), TLCP establishes an organized information delivery system for inter-project phase processes, thereby minimizing the risk of unnecessary iteration loops and information gaps. The forthcoming dissemination of LOIN use cases through web-based platforms will adhere to established requirements, foster interdisciplinary collaboration, and ensure consistent scrutiny of information delivery procedures, which are both integral to sustainable project management objectives.

In summary, the Total Life Cycle Process (TLCP) emerges as a fundamental framework for seamlessly integrating sustainability considerations into project management practices within the Architecture, Engineering, and Construction (AEC) industry. By facilitating efficient information management, fostering collaboration, and enabling informed decision-making, TLCP empowers stakeholders to navigate the intricacies of complex project lifecycles while upholding sustainable principles. This comprehensive approach not only ensures the development of environmentally responsible projects but also contributes to the creation of socially beneficial built environments.

Looking ahead, the forthcoming web-based dissemination of Level of Information Need (LOIN) use cases represents a significant advancement in enhancing information exchange efficiency within the AEC sector. Through TLCP, this initiative not only enhances the precision of information delivery but also encourages interdisciplinary collaboration across project phases. By leveraging

systematic derivation of Information Requirements, cross-phase methodology, agile transition, process optimization, and fostering discussion and exploration.

In the course of the study, it emerged that various approaches have developed in the different phases, which require different information requirements. One example from infrastructure planning depicted in Table 2 is that the information from the planning phase must be prepared in such a way that authorisation is possible based on the applicable technical and legal requirements.

Efficient integration of the TLCP process requires a meticulous consideration of not only the primary criteria for information requirements but also the often-overlooked sub-criteria and sub-sub-criteria. Failure to inspect these finer details can result in significant time loss. By adopting this comprehensive approach, TLCP has the potential to evolve beyond a mere project standard to become a standardized practice within the entire organization. This shift ensures a thorough and consistent application of TLCP principles, enhancing overall efficiency and effectiveness across diverse projects and initiatives.

Implementation and integration of TLCP in a web-based information modelling tool

Implementing and integrating life cycle processes in a web-based information-modelling tool is achieved following the steps shown in the process picture (Figure 3) below:

The first step towards implementation of TLCP in a web-based information modelling tool involves identifying and defining all the life cycle processes that are relevant to the project, tailor-made for each life cycle phase, such as preliminary, conceptual, and detailed design, construction, maintenance, renovation, and demolition, among others. In this step, as listed in Table 2, the impacts, input and output, object types and actors with all inter- and outer relations must be defined in data and rule catalogues. Once the processes have been defined, they are to be modelled in the information modelling tools. After this stage, the data and information throughout the entire life cycle of the project would typically look like various islands, which in some cases are connected, and, in some cases, not, in which the data and information flow is not guaranteed during the entire life cycle.

The next step is to establish connections between the different processes. This is typically done through a process flow diagram or a similar tool that shows how the processes interact with each other and with the building model in so-called process catalogues.

Therefore, in the next step, an integration platform should be designed to establish integration between different systems, such as project management systems or cost estimation tools for additional functionality. Web-based platforms are one of the best examples of such smart services in which nested information can be kept, interconnected and analysed based on integration catalogues.

Such web-based information modelling environments are from one side capable of parsing data and models, and interlinking information, and from the other side connect them to resources from manpower to machinery and collaboration between human and technology using Artificial Intelligence (AI) methods. By establishing such connections between various sources, a comprehensive overview of the project lifecycle not only in the planning and design stages but also in the construction, operation and maintenance is ensured. Once the integration platform is set, different layouts can be designed to host the query and filter catalogues and apply them to the nested information based on the profile specs along phases or in the transition processes. These layouts define which data are to be shared with whom and when based on what expectations and with which results. Thus, information sharing between stakeholders and parties involved in the project is in a targeted way so that process-oriented information flow, collaboration and relevant communication is guaranteed. Such layouts can be purposed to target design, construction, operation and maintenance processes; transition processes; machine, human, and technology processes; or sustainability processes. The smart platforms have the capability to keep the information regularly updated and maintained to reflect changes over time with versioning and archiving mechanisms which is crucial especially in the operation, maintenance or renovation processes

In conclusion implementing and integrating TLCP in a web-based information modelling tool improves collaboration, provides more efficient processes, and better management of the project over its entire life cycle.

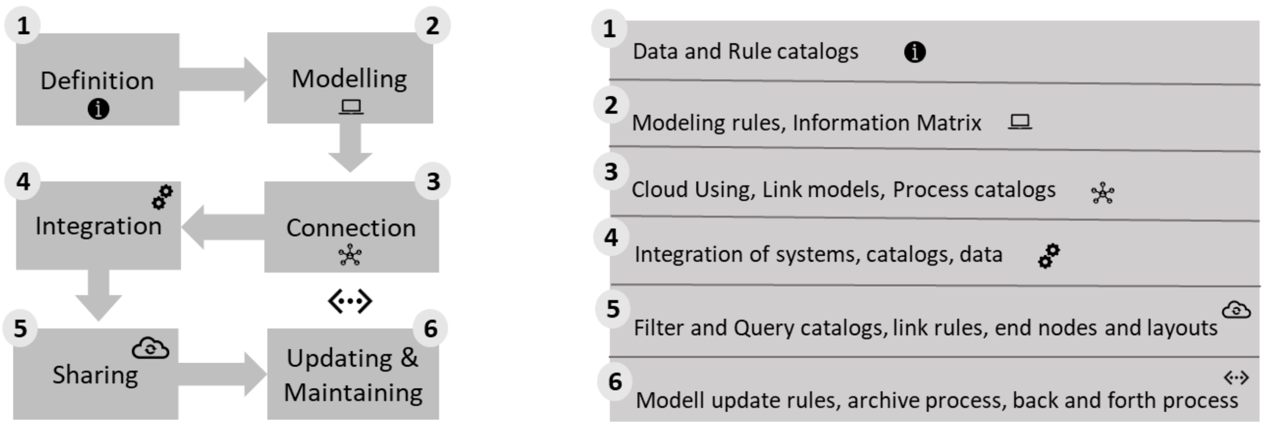


Figure 3: Integration of TLCP in a web-based information-modelling tool

Conclusions and Outlook

The integration of different data sources and information models based on the TLCP is essential to increase efficiency and ultimately promote sustainability in construction. The information management applied enables systematic, evidence-based decision making across the design, construction and management domains, ensuring seamless data exchange and communication between the various project stakeholders across all phases. To improve the flow of information and ensure system interoperability, industry players need to collaborate more and establish common standards and data exchange formats. One promising option is the incorporation of artificial intelligence and machine learning, which could provide the potential for automating analysis and decision support systems in the future. Further research and pilot studies, as well as more widespread use in practice, will drive the development of this methodology and increase its usefulness and practicality.

References

United Nations. (2023, December 05). The Sustainable Development Goals Report (Special edition) [Online]. Available: https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023.pdf

European Commission. (2023, December 05). Der europäische Grüne Deal [Online]. Available: https://eur-lex.europa.eu/resource.html?uri=cellar:b828d165-1c22-11ea-8c1f-01aa75ed71a1.0021.02/DOC_1&format=PDF

Norddeutscher Rundfunk. (2023, December 05). Marode Infrastruktur: Deutschland bröckelt [Online]. Available: https://www.ndr.de/fernsehen/sendungen/extra_3/Marode-Infrastruktur-Deutschland-broeckelt,extra21932.html

von Hauff, M., Kleine, A. (2005) Methodischer Ansatz zur Systematisierung von Handlungsfeldern und Indikatoren einer Nachhaltigkeitsstrategie - Das Integrierende Nachhaltigkeits-Dreieck, Discussion Report, Technical University of Kaiserslautern, Germany

Flemisch, F., Preutenborbeck, M. Wasser, J., Herzberger, N. (2023) Vom Teufelsquadrat zum nachhaltigen Vieleck (Engelsdiamant) und holistischen Bow-Tie-Modell: Denkanstöße zur Methodenentwicklung für die Analyse, Gestaltung und Entwicklung von nachhaltigen Arbeitssystemen, in Proceedings of the Frühjahrskongress, Hannover, Germany

Weber-Lewerenz, B. (2021) Die unternehmerisch verantwortungsvolle Digitalisierung im Bauwesen/Corporate Digital Responsibility (CDR) in Construction Engineering, Bauingenieur, vol. 2/2021, pp. 20-21

DIN e.V. (2023, December 05) DIN 69901-5 Project management - Project management systems - Part 5: Concepts, Beuth Publications

Silvius, G., Schipper, R., Planko, J., van den Brink, J. (2017) Sustainability in Project Management. Farnham, England: Gower Publishing

Project Management Institute (2021) A Guide to the Project Management Body of Knowledge (PMBOK® Guide). Project Management Institute Inc.

Holzbaur, U., Fierke, M. (2021) Nachhaltiges Projektmanagement. Wiesbaden, Germany: Springer Gable

Stumpf, M., Brandstaetter, M. (2012) Nachhaltigkeit im Projektmanagement – Bedeutung der Integrierten Kommunikation in der Innen- und Außendarstellung von Projekten. Germany: Springer Publications

Mohan, N., Ebbers, M., Gross, R., Theis, F. (2021) Design-Build-Verfahren aus Sicht der TGA – Wie können klare Verantwortlichkeiten und die konsequente Anwendung der BIM-Methodik helfen, Planungs- und Bauleistungen besser ineinander zu integrieren?, Build-Ing, vol. 2/2022

Vaatz, A., Hamdan, A., Al-Sadoon, N., Wogan, M. Menzel, K. (2023) Integration of semantic temporal information in BIM using ontologies, in Proceedings of the European Conference on Computing in Construction 2023, Crete, Greece

Mohan, N., Wolf, G., Wogan, M., Beilfuss, J., Gross, R. (2023) Auf dem Weg zu einem ganzheitlichen Lebenszyklusmanagement von Bauwerken für nachhaltiges Bauen: Eine vorgeschlagene Methodik zur digitalen Transformation durch datenbankbasierte Informationsmodelle, in Proceedings of the 34. Forum Bauinformatik, Bochum, Germany

DIN e.V. (2023, December 05). DIN 17412-1 Building Information Modelling - Level of Information Need - Part 1: Concepts and principles; German version EN 17412-1:2020, Beuth Publications

Nejat, N., (2023) Effiziente Infrastrukturprojekte," Intergeo 2023, Infrakit, Germany

DEVELOPMENT OF A PROCESS MODEL-BASED TOOL FOR CIRCULAR CONSTRUCTION SUPPLY CHAIN INTEGRATION

Lu Ding[1], Tong Wang[1], Wim Gielingh [2], and Paul Chan[1]
[1]Delft University of Technology, Delft, NL
[2]Real Capital Systems B.V., Nijkerk, NL

Abstract

The recent research focus on circular construction supply chains challenges the integration of supply chain processes. There is a need to build simulation tools for construction, facility management, deconstruction and reuse of building components. This paper introduces a model-based tool to enhance supply chain process integration developed with Business Process Mapping and Notation (BPMN). The model captures the remanufacturing supply chain for façade products that are demounted and reapplied in a different construction project. The paper highlights the tool's potential for predictive analysis and decision support based on supply chain performances. The demonstration shows an innovative approach to CE-driven supply chains.

Introduction

The construction industry is gradually adopting the concept of circular economy (CE) as one of the major transitional goals to reduce greenhouse gas emissions and waste throughout the full building life cycle. Many recent studies have reached a consensus that supply chain wide innovation is a prerequisite for this transition, as well as a focused research agenda for this topic in construction (Akinade and Oyedele 2019; Migliore 2019; Nasir et al. 2017). Construction supply chain management (CSCM), a field of research and practice that has been active for several decades, was introduced because of the growing demand of efficient construction processes and lessons learned from LEAN manufacturing (Eriksson 2010; O'Brien, London, and Vrijhoef 2002). In the background of CE, more studies have revealed strong interest in developing more circular and closed-loop construction supply chains (Adi Wijaya and Armyn Machfudiyanto 2023; Akinade and Oyedele 2019; Incorvaja et al. 2022).

Managing and coordinating construction supply chains is an ambitious task given its 'complex', 'project-based', 'fragmented' and 'contingent' nature (Koolwijk et al. 2018; Koskela, Vrijhoef, and Dana Broft 2020; Vrijhoef and Koskela 2000). Hence, supply chain modelling could provide support for the simulation evaluation, monitoring and prediction of processes (Cheng et al. 2010; O'Brien, London, and Vrijhoef 2004). Considering that CSCM is a vast subject that has been explored by multiple disciplines from various perspectives, it is difficult to find a single model that serves all needs. Both highly abstract mathematical models and more realistic simulation models have been extensively used for the purpose of analysis and optimization in CSCM (Chen and Hammad 2023). As the new digital systems such as digital twins are further developed, more integrated models and applied tools have been proposed to depict the complex supply chain networks and business processes (Cheng et al. 2010; Neuhäuser et al. 2023). Besides, there is also a clear trend of developing CSCM models for sustainability and CE related objectives. Recently, scholars have particularly highlighted the lack of models that integrate the R strategies such as reuse, recycling and remanufacturing (Chen, Feng, and Garcia de Soto 2022; Chen and Hammad 2023; Hussein et al. 2021). Despite that many studies have already addressed various types supply chain models for different challenges of CSCM, they are in most cases based on linear economy scenarios, characterized by forward flows of products and materials. What is missing, is the modelling of circular construction supply chain operations, that, according to the literature review by Ding, Wang, and Chan (2023), shall involve reverse logistics (RL) and process integration of both forward and reverse flows.

To expand knowledge and practices in modelling circular CSCM operations, this study proposes a process centric supply chain model that represents the remanufacturing supply chain for construction products. The model partially captures and simulates the complex remanufacturing procedures in the circular supply chain of recovered construction products. In the project case, wooden frame window and door components are recovered from one demolition site and remanufactured into new façade elements for another construction site in the city of Amsterdam. The framework of Supply Chain Operations Reference Model (SCOR) is used as the core structure to model and elaborate the supply chain processes and performance indicators. And further, the business process model and notation (BPMN) language is utilized to transform the supply chain processes with more detail levels into a digital representation with the potential to build executable software with the Camunda API. A demonstration tool is designed to be accessed by different supply chain stakeholders to interact with input/ output data and gain more insights for critical decision-making.

Theoretical Base for Modelling

Modelling Supply Chain Integration for CE

The efficiency of reverse supply chains can be improved by standardizing processes and procedures, even when the buildings an components are entirely different, and even when the disciplines involved are different. This may bring more commitment and capacities from the supply chain to CE (Koolwijk et al. 2018; Teixeira and Borsato 2019). It is particularly challenging for CE oriented construction projects, given that most of these projects are still experimental pilots There is limited evidence that such practices could be repeated at greater scales. Thus supply chain integration templates for circular projects, that support repeatable processes and continuous improvement, will be highly relevant for tackling the challenges of CE in construction.

A number of studies in the field of supply chain management connected the concept of supply chain integration with the maturity of business processes (Lockamy and McCormack 2004; Trkman et al. 2007). In these frameworks, the more 'mature' supply chain processes evolve from 'Ad Hoc' business operations, where processes are not well documented or structured, to more integrated and standardized processes. The collaboration between companies improves. More repeatable and predictable procedures are set up to reduce overall costs. The model in Figure 1 by Lockamy and McCormack (2004) defines the maturity of supply chain processes in five levels based on the reference to the key supply chain processes defined in the Supply Chain Operations Reference model (SCOR). SCOR is one of the most commonly applied standards to map supply chain configurations and processes (Teixeira and Borsato 2019). The framework defines the following distinctive processes in the supply chain: plan, source, make, deliver, and return. The SCOR features a hierarchical structure of processes, which consists of four levels of detail in process descriptions. The framework provides an appropriate starting point for practitioners to define the key business processes and sub-processes in a supply chain system (Supply Chain Council (SCC) 2017).

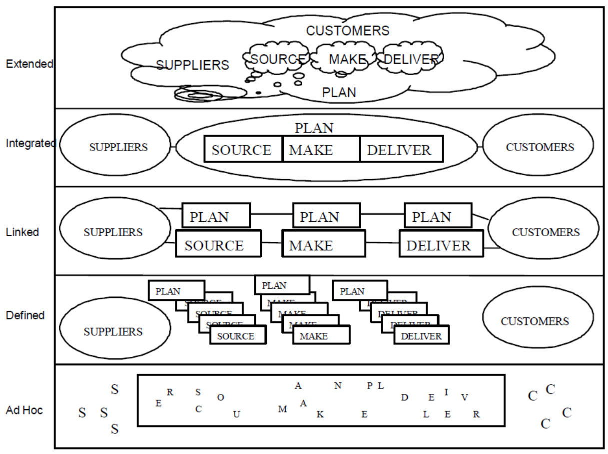


Figure 1: supply chain process maturity model, source: (Lockamy and McCormack 2004)

It would be a rather new experiment to utilize the theory of supply chain process maturity model and SCOR for CE driven initiatives, but there is clearly already a growing discussion on the potential of supply chain integration in circular and sustainable transition. Di Maria, De Marchi, and Galeazzo highlighted that the integration enabled by industry 4.0 technologies such as smart manufacturing play an mediation role for CE performance, because the enhanced links between actors in the supply chain positively affects the utilization of resources in a circular manner (2022). Similar conclusions are drawn by studies in the construction industry, addressing needs for integrating supply chain and logistics processes for the goal of CE in construction projects (Chen et al. 2022; Incorvaja et al. 2022). Modelling the processes of integrated supply chains will be a prerequisite for achieving industrialized CE.

Performance Goals for Circular CSCM

The Ellen MacArthur foundation (Ellen MacArthur Foundation 2013) makes a distinction between the technical cycle, which includes materials that have to be maintained and reprocessed as much as possible as part of a finite material stock, and the biological cycle that may be regenerated by the biosphere. In principle, circular supply chain management for the construction sector will incorporate both loops and should integrate forward and reverse flows of products and services in the full building life cycle (Ding et al. 2023; Farooque et al. 2019). In the ideal condition, a circular supply chain will require the minimum amount of primary resources for production and output zero waste, emissions or hazardous materials (Farooque et al. 2019). However, the more realistic goal definition of circular CSCM will be based on current frameworks of CE and CSCM concepts, that is to systematically integrate CE thinking and R-ladder strategies, such as recycling, remanufacturing and reusing, into the management of resources in construction supply chains, which involves all stakeholders in the construction project life cycle, including construction organizations, manufacturers and waste management agencies (Chen et al. 2022; Ding et al. 2023; Farooque et al. 2019).

The SCOR standard defines 5 main performance attributes: reliability, responsiveness, agility, costs and asset management efficiency. The standard also elaborates more performance metrics and best practices to describe and measure performance both quantitatively and qualitatively. Part of the performance aspects of SCOR already aligns with the performance goals of CE. Especially the cost and asset allocation performance is critical to circular CSCM, because currently the more desired strategies of reuse and remanufacturing of construction products are often hindered by higher material handling cost and increased logistics cost and mismatch between the moments of materials supply and demand. Therefore, in the current context of CE transition, it is reasonable to refer to the SCOR defined performance indicators and prioritize the cost performance while modelling circular construction supply chain processes, hereof the urgent practical goal is that more circular strategies may become more competitive while positioned against traditional linear supply chains.

BPMN

There are multiple ways to model business processes in a supply chain system. Languages such as BPMN and IDEF0 are upon those common standards used for business process management. BPMN is the de facto method for modelling supply chain processes due to its clear logic and visual capabilities. The language is also available in a number of commercial and open source software platforms for customizable development. The combination of SCOR and BPMN has been used by a few previous studies in the manufacturing industry for supply chain modelling (Cheng et al. 2010; Teixeira and Borsato 2019), where the potential to create dynamic and interactive models which industrial participants can understand is demonstrated.

BPMN models of the supply chain system could be approached in different configurations based on how organizational relationships are modelled. The BPMN orchestration model is the basic form that contains a single point of coordination. A collaboration model treats all processes as internal and focuses on only message flows between organizations for inter-organizational processes. The choreography model is designed to highlight interactions between organizations in contrast to internal processes (Eisner 2021). The purpose of modelling for the supply chain determines the method used in BPMN.

Model Construction

This paper proposes an explorative approach to model circular construction supply chain processes, based on demonstration cases of façade element remanufacturing. A few completed construction projects in the Netherlands already demonstrated as pilot projects to recover secondary building façade components and remanufacturer them into new building components in other projects. Based on informal interviews and workshops with the Dutch façade industry stakeholders, some key configurations of the processes are captured for this model. For the purpose of demonstrating the applicability of the model, one pilot case of façade element remanufacturing is used as an example. The case consists of two building sites, project A is a demolition site located in Amsterdam, from which wooden and glass elements from windows and doors have been demounted and then disassembled in factories to be made into new doors and windows for construction project B, which is at another location in the city. Because the part of the project related data is internal for the companies, some assumptions based on the interview and workshops and generated numerical data is used to deliver the proof of concept for the modelling and simulation process as shown in Figure 2.

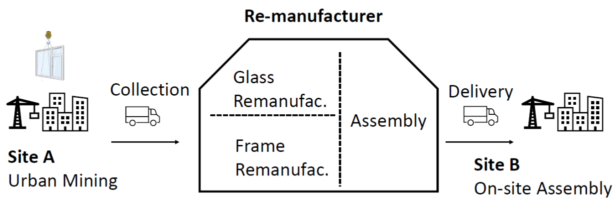


Figure 2: Illustration of the re-manufacturing supply chain

The hierarchical process model defined by SCOR 12.0 is used as the guiding standard for the modelling procedures as shown in Figure 3 (Supply Chain Council (SCC) 2017). The top-level supply chain configuration is modelled to the detail level 1 and 2 defined by SCOR to describe the supply chain actors' goals, responsibilities and process categories. For each actor or stage, key supply chain processes and sub-processes are modelled to the detail level 3 and 4 defined by SCOR and to be ultimately represented in executable BPMN models which could be used to monitor, improve or simulate supply chain processes and performances.

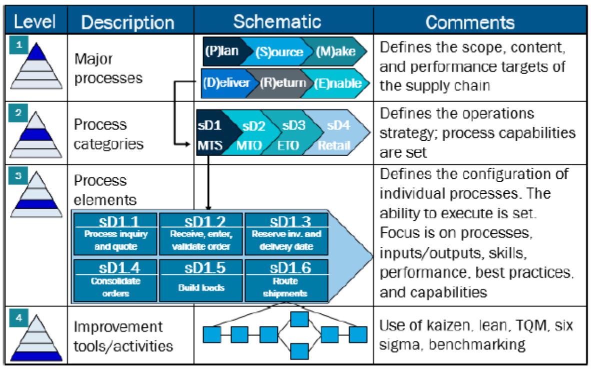


Figure 3: Hierarchical process model defined by SCOR
Source: (Supply Chain Council (SCC) 2017)

Furthermore, to incorporate the performance aspect of the supply chain model, several items from SCOR are chosen to be integrated into the supply chain processes as performance goals for the modelled system. As it will be too complex to implement all performance factors of the SCOR standard, only some of the cost and other matrix related to sourcing and delivery of products and the practice of just-in-time production (JIT) are included in this demonstration model. The reason to emphasize those aspects is that the cases from the pilot project particularly highlighted the challenges of the mismatch between the moments in demolition and construction projects and the increased cost and risk factors due to the increased logistics operations for remanufacturing. The performance indicators are embedded into the BPMN task elements for the evaluation stage of the project. Table 1 shows examples of the performance indicators, selected from SCOR 12.0, to be considered in this demonstration model.

Table1: Performance indicators taken from SCOR

SCOR	Name	Parameters
CO.3.16	Cost to Source Return	On-site treatment and storage cost
CO.3.17	Cost to Deliver Return	Reverse logistics cost to remanufacturer
CO.3.15	Order Delivery Cost	Transportation to construction site
BP.1.10	JIT production	Time of warehousing for products
…	…	…

As the core back-end structure of the model is defined in BPMN, the processes, tasks and required input/output resources are carefully modelled according to the structure defined by SCOR. For the Level 1&2 model, the supply chain for the remanufacturing process consists of four main stakeholders, the project coordinator (often the main contractor), the demolition contractor (urban miner) for project A, the remanufacturer (factory) and the building contractor for the new project B. In this case, the main configuration of the supply chain processes includes project planning, conducting urban mining, delivery to remanufacturer, remanufacturing, delivery to construction site, installation on site. For the Level 3&4 model, sub-processes are created as BPMN tasks to be assigned to the supply chain actors and are detailed for specific operations. The BPMN orchestration model is used to connect the tasks and information flows between different actors. As the remanufacturing supply chain follows a sequential order and the different actors are coordinated as sub-processes of a single project, the model type is suitable for demonstrating the overall supply chain configuration.

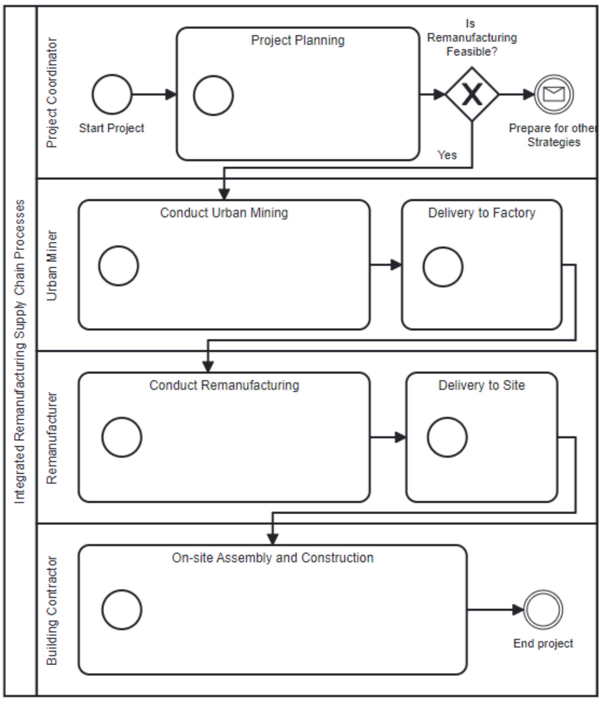


Figure 4: illustration of the main task flows in BPMN

Figure 4 illustrates how the three actors who are responsible for the production processes are controlled by a single point of coordination, which is in this case, the main contractor. Nevertheless, the orchestration model is not fully representative of the real timeline of production. as the four main actors conduct their activities simultaneously and the spatial-temporal coordination between those actors is much more complex than the diagram shows. In this study, the inter-organizational communication between the detailed sub-processes is left out for simplification. The main actors and key processes shown in this overall process diagram is a representation of the detail level 1&2 SCOR process model. The intent wasn't to strictly adhere to the pattern of SCOR for all detailed process definitions but instead to leverage the method and its principles to showcase a proof of concept in remanufacturing operations.

Based on the abovementioned supply chain structures, main process configurations and performance indicators the final BPMN model with detailed sub-tasks is done using Camunda modeler. By using the Camunda API for development, it is possible to attach forms or back-end JAVA classes to certain categories of tasks in the model. Thereby, the supply chain actors are given the possibility to interact with the model through input/output parameters and pre-defined calculations based on the data. Furthermore, the process model could also be executed via the "Deploy" icon within the Camunda portal, and to be accessed by actors through the task list in a web portal interface. More possibilities are allowed by the software architecture of the Camunda API as shown in the diagram Figure 5.

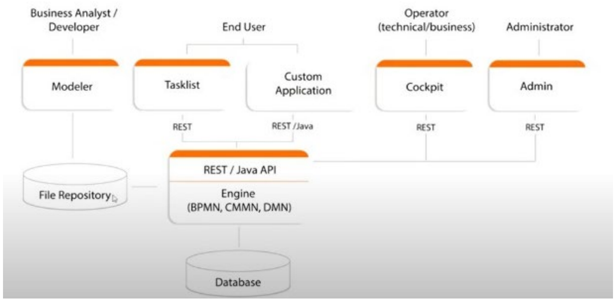


Figure 5: Camunda software architecture, source: (Techbuzz Blog 2022)

Results and Discussion

In this demonstration case, the project-based process of remanufacturing of wooden and glass façade elements is used as both the reference to model the supply chain structure, and also as the application example for potential model utilization. A process-centric model based on the real case is made because most supply chain models without real-world reference will be often quite abstract and could not possibly depict the complexity of processes involving multiple actors and dynamics in CE driven construction activities. On the other hand, the remanufacturing of building products is becoming a more popular and repeated practice in the Dutch construction industry, and the model could potentially be a reference model for other projects similar to the selected case.

Therefore the model is both case-informed and industry-wide generalizable.

The model demonstrates how the circular CSCM operations are modelled as a process based supply chain to achieve more 'maturity' in the form of an integrated supply chain system. The modelling procedure is based on the SCOR standard and BPMN. The execution of the BPMN process is done using the Camunda API which enables web-based user interfaces (UI). It could be accessed by different supply chain stakeholders to create input information and get insights into the supply chain operations in real time. The screenshot in Figure 6 shows how the stakeholders' can login the UI to get access to the executed processes and how the coordinator starts the BPMN process in his/her own dashboard. Despite incorporating limited quantitative data from the real project, the model comprehensively captures the processes for remanufacturing of the construction products in the case.

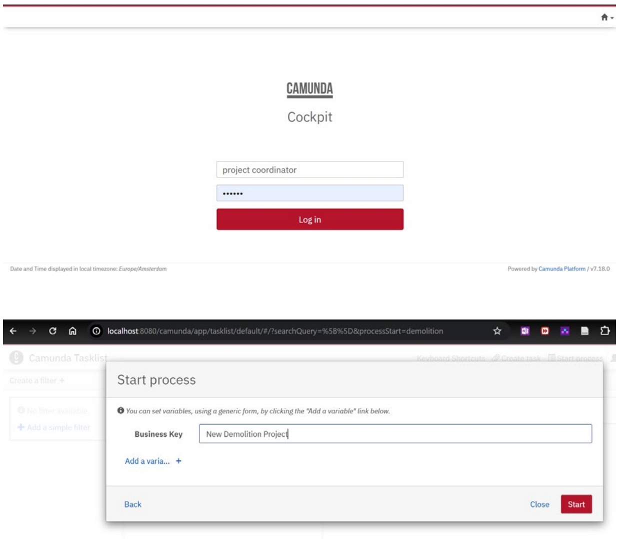


Figure 6: screenshots of the login page and start process

Afterwards, the integration of task parameters and performance indicators is enabled in this model through the Camunda API. This is achieved by creating forms in the Camunda web portal that gets input of variables for the setting of parameters of the model. For example, the transportation distance between the demolition project site, the manufacturer and the new project site; as well as the storage capacity for the locations. Those spatial and resource factors are required to make informative logistics planning between the project sites. As the new construction project may start much later than the deconstruction of the old building, the temporal dimension within the model is also crucial. Whether the schedule of different projects could fit into the resource planning, will drastically influence the performance outcomes such as JIT production and logistics cost. Therefore, the scheduled time of deconstruction, remanufacturing and new construction, are also included as model data and are acquired by the coordinator when initiating the projects. In this way, the model could already provide assumptions of logistics cost and JIT performance scores before going through the other processes. This demonstrates the prediction and simulation potential of the BPMN model that supports decision making before the execution of real processes. With the models and defined tasks, actors, and forms in place, the process can be "deployed" through the web portal. Beginning at the portal's "Start the Project" option, execution aligns with the process model.

In this demonstration case, the coordinator is able to check the feasibility of remanufacturing based on the basic input data about the project, while feasible, the process will be continued to initiate the multi-actor coordination for remanufacturing processes. In this model, the different actors will conduct detailed operations of the sub-processes that are pre-defined in the BPMN. Figure 7 shows screenshots of part of the executable BPMN model with embedded feasibility check for the selected R strategy, Arrow a highlights the tasks where the coordinator put in plannings and get results for feasibility simulation. Arrow b highlights the form interface for step by step data input of logistics information. Arrow c highlights the demonstration results as decision support reference for continuing the remanufacturing processes. In this model, the estimated cost to conduct RL of recovered products and deliver the new products, as well as the score for JIT production (logistics resource utilization) is calculated. The design intention of the model based tool is that the model could then integrate the different actor's detailed operational data and provide more accurate assumptions of the performance in real project timeline. However, as this study is a conceptual demonstration, the functions are not fully detailed to be tested in real projects. Different functionalities that are demonstrated in this study, and possible future development directions are listed in Table 2.

Table 2: Covered functionalities by the tool and progress of development by this study

Functionality & Applications	Progress
Standardize supply chain models of CE strategies in construction with SCOR	Demonstrated
Model supply chain processes in easy-to-understand manner (BPMN)	Demonstrated
Represent integration of different actors in the supply chain system	Demonstrated
Simulate supply chain cost and JIT delivery possibility	Partial Demo.
Allow actors to input and gain key data for model update	Partial Demo.
Simulate and compare performance of different CE strategies (e.g. reuse vs. remanufacturing)	To be Developed
Simulate different configurations of as-is and to-be models	To be Developed

The demonstration of the tool suggest multiple potential areas to apply the model-based tool to CE-driven construction activities incorporating the simulation and predictive capabilities: (1) to improve cost and planning of different CE strategies and reverse logistics operations by including the detailed process cost and time predictions in the task elements in BPMN. (2) to compare as-is and to-be scenarios when designing supply chain re-configuration (Trkman et al. 2007). (3) to continuously monitor supply chain performance by collecting key performance data from more projects where the tool is used.

centric approach brings the model one step closer to real life applications, because it brings more ease of interaction between stakeholders and the model through a process-automation software platform that executes the process model to refined details. This study has selected the commercial software Camunda platform for development and demonstration, similar tools are also

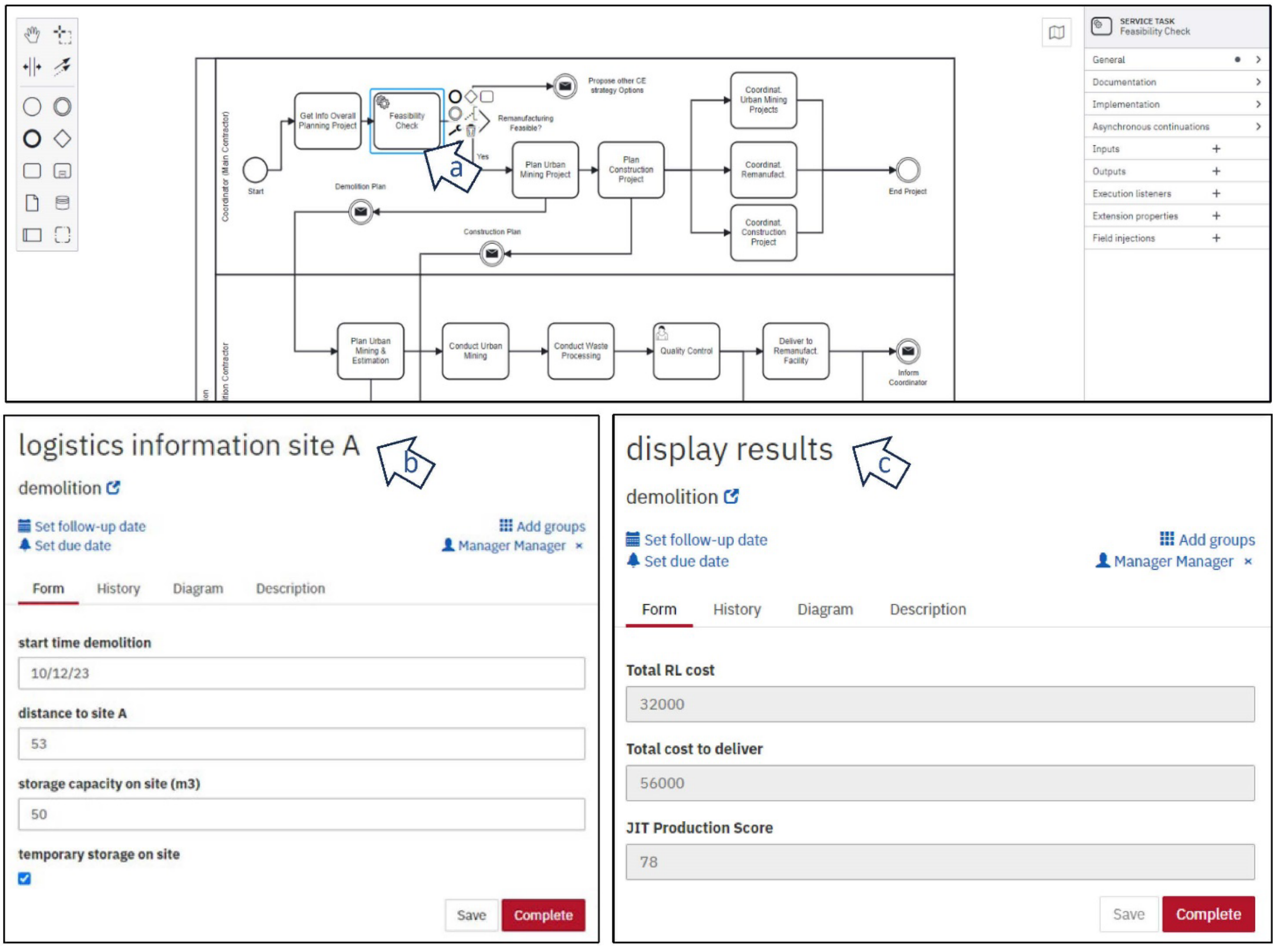


Figure 7: Example BPMN process diagram and the UI for input/ output data

Through post design workshops with experts in construction management from higher education, technical specialists from façade manufacturing companies and consultants for information systems, the tool for process centric supply chain modelling is further evaluated. The innovative approach and potential usefulness of the process simulation and execution tool for CE driven goals have been acknowledged by the potential users. The experts also highlighted that the feature of including cost estimation and monitoring in BPMN will be useful for making more informed decisions about CE in construction. More desired features could be added in the future to the model such as LCA / environmental impact of the processes.

The tool provides the simplicity of having supply chain stakeholders involved to have access to the processes, tasks and relevant parameters in one user interface (UI) that is executed with the BPMN. Compared to other methods to model supply chain operations, this process-available such as Bonita Soft or Eclipse.

One clear limitation of this study is that it is based on tools and standards previously applied mainly to manufacturing, which contains many features that are complex to apply to the project-based supply chain systems in the construction industry and circular economy. And further, the compatibility of the models with the different CE concepts and strategies is also limited. Moreover, the supply chain model tackles only one R-strategy of CE in practice. The more ambitious visions of CE driven supply chain models that may incorporate multiple R strategies are yet to be developed and tested.

Conclusions

The study proposes a process centric circular supply chain model for construction products based on the concept of process maturity in supply chain integration (supported by the SCOR standard) and the BPMN modelling language. It is a proof of concept to demonstrate how such models may act as new tools to improve supply chain integration for CE driven strategies in construction. In this case, the project case of remanufacturing façade elements is taken for the demonstration of their capability to coordinate supply chain operations beyond traditional construction and contractor-supplier coordination.

The model aims to capture the details of the remanufacturing supply chain for construction products, focusing on wooden frame and glass components recovered from a demolition site which are repurposed for a new project in Amsterdam. Four main actors are involved in the modelled case, the coordinator, the demolition contractor for the old building, the remanufacturer and the builder of the new project. For each actor, key supply chain processes and sub-processes are modelled according to the four hierarchies structure defined by SCOR and translated in BPMN for digital representation. Performance metrics such as JIT production and logistics costs from SCOR are integrated in the evaluation matrix for the supply chain model. Afterwards, the BPMN based model is then further developed into a tool for the design of supply chain collaboration, with a web-based interface using the Camunda platform. Through Camunda API, the model enables execution of BPMN tasks and allows for model deployment through a web portal for easy access by supply chain actors. It demonstrates the exchange between supply chain actors and the process model through a range of forms and parameters. Besides, the tool emphasizes the model's predictive and simulation potential, aiding key decision-making points before actual process execution. The study shows how process models are able to capture the complexities of interaction between demolition, remanufacturing and re-construction projects in CE, which traditional CSCM models and tools have not yet effectively addressed. Additionally, the interactive UI developed with Camunda API in this case, shows more multi-stakeholder decision support and collaboration possibilities beyond the process model itself, that may potentially accelerate the adoption of CE strategies by supply chain actors and help in upscaling the practices with more empirical evidence in future cases.

The key innovative feature of the tool is that it applies more mature supply chain process models from the manufacturing industry to cases in circular construction, and utilizes the advanced features from a BPMN software for process simulation and execution. The study tackles an important gap in CE and CSCM literature, which is the lack of understanding and experiments on process integration that connects the overall CE initiatives and strategies with the actual outcome of CE performance in construction supply chains. The study contributes to the methodological advancement for future research to explore modeling techniques for circular construction supply chains, highlighting the importance of standardization of supply chain processes and BPMN's suitability for process execution. By introducing a pilot case of façade element remanufacturing also demonstrates the applicability of the proposed framework to more practical situations, which may lead to more repetitive practices of such CE driven strategies in future construction projects.

The current model in this tool is still limited by the lack of more real-world project data to enhance its functions and improve the model configuration to track and predict the performance more accurately. Future works shall focus on elaborating the process models to suit more practical scenarios and further explore the information models and data integration with existing systems which may support more advanced process execution in more complex cases. Moreover, the simulation approaches of the process model shall be further developed to include more CE scenarios and more comprehensive performance indicators such as carbon emissions and environmental impact.

Authors' Contribution

First draft, L.D.; conceptualization, L.D. and T.W.; tool development, L.D. and T.W.; project coordination, supervision, W.G. and P.C.; review and editing, T.W. and W.G.;

Acknowledgments

The research started as part of the PerpetuAL project, funded by the Ministry of Economic Affairs in the Netherlands, and continued in the project Toekomstbestendige Leefomgeving (Future-proof Living Environment) funded by the Nationaal Groeifonds (Dutch National Growth Fund).

References

Adi Wijaya, P., and R. Armyn Machfudiyanto. 2023. 'Conceptual Framework of Green Supply Chain Management Strategy Selection in the Indonesian Construction Industry'. in Vol. 405.

Akinade, O. O., and L. O. Oyedele. 2019. 'Integrating Construction Supply Chains within a Circular Economy: An ANFIS-Based Waste Analytics System (A-WAS)'. *Journal of Cleaner Production* 229:863–73. doi: 10.1016/j.jclepro.2019.04.232.

Chen, Qian, Haibo Feng, and Borja Garcia de Soto. 2022. 'Revamping Construction Supply Chain Processes with Circular Economy Strategies: A Systematic Literature Review'. *Journal of Cleaner Production* 335. doi: 10.1016/j.jclepro.2021.130240.

Chen, Zixuan, and Ahmed W. A. Hammad. 2023. 'Mathematical Modelling and Simulation in Construction Supply Chain Management'. *Automation in Construction* 156:105147. doi: 10.1016/j.autcon.2023.105147.

Cheng, Jack C. P., Kincho H. Law, Hans Bjornsson, Albert Jones, and Ram D. Sriram. 2010. 'Modeling and Monitoring of Construction Supply Chains'. *Advanced Engineering Informatics* 24(4):435–55. doi: 10.1016/j.aei.2010.06.009.

Di Maria, Eleonora, Valentina De Marchi, and Ambra Galeazzo. 2022. 'Industry 4.0 Technologies and Circular Economy: The Mediating Role of Supply Chain Integration'. *Business Strategy and the Environment* 31(2):619–32. doi: 10.1002/bse.2940.

Ding, Lu, Tong Wang, and Paul W. Chan. 2023. 'Forward and Reverse Logistics for Circular Economy in Construction: A Systematic Literature Review'. *Journal of Cleaner Production* 388:135981. doi: 10.1016/j.jclepro.2023.135981.

Eisner, Michael. 2021. 'Orchestrating People and Systems with BPMN 2.0'.

Ellen MacArthur Foundation. 2013. *Towards the Circular Economy*.

Eriksson, P. E. 2010. 'Improving Construction Supply Chain Collaboration and Performance: A Lean Construction Pilot Project'. *Supply Chain Management* 15(5):394–403. doi: 10.1108/13598541011068323.

Farooque, Muhammad, Abraham Zhang, Matthias Thürer, Ting Qu, and Donald Huisingh. 2019. 'Circular Supply Chain Management: A Definition and Structured Literature Review'. *Journal of Cleaner Production* 228:882–900. doi: 10.1016/j.jclepro.2019.04.303.

Hussein, Mohamed, Abdelrahman E. E. Eltoukhy, Ahmed Karam, Ibrahim A. Shaban, and Tarek Zayed. 2021. 'Modelling in Off-Site Construction Supply Chain Management: A Review and Future Directions for Sustainable Modular Integrated Construction'. *Journal of Cleaner Production* 310:127503. doi: 10.1016/j.jclepro.2021.127503.

Incorvaja, D., Y. Celik, I. Petri, and O. Rana. 2022. 'Circular Economy and Construction Supply Chains'. Pp. 92–99 in.

Koolwijk, Jelle Simon Jowan, Clarine Joanne van Oel, Johannes Wilhelmus Franciscus Wamelink, and Ruben Vrijhoef. 2018. 'Collaboration and Integration in Project-Based Supply Chains in the Construction Industry'. *Journal of Management in Engineering* 34(3):04018001. doi: 10.1061/(ASCE)ME.1943-5479.0000592.

Koskela, Lauri, Ruben Vrijhoef, and Rafaella Dana Broft. 2020. 'Construction Supply Chain Management through a Lean Lens'. Pp. 109–25 in *Successful Construction Supply Chain Management*. John Wiley & Sons, Ltd.

Lockamy, Archie, and Kevin McCormack. 2004. 'The Development of a Supply Chain Management Process Maturity Model Using the Concepts of Business Process Orientation'. *Supply Chain Management: An International Journal* 9(4):272–78. doi: 10.1108/13598540410550019.

Migliore, M. 2019. 'Circular Economy and Upcycling of Waste and Pre-Consumer Scraps in Construction Sector. The Role of Information to Facilitate the Exchange of Resources through a Virtual Marketplace'. *Environmental Engineering and Management Journal* 18(10):2297–2303.

Nasir, Mohammed Haneef Abdul, Andrea Genovese, Adolf A. Acquaye, S. C. L. Koh, and Fred Yamoah. 2017. 'Comparing Linear and Circular Supply Chains: A Case Study from the Construction Industry'. *International Journal of Production Economics* 183:443–57. doi: 10.1016/j.ijpe.2016.06.008.

Neuhäuser, Thomas, Johannes Dimyadi, Christian Johannes Eckart, Franziska Wagner, Andrea Hohmann, Robert Amor, and Rüdiger Daub. 2023. 'Systematic Integration of Building Information and Simulation Models for the Automated Evaluation of Factory Layout Variants'. Pp. 0–0 in Vol. 4, *Computing in Construction*. European Council on Computing in Construction.

O'Brien, W., Kerry London, and R. Vrijhoef. 2004. 'Construction Supply Chain Modeling : A Research Review and Interdisciplinary Research Agenda'.

O'Brien, William, Kerry London, and Ruben Vrijhoef. 2002. 'Construction Supply Chain Modeling: A Research Review and Interdisciplinary Research Agenda'. *Proceedings IGLC 10.*

Supply Chain Council (SCC). 2017. 'Supply Chain Operations Reference Model (SCOR) Version 12.0'.

Techbuzz Blog, dir. 2022. *Camunda Architecture | TECH BUZZ BLOGS*. TECH BUZZ BLOGS.

Teixeira, Kellyn Crhis, and Milton Borsato. 2019. 'Development of a Model for the Dynamic Formation of Supplier Networks'. *Journal of Industrial Information Integration* 15:161–73. doi: 10.1016/j.jii.2018.11.007.

Trkman, Peter, Mojca Indihar Štemberger, Jurij Jaklič, and Aleš Groznik. 2007. 'Process Approach to Supply Chain Integration'. *Supply Chain Management: An International Journal* 12(2):116–28. doi: 10.1108/13598540710737307.

Vrijhoef, Ruben, and Lauri Koskela. 2000. 'The Four Roles of Supply Chain Management in Construction'. *European Journal of Purchasing & Supply Management* 6(3–4):169–78. doi: 10.1016/S0969-7012(00)00013-7.

DEFINING VALUABLE DATA AND STAKEHOLDER ENGAGEMENT IN DIGITAL BUILDING LOGBOOKS: A FRAMEWORK FOR EFFECTIVE DATA STORAGE

Andrea Ibba[1,3], Sun-Ah Hwang[2], Rubén Alonso[3,4], Rosamaria Olivadese[3] and Diego Reforgiato Recupero[1,3]
[1]Università degli Studi di Cagliari, Cagliari, Italy
[2]TU Delft | Delft University of Technology, Delft, Netherlands
[3]R2M Solution s.r.l., Pavia, Italy
[4]Doctorado. Centro de Automática y Robótica, Universidad Politécnica de Madrid-CSIC, Madrid, Spain

Abstract

This paper introduces a comprehensive Digital Building Logbook (DBL) framework aimed at addressing challenges in managing extensive data generated throughout the lifecycle of modern buildings. The framework establishes a fundamental structure, defining primary valuable data for DBLs to enhance the value of buildings or dwellings. It delineates essential data elements, guides stakeholder prioritization and engagement, and outlines effective storage methods. Additionally, we propose potential applications of the framework in real-world scenarios, extracted from the analysis of existing DBL implementations. This contribution significantly advances the Architecture, Engineering, Construction, and Operation (AECO) sectors by providing a structured approach to DBL development, offering early insights for future implementations.

Introduction

Data is generated at every phase of the building lifecycle, with each phase presenting distinct data requirements (European Commission et al., 2020). Stakeholders naturally seek precise data to address their specific needs at various points in the building lifecycle. However, the AECO sectors traditionally face three technical challenges to perform effective data collection and management within the sector (European Commission et al., 2020).

- The absence of a systematic approach for capturing, storing, analyzing, and organizing valuable data and information, resulting in data loss.
- Fragmented and dispersed data storage across multiple organizations and even departments within the same organization.
- Limitations in the accessibility to and interoperability of the data that is collected and stored by one actor with other stakeholders across the construction market value chain.

Essential data required by the various stakeholders often lacks a centralized repository, and a systematic approach to data collection, organization, and management is absent. Addressing the needs of all stakeholders thus proves to be a challenge. In response, the Digital Building Logbook (DBL) emerges as a proposed solution. As articulated by the European Commission et al. (2020), *"the capturing and maintenance of data and information is the backbone of the DBL."* However, a critical gap persists in understanding what constitutes valuable data for the diverse involved stakeholders. This paper therefore aims to bridge existing knowledge gaps and build a coherent understanding of valuable data in the DBL domain and examine how different stakeholders can engage with that data. The proposed framework adopts the perspective of a local, possibly multi-building, and restricted DBL, emphasizing a constrained data flow and excluding the consideration of national or European databases. The presented framework seeks to establish a versatile DBL capable of accommodating the diverse requirements of various stakeholders rather than being confined to a specific category. Recognizing that stakeholders may possess distinct needs and perspectives on valuable data, this paper endeavours to define valuable data by linking it not directly to stakeholder's needs but rather to indicators outlined in prior studies. These indicators are intended to furnish data or processed data calculations, thereby enhancing the DBL to deliver a multitude of crucial functionalities for all interested stakeholders. Despite this localized focus, the work is forward-looking, aspiring to establish a repository that, while initially limited in scope, is designed to be expansive and interoperable in the near future. Therefore, the main contribution of this article focuses on addressing the gaps that we have identified. To achieve this, the paper will complement the list of indicators proposed in the survey paper by Alonso et al. (2023) with a framework that defines a specific workflow and data flow for a potential DBL framework applicable in various real-world cases. This need constitutes the problem statement of this article. This innovative approach guides our research and aims to provide early insights for implementing DBLs that can adapted to different contexts and enhance their effectiveness in addressing stakeholder needs. In starting the analysis, the subsequent section introduces several recent works around DBLs. Here, particular emphasis is placed on comprehending the present landscape concerning the formulation of DBL frameworks and the comprehension of DBL data. The succeeding sections then present a theoretical framework for storing valuable data in DBLs, beginning with the strategies employed by the authors to elucidate the meaning of valuable data in the context of DBLs, followed by a set of guidelines on how to store this valuable data. The section thereafter presents the hypothetical application and validity of our framework in selected real-world scenarios observed in practice.

Related Works

In the realm of DBLs, literature is relatively nascent, and consequently, frameworks dedicated to DBLs are still in their early stages of development. While the body of work is emerging, several pioneering studies have laid foundational insights into the domain. Within these studies is worth mentioning Malinovec Puček et al. (2023), where the DBL data structures within the context of the EUB SuperHub project are defined. The framework presented in the article aims to provide an essential set of input data throughout a building's life cycle to compute the passport rating across three areas: energy efficiency, sustainability, and smartness; aligning with EU legislation requirements and anticipating future legislation. One considered research is by Mêda et al. (2022), where the authors focused on proposing a process-based framework for DBLs to address the lack of clear conceptualization of DBL processes, data requirements, and stakeholder interactions. The framework is developed using Business Process Modeling Notation (BPMN) flowcharts and aims to support the EU framework for data collection, management, and interoperability.

Our paper aligns itself with the current trajectory of DBL research, seeking to contribute to the existing body of knowledge by tracking and building upon recent advancements in DBL frameworks and their applications. Another notable scientific publication cited in the proposed paper is a comprehensive survey carried out in 2023 (Alonso et al., 2023), which encompasses all publications on the DBL indexed in prominent scientific search engines such as Scopus, Web of Science, and Google Scholar. This work holds significance as it provides a comprehensive overview of the existing studies on the DBL topic at large, also giving a focus on data storage and DBL indicators.

The publications of Gómez-Gil et al. (2022, 2023) were crucial in defining the parameters of valuable data in DBLs. In their series of scientific works, they first analyze the four most developed European DBL models to provide a defined overview to what constructs a DBL —including the identification of the key stakeholders in the DBL and potential user needs, as well as a proposed structure of indicators, data sources and functionalities— to identify gaps for future developments (Gómez-Gil et al., 2022). Their research then progresses to explore the capacities of DBLs as a gateway to link existing national databases in Spain and Italy, further exploring the available data sources and types that can support the services of a DBL (Gómez-Gil et al., 2023).

For a clearer insight into how valuable data can be employed in the services of DBLs, an understanding of what indicators are and how they process data is needed. In their work, Tupenaite et al. (2017) developed an integrated, hierarchically structured system of sustainability indicators that can be used in the sustainability assessments of new residential projects in the Baltic States. Similarly, Møller et al. (2018) developed the Sustainable Tool for Assessment, Planning, Learning, and Engaging (STAPLE) to create a single common method for a comprehensive and inclusive design and building process for sustainable buildings. How the two publications define sustainability indicators was crucial when building the theoretical framework for the proposed paper.

Apart from scientific publications, two recent European studies on DBLs are fundamental to this paper. First is the European tender [1] led by Ecorys et al.. The study aims to develop a common European model for DBLs (1) to promote tools and protocols that ease data use and sharing throughout the construction ecosystem and (2) to implement DBLs in a harmonious and interoperable way across the EU27 Member States. Findings from the project that are of high relevance to this paper include their final report on *Technical Guidelines for Digital Building Logbooks* (Ecorys et al., 2023), where the guidelines on setting up and operating DBLs under a common EU framework are discussed. The other European project referenced in the proposed paper is Demo-BLog [2] led by TU Delft et al.. Short for *Development and Demonstration of Digital Building Logbooks*, Demo-BLog is a Horizon Europe-funded project that seeks to exhibit how DBL data can serve to advance the current evaluation strategies for climate and energy transition implemented at various levels in Europe. Findings from the project that are relevant to this paper include the state-of-play analysis of selected best practices, where the potential role and scope of the DBL, including the central features, as well as data handling and governance issues are explored with stakeholders. The proposed paper combines findings from both literature and ongoing European studies to delve into the specific data elements that should be included in DBLs, how stakeholders should engage with this data, and how the data should be effectively stored and organized. This approach adds depth and specificity to the previous research on DBLs, providing practical guidance for implementing and leveraging DBLs in real-world scenarios.

Defining Valuable Data in DBLs

In the context of this paper, 'valuable' data is defined as essential files derived from different phases of the building or dwelling lifecycle. These files are deemed crucial for storage in a secure repository, as they contain data that can significantly contribute to enhancing specific categories of indicators.

The value of both, data and indicators, lies in their ability to assist in the decision-making processes and facilitate actions aimed at optimizing building performance and sustainability. Due to the dynamicity of the data requirements that the construction sector may have, our list of valuable data does not aim to present a comprehensive compilation of all possible data that can be gathered during the building lifecycle, nor are we excluding any type of data that

[1] See https://www.ecorys.com/case-studies/technical-study-for-the-development-and-implementation-of-digital-building-logbooks-in-the-eu/

[2] See https://demo-blog.eu/ for more information

may become valuable in the future. Our list and definition of valuable data, at this stage, is formed by an analysis of the state of the art, documentation from policymakers, and an examination of commonalities in similar approaches. It aims to provide and present our process of defining what is the valuable data as a foundation that can be used as a reference to characterize the information stored in DBLs. This paper departs from the notion highlighted in the conclusion by Alonso et al. (2023) that a DBL should avoid overlapping with other tools, including Building Information Models (BIMs) or Common Data Environments (CDEs). In essence, BIM focuses on detailed 3D modelling and information management during the building process and a CDE serves as a collaborative platform or centralized digital hub where construction project stakeholders can access, share, and manage any project-related data and documents. In contrast, a DBL encompasses a comprehensive repository of essential files from the entire building lifecycle, offering a secure storage solution and serving as a foundation for diverse stakeholder needs by providing enriched functionalities based on specific indicators. The proposed framework therefore aims to establish a foundational structure that helps to delineate the primary valuable data to be stored within the DBL to positively enhance the value of the building or dwelling based on different stakeholder needs.

In the realm of DBLs, data plays a crucial role in assessing opportunities and challenges across diverse optimization scenarios for buildings and dwellings. The services built around the DBL then process the relevant data to encourage energy renovation by providing relevant information to (1) design renovation roadmaps, (2) carry out maintenance strategies and (3) monitor and assess the decarbonisation progress of the building (Gómez-Gil et al., 2022). Such services are called "functionalities" (European Commission et al., 2020).

There are various functionalities available in DBLs to support optimization practices of buildings, from providing building analysis and status to conducting energy performance assessments (Gómez-Gil et al., 2022). To enable these functionalities, several necessary indicators must be defined.

An indicator refers to a specific metric or measurable value derived from the data collected, output, or processed within the DBL. These indicators are intricately linked to the valuable files stored in the DBL, as the data within these files can be extracted or processed to identify meaningful values. These values, in turn, serve as indicators that provide valuable insights into various aspects of a building's lifecycle.

For example, consider an EPC stored in the DBL. The data within this file, related to a building's energy efficiency, can be processed to calculate indicators such as the energy consumption rate, carbon footprint, or overall energy performance index. These indicators then offer valuable information that can be used to provide functionalities such as energy efficiency recommendations, comparison with industry benchmarks, or compliance checks with environmental standards. In this way, indicators play a crucial role in unlocking diverse functionalities within the DBL, tailored to meet the needs of different stakeholders involved in the building lifecycle.

At the moment, the European Union (EU) does not provide an official list of minimum indicators that should be included in the DBL, and there is no consensus among the existing initiatives and proposals (Gómez-Gil et al., 2023). The role of a DBL can be seen to serve as a database for storing and processing diverse indicators, shedding light on the conditions and status of a building. These indicators are contingent on the type of data collected or available within the DBL, influencing the outputs derived. A range of indicator categories that aligns with the various data types should therefore be integrated into a DBL to offer information about the building that is of value. Based on an analysis of all the scientific publications on the DBL and its indicators in 2023, Alonso et al. (2023) derived nine indicator categories in their survey, as can be seen in Table 1 (see next page).

The identified indicators require various data to provide the functionalities they can support. To ease the identification of the necessary data, the Demo-BLog project developed a data template that also documents the data types, formats, and sources. The template was initially created to identify a common data language and/or categorisation method that can be applied to the five currently operational DBL initiatives studied in the project: **Chimni** that operates in the United Kingdom, **Woningpas** in Flanders, Belgium, **CLÉA** in France, **CAPSA** in Germany and **CIRDAX** in the Netherlands. The data fields identified in this template can thus be seen as the result of studying every functionality and data field employed across the five initiatives studied, with the terms and scope generalised thereafter to encompass the varying approaches into one universal table. In total, Demo BLog identified 79 data fields across the five cases studied that can be considered valuable to a DBL.

However, a practical challenge identified in Demo-BLog is the inconsistencies in the data structure between various data sources. For instance, while all utility providers furnish data on the energy consumption of a building or dwelling, they do so in different formats and methods. Though valuable data is abundant, the absence of a standardized format or scheme to represent the data remains a barrier for DBLs when coordinating with more than one data source. While this may lead to challenges in data organization and interoperability within a DBL, it does not necessarily depreciate the value of the data stored therein. This paper thus sees the need to compile a catalogue of pertinent files and/or documents associated with a building or dwelling, essential for storage within the DBL. These documents should be accessible to various stakeholders or services, facilitating the computation, storage, or retrieval of indicators crucial for building assessment.

As a starting point, the framework initially emphasizes the

Table 1: Indicators categories, indicators role and metrics as identified by Alonso et al. (2023)

Indicator Category	Indicator Role	Indicator Metrics
Building Information Indicators	Measures the efficiency of the building's information management practices	Accuracy, completeness, ease of access, and security of information
Energy Consumption Indicators	Reveals the building's energy usage, aiding in identifying areas for improvement	Energy consumption over time, meter readings, real-time energy usage
Indoor Comfort Indicators	Provides information on indoor environment aspects	Temperature, humidity, air quality, and lighting levels
Maintenance Indicators	Assists in tracking maintenance tasks such as cleaning, HVAC checks, and equipment replacement	Schedules, checklists, real-time maintenance notifications
Occupancy Indicators	Displays the number of people in the building, optimizing energy use and resource allocation	Number of people
Safety Indicators	Offers information on building safety	Fire alarm status, emergency lighting, evacuation routes
Financial Indicators	Shows financial performance	Operating costs, revenue, return on investment
Sustainability Indicators	Provides insights into sustainability performance	Carbon footprint, water usage, waste generation
Smart Readiness Indicators	Reflect a building's "smartness"	Metrics based on a common EU scheme

inclusion of files or documents that inherently contribute value to the DBL and the building or dwelling. The focus is on saving files without immediate data extraction. Each stored file should be associated with a record entity, allowing users to input essential information, for instance, the issued date for an Energy Performance Certificate (EPC). This record entity serves as a dynamic repository, capable of being updated with information extracted from the files or additional services, thus laying the groundwork for diverse functionalities.
In the following, a non-exhaustive list of valuable files/documents is identified for storage purposes in a DBL for future reference, data extraction, or data processing:

- **Baseline Industry Foundation Classes (IFC) Files**: These files contain the foundational information about a building or dwelling in IFC format that can serve as a reference point for subsequent changes or evaluations.

- **Energy Performance Certificates (EPCs)**: EPCs provide a detailed overview of the energy efficiency of a building, indicating its environmental impact and offering insights into potential improvements.

- **Carbon Bill**: This document outlines the carbon footprint associated with the building or dwelling, detailing the carbon emissions generated through its lifecycle.

- **Building Renovation Passport**: A comprehensive record detailing the history of renovations, upgrades, and modifications to the building, offering a holistic view of its evolution.

- **Selected Renovation Scenario IFC Files**: Similar to baseline IFC files, these documents specifically capture information on selected renovation scenarios, providing a targeted perspective on planned changes.

- **Post-Renovation Analysis**: This file contains assessments, data, and analyses conducted after a renovation, offering insights into the impact of the interventions.

- **Facility Management Reports**: Detailed records on facility maintenance, operations, and management activities.

- **Sensor Data Logs**: Selective storage of sensor data in a DBL, focusing on key metrics such as occupancy and temperature that are relevant for building operations and efficiency, rather than comprehensive data like in a digital twin.

- **Fire Alarm System Reports**: Documentation and analysis of the status, functionality, and events related to the fire alarm system within a building.

Table 2 links each of the specified valuable files with the associated indicator categories that are previously outlined. This establishes a direct correlation between the valuable files and the pertinent data they contain that contribute to the indicators, and further, functionalities, within the DBL. Consequently, this connection aligns with the overarching objectives of the DBL, enhancing the overall value proposition for the building or dwelling.
In conclusion, is worth noting that the valuable files outlined in this section do not constitute an exhaustive list of all possible files that can be stored in a DBL. Instead, they

Table 2: *Association of Valuable Files with Indicator Categories. BII = Building Information Indicators, ECI = Energy Consumption Indicators, ICI = Indoor Comfort Indicators, MI = Maintenance Indicators, OI = Occupancy Indicators, SI= Safety Indicators, FI = Financial Indicators, SuI = Sustainability Indicators, SRI = Smart Readiness Indicators*

	BII	ECI	ICI	MI	OI	SI	FI	SuI	SRI
Baseline IFC files	X			X					X
EPC	X						X		
Carbon Bill		X					X	X	
BRP	X							X	
Renovation Scenario IFC	X			X					
Post-Renovation Analysis	X	X		X					
Facility Management Reports				X		X			
Sensor Data Logs		X	X		X				
Fire Alarm System Reports						X			

represent foundational information that is present in real-world scenarios. These files are instrumental in generating valuable data, thereby enhancing the overall value of the building and augmenting the utility of the DBL.

Valuable Data and Stakeholders

Stakeholders along the construction and built environment value chain are essential in gathering, extracting, and processing valuable data intended for storage in the DBL. How this paper defines a DBL stakeholder is grounded on the findings of the study on the development of DBLs in the EU, funded by the European Commission, where five main user groups are defined (Ecorys et al., 2023):

- **Governmental agencies** that need data for evidence-informed policy-making, the issuing of licences and the enforcement of regulation and disaster management. This user group comprises all levels of governance, including national, regional and local levels.
- **Construction companies** need data to obtain and report building-related information during the design and execution phases of a building lifecycle. The sector encompasses various building professionals, including architects, engineers, contractors and real estate developers, and the collaboration between the relevant stakeholders is crucial to identifying the varying data needs and responsibilities.
- **Building owners and users** who need data on their buildings for monitoring, maintenance and necessary interventions. Building owners have an added responsibility to provide direct or indirect access to building information to not only the building users but also to other relevant stakeholders of the DBL.
- **Financial institutions** that need data to perform various analyses on the assets market and its developments, to gain a better understanding of building transactions. Here, increased transparency and data quality provided by the DBL is key.
- **Utility companies** that need data that provides information on the connection of utilities and the analyses on how they are used to provide detailed information with regards to the building use or performance.

To contextualize stakeholders' interests in data types with real-world DBLs, we have compiled a table that correlates each stakeholder group with the valuable data types that can effectively add value for them. This association is derived from the description of the stakeholders' groups and the data commonly used in their activities. Additionally, the table indicates the indicator groups related to each stakeholder group. These associations are detailed in Table 3.

Moving forward, in the subsequent part of this section, we will examine three real-world DBL implementations among the five studied in the Demo-Blog project, as presented in Section *Defining Valuable Data in DBLs*, to explore their operational mechanisms, identify the most valuable data for each, and present realistic use cases that use some valuable data previously identified.

- The **Woningpas** DBL initiative in Flanders, Belgium, is automatically available to building owners, including individuals and housing companies. Owned by four Flemish Government bodies, it aims to enhance energy efficiency in housing stock. Key stakeholders, including government agencies like VEKA and OVAM, collaborate to ensure energy labels meet regional standards. The Energy Performance Certificate (EPC) is a crucial document supporting this goal by providing essential data for planning energy efficiency interventions and continuous improvement efforts in homes.
- The **CLÉA** DBL, owned by Qualitel, emphasizes user-friendliness. It offers an equipment module with user guides and maintenance alerts for HVAC systems. This feature connects building owners with utility companies to ensure home comfort. Facility management reports tracking service changes are vital for this functionality.
- The **CIRDAX** DBL is a commercial digital materials database operating in the Netherlands and Belgium. It catalogues building components and materials obtained through 3D scanning and manual services. Its focus lies in creating a marketplace for secondary materials, connecting demand with supply for building

Table 3: Association of Stakeholder Groups with Indicators and Valuable Data from Table 2

Stakeholders	Indicators	Valuable Data
Governmental Agencies	BII, ECI, MI,Sul, SRI, FI	IFC files, EPC, BRP
Construction Companies	BII, ECI, MI, OI, SI, FI, Sul, SRI	IFC files, EPC, BRP, Renovation scenario IFC, Post renovation analysis, Fire alarm system report
Building owners and users	BII, ECI, ICI, MI, OI, SI, FI, Sul, SRI	All valuable data in table 2
Financial institutions	FI, Sul, SRI	IFC files, EPC, Carbon bills
Utility companies	SI, SRI, BRP	EPC, Facility Management reports

reuse. Key documents like the baseline IFC file, detailing building construction, and the carbon bill, calculating embodied carbon, engage the construction sector in circularity goals by providing crucial data for material reuse and carbon impact assessment.

The examples highlight the importance of considering the unique context of each DBL — including its founding goals, stakeholder domains, and primary target market — when designing and testing data storage, extraction, processing, and transfer methods. Considering the DBL as a tool that can accommodate various perspectives, priorities, and challenges is crucial in this process. With this in mind, the following sections present a framework that could be utilized in these scenarios. Subsequently, in the next section titled *Framework Alignment*, we will illustrate how our framework could be applied to these case examples, leveraging the valuable data previously presented to provide insights to different stakeholder groups.

Data Storage Guidelines

In this section, we anchor our approach in the context of a local DBL. These guidelines are crafted to establish a framework applicable to individual buildings or local DBLs, laying the foundation for potential expansion toward interoperable or national DBLs. Recognizing the dynamic nature of the DBL landscape, our focus is on constructing a framework that leverages existing technologies and data processes, minimizing the costs associated with creating new ones.

The suggested strategy operates under the premise that the DBL functions as a cloud-based solution. Consequently, the data storage for the DBL is also envisioned as cloud-based, ensuring data availability.

Our approach conceives the DBL as a “sink” for data, serving as a central storage repository where building files containing valuable data are stored. Consequently, our approach concentrates on DBLs focused on data archiving and not on having pointers or links to data in external tools. Another important point, crucial for security, is to encrypt every file in the DBL. Access to these encrypted files will be strictly controlled by an access management system, ensuring that only authorized stakeholders can retrieve them. Access can be facilitated through either a user-friendly interface or an Application Programming Interface (API). This solution needs to accommodate three primary types of files: (1) documents and binary data formats, (2) computer-readable data, and (3) key values. By encompassing these categories, we aim to cover the entire spectrum of considered valuable data.

To support these types and formats of data and to be flexible in both formats and semantics, an architecture with a data layer that manages these different options is considered. This layer should allow higher layer components (e.g., the aforementioned semantic analysis) to manage and handle this data, before its presentation to the user.

In this framework, files remain unaltered once uploaded, accompanied by a record detailing vital file characteristics. Stakeholders responsible for uploading, whether users or services, must store this record simultaneously with the file, enabling other services to access files, extract crucial information, and insert it within the record or as an output calculated with that information.

To enhance interoperability, the record entity is intricately linked to one or more files, allowing users or authorized services to systematically append relevant information. To address the absence of a unanimous consensus on defining a clear data structure for DBL data, we propose storing significant files alongside a flexible Record structure. This structure, while inherently flexible and customizable for various needs, should incorporate standardized essential metadata, such as the associated building, upload date, user responsible for the upload, and a log documenting any modifications.

Embedded within this framework, a fundamental entity representing the building or dwelling will encapsulate key information, including address, the number of floors, and other pertinent details. This central entity will serve as the linchpin connecting each record to its associated building. By referencing the building entity in all records, we establish a cohesive structure that facilitates the organization of valuable data and files in a scalable manner. This design choice not only enhances the coherence of the DBL but also contributes to its scalability, accommodating a broad spectrum of building types and configurations.

For instance, imagine a scenario that highlights the flexibility of the record structure within a DBL, particularly with storing a BRP. Stakeholders uploading records have the option to include additional details such as issuance and expiration dates, or they may choose to provide only mandatory information like the associated building. This dynamic structure allows for various use cases and enables modifications to align with semantic model structures.

In essence, our data storage guidelines are rooted in practicality to provide a scalable and adaptable framework that aligns with existing technologies, while simultaneously laying the groundwork for potential future developments.

Framework Alignment

In this section, we will explore the hypothetical application and validity of our framework in the use cases introduced in the *Valuable Data and Stakeholders* section.

Aligned with the proposed framework, the DBL employs a robust data storage strategy. Each file, including the EPC, the facility management report, and the building IFC is encrypted to ensure security. Access is meticulously controlled, permitting only authorized stakeholders – homeowners, government agencies, energy consultants, or service providers – to retrieve these files. Stakeholders can access files through a user-friendly interface or APIs, emphasizing accessibility.

As previously mentioned, to maintain data integrity, uploaded files remain immutable, and stakeholders or applications responsible for uploading the data are required to accompany each file with a dynamically structured record. This record, linked to the file, is populated with metadata providing essential details such as the date of issue or expiry.

Stakeholder engagement is a pivotal aspect. In the **Woningpas** use case, homeowners, responsible for storing documents, enjoy flexibility in contributing additional details to the associated record. Although the identity of the stakeholder uploading the EPC might not be certain, the process remains straightforward, facilitated through either the user interface or APIs. Upon EPC upload, a record entity is created, detailing the associated building, the uploader, and the upload date.

This structured framework facilitates collaboration. Stakeholders with access can download the EPC and utilize it for various calculations. The outputs, which could range from Key Performance Indicators (KPIs) to specific energy values or recommended interventions, can be stored as files with a record linked as the case of valuable data or even only with a record. This process, whether through user interaction or automated APIs, reinforces the scalability and adaptability of the framework, promising future expansions in the DBL landscape.

Within the **CLÉA** scenario, a facility management report is primarily generated by utility companies that provide services and products such as energy provision and HVAC equipment to homes. The collected data, encompassing both their professional insights and meter-recorded utility usage, can be efficiently uploaded to the DBL through APIs or the user interface. The data may take a static form, such as documents containing processed information handed over to homeowners as statements or transaction receipts, effectively serving as the facility management report.

In our DBL framework, for simplicity, we focus on static data uploaded as files, which will not change over time. From this file, key stakeholders can perform calculations to determine when the next services, possibly maintenance tasks or updates, are scheduled to take place. This process, seamlessly integrated into the DBL, aligns with our framework's emphasis on practicality, scalability, and adaptability. It provides a robust foundation for potential expansions in the DBL landscape while ensuring valuable data remains organized and accessible.

Lastly, in the **CIDRAX** scenario, the baseline IFC file stored in the DBL is created collaboratively by stakeholders involved in the construction or renovation processes. The document serves as a comprehensive snapshot of the building's construction, offering a thorough understanding of its structure. It includes intricate details about the type and quantity of materials used, their spatial distribution within the building, and other relevant specifics. The active participation of stakeholders such as architects, material providers, and contractors is thus essential during the construction or renovation stages to ensure the accuracy of this data. This information is then stored in the DBL, making it accessible to authorized stakeholders.

Authorized stakeholders, including companies like CIRDAX, can retrieve this stored baseline IFC file from the DBL. The document becomes a valuable resource for CIRDAX's objective of promoting circular building practices. By analyzing the detailed information within the baseline IFC file, stakeholders can identify the availability of reusable materials, quantification of materials and optimization of construction practices. After identifying this information, it can be incorporated into the DBL as an immutable file, accompanied by a record entity. This ensures accessibility for all stakeholders while being securely stored in a protected repository.

This section lays the groundwork for conceptualizing DBLs as repositories of pertinent information. Starting from such realistic scenarios, derived from empirical observations of the data types already used in real-world DBLs and their purposes, we generalized the processes to establish paradigms on how a DBL should function. This information will be the source for the definition of an architecture and the elaboration of a platform that can be validated by using empirical data and quantitative and qualitative indicators. Qualitative processes such as user satisfaction surveys and stakeholder interviews can provide insights into the usability and effectiveness of the framework. On the quantitative side, metrics such as the average time taken to retrieve information and analysis of API usage patterns, including the frequency and volume of API calls and the data used by different stakeholders, can measure the efficiency and relevance of the data.

In conclusion, the hypothetical application of our proposed framework across diverse real-world use cases demonstrates its robustness, scalability, and adaptability within the dynamic landscape of DBLs. Our emphasis on data security, controlled access, and the integration of dynamic record entities ensures the reliability and integrity of valuable data. Stakeholder engagement remains at the forefront, providing flexibility for various users to contribute and access information through user-friendly interfaces or APIs. The structured framework not only supports current needs, as illustrated in the presented use cases but also promises a foundation for future expansions in the evolving DBL landscape. Through this approach, we aim to facilitate collaboration, enhance data organization, and uphold accessibility, fostering a more effective and sustainable management of building-related information.

Conclusion

Buildings generate increasing amounts of data and information, which end up scattered in disparate databases. This volume and distribution of data lead to difficulties and challenges in making informed decisions based on it. DBLs have emerged as a key solution to bridge these data availability and archiving challenges.

However, not all data is equally important for making decisions during the building life cycle. Plus, there's still no clear agreement on what DBLs should do, what data they need, and how they fit with other construction software.

It is important to recognize that certain aspects, such as the detailed structure, components, and features of the DBL that will implement our framework, require further elaboration and exploration. These areas should be considered as part of future work to enhance the understanding and implementation of DBLs in real-world scenarios. By acknowledging the need for continued research and development in these areas, we aim to contribute to the ongoing advancement of DBL frameworks and their effective utilization in the construction industry.

The analysis done in this article also allows us to identify some research gaps, which will enable us to frame future work. Specific aspects of these research gaps include:

- **An Undefined Criteria for Valuability**: The absence of universally accepted criteria for categorizing data as "valuable" in the context of DBLs. This article has attempted to find some common ground and offer suggestions on this topic.
- **Inconsistencies in Data Structure**: The absence of a standardized format or scheme to represent valuable data, leading to inconsistencies in data organization and challenges in interoperability between different stakeholders.
- **Limited Stakeholder Engagement**: A knowledge gap exists among various stakeholders, including architects, engineers, and maintenance personnel, regarding a standardized approach to systematic data collection and storage.
- **Overlapping and Interoperability Challenges**: The lack of clear differentiation between distinct solutions in the construction ecosystem (e.g. Common Data Environments, Digital Twins, Construction Management Software), what data each manages, processes or stores and how they interoperate.

Our future work is in line with addressing these gaps and defining an optimal architecture to support the data mentioned in the article. This, together with research on ontologies, such as Smart Applications REFerence (SAREF)[3], which can be used to enrich the data, will allow us to design an efficient and adaptable reference DBL, and properly compare and evaluate the available DBLs.

Acknowledgements

This research has been partially funded by the European Union's Horizon research and innovation programme, as part of CHRONICLE project - Building Performance Digitalisation and Dynamic Logbooks for Future Value-Driven Services (Grant agreement ID 101069722).

References

Alonso, R., Olivadese, R., Ibba, A., and Reforgiato Recupero, D. (2023). Towards the definition of a european digital building logbook: A survey. Heliyon, 9(9).

Ecorys, TNO, Arcadis, and Contecht (2023). Technical study for the development and implementation of digital building logbooks in the eu.

European Commission, Executive Agency for Small and Medium-sized Enterprises, Volt, J., Toth, Z., Glicker, J., De Groote, M., Borragán, G., De Regel, S., Dourlens-Quaranta, S., and Carbonari, G. (2020). Definition of the digital building logbook : report 1 of the study on the development of a European Union framework for buildings' digital logbook. European Commission Publications Office.

Gómez-Gil, M., Espinosa-Fernández, A., and López-Mesa, B. (2022). Review and analysis of models for a european digital building logbook. Energies, 15(6).

Gómez-Gil, M., Maria Sesana, M., Salvalai, G., Espinosa-Fernández, A., and López-Mesa, B. (2023). The digital building logbook as a gateway linked to existing national data sources: The cases of spain and italy. Journal of Building Engineering, 63.

Malinovec Puček, M., Khoja, A., Bazzan, E., and Gyuris, P. (2023). A data structure for digital building logbooks: Achieving energy efficiency, sustainability, and smartness in buildings across the eu. Buildings, 13(4).

Møller, R. S., Rhodes, M. K., and Larsen, T. S. (2018). Dgnb building certification companion: sustainability tool for assessment, planning, learning, and engaging (staple). Towards Energy Sustainability, page 135.

Mêda, P., Calvetti, D., Kifokeris, D., and Kassem, M. (2022). A process-based framework for digital building logbooks. page 473 – 480.

Tupenaite, L., Lill, I., Geipele, I., and Naimaviciene, J. (2017). Ranking of sustainability indicators for assessment of the new housing development projects: Case of the baltic states. Resources, 6(4).

[3]See `https://saref.etsi.org/` for more information

2024 European Conference on Computing in Construction
Chania, Crete, Greece
July 15-17, 2024

DIGITAL TWIN ENABLED CONSTRUCTION PROGRESS MONITORING

Alwyn Mathew[1], Shuyan Li*[1], Kacper Pluta[2,3], Rahima Djahel[3], and Ioannis Brilakis[1]
[1]University of Cambridge, England, United Kingdom
[2]Université Gustave Eiffel, CNRS, LIGM, France
[3]Inria, Université Côte d'Azur, France

Abstract

Digital Twin technology has revolutionised overseeing newly built structures. This study suggests employing digital twin-based automatic progress monitoring on construction sites, comparing 3D point clouds with their Building Information Modelling to track progress and predict completion. It highlights integrating semi-continuous monitoring with a building's digital twin for efficient construction management. Leveraging precise data enhances understanding and identifies schedule deviations, enabling timely actions. Demonstrated through real-world construction data, visualised Gantt charts showcase its efficacy, offering insights into task status, schedule deviations, and projected completion dates. This underscores digital twin technology's potential to transform construction oversight.

Introduction

Scan-to-BIM provides a comprehensive assessment of the as-built state in the context of performance improvement (Bosché et al., 2015; Drobnyi et al., 2023b, 2024). Despite its impact, laser scanning is somewhat limited to surface recognition, lacking depth in reflecting built element quality (Hoiem et al., 2022). Researchers explore embedded sensing for conditional data assessment (Alizadehsalehi and Yitmen, 2016). Integrating laser scanner technologies and advanced wireless sensors offers opportunities for comprehensive project exploration, enhancing performance control and project management by merging as-planned models with data-capturing reality. Effectively managing vast and complex data for real-time progress monitoring requires an intelligent system continuously learning from various sources, including historical archive data (Boje et al., 2020). IoT technologies and related systems, combined with digital twin and cognitive computing, collect real-time data (Dawood et al., 2020). Visualizing digital data for stakeholders in different project stages is crucial, with XR technologies providing multidimensional perspectives (Alizadehsalehi et al., 2020).

As per the Project Management Body of Knowledge (PMBOK), controlling and monitoring a construction project encompasses processes to oversee progress and performance, identifying areas requiring plan adjustments, and initiating corresponding changes (Guide, 2001). These processes entail measuring progress through inspections (as-built) and comparing it with the project plan (as-planned) to validate predicted performance. The overarching goal of monitoring is to ensure effectively managed results and outputs by measuring and assessing project performance (Lin and Golparvar-Fard, 2020). Measuring work in progress on construction sites is crucial for project management, impacting various aspects such as time, cost, quality, and safety. This task is particularly challenging due to the complexity and interdependency of activities (Arif and Khan, 2021).

Traditional progress tracking practice depends on visual inspections and daily or weekly reports created based on those inspections to ensure that work meets contract specifications and schedule (Golparvar-Fard et al., 2009). These traditional practices are often slow and rely heavily on the inspectors' personal judgment, observational skills, and weekly expert follow-ups with a high probability of complete and accurate reports. Effective monitoring is crucial for project success; however, even the most robust monitoring systems are insufficient if the project is poorly designed or built on flawed assumptions. Building Information Modelling is a digital representation of a building, capturing 3D geometry and semantic descriptions of components (Kim et al., 2020b). The AEC industry-accepted BIM provides a suitable basis for automated construction progress monitoring. It serves three essential purposes: providing as-planned data, as-built data, and enabling their comparison (Machado and Vilela, 2020). As the baseline for construction progress monitoring, BIM binds AEC contract information, facilitates access to geometric data, allows for special visualization of schedules, and manages progress-related information. Recognized as a rich data source, BIM is pivotal for automated project progress monitoring (Kim et al., 2020a). A well-designed BIM model analyses operations in construction, aiding site management, enhancing communication, coordinating contractors, and planning logistics (Kopsida and Brilakis, 2020). While traditional Building Information Modelling (BIM) and construction schedules effectively capture the as-designed and as-planned phases, they inherently lack timely updation of the as-built and as-performed states during construction progression. This deficiency arises due to the static nature of BIM and schedules, which do not dynamically update as construction activities unfold. Construction sites' inherent heterogeneity and temporal dynamics present formidable challenges to accurate progress monitoring.

This paper proposes using DT to monitor the progress of large-scale construction sites. The major contribution is the tight integration of progress monitoring to DT, facilitating timely progress monitoring of a real construction site. The pipeline of the DT-based progress monitoring is as follows: The workers capture laser scans at the site and are passed to the DT platform. Then, the detection

*Corresponding author

model in the platform automatically detects as-built elements and pushes the information generated from raw data to the graph. This will trigger the progress monitor and traverse the graph to get the progress information at the activity level. Here, an activity refers to a grouping of tasks, where each task represents a specific job tied to a particular element on the construction site. We calculate the current status of the progress (on schedule, ahead of schedule, and behind schedule) of each activity and estimate the finished date if the progress is behind schedule. We visualize these results in a Gantt table via a dashboard, which managers can quickly investigate.

Background

In response to these challenges, Digital Twins Construction (DTC) is emerging as a focal point of attention, offering a reliable information source for continuous production planning and adaptive product design throughout the construction lifecycle. However, the successful integration of DTC faces challenges – industry-wide adoption, technical intricacies ensuring the precision and accessibility of data (Sacks et al., 2020). Crucially, DTC should not be perceived merely as an extension of existing tools like BIM but as a transformative approach to construction production management, emphasizing a closed control loop (Sacks et al., 2020). The concept of DTC, coupled with automated data acquisition, establishes a framework for automatic progress monitoring, obviating the limitations of conventional techniques. This study directs readers to a comprehensive exploration of digital twins in construction through a recent review paper authored by (Opoku et al., 2021), providing invaluable insights into the evolving landscape of construction technology and methodologies.

To streamline and enhance the efficiency of progress monitoring processes, the initial step involves identifying newly constructed objects since the last data acquisition. This task is inherently challenging during the construction cycle due to noise, missing data, and local disparities in the position of as-built elements compared to their as-designed counterparts. The proposed solution involves locating instances of the as-designed model within LiDAR data acquired on construction sites. Current state-of-the-art approaches for object detection fall into two primary categories: traditional and deep learning (Chu et al., 2023; Lan et al., 2024). Traditional computer vision algorithms for object detection rely on deterministic procedures involving primitive shape fitting and statistical analysis (Drobnyi et al., 2023a). The most established methods within this category include RANSAC, Hough transform, and region growing. For instance, the efficacy of the Hough transform in detecting pipes within noisy 3D point clouds was demonstrated by (Ahmed et al., 2014). RANSAC-based methods have gained popularity due to their robustness in automatically segmenting building object instances represented by basic shapes such as cuboids and cylinders, enabling the detection of slabs and pipes. (Anagnostopoulos et al., 2016) applied RANSAC to detect wall surfaces, facilitating the reconstruction of rooms from 3D point cloud data (Anagnostopoulos et al., 2016). The second category employs deep learning techniques, with deep neural networks emerging as the predominant method for object detection. Notably, the PointNet architecture, a deep neural network specifically designed for point clouds, was introduced by (Qi et al., 2017). PointNet predicts the class label for each object segment, receiving a cluster of points as input and outputting a category prediction among 13 classes (Chen et al., 2019).

(Hu and Brilakis, 2024) proposed an automatic clustering method to segment the points corresponding to the as-designed instance. The workflow contains (1) Instance descriptor generation, (2) PROSAC (Progressive Sample Consensus) based shape detection, and (3) DBSCAN (Density-Based Spatial Clustering of Applications with Noise) based cluster optimization. The method matches design-intent planar, curved, and linear structural instances in complex scenarios, including (1) the as-built point cloud is noisy with high occlusions and clutter; (2) deviations between as-built instances and as-designed models in terms of position, orientation, and scale; (3) both Manhattan-World and non-Manhattan-World instances.

Methodology

In this work, we develop our progress monitoring method based on a holistic cloud-based Digital Twin Platform (DTP). This platform operates on a structured ontology, facilitating storing both the as-designed and as-built information for every element within a construction site. This platform intricately captures and retains the element-level status of each component present on the construction site. The status information is derived through various DT services, which meticulously process raw 3D point clouds from routine construction site surveys with laser devices. We use a BIM-assisted 3D object detection algorithm to ascertain the presence of each element within the as-built data. Such information is transmitted and systematically stored in the DT platform as element-level status, forming a comprehensive repository of the construction site's dynamic and evolving conditions. By retrieving and processing the information in the DTP, we calculate the activity-level progress status and estimate the finish date of the activity.

Ontology

The pertinent partial ontology for this study is illustrated in Figure 1; the complete ontology can be found in (Schlenger et al., 2022). Construction information is organized within a graph-based database, aligning with the structure defined by the ontology. The ontology bifurcates into as-planned and as-performed segments, encompassing `Work package`, `Activity`, `Task`, and as-designed elements under the as-planned side. In contrast, the as-performed side includes `Construction`, `Operation`, `Action`, and as-built elements. These nodes adhere to a hierarchical

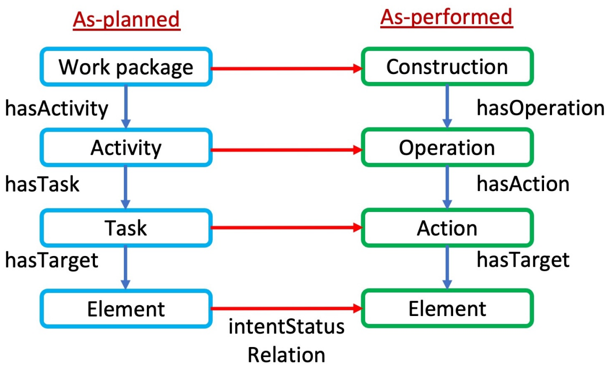


Figure 1: Visualization of a segment of the Digital Twin ontology employed in the progress monitoring pipeline.

order and establish connections through various relationships. Notably, the work package node, functioning as an aggregation of activities, is linked to respective activity nodes through the `hasActivity` relationship. The as-planned and as-performed sides of nodes are connected with `intentStatusRelation`. Recognizing that construction plans often operate at the activity level, the as-planned schedule is stored within the `Activity` node, capturing start and end dates. The schedule reflecting the as-performed activities is archived within the `Operation` node, mirroring the data in the `Activity` node. In instances where no prior surveys have been conducted for the target building, the operation's start date is presumed to align with the as-planned start date. Alternatively, if previous surveys have been undertaken, the operation's start date is designated as the most recent scan date. The conclusion of the operation is determined by extracting the latest update date at the element level. The as-built element node retains the progress at the element level, computed using the BIM-assisted 3D object detection algorithm (Hu and Brilakis, 2024).

Object detection

Automating progress monitoring necessitates the initial step of detecting constructed objects on construction sites, a task burdened with challenges, as outlined in the preceding background section. Streamlining this process requires a global registration between the BIM model and 3D point cloud data. We utilize a global registration method to align the coordinates of the BIM model and LiDAR data (Monasse et al., 2023). This efficient algorithm optimizes global robust energy between two line segments extracted from the BIM and LiDAR data. Once registered, the Region-Of-Interest can be confined to the upscaled bounding box of the query as-designed element, provided the BIM model and LiDAR data share the same coordinate system. This region of interest will be an input to a filtering step to remove clutter, if they exist. Given the geometric richness of the construction environment, we propose a novel solution based on geometric features, with a specific focus on planar polygons as a robust data abstraction. Our method involves detecting and clustering planar polygons in each dataset, followed by a matching step to compare planar polygons within associated clusters. This comparison allows us to identify local discrepancies in position, if they exist, ultimately eliminating false detection. We not only determine if there is a local discrepancy but also calculate the corresponding geometric transformation. This can avoid false positive detection when facing significant local discrepancy.

Integration with DTP

Specific Application Programming Interfaces (APIs) have been developed to facilitate seamless communication with the DTP[1]. These APIs enable the retrieval, creation, and updating of nodes within the DTP. The comprehensive progress monitoring pipeline is visualized in Figure 4. Given the assumption that the DTP is current with both as-built element progress and operation start and end dates, an initial fetch request is initiated to retrieve all `activity` nodes from the DTP. Subsequently, each activity's as-planned start and end dates are extracted from the corresponding `activity` nodes. Following the hierarchical structure, all as-designed element nodes linked to each `activity` node through relationships like `hasTask` and `hasTarget` are retrieved from the DTP. Leveraging the intent-status relation, with `intentStatusRelation` corresponding to the as-built node for each as-designed node is fetched. The element-level as-built progress is then aggregated from the as-built element nodes. Employing reverse relationships with `hasTarget` and `hasAction`, `operation` nodes are fetched, and the associated as-performed schedules are compiled. Once the as-planned schedule, as-performed schedule, and element-level as-built progress are at our disposal, the groundwork is laid for the computation of progress at the activity level.

Progress calculation

The determination of activity status in relation to the schedule is outlined in Table 1. An activity is marked as ahead of schedule if the element-level progress exceeds zero and the end date of the operation precedes the corresponding activity end date. The cumulative assessment considers an activity as ahead of schedule if a majority of its elements exhibit this characteristic. Simultaneously, the percentage of completed tasks within an activity is computed. In the case of an activity falling behind schedule, the maximum delayed task determines the extent of the delay. To further enhance understanding, the calculated percentage of tasks completed, and the determined delay duration are employed to estimate the revised end date using a projection function. Presently, the projection utilizes an S-shaped function, closely resembling actual construction progress (San Cristóbal, 2017).

Results & Discussion

The experimentation encompassed utilising both as-designed data and 3D point cloud data obtained from a

[1]`https://github.com/BIM2TWIN-Team/DTP_API`

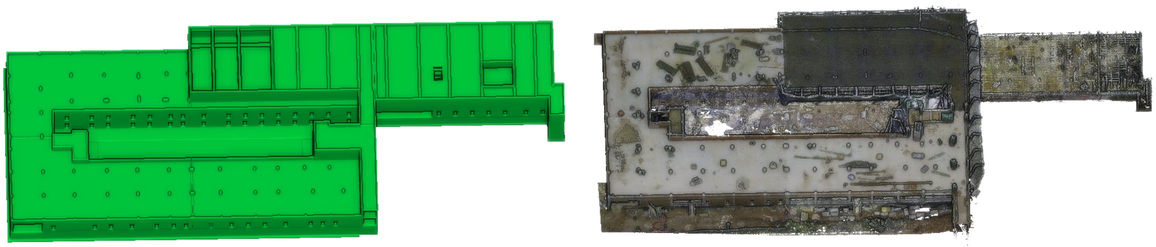


Figure 2: As-designed and point cloud data from a construction site in Spain.

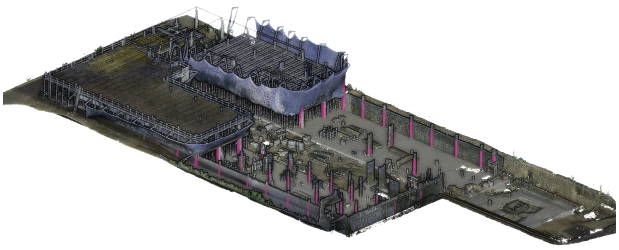


Figure 3: Visualization of columns detected on a construction site. The detected columns are marked in pink.

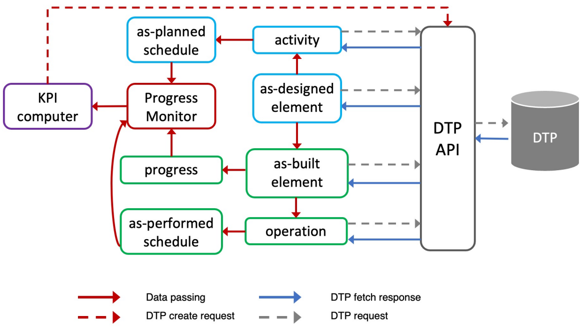


Figure 4: Visualization of the pipeline depicting the progress monitoring algorithm execution flow

Table 1: Criteria for assessing activity status concerning schedule compliance.

Condition	Decision
Element level progress > 0	
Activity end time > Operation end time	Ahead
Activity end time < Operation end time	Behind
Activity end time = Operation end time	On
Element level progress = 0	
Activity end time > Operation end time	On
Activity end time < Operation end time	Behind
Activity end time = Operation end time	On

construction site located in Spain, as shown in Figure 2. The detection results of columns in shown in Figure 3. As can be seen, our method can detect corresponding elements in the whole point cloud building. The output of the progress monitoring encompasses key metrics, including the percentage of tasks completed, progress status, the number of days ahead or behind the schedule, and the projected completion date for each activity. To enhance the interpretability for construction managers, the results are visually presented in the form of a Gantt chart, as depicted in Figure 5. Gantt charts were chosen due to their widespread usage in construction scheduling, making them a familiar and effective visualization tool for managers. Each activity is graphically represented by two horizontal bars: one reflecting the as-planned schedule and the other the as-performed schedule. The as-planned schedule is denoted by a grey bar, while the as-performed schedule is illustrated with a coloured bar. In the chart, a dark red bar signifies that the activity is complete but was delayed, whereas a dark green bar indicates that the activity is not complete but is on schedule. Light red signifies that the activity is behind schedule and has not yet been initiated, while light green indicates that the activity is on schedule and has yet to commence. The textual information overlaid on the grey bar corresponds to the name of the activity assigned by the construction company. Additionally, text overlaid on the coloured bar details the progress status, the number of days the activity is ahead or behind schedule, and the projected completion time.

Following this, the Key Performance Indicators (KPIs), including the *percentage of tasks delayed per activity* (KPI1) and the *percentage of delay in days per activity* (KPI2), are systematically computed. KPI1 is the quantitative measure obtained by dividing the number of delayed tasks for each activity by the total number of tasks scheduled for that specific activity. This ratio provides a nuanced understanding of the prevalence of task delays within individual activities, contributing valuable insights into the project's task-level performance. Simultaneously, KPI2 is calculated by determining the ratio between the number of de-

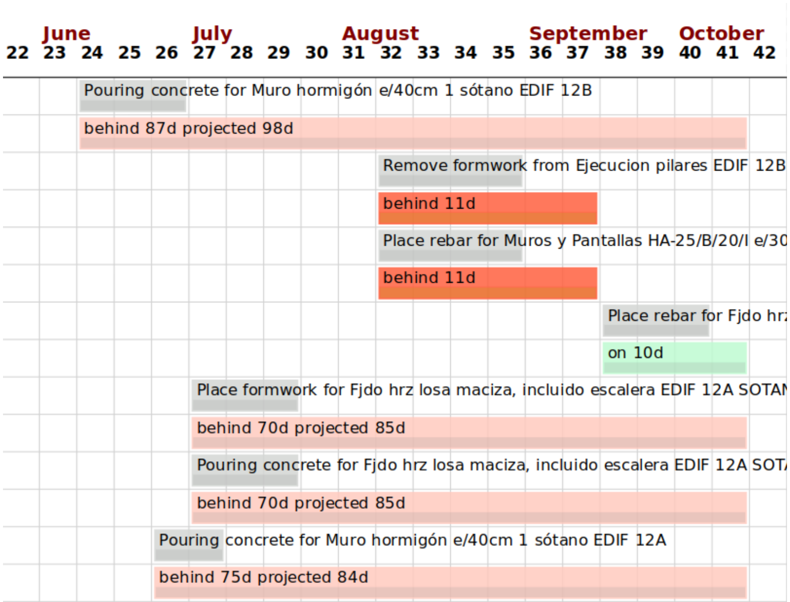


Figure 5: Graphical representation of progress monitoring algorithm output in Gantt chart format for enhanced comprehension

layed days for each activity and the planned duration (as-planned days) for that respective activity. By doing so, KPI2 offers a comprehensive overview of the temporal aspects of project performance, highlighting the extent of delays compared to the initially envisaged project timeline.

These calculated KPIs play a pivotal role in assessing the efficiency and timeliness of project activities. Once computed, the KPI values are securely stored within the DTP, forming an integral part of the platform's repository of project performance data. This centralized storage ensures that historical KPI data is readily available for analysis, comparison, and continuous improvement efforts. Furthermore, to enhance accessibility and visibility, these KPIs are dynamically displayed on the project's dashboard. This strategic placement on the dashboard gives project stakeholders timely, at-a-glance insights into key performance metrics. The dashboard becomes a central hub for monitoring and understanding project delays, fostering a proactive approach to decision-making based on the current project status. The source code for this paper is accessible on GitHub at the following URL: `https://github.com/BIM2TWIN-Team/WP3-progress-monitor`.

As highlighted in a recent comprehensive review on computer vision (CV) aided progress monitoring in construction (Sami Ur Rehman et al., 2022), the conventional progress monitoring methods are characterized by slowness, tediousness, and susceptibility to errors. However, the existing body of literature on CV-based progress monitoring is dispersed across various domains, lacking a cohesive focus on methodologies and processes throughout the entire CV-based progress monitoring workflow. This article addresses this gap by presenting a holistic approach that provides timely information and knowledge through Key Performance Indicators (KPIs). Such data availability is critical, enabling simulations of alternative execution plans that prove invaluable at different construction stages. This capability is instrumental in minimizing delays and optimizing equipment usage, as discussed by Yeung et al. (Yeung et al., 2022). By consolidating insights and methodologies, our approach aims to streamline CV-based progress monitoring into a coherent and efficient process.

Conclusion

This study emphasizes the profound impact of DT technology on monitoring and managing newly constructed buildings. The proposed DT-based method marks a substantial leap in automatic progress monitoring for real-world construction sites. Utilizing scanners to collect 3D point clouds and Digital Twin Platforms (DTP), this approach allows detailed construction status analysis and completion timeline prediction. The seamless integration of semi-continuous monitoring with the building's DT underscores the pivotal role of DT tech in efficient construction management. Leveraging precise data not only enhances project understanding but also enables timely deviation identification, empowering stakeholders to implement corrective actions and proactive strategies, enhancing project efficiency. Extensive real-world experiments validate this method's effectiveness, highlighting DT's potential to transform construction monitoring and foster adaptive project management. As DT continues evolving, its integration into construction processes will be vital for achieving optimal efficiency, accuracy, and proactive decision-making.

Acknowledgments

This work is supported by the EU Horizon 2020 BIM2TWIN: Optimal Construction Management & Production Control project under agreement No. 958398.

References

Ahmed, M. F., Haas, C. T., and Haas, R. (2014). Automatic detection of cylindrical objects in built facilities. Journal of Computing in Civil Engineering, 28(3):04014009.

Alizadehsalehi, S., Hadavi, A., and Huang, J. C. (2020). From bim to extended reality in aec industry. Automation in Construction, 116:103254.

Alizadehsalehi, S. and Yitmen, I. (2016). The impact of field data capturing technologies on automated construction project progress monitoring. Procedia engineering, 161:97–103.

Anagnostopoulos, I., Pătrăucean, V., Brilakis, I., and Vela, P. (2016). Detection of walls, floors, and ceilings in point cloud data. In Construction Research Congress 2016, pages 2302–2311.

Arif, F. and Khan, W. A. (2021). Smart progress monitoring framework for building construction elements using videography–matlab–bim integration. International Journal of Civil Engineering, 19:717–732.

Boje, C., Guerriero, A., Kubicki, S., and Rezgui, Y. (2020). Towards a semantic construction digital twin: Directions for future research. Automation in construction, 114:103179.

Bosché, F., Ahmed, M., Turkan, Y., Haas, C. T., and Haas, R. (2015). The value of integrating scan-to-bim and scan-vs-bim techniques for construction monitoring using laser scanning and bim: The case of cylindrical mep components. Automation in Construction, 49:201–213.

Chen, J., Kira, Z., and Cho, Y. K. (2019). Deep learning approach to point cloud scene understanding for automated scan to 3d reconstruction. Journal of Computing in Civil Engineering, 33(4):04019027.

Chu, Q., Li, S., Chen, G., Li, K., and Li, X. (2023). Adversarial alignment for source free object detection. In Thirty-Seventh AAAI Conference on Artificial Intelligence, pages 452–460.

Dawood, N., Pour Rahimian, F., Seyedzadeh, S., and Sheikhkhoshkar, M. (2020). Enabling the development and implementation of digital twins: Proceedings of the 20th international conference on construction applications of virtual reality.

Drobnyi, V., Hu, Z., Fathy, Y., and Brilakis, I. (2023a). Construction and maintenance of building geometric digital twins: State of the art review. Sensors, 23(9):4382.

Drobnyi, V., Li, S., and Brilakis, I. (2023b). Deep-learning guided structural object detection in large-scale, occluded indoor point cloud datasets. In Proceedings of the 2023 European Conference on Computing in Construction and the 40th International CIB W78 Conference, volume 4 of Computing in Construction.

Drobnyi, V., Li, S., and Brilakis, I. (2024). Connectivity detection for automatic construction of building geometric digital twins. Automation in Construction, 159:105281.

Golparvar-Fard, M., Peña-Mora, F., and Savarese, S. (2009). D4ar–a 4-dimensional augmented reality model for automating construction progress monitoring data collection, processing and communication. Journal of information technology in construction, 14(13):129–153.

Guide, A. (2001). Project management body of knowledge (pmbok® guide). In Project Management Institute, volume 11, pages 7–8.

Hoiem, D., Bretl, T., Degol, J. M., Fard, M. G., Lin, J. J.-C., Kataria, R., Han, K. K. I., and Tsoi, K. W. (2022). Computation of point clouds and joint display of point clouds and building information models with project schedules for monitoring construction progress, productivity, and risk for delays.

Hu, Z. and Brilakis, I. (2024). Matching design-intent planar, curved, and linear structural instances in point clouds. Automation in Construction, 158:105219.

Kim, S., Kim, S., and Lee, D.-E. (2020a). 3d point cloud and bim-based reconstruction for evaluation of project by as-planned and as-built. Remote Sensing, 12(9):1457.

Kim, S., Kim, S., and Lee, D.-E. (2020b). Sustainable application of hybrid point cloud and bim method for tracking construction progress. Sustainability, 12(10):4106.

Kopsida, M. and Brilakis, I. (2020). Real-time volume-to-plane comparison for mixed reality–based progress monitoring. Journal of Computing in Civil Engineering, 34(4):04020016.

Lan, L., Wang, F., Li, S., Zheng, X., Wang, Z., and Liu, X. (2024). Efficient prompt tuning of large vision-language model for fine-grained ship classification. arXiv preprint arXiv:2403.08271.

Lin, J. J. and Golparvar-Fard, M. (2020). Construction progress monitoring using cyber-physical systems. Cyber-physical systems in the built environment, pages 63–87.

Machado, R. L. and Vilela, C. (2020). Conceptual framework for integrating bim and augmented reality in construction management. Journal of civil engineering and management, 26(1):83–94.

Monasse, P., Djahel, R., and Vallet, B. (2023). Registration for urban modeling based on linear and planar features. In The 11th European Workshop on Visual Information Processing (EUVIP).

Opoku, D.-G. J., Perera, S., Osei-Kyei, R., and Rashidi, M. (2021). Digital twin application in the construction industry: A literature review. Journal of Building Engineering, 40:102726.

Qi, C. R., Su, H., Mo, K., and Guibas, L. J. (2017). Pointnet: Deep learning on point sets for 3d classification and segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 652–660.

Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., and Girolami, M. (2020). Construction with digital twin information systems. Data-Centric Engineering, 1:e14.

Sami Ur Rehman, M., Shafiq, M. T., and Ullah, F. (2022). Automated computer vision-based construction progress monitoring: A systematic review. Buildings, 12(7):1037.

San Cristóbal, J. R. (2017). The s-curve envelope as a tool for monitoring and control of projects. Procedia computer science, 121:756–761.

Schlenger, J., Yeung, T., Vilgertshofer, S., Martinez, J., Sacks, R., and Borrmann, A. (2022). A comprehensive data schema for digital twin construction. In 29 Th International Workshop on Intelligent Computing in Engineering. https://publications. cms. bgu. tum. de/2022_Schlenger_EGICE. pdf.

Yeung, T., Martinez, J., Sharoni, L., and Sacks, R. (2022). The role of simulation in digital twin construction. In Proceedings of the 29th EG-ICE international workshop on intelligent computing in engineering, pages 248–258.

HOW CAN DIGITAL TWINS BE USED IN HIGHWAY MAINTENANCE? A QUESTIONNAIRE SURVEY FOR INDUSTRY PRACTITIONERS

Mengtian Yin[1], Varun Kumar Reja[1], Ran Wei[1], Brian Sheil[1], and Ioannis Brilakis[1]
[1]Department of Engineering, University of Cambridge, UK

Abstract

The digital twin (DT) concept can potentially bring about a systematic approach to digitalizing information for road infrastructure systems. However, it is currently unclear how the DT concept can be applied in the highway maintenance domain and what the benefit is. This study conducts a questionnaire survey to explore what industry professionals view as the main obstacles in current practice and what areas of highway maintenance can improve if using a road DT. A total of 183 responses were obtained from highway industry practitioners. The results reveal the main issues in maintenance processes and potential use cases of DTs.

Introduction

Highway infrastructure systems play an essential role in society, the economy, and security. However, with the increasing number of vehicles travelled on the road network, real-world road assets are experiencing rising traffic load and utilization (Sinha, Labi and Agbelie, 2017), causing risks of structural failures. Therefore, maintaining highway assets consistently above a standard level is challenging, requiring a comprehensive acquisition of dynamic road condition data and effective data analytics tools to forecast road states as well as make sound treatment strategies. Unfortunately, highway authorities still heavily rely on manual inspection to assess road condition (Koch *et al.*, 2015) and make treatment plans based on the experience of pavement experts. The entire highway maintenance process may be changed by the newly emerging digital twin (DT) approach, which might enhance data acquisition, integration, analysis, and decision-making in a systematic way.

A DT represents a digital replica of physical assets and processes involved in the built environment lifecycle, enabled by a twin-based system (Brilakis *et al.*, 2019). In the highway maintenance context, DTs provide a detailed virtual model that simulates the behaviour and operations of physical highway infrastructure systems. Based on the dynamic virtual model, DTs bring about a new information management (IM) approach that can potentially improve data interoperability, integration, and analytics. In this way, highway experts may be able to efficiently retrieve useful information from DTs and make more appropriate decisions on maintenance interventions. So far, the DT concept has been widely applied in the fields of aerospace (Li *et al.*, 2021), manufacturing (Kritzinger *et al.*, 2018), and recently gained popularity in building construction (Sacks *et al.*, 2020). However, implementation of DT approach in highway maintenance is still scarce. It is currently unclear how the DT approach can be applied in this area and what actual benefits a DT can provide. Without knowing the purpose of using a DT, it is impractical to build a road DT system with a decisive scope of meta modelling, level of detail (LOD) for model instance generation, and frequency of data updating.

To tackle these issues, this study describes results from a questionnaire survey for highway industry practitioners to understand the main barriers in their professional workflows and the potential improvements that can be realised using a road DT. This questionnaire study is designed based on (a) a state-of-the-art review of DTs in road domain; and (b) interviews of 20 industry experts from the UK National Highways company, which manages England's motorways and major A roads (National Highways, 2021). The survey was further disseminated to highway industry practitioners internationally through different channels, such as email invitations. As a result, 183 valid responses have been obtained from different regions, mainly the UK, China, US, India, and Australia. The survey unveils crucial phases, applications, and use cases that a road DT can potentially take effect on. It also shows how practitioners perceive the challenges of DT deployment in the highway industry.

In the following sections, we first review relevant studies about DTs of road infrastructure. Then, we demonstrate the research methodology before presenting and discussing the survey results. At last, we draw conclusions by summarizing key findings and contributions.

Background research on DTs for road infrastructure

In the planning and design phases, a road DT is mainly used to assess the impacts of new road designs on existing networks and natural environments. Machl et al. (2019) present a DT-based method to assist road planners in assessing design impacts on agricultural core road networks. Jiang et al (2022; 2022a) propose a method to automatically generate DTs of road networks from open map data and use the DT to check the clearance of underpass roads in road widening projects. Jiang et al. (2022b) propose a DT-enabled sustainable urban road planning framework that integrates multi-criteria decision-making and Geographic Information System (GIS), where DTs provide the routes and traffic congestion data.

In the road construction phase, the DT is indirectly used to assess the performance of new road materials. Meza et al. (2021) developed a road DT to examine the use of secondary raw materials (SRM) in real road construction projects. Manually developed road Industry Foundation Classes models are imported into a common data environment platform that allows the integration of sensor data installed in various pavement layers. In addition, some existing studies integrate building information models (BIM) and the Internet of Things (IoT) for road construction monitoring, which is close to the concept of DT. For example, Han et al. (2022) propose a BIM-IoT and intelligent compaction (IC)-integrated framework for road compaction quality monitoring and management.

In the operation and maintenance (O&M) phase, many road DTs are created in different ways without a specific purpose. Marai et al. (2021) present a road DT system using a 360° camera and IoT devices. The live stream, GPS location, and measurements of the temperature and humidity are dynamically sent to the DT, with interested objects (e.g., vehicles) detected and tracked. Wang et al. (2021) use 3D GIS technology to develop a highway DT, where the Computer-generated Architecture language is employed to build the road geometric model. Steyn and Broekman (2022) develop a DT prototype of a local road network that uses LIDAR, unmanned aerial vehicles, and sensor systems to capture road geometry, traffic flow data, and road environment data, respectively. In contrast, some studies discuss some potential applications of DT. For instance, Chen et al. (2022) propose a DT-based framework for road condition prediction that inputs historical and real-time data from the whole road lifecycle into machine learning algorithms to predict future road performance. Agrawal et al. (2022) invite highway experts to plan the adoption of DTs to detect pavement defects, predict future defects, and make maintenance work orders.

From the above review, it can be summarized that although there are some implementations of DT in the road maintenance phase, the application scenarios and benefits of DTs have not been systematically explored using surveys of realistic problems encountered by the highway industry. Furthermore, the existing studies primarily focus on DT generation and DT-enabled data analytics, whereas few research has investigated the requirements and challenges for DT implementations considered by the industry practitioners. Understanding the specific use of road DTs and value points could help the highway industry develop DTs as a new IM approach to solve the existing issues.

Research methodology and questionnaire design

Figure 1 shows an overview of the research design. This study first interviewed 20 highway experts from UK National Highways. The objective is to understand existing maintenance processes and the problems that underpin them. The Information Delivery Manual (IDM) (29481-1, 2010) approach is employed to represent different levels of maintenance processes along with the related actors, phases, activities, and data exchange items.

Having documented all relevant existing maintenance processes, the authors analyse the IDM documents to identify any inefficiencies in these processes. Based on the findings, together with the interview conversations and literature review, a questionnaire is designed, covering the following aspects: (a) participant role and experience; (b) barriers in current practices; (c) current use of DT; (d) perceived opportunities of DT; (e) challenges of DT implementation. Most questions are multiple-option questions, where the options come from the authors' findings (e.g., identified inefficiencies) or proposals (e.g., proposed DT applications). The questionnaire also allows participants to input their comments for most questions, in case the predefined options are not within their considerations. The questionnaire form can be accessed through the link: https://forms.gle/Nj7g7BT9D3VHo3x4A .

Finally, we disseminated the questionnaire survey through various channels, such as email invitations, LinkedIn, and Weibo. To allow relevant people to complete this questionnaire, we included prompts in the invitation that state "*the target population is individuals who work in the highway industry, especially in the following areas: (a) road survey and inspection; (b) road maintenance; (c) information system/database operation; (d) strategy making and research innovation; (e) asset management; and (f) safety management*".

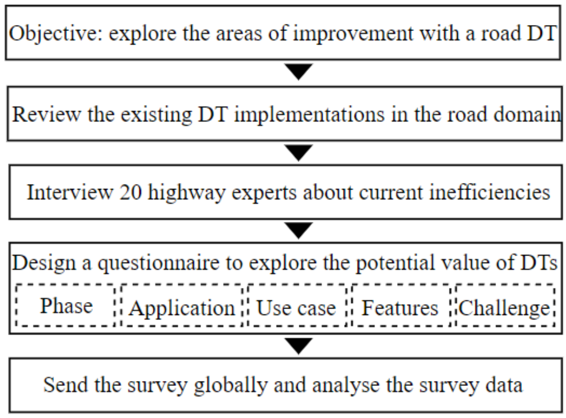


Figure 1. Overview of the research design.

Survey results and discussion

Overview

The survey data collection period covers approximately 40 days, from October to November in 2023. The number of email invitations is over 500. As a result, 183 valid responses have been obtained as valid samples for data analysis. This sample size enables a 95% confidence level with an 8% margin of error for the survey results, as calculated using Slovin's formula (Tejada and Punzalan, 2012). Given the exploratory nature of this study and the limited pool of respondents, these parameters can be deemed appropriate.

Table 1 shows the demographics of all respondents. Maintenance managers and engineers (27.9%), road inspectors (20.8), and technology developers (19.1%) occupy the largest proportions, probably because this topic gains interest from these groups. It is worth noting that some respondents selected multiple roles in this survey.

Table 1: Demographics of the 183 respondents.

Roles	Proportion
Road surveyor	14.2%
Road inspector	20.8%
Maintenance manager and engineer	27.9%
Scheme maker	10.4%
Asset manager	14.2%
Technology developer	19.1%
Network-level strategy maker	11.5%

In terms of participants' experience in the industry, 34.4% of participants have 5-10 years of working experience. People who have worked for over 10 years account for 32.8%. In comparison, practitioners who have started their careers (0-2 years) and have 2-5 years' experience make up 9% and 24%, respectively.

Main barriers in current practice

Figure 1 presents the responses from the participants regarding question 3, which is about selecting the main barriers faced by their organization in highway maintenance. It shows that "*hard to derive optimal maintenance decisions under budget constraints*" is the most selected choice (55.2%), which implies that the main limitation for the highway industry is still budget constraints. In this situation, practitioners must carefully evaluate the impacts of different decisions and prioritize the importance of different tasks, which is difficult to achieve due to several reasons. First, as the ranked second (50.8%) choice points out, much of the asset information stored in the existing maintenance management systems (MMS) and asset management systems (AMS) is inaccurate, incomplete, and unreliable, especially for old and underground assets (e.g., electricity). Second, it is still difficult to obtain comprehensive road condition data, which is also acknowledged as the third barrier. If maintenance planners do not know what happens on the entire road network, it is difficult to make informed decisions. Furthermore, 46.4% of participants agree that there is a lack of precise evaluation of the costs, benefits, and risks of alternative strategies. This is because most existing decision-support tools can only optimize short-term plans with controllable factors rather than forecasting long-term, multi-year network evolution with uncertainties incorporated.

Some respondents also identified specific problems they faced. For example, two answers illustrate that there is a lack of "*consistent operational approach*" and "*maintenance strategy based on needs*". These problems relate to other written answers that demonstrate a lack of knowledge of how asset handover works and the complexity of the data specifications of the current asset systems. The barriers occur because an increasing number of highway authorities are adopting asset information models (AIM) for managing highway assets, but maintenance people do not know how to use these asset data in their professional work. For example, the National Highways publishes an Asset Data Management Manual (ADMM) (National Highways, 2023) as their data handover specifications for all road construction projects, but they currently use another system with disparate data schemas to schedule routine maintenance.

Apart from the asset data usage problem, other written answers mention some challenges in lack of skilled workforce, lack of trust in analytical processes, communication with stakeholders, inefficient contract preparation, and difficulty obtaining road data without traffic disruption.

Current use of relevant technologies

In this survey, we also investigate the IM approaches and technologies adopted by highway organizations in the current practice. In the questionnaire, we first clarify what a DT is in the highway maintenance context: "*a digital replica of the physical highway system, encompassing road infrastructure assets, traffic elements, environmental factors, and more*." Afterward, the fourth question first surveys the participants' familiarity with the DT concept. As a result, 63.9% of participants responded

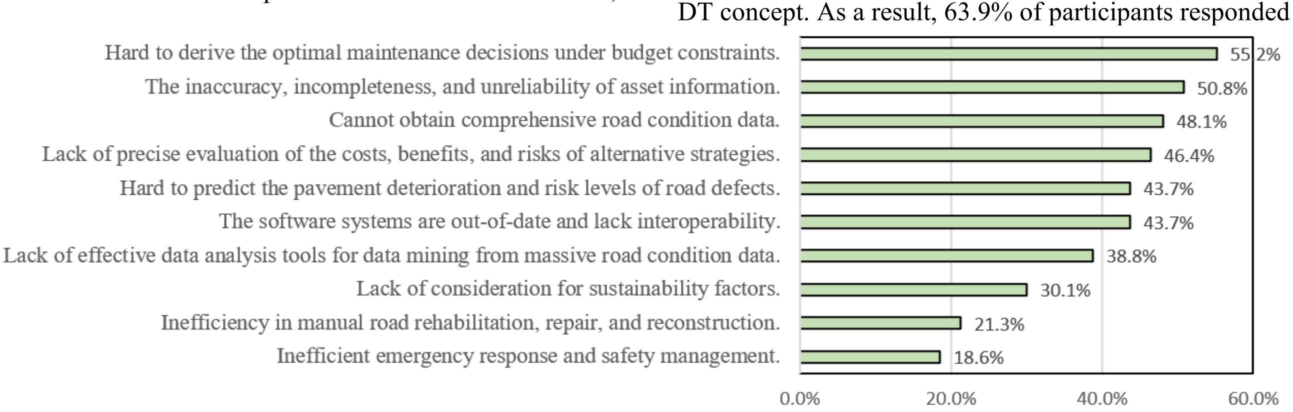


Figure 1. Responses to Question 3: "What are the main challenges and barriers faced by your organization in highway maintenance operations?"

"somewhat familiar," and 22.4% of participants chose "very familiar". In contrast, only 13.7% of participants acknowledge that they are not familiar with DTs at all.

The fifth question sought to clarify what technological approaches are currently used by participants' organizations. As shown in Figure 2, GIS is the most used technology for highway maintenance, which is intuitive because nearly all highway assets should be georeferenced in large road networks. Moreover, it was found that BIM approach is increasingly used in the road engineering domain (54.1% of participants). How to best use the road BIM files from the design and construction phases in the maintenance phase requires further studies. One potential solution is to make BIM models exchangeable with twin-based systems to realise their maximum value. In addition to the predefined options, participants also mention drone-based real-time data collection techniques and material deterioration models.

The sixth question aims to explore the contexts and projects in which highway organizations adopt DT technology and how it is used. From the answers obtained, it was found that most applications of DTs are 3D visualization to assist maintenance planners in making decisions. However, there are some noteworthy other applications, as listed in Table 2.

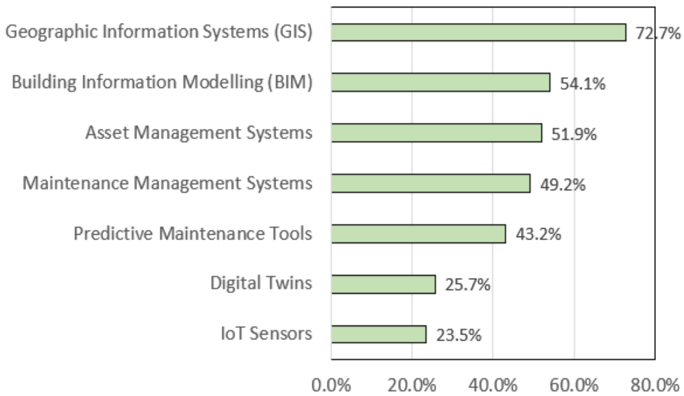


Figure 2. Responses to Question 5: "Does your organization currently use digital technologies or solutions for highway maintenance? "

Table 2: Selected answers for Question 6: "briefly describe the context or project where digital twins were/are utilized."

Current use of DT in organizations
1. Behavior/deterioration modeling of road assets.
2. Combining survey and as-built data for digital design.
3. Decision-support tools to predict maintenance needs.
4. Analysis of budget, safety, environment, flood risk, driver behavior, and fleet composition.
5. Comprehensive data storage with road asset definition, road construction, and condition data.
6. Measurement of complex road segments.
7. Real-time weather detection of the highway environment and refined weather forecast.
8. Emergency response and traffic jam management.
9. Supporting the development of autonomous vehicles.
10. Education and training.

As can be seen from the above table, some highway organizations have started to leverage the rich data in digital twins for behaviour prediction, road renovation, maintenance need estimation, and multi-factor analysis for maintenance planning. Moreover, it is worth noting that the integration of road DTs and connected autonomous vehicles (CAV) has the potential to create huge benefits for both the highway industry and the automotive industry: on one hand, the road DT can provide high-fidelity data about road conditions, providing better navigation to CAVs to avoid dangerous road defects (e.g., potholes). On the other hand, the rich data collected from CAVs driving on the network, such as images and sensing data (e.g., temperature, roughness), can be used to update the road DT in real time, which could overcome the critical problems regarding the DT data collection for the huge road network.

Potential use cases of road DT

This section of the questionnaire aims to understand how highway industry practitioners perceive the opportunities for DTs in highway maintenance. Specifically, we try to identify (a) important phases for applying DT; (b) important features that a road DT must have to take effect; (c) valuable applications of road DTs; (d) potential use cases; and (e) main benefits of road DTs.

Based on our interview, the process of highway maintenance can be generally divided into five stages:

- *Scheme making and prioritization*: experienced engineers receive annual network-level survey data and make maintenance schemes for a specific road section. All schemes must be prioritized based on severity levels.
- *Road assessment*: road inspectors physically arrive at road sites to inspect road defects. Various techniques can be used to investigate the deterioration of pavements, such as coring, deflectograph, and ground penetration radar (GPR).
- *Maintenance planning*: based on the results of inspections, maintenance planners make treatment plans with consideration of budgets, costs, labour, and risk level.
- *Maintenance execution*: based on the maintenance plans, contractors of highway authorities conduct road repair and rehabilitation using both humans and machines.
- *Quality assurance and control*: the process of quality inspection to ensure that the maintenance work reaches the required standard.

Table 3 presents the results of Question 7, showing that the stage of *Scheme making and prioritization* is the most commonly agreed phase in which a road DT can take effect. This might be because the DT can be used to check road conditions in 3D views and simulate the effects of different strategies. In addition, *Maintenance planning* is also a popular response (74.3%). By contrast, fewer participants recognize that a DT can be useful in the phases of *Maintenance execution* and *Quality assurance and control*, suggesting that the DT alone may not be useful in fieldwork that involves mainly manual work. Finally, some participants also wrote answers, such as operational planning and control, long-term investment optimization, and maintaining inventory data for soft estate assets. One answer demonstrates that the DT can enable new intelligence-led proactive maintenance that could totally change the existing processes.

Table 3: Responses to Question 7: "In which phases of highway maintenance do you see that digital twins can take effect?"

Phase	Count	Proportion
Scheme making and prioritization	148	80.9%
Road assessment	122	66.7%
Maintenance planning	136	74.3%
Maintenance execution	108	59%
Quality assurance and control	99	54.1%
None of above	2	1%

Question 8 was designed to investigate what features of road DTs are perceived as important for industry practitioners in highway maintenance. From the literature review, we summarize several key features of DTs: data transparency, data security, autonomous decision-making, prediction, simulation, data diagnostics, visualization, and data storage and integration. In the domain-specific context, we provide explanations for the following features in the questionnaire:

- Data diagnostics: identify the anomalies, risks, and their causes in the condition data.
- Simulation: simulate the operations of alternative strategies and assess their effects.
- Prediction: predict future states of roads, such as deterioration and structural failure.
- Autonomous decision-making: generate suggestions to decision-makers based on performance indicators.

Figure 3 presents the results for Question 8. As can be seen, prediction and simulation are two of the most important functional features that a road DT is expected to have. In comparison, fewer participants selected the feature of autonomous decision-making (54.6%). This implies that industry practitioners tend to make decisions on their own based on the intermediate results from DT-enabled simulation and prediction, rather than leaving them entirely with DTs. Furthermore, Figure 3 shows that the features of data transparency and data security are the least chosen (32.8% and 30.1%), suggesting that industry practitioners care more about the practical functions that a DT could bring about. In contrast, system architects and DT developers might consider these two features to be crucial.

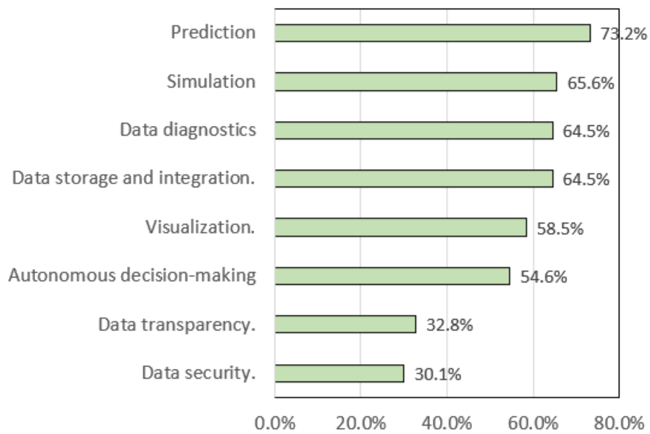


Figure 3. Responses to Question 8: "What features of a digital twin system do you perceive are important in highway maintenance?"

The subsequent question is about the application of road DTs. Based on the review and interview, the authors list several applications for road DTs, and the participants are asked to make selections of applications that they think are valuable. The result of Question 9 is presented in Figure 3. The most chosen DT application (73.2%) is *Predicting the deterioration and structure failures of road assets based on historical and dynamic road data*. In essence, deterioration prediction is a longstanding research problem in the field of highway maintenance. The role of DTs in this task is to be the data holder of historical and real-time road condition data, road construction data (e.g., material), traffic data, and climate data. The prediction of deterioration and failure depends on analytical models that process these DT data and predict the future states of pavements and other assets.

The second-ranked application is *Helping routine maintenance planners make better maintenance strategies, considering multiple factors*. From our interview, routine maintenance and reaction is one of the crucial processes to ensure the operation of the UK highway infrastructure, in which road inspectors drive on the entire network every week to visually check abnormal conditions. Maintenance planners have trouble making the best decisions in prioritizing the repair of the damaged assets and sending work orders that can minimize costs and risks. Therefore, a DT can be used to help maintenance planners make sound strategies in consideration of multiple factors. For instance, a DT may provide cost information for an asset or estimate the risks of the observed defects in association with the traffic environment. The third-ranked application is *Generating optimal plans for road investigation and repair within a* *scheme-level maintenance project.* The UK highway authorities use scheme-level road investigation and maintenance as a key strategy to systematically repair, rehabilitate, or reconstruct road sections. Within a maintenance scheme, scheme makers gather different kinds of survey data and then use their knowledge and experience to determine the treatment strategies. A road DT could aid in the creation of optimized treatment plans with less bias due to human experiences.

Question 10 is an open question that asks participants to describe at least two use cases of DTs that are valuable for highway maintenance. A total of 125 written responses were obtained. While some answers overlap with the applications listed in Question 9, many responses provide valuable insights from the viewpoints of highway experts. Table 4 presents 10 selected answers.

Table 4: Selected answers for Question 10: "Please describe the potential use cases of digital twins in highway maintenance that you believe hold significant promise."

Potential use cases of DT for highway maintenance
1. Supporting risk-based proactive maintenance and automated, human-free road inspection.
2. Live highway monitoring (traffic and pavement structure conditions) to generate instructions for emergency repair and prioritization of repair.
3. Improve predictability of maintenance need, simplicity of maintaining, and up-to-date asset record (location, condition and interventions delivered)
4. Provide data for autonomous cars/machines for snow-plowing the road; provide data for autonomous marking/paining on the road.
5. Data from inspections and maintenance activities is visualized and pinned to the asset to show it "service" history current condition and is then used to predict future maintenance interventions.
6. Understanding 3rd party productivity risks to site works e.g., PROW (Public Right of Way) or ecology; understanding 3rd party safety risks including where to park, traffic levels, works access.
7. Providing detailed incident tracking to understand real-world risks on the network and comparing with asset condition; CAV routing for maintenance and wear.
8. Highlight underground drainage defects.
9. Understanding interaction between different asset classes to help aid intelligence-led predictive modelling for best asset interventions taking account of whole life asset management.
10. Prescribing future designs and built outcomes that impact maintenance.

By analysing the collected responses, some use cases of road DTs can be realized in the short term, while others can only be achieved in a long-term period (over 10 years). For the former, the feasible use cases include integration of highway asset information with a single source of truth, visualization of inspection and maintenance activities, and decision-making for maintenance interventions. For the latter, predicting future states of roads and proactive maintenance may require the collection of road condition data and maintenance records over many years to develop usable deterioration models and approaches for risk-based predictive maintenance.

The last question (Question 11) in this section allows participants to select the potential benefits of DTs in highway maintenance. The results are presented in Figure 5. Better asset lifecycle management, cost savings, enhanced operational efficiency, and improved decision-making are recognized as the main benefits that digital twins can bring to the field of highway maintenance. It suggests that a road digital twin should incorporate asset management principles to facilitate better operational and maintenance planning. Hence, a DT not only serves as a data hub to store and integrate road data but also provides processes and workflows for lifecycle management, from data collection and data updating to the usage of asset data.

Perceived challenges of DT implementation

In the previous section of the questionnaire, the potential use cases and benefits of road DT are studied. However, it is also important to be aware of the challenges of adopting DT approach for highway organizations to better understand which use cases can be implemented in the short term and which use cases require long-term investment and development, thus deriving a reasonable technical roadmap for DT standardization, development, and promotion. Therefore, the last section of the questionnaire investigates the difficulties perceived by

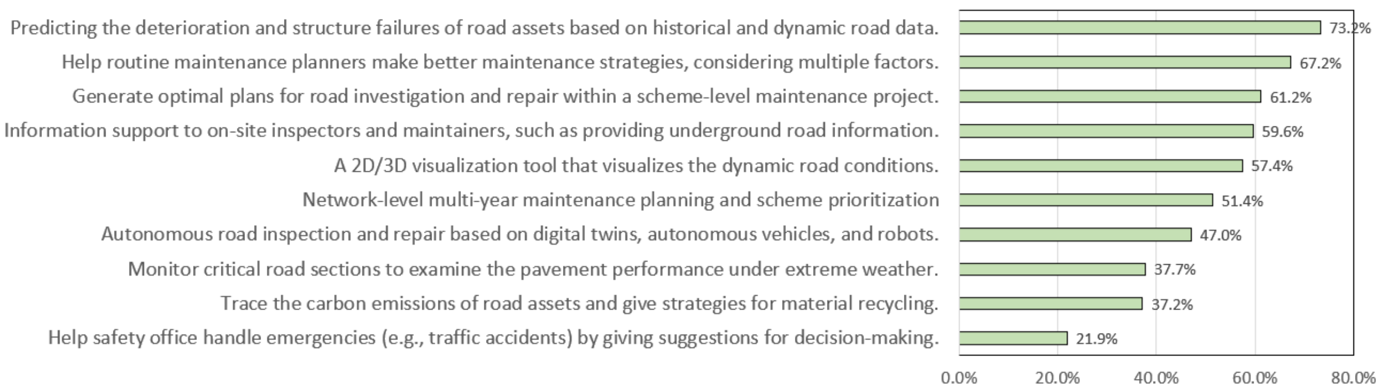


Figure 4. Responses for the Question 9: "Which of the below digital twin-based applications do you perceive are important and valuable in highway maintenance?"

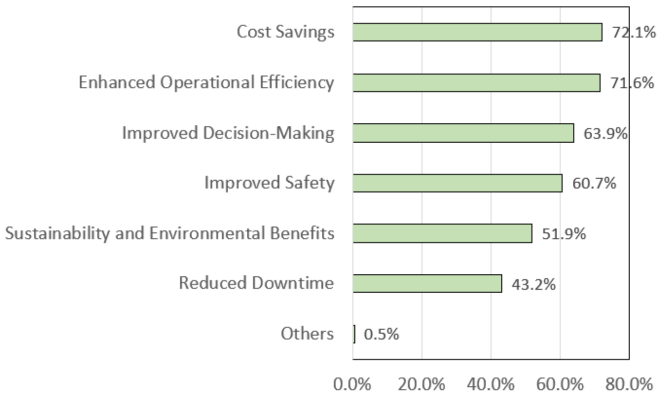


Figure 5. Responses for the Question 11: "Please select the potential benefits you expect digital twins to bring to highway maintenance."

industry practitioners when it comes to implementing DTs for their organizations.

As presented in Figure 6, the responses for Question 12 show that the top 3 challenges of DT implementation are (a) skills and training gaps (61.2% selected); (b) lack of data integration (55.7% selected); and (c) lack of standardization (51.9% selected). It suggests that training and education for highway practitioners to use new DT technologies and systems will be a significant issue because they become accustomed to the existing tools (e.g., GIS). Hence, the design of user-friendly human-DT interfaces is important to reduce the negative impacts of training costs and user reluctance towards new technologies.

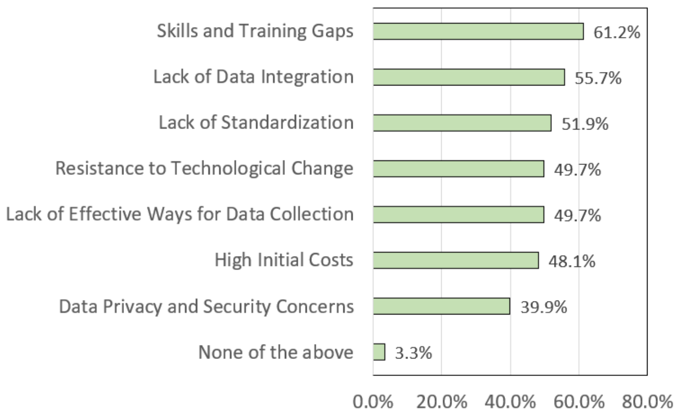


Figure 6. Responses to Question 12: "Please select the primary challenges or barriers you foresee in implementing digital twins for highway maintenance."

The challenge of data integration stems from several aspects. First, the connection of the new DT system with the existing systems and databases used by highway organizations is difficult, requiring a heavy workload of manual alignment between different data schemas. Second, when the road DT has new data requirements about road construction and condition, it is challenging to update the old asset data according to the new requirements, which would induce much missing or mismatched information. Third, designing a data schema for road DT that has sufficient interoperability and scalability is itself an intricate task. This also relates to the 3rd-ranked challenge: lack of standardization. A well-defined data standard is crucial for DT instance creation and information exchange between different DTs.

Apart from the options shown in Figure 6, some respondents also input other problems for implementing DTs in the highway industry, as shown in Table 5.

Table 5: Selected written answers about the challenges of DT implementation.

Other challenges of road DT implementation
1. Lack of political understanding of the urgency.
2. Silo mentality: different sectors or organizations involved in highway maintenance operate independently and without much coordination or information sharing.
3. Technical integration of emerging technologies.
4. Leadership challenge.
5. The cost of maintaining DTs may exceed the values they create.
6. Lack of trust in the analytical processes
7. The DT data obtained from existing databases may have poor levels of accuracy.
8. Procurement of the DT systems and meshing with the existing maintenance systems.

From the above table, policy support, coordination between organizations, and cost estimation for maintaining DTs are also issues for implementing DTs in the highway industry.

Conclusions

Maintaining highway assets at the required service level is critical to ensuring the efficiency and safety of road networks. The digital twin (DT) concept has the potential to be applied in highway maintenance and improves existing processes based on dynamically updating digital road models. This paper presented a questionnaire survey for highway industry practitioners to identify the main barriers in current practice and potential use cases of DTs in this field. The questionnaire was designed based on previous interviews with highway experts in the UK industry, and it was then disseminated globally, with 183 responses obtained from professionals across different roles in the highway industry.

The findings reveal a consensus on the main inefficiencies in highway maintenance: the complexity of making optimal maintenance decisions under constraints, the poor quality of asset inventory information, and the difficulty of acquiring comprehensive road condition data. The survey also indicates that the DT can be potentially used for deterioration prediction, maintenance planning, and scheme development, taking advantage of its well-structured, dynamically updated road asset information in cyberspace.

This study contributes to the knowledge of how industry practitioners perceive the prospects and challenges of road DT in highway maintenance, which provides valuable insights for the design and implementation of road DTs. The results of this survey can be effectively used by academia and industry to develop a purpose-driven approach for applying DTs in the highway maintenance field, where the most valuable use cases with fewer technical and managerial challenges could be implemented first and DTs can be progressively improved to realise more application values. In future work, the survey data will be further analysed based on the roles and experiences of participants.

Acknowledgments

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 101034337. We wish to thank Dr. Pieter Pauwels, Dr. Stefano Cavazzi, Yogesh Patel, Matt Peck, Katrin Johannesdottir, and Simon Hayton for their contributions to this paper.

References

29481-1, I.S.O. (2010) 'Building information modeling–information delivery manual–Part 1: methodology and format'.

Agrawal, A., Fischer, M. and Singh, V. (2022) 'Digital Twin: From Concept to Practice', *Journal of Management in Engineering*, 38(3), pp. 1–13. Available at: https://doi.org/10.1061/(asce)me.1943-5479.0001034.

Brilakis, I. *et al.* (2019) 'Built environment digital twining', in. International Workshop on Built Environment Digital Twinning presented by ….

Chen, K. *et al.* (2022) 'Identifying the most suitable machine learning approach for a road digital twin', *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, 174(3), pp. 88–101.

Han, T. *et al.* (2022) 'A BIM-IoT and intelligent compaction integrated framework for advanced road compaction quality monitoring and management', *Computers and Electrical Engineering*, 100(April), p. 107981. Available at: https://doi.org/10.1016/j.compeleceng.2022.107981.

Jiang, F., Ma, L., Broyd, T., Chen, W., *et al.* (2022a) 'Building digital twins of existing highways using map data based on engineering expertise', *Automation in Construction*, 134(April 2021). Available at: https://doi.org/10.1016/j.autcon.2021.104081.

Jiang, F., Ma, L., Broyd, T., Chen, W., *et al.* (2022b) 'Digital twin enabled sustainable urban road planning', *Sustainable Cities and Society*, 78, p. 103645.

Jiang, F., Ma, L., Broyd, T., Chen, K., *et al.* (2022) 'Underpass clearance checking in highway widening projects using digital twins', *Automation in Construction*, 141(May), p. 104406. Available at: https://doi.org/10.1016/j.autcon.2022.104406.

Koch, C. *et al.* (2015) 'A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure', *Advanced Engineering Informatics*, 29(2), pp. 196–210. Available at: https://doi.org/10.1016/j.aei.2015.01.008.

Kritzinger, W. *et al.* (2018) 'Digital Twin in manufacturing: A categorical literature review and classification', *Ifac-PapersOnline*, 51(11), pp. 1016–1022.

Li, L. *et al.* (2021) 'Digital twin in aerospace industry: A gentle introduction', *IEEE Access*, 10, pp. 9543–9562.

Machl, T., Donaubauer, A. and Kolbe, T.H. (2019) 'Planning agricultural core road networks based on a digital twin of the cultivated landscape', *J. Digit. Landsc. Archit*, 4, pp. 316–327.

Marai, O. El, Taleb, T. and Song, J. (2021) 'Roads Infrastructure Digital Twin: A Step Toward Smarter Cities Realization', *IEEE Network*, 35(2), pp. 136–143. Available at: https://doi.org/10.1109/MNET.011.2000398.

Meža, S. *et al.* (2021) 'Digital twins and road construction using secondary raw materials', *Journal of Advanced Transportation*, 2021, pp. 1–12.

National Highways (2021) *Written evidence submitted by National Highways (RDF0027)*. Available at: https://committees.parliament.uk/writtenevidence/41957/ html/.

National Highways (2023) *Asset Data Management Manual (ADMM) and other management and maintenance guides*. Available at: https://nationalhighways.co.uk/suppliers/design-standards-and-specifications/admm-and-other-management-and-maintenance-guides/.

Sacks, R. *et al.* (2020) 'Construction with digital twin information systems', *Data-Centric Engineering*, 1. Available at: https://doi.org/10.1017/dce.2020.16.

Sinha, K.C., Labi, S. and Agbelie, B.R.D.K. (2017) 'Transportation infrastructure asset management in the new millennium: continuing issues, and emerging challenges and opportunities', *Transportmetrica A: transport science*, 13(7), pp. 591–606.

Steyn, W.J. and Broekman, A. (2022) 'Development of a Digital Twin of a Local Road Network: A Case Study', *Journal of Testing and Evaluation*, 50(6), p. 20210043. Available at: https://doi.org/10.1520/jte20210043.

Tejada, J.J. and Punzalan, J.R.B. (2012) 'On the misuse of Slovin's formula', *The philippine statistician*, 61 (1), pp. 129–136.

Wang, S., Zhang, F. and Qin, T. (2021) 'Research on the construction of highway traffic digital twin system based on 3D GIS technology', in *Journal of Physics: Conference Series*. IOP Publishing, p. 42045.

RENOVATION DIGITAL TWIN FOR BUILDING RETROFIT MONITORING: A SOFTWARE PRODUCT AND AN ORGANIZATIONAL ECOSYSTEM

Omar Doukari* and Ala Suliman
Faculty of Engineering and Environment, Northumbria University, NE1 8ST, UK

Abstract

To meet the European energy-saving and decarbonisation goals, the delivery of renovation projects must be accelerated. Automated processes and software tools are crucial to facilitate and optimise the project execution and onsite monitoring by providing early warnings and notifications to the project team. Digital twins are one of the most promising approaches as they provide real-time data and use simulation to inform and speed up decision-making processes. This paper presents the design, development, and deployment of a Renovation Digital Twin (RDT) that enables monitoring of the retrofitting progress and the provision of feedback to the project stakeholders. The RDT aims at addressing a set of seven design requirements including compliance with the General Data Protection Regulation (GDPR) while sensing building data. The RDT integrates 13 Key Performance Indicators (KPIs) allocated to five application domains: quality, cost, scheduling, safety, and environment.

Introduction

To achieve the energy-saving and decarbonisation goals of the European Union, the delivery of renovation projects must be accelerated. The adoption of digital twining technologies in construction and renovation project is one of the most promising approaches. Digital twins (DTs) refer to the creation of exact digital replicas or virtual representation of an object or system that spans its lifecycle. Each replica possesses the properties and/or behaviours of a system in the physical world. These digital replicas are updated from real-time data, and use simulation to optimise decision-making processes (Tuhaise *et al.*, 2023).

DTs offer several advantages in the construction industry. They enable improved resource management by optimising energy usage, reducing the carbon footprint of buildings, and providing real-time access to as-built models, thus fostering harmonisation between current progress and future deliverables. Moreover, DTs enhance collaboration by enabling all stakeholders to access a shared digital representation of a project, improving communication and collaboration between engineers, architects, contractors, and other team members. Furthermore, they help minimise waste, reduce construction costs, and identify and mitigate potential safety hazards on construction sites, ultimately improving the efficiency, quality, and safety of construction projects while reducing costs (Tao *et al.*, 2018). Hence, in construction industry, DTs can be leveraged to reduce costs, accelerate delivery, enhance the safety, and provide efficient monitoring of projects throughout their lifecycle (Saback de Freitas Bello *et al.*, 2023).

Monitoring activities is a critical task in construction project management. DTs can be utilized to monitor onsite construction activities to offer real-time insights and predictive capabilities. By integrating various data sources such as 3D models, IoT (Internet of Things) sensor data, and building information, DTs enable the continuous monitoring of construction processes, allowing for the early detection of potential issues and the optimization of resource utilization. This proactive approach enhances safety, risk management, and contributes to improved efficiency and quality of construction project (Greif *et al.*, 2020). As a result, the adoption of DTs for onsite monitoring has a great potential to be a valuable tool for construction project management and execution.

Decarbonization and energy-saving are global and urgent goals of the construction industry. Onsite monitoring of renovation activities is crucial to approach these goals as it supports accelerated project delivery. Although DTs is a promising technology in approaching these goals, only a few research has been conducted so far, to effectively demonstrate the design, development, and usage of such technologies. The situation is more challenging in building retrofit where existing tools are usually lacking or adapted from new build construction tools. Hence, this paper aims to present the design, development, and deployment of a Renovation Digital Twin (RDT) that enables monitoring of the retrofitting progress and the provision of feedback to the project stakeholders.

The remainder of the paper is structured as follows. Section 2 summarises research works and gaps related to the topic being addressed by the research. Section 3 presents the research methodology implemented to achieve the research objectives, and introduces the solution proposed which includes the digital twin developed for renovation projects monitoring, integrating a set of KPIs identified with the help of the RINNO project's industrial partners. Conclusions, limitations and future works are finally outlined in Section 4.

Previous Works

The concept of DTs in the context of building renovation and onsite retrofitting works has garnered increasing attention within the architecture, engineering, construction, and operations (AECO) industry. Pan, et al. (Pan *et al.*, 2023) discuss the emerging and anticipated benefits of DTs in the building deep renovation life cycle, emphasizing their potential to reshape the process of assets' construction and maintenance through decision encoding. The authors highlight that while the applications of DTs are still at an early stage, they hold promise for various stakeholders involved in AECO, provided that their full potential is effectively exploited (Pan *et al.*, 2023).

In a similar vein, a systematic review by Opoku et al. (Opoku *et al.*, 2021) delves into the status, evolution, and key applications of digital twins in the construction industry. The review provides insights into six areas of application in the lifecycle phases of a construction project, shedding light on the diverse contexts in which DTs can be leveraged within the construction domain. These application areas include building information modelling, structural system integrity, facilities management, monitoring, logistics processes, and energy simulation (Opoku *et al.*, 2021). Similarly, Tuhaise et al. (Tuhaise *et al.*, 2023) identified key technologies, research gaps, and future research directions, focusing on technologies in data transmission, interoperability, data integration, data processing, and visualization. From a different perspective, Ammar et al. (Ammar *et al.*, 2022) conducted a study that explored the applications and challenges in construction through interviews. Their analysis identified 40 applications and uncovered 34 implementation challenges. The study also presents a case study that exemplifies the practical implementation of DTs in a construction setting, underscoring the potential for different information pipelines from the site to decision-making processes (Ammar *et al.*, 2022).

Research Gaps and Contribution

Managing a renovation project presents several challenges for project stakeholders and policymakers, such as disruption to, and by occupants, which usually lead to more time and cost overruns, more health and safety hazards, and worse quality performance when compared to new-built construction projects (Doukari, Wakefield, *et al.*, 2024). The literature review conducted collectively underscores the transformative potential of DTs in synchronising and enhancing the performance of construction activities and ensuring better control of the cost, time, safety, and quality of construction projects. Despite that, the use of DTs in building renovation and onsite retrofitting works was noticed to be limited in this literature body, and no comprehensive framework has been yet proposed. Hence, implementing this technology in building renovation should offer valuable insights into their evolving role and improve the efficiency of monitoring onsite retrofitting activities.

The contribution of this paper is twofold: (i) present a novel RDT solution (i.e., product system) that enables monitoring of the retrofitting progress and the provision of feedback to the project stakeholders; and (ii) provide detailed descriptions of the Research and Development (R&D) methodology implemented (i.e., organizational system) which includes the design, development, and deployment processes as well as the technological ecosystem required for the software product development.

Renovation Digital Twin – RDT

The RDT platform is developed to enable monitoring of the retrofitting progress and the provision of feedback to

Table 1: Renovation Project KPIs

Application	Description	Calculation	Representation
Quality	1- Number of quality incidents 2- Monitoring the quality controls 3- Number of customer complaints 4- Identification and alerts on recurring quality issues	1- Number of opened quality incident forms 2- Quality controls to be done in the next 10 days 3- Number of customer complaints 4- The 5 most recurrent quality issues	1- Line chart 2- List 3- Line chart 4- List
Cost	5- Cost savings 6- Cost overruns	5- Sum of the registered savings 6- Sum of the registered overrun costs	5- Line chart 6- Line chart
Scheduling	7- Delay monitoring 8- Milestones monitoring 9- Duration for resolving issues	7- Difference between days worked and days scheduled 8- % of achieved, ongoing and upcoming tasks 9- Average duration between opening and closing issues	7- Line chart 8- Pie chart 9- Line chart
Safety	10- Identification and alerts on recurring safety issues 11- Number of safety incidents 12- Monitoring safety incidents control	10- The 5 most recurrent safety issues 11- Number of safety issues 12- Stakeholders involved in safety issues	10- List 11- Line chart 12- List
Environment	13- Monitoring waste	13- Number of recorded waste containers	13- Line chart

the project stakeholders. Its main role is to: (i) gather onsite information about weekly project safety, quality, cost, completion of tasks and delays, and information related to waste management; and (ii) provide the project stakeholders with timely insights to take appropriate actions if needed.

Methodology

The research and development methodology implemented to achieve the research objectives is based on a 5-step process (Figure 1):

- First, workshops with onsite construction workers, project and site managers and engineers are organised in order to identify the most relevant information they need to be regularly reported during the retrofitting phase of a renovation project. For two days, fourteen participants from the RINNO project's industrial partners contributed to the identification of renovation project stakeholders' needs and system requirements in terms of onsite information and progress notification;
- Second, this input was categorised into different application domains, including quality, cost, scheduling, safety, and environment to facilitate KPIs identification;
- Third, for each application domain, a list of KPIs for retrofitting works monitoring is defined, and their sources of data, calculation formulas, representation formats, and frequency of measurement are detailed (Table 1);

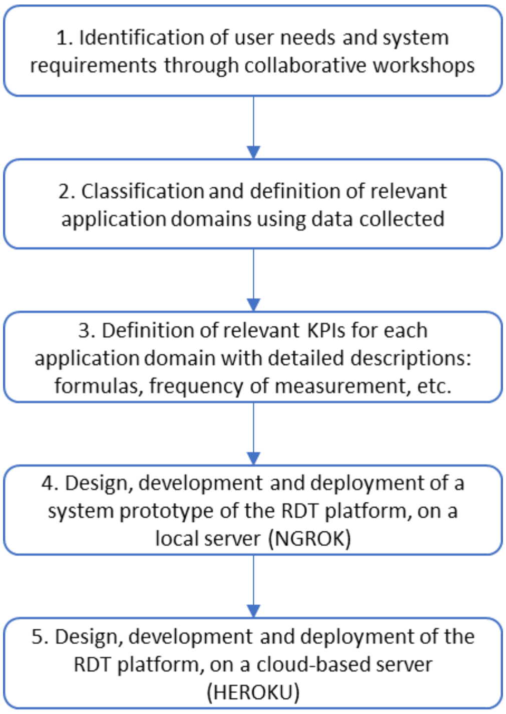


Figure 1: Research and development methods.

- Fourth, a system prototype, including main functionalities and user interfaces, is developed so as to enable both testing and validation with site managers and workers. This first version is deployed on a local machine, and remote access for testing purposes, is enabled using NGROK (NGROK, 2024);
- Fifth, a final version of the RDT platform is developed enabling the integration of the BIM model of the building being renovated along with a set of friendly graphical user interfaces (GUIs) for project KPIs visualisation and user notifications. This final version is deployed on HEROKU (HEROKU, 2024).

Design Requirements

The starting point of the RDT platform development is the set of design requirements (DR), including components and high-level functionalities, that have been identified with the RINNO project's industrial partners (Point 1, Figure 1). As presented in Table 2, seven requirements must be met to develop a suitable digital twin platform for renovation projects.

Table 2: Design requirements of the RDT platform.

DR1: Regularly provide relevant project KPIs regarding quality, cost, scheduling, safety, and environment (Table 1).
DR2: Store historical project KPIs during the retrofitting phase.
DR3: Simulate project KPIs and their progress overtime.
DR4: Provide a user-friendly interface through integrating BIM and relevant charts and/or graphs.
DR5: Enable stakeholders' notification when project KPIs exceed specific thresholds.
DR6: The RDT platform should be accessible anywhere and at any time for all project stakeholders through internet.
DR7: Comply with regulations, such as General Data Protection Regulation (GDPR), while sensing building data.

System Architecture

The RDT platform architecture presented in Figure 2 shows seven main components that help meet and implement the seven design requirements identified in Table 2. The RDT components and their relationships to the design requirements can be summarised as follows:

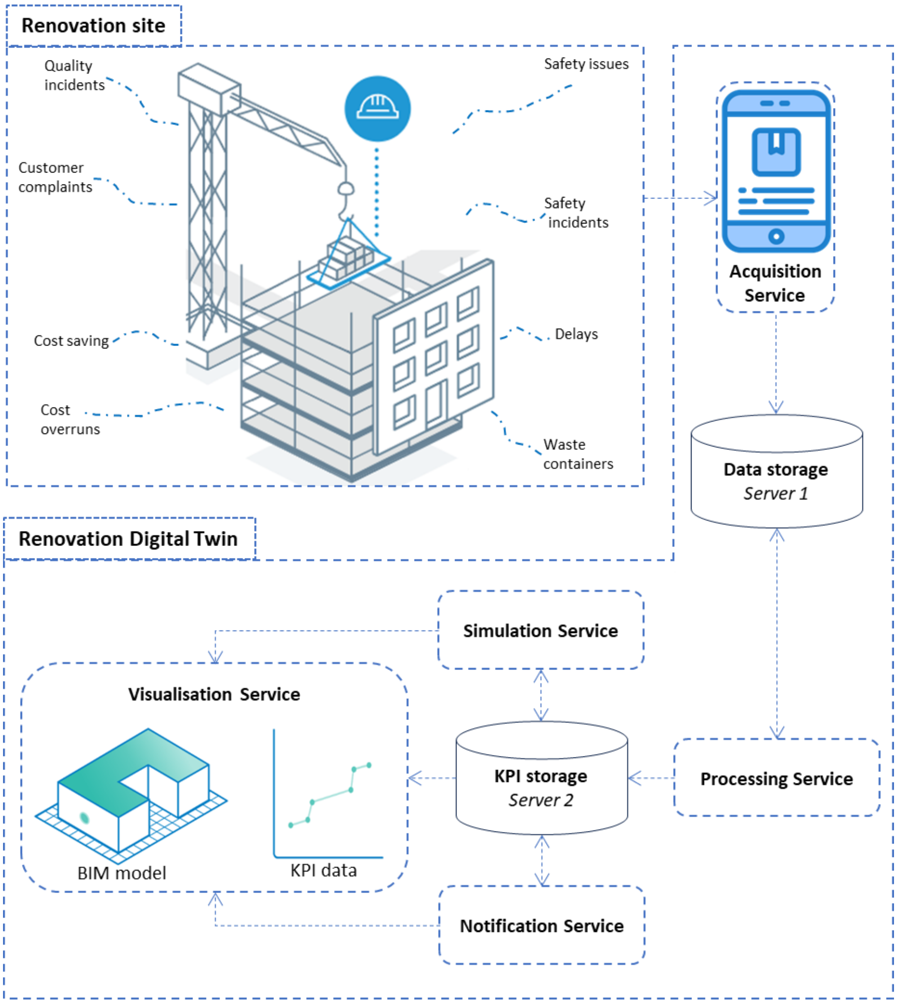


Figure 2: Renovation Digital Twin – RDT System Architecture.

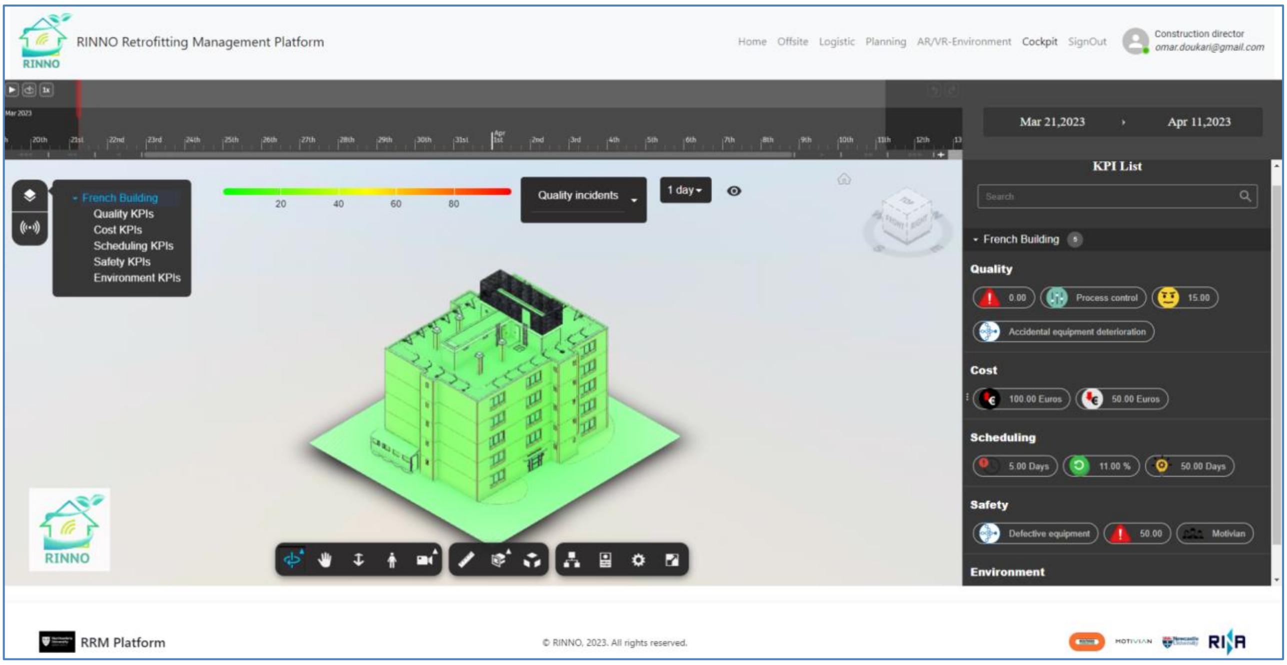


Figure 3: Renovation Digital Twin – RDT main UI.

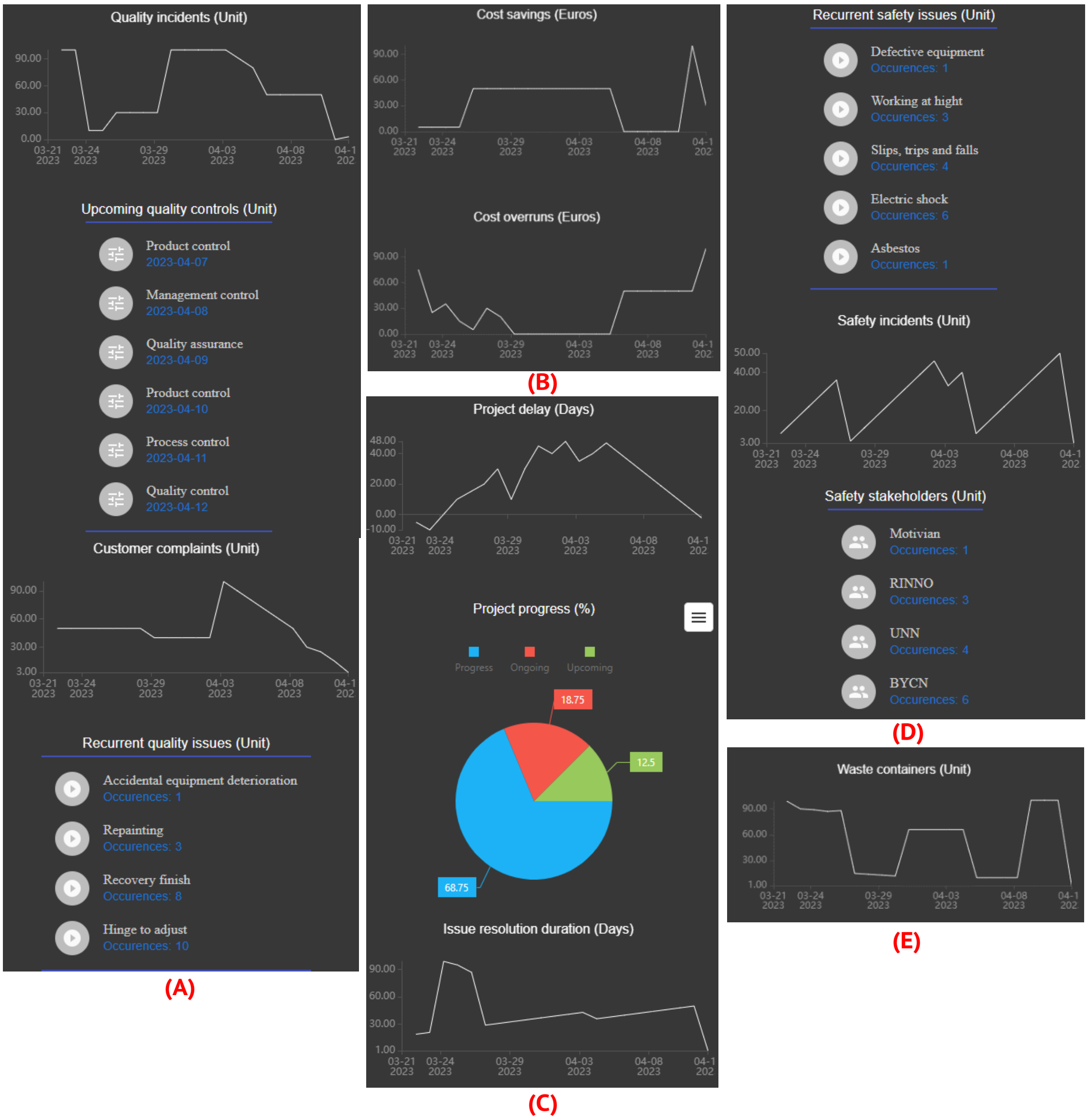


Figure 4: The RDT platform KPIs' dashboard for (A) Quality (B) Cost (C) Scheduling (D) Safety (E) Environment.

- **Acquisition service**: enables collecting raw data (ex. quality incidents, complaints, safety issues, etc.) from the onsite execution of the renovation project, and sending them back to be stored into the Data storage server. To comply with the GDPR regulation as stipulated by **DR7**, the Acquisition service is implemented as a mobile application instead of installing smart sensing devices such as sensors, which may need occupants and operators consent to data access (Doukari, Seck, Greenwood, *et al.*, 2022). The mobile application is used by the project manager to collect daily data related to the project progress and execution.

- **Data storage**: consists of a Structured Query Language (SQL) Server that stores daily project data collected through the Acquisition service. It also provides raw data for the 'Processing service' so that project KPIs can be calculated and stored as required by **DR1**.

- **Processing service**: enables calculating project KPIs using project's raw data stored into the Data storage component. It uses simple SQL queries to fetch data from Server 1, calculates the project KPIs, and then stores them into the KPI storage component (Server 2). As such, the Processing service helps meet requirement **DR1**.

- **Notification service**: consists of a 'background process' that runs continuously as a 'KPI Watcher' and triggers notifications when KPI threshold values are reached or exceeded. For each KPI in Table 1, min and max values are defined. While the min values are set to check data quality and consistency, the max values help trigger notifications via color-coded information displayed through the Visualisation service. The Notification service addresses requirement **DR5**.

- **Simulation service**: consists of simulating historical KPI values which are stored in Server 2 and their progress overtime. The result is visualised through a 'Timeliner' implemented and integrated into the Visualisation service. The KPI values are also simulated as shading and color-coded data into the BIM model elements using Forge APIs (Application Programming Interfaces). As such, the Simulation component enables '4D BIM simulation' (Doukari, Seck and Greenwood, 2022) aiming to meet requirement **DR3**.

- **KPI storage**: consists of a second SQL Server that stores project KPIs calculated by the Processing service. It feeds all RDT services (i.e., Simulation, Notification, and Visualisation) with KPI values through responding to SQL queries. This component implements **DR2**.

- **Visualisation service**: implements the main component where all visualisation, simulation and notification are displayed. The service integrates the BIM model of the building being renovated through Forge APIs, so as to allow end-users to access and easily visualise data related to current and/or historical project KPIs regarding quality, cost, scheduling, safety and some environmental aspects such as waste. This includes visualisations as shading created on top of the BIM model, 4D BIM simulations using the Timeliner, and through charts and graphics as illustrated in Figures 3 and 4. The Visualisation service answers requirement **DR4**.

Finally, to comply with requirement **DR6** and enable project stakeholders to access data anywhere and at any time, the RDT platform's architecture is designed and developed as a web-based application.

Development & Deployment

The RDT is developed using JavaScript for the backend, and React library (React, 2022), CSS and HTML technologies for the frontend (Figures 3 and 4). Particularly, MUI library (MUI, 2022) is used in order to ensure simplicity, clarity and responsiveness of the user interface (UI) components implemented, and so enable visualisation in different contexts (onsite and offsite) and on several device types with different hardware and software specifications, including smartphones, tablets and laptops. The RDT integrates Forge APIs and its main

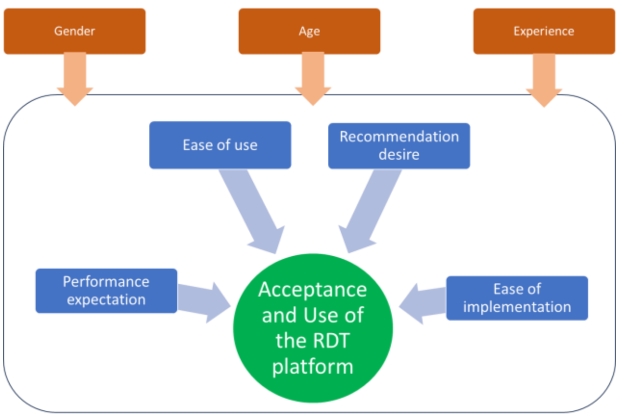


Figure 5: UTAUT model adapted to test and evaluate the acceptance of the RDT platform.

UI was inspired by a previous prototype version developed and shared by Autodesk (Sample Application, 2023).

As for the RDT deployment, it is processed following a 2-step process. First, for testing purposes, the platform is deployed on a local machine through NGROK (NGROK, 2024) and made accessible to the RINNO project's industrial partners so as to demonstrate its design and validate its functionalities. Second, the RDT is finally deployed on a permanent basis through HEROKU which is a cloud-based platform providing publication services for third-party web-based applications (HEROKU, 2024). A short demonstration video of the RDT platform in action can be visualised at (Doukari, 2024). The developed platform integrating the BIM model of the French demonstration site, which is one of the RINNO project's case studies (Doukari *et al.*, 2021), can be accessible online at (Doukari, 2023).

Conclusion, Limitations, and Perspectives

Despite the clear and obvious added values that DTs can provide to the construction industry, only little research have been conducted to showcase practical development and use of such technologies. The situation is more complex in renovation projects because of the dearth of digital tools and automated processes which are often adapted from new build construction tools, while such technologies are much more needed in building retrofit in order to expedite the delivery of renovation projects and meet the European energy-saving and decarbonisation goals by 2050 as per the EU commitment. In addition, specific requirements, such as complying with the GDPR regulation, may need to be addressed while conducting retrofitting works. To overcome these challenges, this paper introduced a Renovation Digital Twin (RDT) platform to enable monitoring of the retrofitting works and so accelerate the rate and amount of renovation projects in Europe. The adopted approach included the design, development and deployment of the RDT so as to document the process implemented and so enable reproducibility of the results. However, the research presents some limitations which can be summarised as follows: (i) the BIM model integration through the 'Visualisation service' is based on a semi-automatic process where the digital mock-up of the building under renovation first needs to be created and uploaded in the cloud, and then related (hosting) data must be updated within the RDT. To enable an efficient lean-based renovation project management as described in (Doukari, Kassem, *et al.*, 2024), a smooth and automatic BIM data preparation needs to be adopted; (ii) the 'Simulation service' enables 4D BIM simulation but only to visualise historical KPI values. There should certainly be much more benefits for the project participants if the simulation capabilities could be extended to predict future KPI trends using machine learning techniques (Rogage and Doukari, 2024) and/or a scenario-driven approach (Doukari, Wakefield, *et al.*, 2024); and finally (iii) future extensions of the research should investigate the acceptance and use of the RDT by end-users through its application and implementation within a real-world renovation project. An adapted version of the UTAUT model (Unified Theory of Acceptance and Use of Technology) (Dwivedi *et al.*, 2019) will be used to evaluate the RDT's performance and overall suitability for the management and delivery acceleration of renovation projects (Figure 5).

Acknowledgments

The authors would like to gratefully acknowledge the useful assistance, help and support received from Boubacar SECK and his MSc. BIM students at CESI Engineering School in Paris (France) during the deployment of the RDT platform (CESI, 2024). This work is supported by the European Union's Horizon 2020 Research and Innovation Programme through the RINNO project (https://RINNO-h2020.eu/) under Grant Agreement Number 892071.

References

Ammar, A., Nassereddine, H., AbdulBaky, N., AbouKansour, A., Tannoury, J., Urban, H. and Schranz, C. (2022), "Digital Twins in the Construction Industry: A Perspective of Practitioners and Building Authority", Frontiers in Built Environment, Vol. 8.

CESI. (2024), "École d'ingénieurs, Bachelors et Mastère Spécialisé® - CESI Paris-Nanterre", Campus de Nanterre, available at: https://paris.cesi.fr/ (accessed 1 January 2024).

Doukari, O. (2023), "RINNO Digital Twin", available at: https://rinno-digital-twin-369ee4493559.herokuapp.com/ (accessed 17 January 2024).

Doukari, O. (2024), "Renovation Digital Twin - RINNO Project", 26 January, available at: https://www.youtube.com/watch?v=iLpntlXfOYE (accessed 26 January 2024).

Doukari, O., Kassem, M. and Greenwood, D. (2024), "A distributed collaborative platform for multi-stakeholder multi-level management of renovation projects", Journal of Information Technology in Construction (ITcon), Vol. 29 No. 11, pp. 219–246, doi: 10.36680/j.itcon.2024.011.

Doukari, O., Lynn, T., Rosati, P., Egli, A., Krinidis, S., Angelakoglou, K., Sougkakis, V., et al. (2021), "RINNO: Transforming Deep Renovation through an Open Renovation Platform", presented at the ICDS The Fifteenth International Conference on Digital Society, Nice, France.

Doukari, O., Seck, B. and Greenwood, D. (2022), "The efficient generation of 4D BIM construction schedules: A case study of the Nanterre 2 CESI project in France", Frontiers in Built Environment, Vol. 8.

Doukari, O., Seck, B., Greenwood, D., Feng, H. and Kassem, M. (2022), "Towards an Interoperable Approach for Modelling and Managing Smart Building Data: The Case of the CESI Smart Building Demonstrator", Buildings, Multidisciplinary Digital Publishing Institute, Vol. 12 No. 3, p. 362, doi: 10.3390/buildings12030362.

Doukari, O., Wakefield, J., Martinez, P. and Kassem, M. (2024), "An ontology-based tool for safety management in building renovation projects", Journal of Building Engineering, Vol. 84, p. 108609, doi: 10.1016/j.jobe.2024.108609.

Dwivedi, Y.K., Rana, N.P., Jeyaraj, A., Clement, M. and Williams, M.D. (2019), "Re-examining the Unified Theory of Acceptance and Use of Technology (UTAUT): Towards a Revised Theoretical Model", Information Systems Frontiers, Vol. 21 No. 3, pp. 719–734, doi: 10.1007/s10796-017-9774-y.

Greif, T., Stein, N. and Flath, C.M. (2020), "Peeking into the void: Digital twins for construction site logistics", Computers in Industry, Vol. 121, p. 103264, doi: 10.1016/j.compind.2020.103264.

HEROKU. (2024), "Cloud Application Platform | Heroku", available at: https://www.heroku.com/ (accessed 25 January 2024).

MUI. (2022), "MUI: The React component library you always wanted", available at: https://mui.com/ (accessed 19 October 2022).

NGROK. (2024), "ngrok for local deployment", 23 January, available at: https://ngrok.com/docs/guides/getting-started/ (accessed 25 January 2024).

Opoku, D.-G.J., Perera, S., Osei-Kyei, R. and Rashidi, M. (2021), "Digital twin application in the construction industry: A literature review", Journal of Building Engineering, Vol. 40, p. 102726, doi: 10.1016/j.jobe.2021.102726.

Pan, Y., Hu, Z. and Brilakis, I. (2023), "Digital Twins and Their Roles in Building Deep Renovation Life Cycle", in Lynn, T., Rosati, P., Kassem, M., Krinidis, S. and Kennedy, J. (Eds.), Disrupting Buildings: Digitalisation and the Transformation of Deep Renovation, Springer International Publishing, Cham, pp. 83–96, doi: 10.1007/978-3-031-32309-6_6.

React. (2022), "React", available at: https://reactjs.org/ (accessed 17 October 2022).

Rogage, K. and Doukari, O. (2024), "3D object recognition using deep learning for automatically generating semantic BIM data", Automation in

Construction, Vol. 162, p. 105366, doi: 10.1016/j.autcon.2024.105366.

Saback de Freitas Bello, V., Popescu, C., Täljsten, B. and Blanksvärd, T. (2023), "Analysis of Digital Twins in the Construction Industry: Current Trends and Applications", Vol. 2, presented at the International Symposium of the International Federation for Structural Concrete, fib Symposium 2023, June 5-7, Istanbul, Turkey, Springer, pp. 1080–1088.

Sample Application. (2023), "Autodesk-Forge/forge-dataviz-iot-reference-app", JavaScript, Forge Platform, 5 November.

Tao, F., Cheng, J., Qi, Q., Zhang, M., Zhang, H. and Sui, F. (2018), "Digital twin-driven product design, manufacturing and service with big data", The International Journal of Advanced Manufacturing Technology, Vol. 94 No. 9, pp. 3563–3576, doi: 10.1007/s00170-017-0233-1.

Tuhaise, V.V., Tah, J.H.M. and Abanda, F.H. (2023), "Technologies for digital twin applications in construction", Automation in Construction, Vol. 152, p. 104931, doi: 10.1016/j.autcon.2023.104931.

THE INNOVATIVE POTENTIAL OF GENERATIVE PRE-TRAINED TRANSFORMERS (GPTs) FOR QUALITY INSPECTIONS IN SWEDISH CONSTRUCTION PROJECTS

Dimosthenis Kifokeris[1], Jan Kohvakka[2], Christian Koch[3,4], and Donia Aslanzadeh[5]
[1]Chalmers University of Technology, Gothenburg, Sweden; [2]Incoord, Stockholm, Sweden; [3]Halmstad University, Halmstad, Sweden; [4]University of Southern Denmark; [5]Robert Dicksons Stiftelse, Gothenburg, Sweden

Abstract

Approaching quality inspection plans in Swedish construction projects as mere checklists and minimizing the clients' involvement, can reduce their value. We propose improving this process through a general cloud service concept for clients, designers, and contractors, utilizing generative pre-trained transformers (GPTs). Methodologically, we synthesize literature insights on GPT uses for construction, and empirical inquiries on developing a quality self-inspection service. We posit that through such a service, project knowledge, known quality defects and lessons-learned from previous cases can be better accessed and shared – potentially leading to time savings, suggesting best practices, and improving the collaboration among clients, designers, and contractors.

Introduction and background

In the Swedish construction sector, quality errors and defects have been annually costing ca. 100 billion SEK (ca. 8.8 billion €) (Boverket, 2018). According to Koch et al. (2020), quality problems derive from a broad set of causes, incl. design failures (e.g., detail mistakes, incomplete documents, and faulty technical aspects) execution errors on site, weather disturbances, and more. Koch and Jonsson (2015) showed that the presence of such quality defects forms a recurrent reality in Swedish construction sites.

The existence and reproduction of quality defects must be understood in context. In Sweden, the responsibility of complying with the relevant regulatory framework is handed over to the contractors and designers through a series of self-inspections and reporting of inspection and testing, labelled as "self-control" (Tolstoy, 2012; Koch and Jonsson, 2015) – statutory in the Swedish Planning and Building Law (PBL 2010:900). Critically, the industry's established practice in connection to quality self-inspections reflects "a sufficiently high quality" logic that has been reported to reduce the process into more of a "check listing" rather than an actual business strategy (Koch and Jonsson, 2015). In some cases, site managers and/or craftsmen take the implementation of quality self-inspection upon themselves (Koch and Jonsson, 2015). This is a practice which, while potentially benefitting from professional experience (Koch and Jonsson, 2015), can indeed be adversely affected by complex and ingrained behaviors and a lack of communication and commitment between both the professionals enacting the self-inspections. The collaboration with the clients can also suffer (Engström and Stehn, 2016) – with the latter often experiencing a reduced understanding of and commitment in the process and its results. Underpinning this problem, Koch and Buser (2020) have also identified that the formal quality control system in Sweden (incl. self-inspections) is unable to bridge the gap of quality understanding between the project and company headquarter levels.

Given the above, and as social practices in construction (incl. reactive and proactive problem solving) can exacerbate the reproduction of quality defects (Koch and Schultz, 2019), it can be derived that the way self-inspections are understood, implemented, and shared by the relevant actors, is connected with the recurrence of quality defects in Swedish projects.

Therefore, in this paper we address the research question of how we can improve quality self-inspection in Swedish construction, by proposing the concept of a general cloud quality self-inspection service for clients, designers, and contractors, based on generative pre-trained transformer (GPT) AI. The core proposition is that professionals performing quality controls can be provided with a much larger knowledge basis for planning and executing inspections through GPT systems. GPT systems, while popularized through ChatGPT, are little understood in the construction context, with only few studies mapping relevant opportunities and challenges (e.g., Ghimire et al. (2024)), sparse early stage attempts of leveraging such systems for construction applications (e.g., Zhen and Fischer, 2023), and virtually no efforts explicitly focused on construction quality.

Following this Introduction, the paper's research method is described. Then, a literature review on the limited GPT state-of-art for construction is conducted. Afterwards, our relevant empirical insights, as derived from four consecutive research projects (three concluded and one ongoing) on the development of a Swedish cloud-based service for quality self-inspection, are summarily offered. The insights from the literature review and our empirical studies are then synthesized into our proposed concept. Finally, we conclude with some discussion points, end remarks, and recommendations for future research.

Research method

This paper builds on the aforementioned background, a literature review, and empirical work, to develop a concept attempting to address the stated research question.

The literature review of this study concerns a summary of standing quality-related considerations within the

Swedish construction sector and its self-inspection processes (forming the background for the introductory section), as well as the (very limited) state-of-art-research on using GPT-based tools for construction. It is based on a concept-centric literature review enhanced by units of analysis that was conducted in iterations (Webster and Watson, 2002). The main searched concepts were "construction quality in Sweden", "quality inspection in Sweden", and "application of GPT in construction". The emerging units of analysis included, indicatively, "GPT use for construction quality". Finally, exclusion and inclusion criteria (e.g., contextual relevance) were applied on the found sources (Dundar and Fleeman, 2017) for finally resulting in the ones featured in the current study.

Our empirical material draws from three development projects in Sweden (concluded, respectively, in 2018, 2019, and 2022), as well as an ongoing one (2023-2024), where we have been working on developing a digital tool as a cloud service (software-as-a-service) for quality assurance and self-inspection in construction. This cloud service is intended as a tool for clients, designers and contractors who actively want to work with quality assurance and self-inspection – as it has been shown that this active and systematic work is expected to contribute to better project quality and a reduction in errors and defects. Our empirical insights were obtained in those projects through various qualitative methods (Bell et al., 2019), including, among others, expert interviews, focus group meetings, market observations, conceptualizations of recommender systems based on different (not GPT) algorithms, and an incorporation of methods of analyzing construction quality defects and their sources, like the "stumbling stone" analysis in Apelgren et al. (2005).

Finally, we synthesize the research units of our literature and empirical results (Bell et al., 2019) into our proposed concept of a cloud quality self-inspection service for clients, designers, and contractors, based on GPT-powered AI, and discuss its potential strengths and weaknesses, as well as opportunities and challenges in its implementation.

Literature review

In order to substantiate this paper's background, accompany its empirical material, and inform the conceptual design of the cloud GPT-based platform, this literature review will go through the state-of-art uses of GPT within construction in general, and with regard to quality control in particular.

In short, generative pre-trained transformers (GPT), as a framework of generative AI, are a type of large language model (LLM) – i.e. artificial neural networks used in natural language processing (NLP) tasks (Abdullah et al., 2022). GPTs are based on the transformer architecture – namely, they are pre-trained on very large datasets of unlabeled text in order to generate human-like responses to queries (Abdullah et al., 2022). According to Saka et al. (2023), while in the 2020s other LLMs like BERT have been experimentally applied in construction for a few cases outside of quality management, GPT models are still relatively new in the field. This is indicated in the few retrieved relevant publications describing dedicated systems – with none of them explicitly addressing quality management and control (Abioye et al., 2021; Amer et al., 2021; Prieto et al., 2023; Uddin et al., 2023; You et al., 2023; Zheng and Fischer 2023; Saka et al., 2024). Moreover, the use of GPT in those studies also exhibits particular limitations per case.

In particular, Amer et al. (2021) used the light version of GPT-2 to integrate master schedules with look-ahead plans, with the main limitation being the intense need for manual data pre-processing. Abioye et al. (2021) discussed how GPT models can assist the project team in performing impact analyses on how change orders can affect the project scope, and identified the appropriate representation of fragmented language as a major limitation. Prieto et al. (2023) used ChatGPT-3.5 for supporting the scheduling of construction tasks, while Uddin et al. (2023) implemented it to facilitate site hazard recognition and support construction safety education. However, none of the latter two studies featured a quantitative evaluation of their results, and the systems exhibited inadequacies with zero-shot learning (i.e., the ability of the learner to correctly predict the class of new observed samples, without the latter belonging to a class observed through training (Xian et al. (2020)). Zheng and Fischer (2023) introduced a prompt-based visual assistant integrating ChatGPT-3.5-turbo with BIM for better information retrieval, but the system was limited to single-turn conversations and no quantitative evaluation of the study's results were offered. You et al. (2023) used ChatGPT-4 for automated construction task sequence planning in robot-based assembly, which was however limited to not being able to utilize visual information. Aladağ (2023) assessed the performance of ChatGPT in construction risk management as moderate, with it providing more accurate knowledge in risk response monitoring, and less so in risk identification and analysis. In a later study, Zhang et al. (2024) used ChatGPT-4 for automated data mining connected to building energy management, which showed the potential of the system in energy management, but also identified its limitation connected, mainly, to its demarcated domain knowledge and almost exclusive reliance on the users' own skills to understand and evaluate its prompt results. Finally, Saka et al. (2024) developed a GPT-based material selection and optimization prototype that integrated BIM data through the OpenAI API, with the limitations of it requiring the identification of the specific component to provide accurate responses.

Underpinning the potential exhibited in systems such as the above, Ghimire et al. (2024) identified relevant opportunities for using GPT in construction, as in, indicatively, document and data management, AI-generated designs, task forecasting, project data synthesis, and material assessment. However, they also identified challenges connected to, among others, the GPT learners' limited domain knowledge, the over-reliance on the users' analytical skills to interpret GPT-generated results with regard to their accuracy, generalizability, and interpretability, a high cost of integration, non-updated regulations on implementation, and potentially unethical data use (Ghimire et al. 2024).

Notably, full publications on utilizing GPT systems for construction quality issues could not be found during this review, with only some initial ideas identified in pre-prints like the one by Rane et al. (2023) – where the authors investigate GPT along with augmented reality (AR), virtual reality (VR), and BIM, for identifying defects and deviations from standards in quality control. Also, at this early-stage study, the aforementioned opportunities and challenges identified by Ghimire et al. (2024) are still apparent. Nonetheless, leveraging GPT for issues identified explicitly within the context of quality self-inspection in *Swedish* construction projects (like the "check listing" and a reduction in the clients' involvement in the process, as described in the Introduction), has not been investigated yet. In the latter case, it could be expected that, to an extent, similar opportunities, and challenges as the above should apply – but with a strong contextualization come much more specific considerations, which would require dedicated investigation.

Empirical insights

Corroborating the background studies presented earlier, a central challenge highlighted by our previous projects is that research-based knowledge of quality assurance and self-inspection is limited. Moreover, there are examples in practice where the relevant actors do not know what should be included in a self-inspection procedure, or what a verified self-inspection is. Developing relevant self-inspection lists for specific projects can also be challenging; self-inspection that is not project-adapting can make the control procedure and experience feedback ineffective.

Furthermore, client organizations that do not have tools, knowledge or a process for active quality control may need help addressing queries about construction technology, material selection, standards, and regulations, as well as getting proposals for solutions to technical challenges, analyzing commonly occurring errors and defects, and proposing effective (self-)control.

Designers and contractors may face similar challenges with every new project, and doing this in a valuable way usually requires a review and analysis of previous experiences. Finally, a proposal for self-inspection must be specifically adapted for the project at hand at least up to a requisite point. It has been observed that, normally, clients approve of time allocated for self-inspections, but this time often ends up being utilized in other activities due to the lack of planning skills. This may even mean that quality self-inspections are maybe considered as "non-value-adding" activities by some actors and are thus not prioritized.

Therefore, developing and implementing a solution to the challenges above can accelerate both the interest and understanding of self-inspection as an aid in value creation within the Swedish context. This could result in a construction sector where the customer receives the product they ordered, and the designers and contractors feel professional pride and commitment when delivering projects of high quality.

This ambition has guided our development process so far, which has already resulted in a prototype of a cloud service for quality control. A well-structured system can have a critical impact on the generation of project specific prompts and self-inspection lists. In this vein, the system should be able to use multiple and reliable data sources (with Swedish examples including BBR, Säker Vatten, GVK, PBL, the clients' own local data sources, etc.) to add context to the prompt and list generation. Letting the system interact with such data sources (e.g., by pulling the relevant data for a specific query) will significantly increase the efficiency of content generation. However, AI as such, least of all GPT, has not been part of this platform up until the latest (ongoing) project. During the latter, we have identified that implementing a GPT AI layer on top of the existing cloud service prototype, can potentially help addressing the aforementioned client needs (especially with regard to knowledge requirements for getting value out of self-inspection reports) in an educational and intuitive way. Moreover, it can also help with time savings for contractors, in relation to retrieving lessons learned from previous experiences and adapting the self-inspection template per project for contextualizing the work ahead.

Therefore, implementing GPT AI in the existing cloud service prototype can potentially not only increase the service's area of implementation, but also increase user value. Moreover, communication among the relevant stakeholders can be improved, and a new type of quality data (incl. textual data in queries) can be collected and analyzed. This insights can also be supported by investigating the way currently available GPT models are utilized to explore possible areas of use. Depending on the case, the models have been trained in large datasets, thus becoming suitable for a range of different tasks – which are, however, interrelated within the domain of the use case. In addition to answering questions and generating content, GPT systems can also analyze text, assess probabilities and relevance, suggest improvements, explain complex contexts, and suggest best practices.

As probably the most popular GPT language models, OpenAI's various versions of ChatGPT are trained on vast datasets connected to multiple topics, in order to generate text and answer user queries in a human-like manner. ChatGPT has only been publicly available for a short time, but interest has grown rapidly. This may indicate that, even in niche services, there is a high demand based on the GPT technology. The cloud support for the improvement of quality self-inspection in (Swedish) construction projects can be considered such a niche service with, as of yet, no use cases featuring the implementation of a GPT. Therefore, the uniqueness of such an implementation should not only lie in the implementation of the GPT itself, but in the value that the GPT along with the cloud service and its infrastructure can create for the users. It has been ascertained that by implementing GPT AI in the intended way, the cloud service has the potential to become very attractive to clients, designers, and contractors, and create value for them when engaging in quality assurance and self-inspection through the service.

The insights above have consequently led to some crucial considerations, which in turn reflect those obtained from the literature review. In particular, the results of a potential GPT-powered cloud service should be evaluated in terms of reliability, consistency, interpretability, generalizability, and data demands and usage. Basically, the results from the language model should be interpreted as advice and recommendations that still need to be evaluated by the user – which would in turn require the user (e.g., the client) to have some competence within the field for which the GPT is used. Finally, the implementation of the GPT AI should be evaluated from an ethical perspective (e.g., regarding privacy and ownership of data), in terms of the risks involved in implementing it as a language model or NLP support, and most importantly, by considering what kind of value streams are created and for whom.

Moreover, standing inquiries have emerged in terms of the system's conceptualization, including, but not being limited to, the following:

- Simulated prompts: From where can the system get its inspiration? Is it going to be agglomerating data related to previous similar projects from the same or another client, known errors and defects, professional roles, working tasks, relevant standards on quality management (like ISO 9001:2015) and/or risk management (like ISO 31000:2018) – or a combination thereof?
- System-generated content: Could it be both the self-inspection checklists themselves (e.g., regarding HVAC installations), and, preliminarily, suggestions for the prompts leading to such checklists (which would in turn assist users with a limited knowledge on the relevant domain)?
- The web app perspective: What will each professional role's experience be when using the system? How may the system's interface and content be adapted for the needs of the designers and contractors (who, based on own experience, may anticipate certain quality problems and self-inspection requirements in the project at hand), or the clients (who may not know and be able to predict quality problems that may emerge)?

Addressing the aforementioned insights, considerations, and standing inquiries is still a work in progress. Nonetheless, the proposed concept described in the following section can act as the relevant basis.

Conceptual design

In this section, we will frame the concept of our proposed GPT-based cloud quality self-inspection service for clients, designers, and contractors, based on the context described in the Introduction and the insights and considerations derived from our literature review and empirical project work.

First, the process graph in Fig. 1 (see next page) conceptualizes the interaction of the system's workflow within a project use case, with the different professional roles of the users who intend to use it. In particular, the process is initiated with the client creating a project, by inputting attributes like the project's name and number, sub-projects (i.e., processes connected to either planning/design or construction) and their names (e.g., SP01 Planning, UP02 Production), and dates for self-inspections (also shared with the designer and/or general contractor via emails). Then, the project is kicked off in the app and an invitation is automatically sent to the designer and/or general contractor. After the latter accept the invitation, their work in the project starts by creating a discipline list related to each sub-project, which can either be selected from a drop-down list, or defined by the user. Examples of such disciplines can indicatively include HVAC, plumbing, or electrical system installations. Then, they can invite other users as the lead of one or more of the disciplines. The leads can then create quality self-inspection checklists, either by themselves or by inviting more users to collaborate. Finally, the leads and their invited users (if any) create checks and complete them; with this completion, a progress report is generated into a PDF file, which can then be shared or downloaded.

Then, in Fig. 2 we schematically represent the concept of how GPT should work in the background for the realization of the platform's process steps described previously. Specifically, we show how the system can potentially build its prompts and generate its suggested checklists and reports by interacting with the user, utilizing relevant SQL and Vector databases for relational creation, storing, updating, and retrieval of data, using auxiliary tools like web search, and, primarily, calling and retrieving data from the LLM through its API. The content generation depends on the user's role, the databases used, and the training of the LLM informing the GPT. Over the graph itself, we include the problematization about whether the user input should follow a different generation rule depending on the user's profession, i.e., whether it should be role-based generation (potentially more suitable for designers and/or contractors) or risk-based generation (namely, returning risk analyses based on project description, which is potentially more suitable for clients). Resolving this problematization is still a work in progress, as, for example, we have so far not engaged in correlating potential risks with the presence of quality defects. However, it can be reasonably assumed that it will not affect the general structure of the concept, but rather the content of the prompts, the suggested lists, the informing databases, and even the LLM.

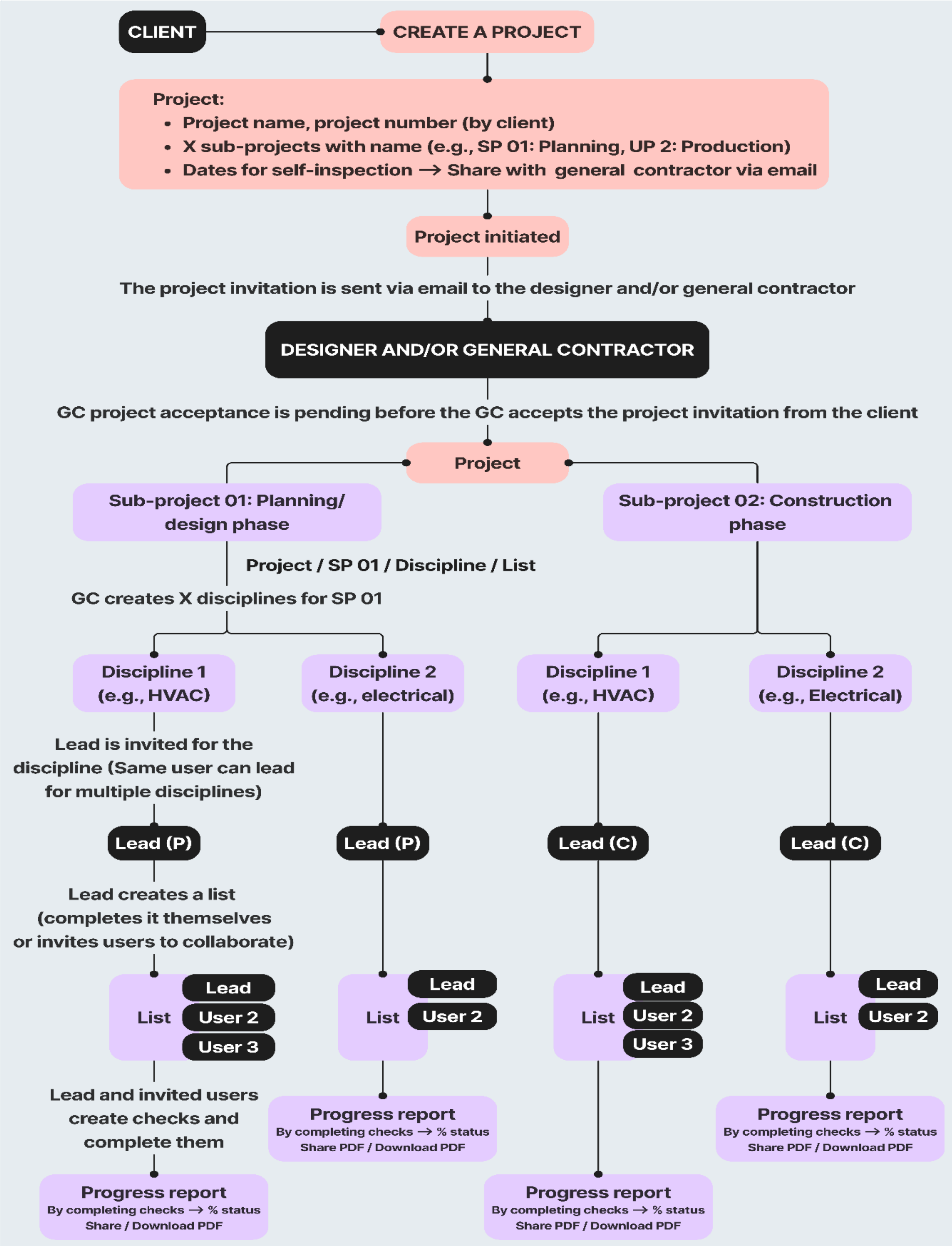


Figure 1: Interaction of the proposed system's workflow with the different professional roles using it

To further enhance the understanding of how the GPT would interface with the user, we offer the following interaction example connected to the project description analysis step:

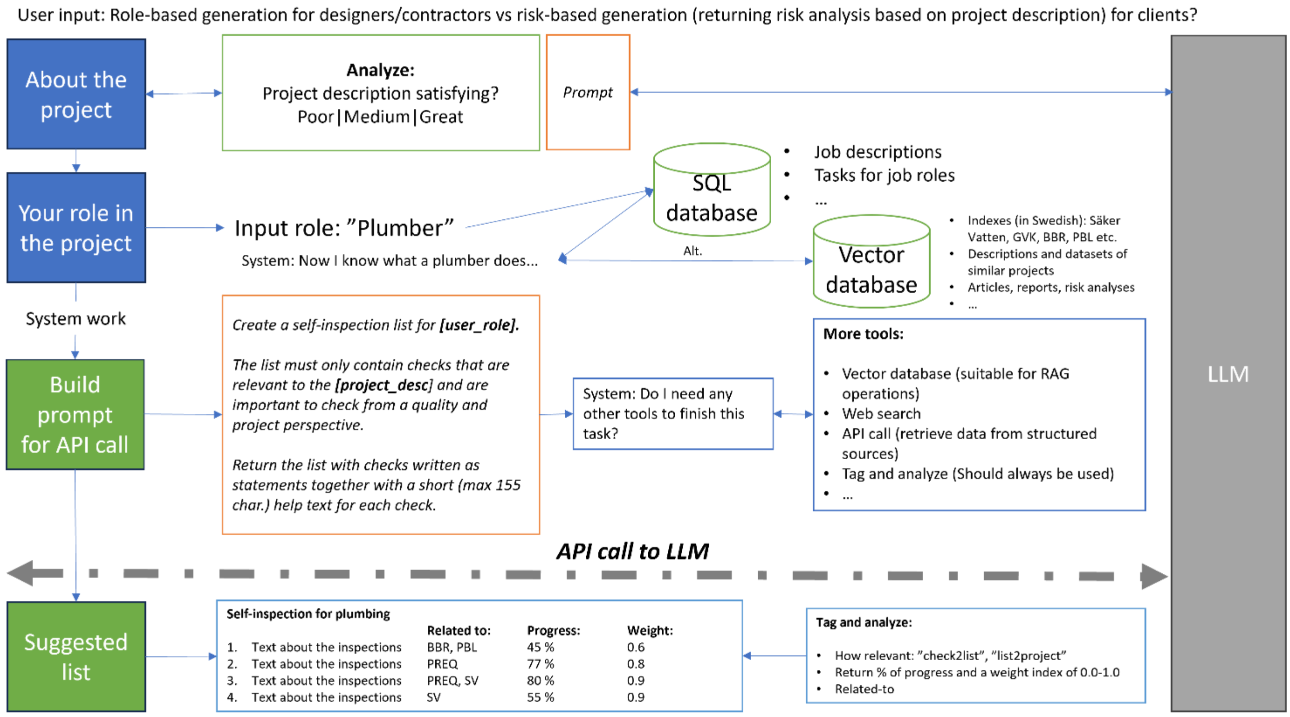


Figure 2: Schematic representation of how GPT should work in the background for the realization of the platform's process steps

You are a semantics and text analyzer, and your task is to understand the meaning of a given text and analyze it.

[text start]

A construction project has been initiated, which involves the raising of a new multi-story residential building commissioned by KKVC. This development will house 50 apartments and will feature comprehensive HVAC and electrical installations to ensure optimal conditions for future residents. The design aims to deliver comfortable, modern living spaces that meet current housing demands. Specific details about the project's size, budget, timeframe, and phases have not been provided, but the core focus will be on delivering a residential building with high-quality internal systems for heating, ventilation, air conditioning, and electrical services.

[text end]

Analyze this text and grade it. You should specifically look for these components:

- *How well does the text describe the project in general?*
- *Is there information in the text about the project budget?*
- *Is there information in the text about the project timeframe?*
- *Is there information in the text about the project and the construction phases involved?*
- *Is there information in the text about the client?*
- *Is there information in the text about the disciplines included in the project?*

Consider all questions above and then grade the text with a qualitative characterization of being Poor, Medium, Great, where:

- *Poor means that the text includes little information addressing the questions above.*
- *Medium means that the text includes a reasonable amount of information addressing the question above.*
- *Great means that the text includes a very adequate amount of information about the project and answers all of the questions above.*

Only return your score and a one-line piece of advice on how to improve the text to get a higher score (i.e., how to provide more information addressing the question above).

In short, Fig. 1 represents a simplified process flow of the cloud service concept, while Fig. 2 summarizes the functionality of the added GPT layer further empowering the service's infrastructure. These schemas are not to be considered separately, as they jointly constitute the concept of our proposed cloud-based service.

Discussion and conclusions

It has been documented that flaws in the way quality self-inspections are practiced can contribute to the recurrence of quality defects in Swedish projects. Addressing this through improving the self-inspection process itself has scarcely been the focus of relevant research. Therefore, in this paper we address the research question of how we can improve quality self-inspection in Swedish construction projects, by proposing the concept of a cloud quality self-inspection service for clients, designers, and contractors. This service is conceived to leverage generative pre-trained transformer (GPT) AI to access, retrieve, share, and create content (in the form of prompts and checklists) about relevant project knowledge and quality issues, on the basis of the understanding that professionals enacting quality controls using GPT can be provided with a much larger knowledge basis for planning and executing quality inspections.

The GPT AI support is expected to help clients, designers, and contractors to address inquiries and generate content for control plans and self-checklists. In addition, this AI could potentially help in calculating the probability and relevance of control plans and self-inspections, as well as suggesting improvement measures and best practices. The effects of this are that the client, designer and contractor can get a tool where qualitative (aggregated knowledge) and quantitative (the number of answers and suggestions) help is available in a cloud service throughout the project. It is even envisioned that, through further development, the users will, through some clicks in the app, get a proposal for a control plan together with a risk assessment and a list of quality-critical elements for both planning and execution. That foundation would be critical when developing functional requirements and requesting documents, as clients, designers and contractors can enter some brief information about the current project and get suggestions for self-checks based on the most quality-critical work steps.

The expected benefits of the cloud service include the addressing of client needs with regard to knowledge requirements for getting value out of self-inspection reports in an educational and intuitive way. Moreover, large time savings for designers and contractors can be potentially manifested, especially in relation to retrieving lessons learned from previous experiences and adapting the self-inspection template per project for contextualizing the work ahead. These benefits can primarily be achieved by having the GPT access, retrieve, share, and create content about project knowledge, experiences, known quality defects, and best practices. The aggregated knowledge to which the language model can potentially have access is in itself almost priceless.

Nonetheless, this research has several limitations. At its current stage of development, our proposed concept is still a work in progress, comprising an existing cloud service base but not a fully formed GPT layer yet. As such, it does not wholly address ongoing challenges connected to, among others, the GPT learners' limited domain knowledge, the over-reliance on the users' analytical skills to interpret and identify potential shortcomings in the content produced by the learner, the results' accuracy, generalizability, and interpretability, the potentially high cost of dedicated integration, non-updated regulations on GPT implementation, and possibly unethical data use. Moreover, there are yet not fully elaborated design considerations related to the knowledge requirements of the system's simulated prompts, the substance of the system-generated content, and the service's different functionalities and interfaces depending on the role of the user accessing it. Furthermore, measures guaranteeing the integrity and quality of the data and information processed by the GPT have not been considered yet. Using the proposed system demands specialized knowledge and expertise, and as such it must be ensured the GPT operates not merely as a linguistic transformer but as an intelligent, knowledge-based quality assessor. Finally, capabilities of assessing images (e.g., of hand-written reports), and data and/or accessor bias have not yet been considered.

Future work on the system will thus comprise, among others, the robustification of the concept for addressing the aforementioned challenges, the development and integration of the GPT layer within the cloud service base (incl. any needed APIs and user interfaces), and several rounds of testing, validation, and re-development, in order to iron out any weaknesses in the system's functionalities (incl. analyzing documents for quality against a human accessor). It is also critical to ensure the repeatability of the application, by comparing different LLM versions and vendors for the suitability informing the GPT layer of the proposed system. The ambition is that, afterwards, the cloud service will be available to all users regardless of geographic or economic conditions. Such an availability is expected to contribute to a more equal social development, as everyone would get the same opportunity and access to the service.

Acknowledgements

We are grateful for the funding received by Formas, the Research Council for the Environment, Land Industries and Community Development (Forskningsrådet för Miljö, Areella Näringar och Samhällsbyggande), as well as Incoord and Robert Dicksons Stifelse, which are jointly financing the project to which this paper is connected.

References

Abioye, S.O., Oyedele, L.O., Akanbi, L., Ajayi, A., Delgado, J.M.D., Bilal, M., Akinade, O.O., & Ahmed, A. (2021) Artificial intelligence in the construction industry: A review of present status, opportunities and future challenges. Journal of Building Engineering, 44, 103299.

Abdullah, M., Madain, A., & Jararweh, Y. (2022) ChatGPT: Fundamentals, Applications and Social Impacts. In: Proc. 9th SNAMS Int. Conference. DOI: 10.1109/SNAMS58071.2022.10062688.

Aladağ, H. (2023) Assessing the Accuracy of ChatGPT Use for Risk management in construction projects. Sustainability, 15(22), 16071.

Amer, F., Jung, Y., & Golparvar-Fard, M. (2021) Transformer machine learning language model for auto-alignment of long-term and short-term plans in construction. Automation in Construction, 132, 103929.

Apelgren, S., Koch, C., & Richter, A. (2005). Snublesten i byggeriet (Stumbling stones in construction). Byg Rapport No. R-107. Lyngby, Denmark, Danish Technical University.

Bell, E., Bryman, A., & Harley, B. (2019) Business Research Methods (5th ed.). Oxford, UK, Oxford University Press.

Boverket (2018) Kartläggning av fel, brister och skador inom byggsektorn (Mapping of errors, deficiencies and damages in the construction sector). Stockholm, Sweden, Boverket.

Dundar, Y., & Fleeman, N. (2017) Applying inclusion and exclusion criteria. In: Boland, A., Cherry, G., and Dickson, R. (eds) Doing a systematic review: a student's guide, pp.79–922. London, UK, Sage.

Engström, S., and Stehn, L. (2016) Barriers to client-contractor communication: implementing process innovation in a building project in Sweden. International Journal of Project Organisation and Management, 8(2), 151-171.

Ghimire, P., Kim, K., & Acharya, M. (2024) Opportunities and Challenges of Generative AI in Construction Industry: Focusing on Adoption of Text-Based Models. Buildings, 14(1), 220.

ISO 9000 (2015). Quality management systems. Geneva, Switzerland, International Organization for Standardization.

ISO 31000 (2018). Risk management. Geneva, Switzerland, International Organization for Standardization.

Koch, C, and Buser, M. (2020). Good Enough Quality: Multiple Quality Cultures in a Swedish Region. In: Scott, L. & Neilson, C.J. (eds) Proceedings of the 36th Annual ARCOM Conference, pp. 465-474. UK, ARCOM.

Koch, C., & Jonsson, R. (2015). Egenkontroll: En nulagesbeskrivning (Self-inspection: A description of the state-of-art). SBUF rapport ID: 1503. Gothenburg, Sweden, Chalmers University of Technology.

Koch, C., & Schultz, C.S. (2019) The production of defects in construction – an agency dissonance. Construction Management and Economics, 37(9), 499-512.

Koch, C., Shayboun, M., Manès, A., & Nordlund, T. (2020). Produktivitetsläget i svenskt byggande 2018: lokaler, flerbostadshus, grupphus och anläggning (The state of productivity in Swedish construction in 2018: premises, apartment buildings, group houses and facilities). SBUF rapport ID: 13642. Gothenburg, Sweden, Chalmers University of Technology.

Plan- och bygglag (PBL) (2010:900). Landsbygds- och infrastrukturdepartementet SPN BB. Sveriges Riksdag.

Prieto, S.A., Mengiste, E.T., & de Soto, B.G. (2023) Investigating the use of ChatGPT for the scheduling of construction projects. Buildings, 13(4), 857.

Rane, N., Choudhary, S. & Rane, J. (2023) Integrating ChatGPT, Bard, and Leading-edge Generative Artificial Intelligence in Building and Construction Industry: Applications, Framework, Challenges, and Future Scope. SSRN. DOI: 10.2139/ssrn.4645597.

Saka, A., Taiwo, R., Saka, N., Salami, B.A., Ajayi, S., Akande, K., & Kazemi, H. (2024) GPT models in construction industry: Opportunities, limitations, and a use case validation. Development in the Built Environment, 17, 100300.

Tolstoy, N. (2012) Kontrollplaner enligt bygglagstiftningen (Control plans according to building legislation). Bygg & tegenkontrollnik 2/12, 12-14.

Uddin, S.M.J., Albert, A., Ovid, A., & Alsharef, E. (2023) Leveraging ChatGPT to Aid Construction Hazard Recognition and Support Safety Education and Training. Sustainability, 15(9), 7121.

Webster J. & Watson R.T. (2002) Analyzing the Past to Prepare for the Future: Writing a Literature Review. MIS Quarterly, 26(2), pp.xiii–xxiii.

Xian, Y., Lampert, C.H., Schiele, B., & Akata, Z. (2020) Zero-Shot Learning – A Comprehensive Evaluation of the Good, the Bad and the Ugly. arXiv:1707.00600.

You, H., Ye, Y., Zhou, T., Zhu, Q., & Du, J. (2023) Robot-Enabled Construction Assembly with Automated Sequence Planning based on ChatGPT: RoboGPT. arXiv:2304.11018.

Zhang, C., Lu, J., & Zhao, Y. (2024) Generative pre-trained transformers (GPT)-based automated data mining for building energy management: Advantages, limitations and the future. Energy and Built Environment, 5(1), 143-169.

Zheng, J., & Fischer, M. (2023) BIM-GPT: a Prompt-Based Virtual Assistant Framework for BIM Information Retrieval. arXiv:2304.09333.

EXPLORING DEEP GENERATIVE MODELS IN BUILDING DESIGN

Soheila Kookalani[1], Erika Parn[1], Ioannis Brilakis[1],
Ning Pan[1], Mudan Wang[1], Hamidreza Alavi[1]

[1]Department of Engineering, University of Cambridge, Cambridge, UK.

Abstract

The complex and time-consuming nature of building design necessitates meticulous attention to detail and adherence to principles. Automation in the design process is crucial with a growing demand for innovative and sustainable structures. This review explores the transformative impact of integrating deep generative models into building design, showcasing their ability to generate realistic 3D models, layouts, and structural designs. These models address challenges in architectural design and sustainability assessment, enabling generative design and energy-efficient building design. The review suggests potential solutions and highlights the role of the models in enhancing sustainability, cost-efficiency, and creativity in the built environment.

Keywords: Design automation, Deep generative models, Generative adversarial networks, Variational autoencoders, Reinforcement learning

Introduction

The building industry faces a critical challenge in enhancing the efficiency, cost-effectiveness, and sustainability of design processes in the built environment. There is an urgent need for research and development focused on integrating automation into design processes to address this challenge (Bourahla et al., 2022; Rigger et al., 2018). The paramount importance of streamlining design procedures, reducing time and costs, and increasing productivity necessitates a thorough investigation of traditional design methodologies.

The potential for the building industry to significantly decrease design timelines and environmental impact underscores the urgency of addressing this issue. In this context, design automation emerges as a promising solution, utilizing computer capabilities to quickly produce diverse design options that meet complex objectives. Therefore, designers can evolve and refine designs efficiently, minimizing material consumption and enhancing efficiency (Karan and Asadi, 2019; Xiang et al., 2021, 2020a, 2020b).

Recent advancements in artificial intelligence (AI) and intelligent optimization algorithms have led to the development of new generative design (GD) algorithms and systems. Intelligent optimization, with its ability to find optimal solutions that maximize or minimize specific objective functions, is ideal for structural design purposes.

The integration of AI techniques in the building industry has been steadily growing (Kookalani et al., 2022a, 2022b, 2021; Kookalani and Cheng, 2022, 2021a, 2021b), with the adoption of GD methodologies and AI-driven optimization techniques poised to transform how structures are designed, constructed, and operated. Deep generative models (DGMs) significantly impact GD by streamlining processes, reducing errors, and enhancing design possibilities.

This paper aims to extensively examine DGMs and their role in optimizing structures. The study highlights DGM applications in the building design industry, emphasizing their ability for intelligent optimization. Complex algorithms enable architects and engineers to efficiently explore design alternatives, providing opportunities for real-time feedback. This allows designers to make informed decisions and create structures that are not only aesthetically appealing but also functionally efficient and sustainable.

Methodology

This section outlines the methodology employed in conducting the literature review for this paper, which delves into the application of DGMs in the field of building design. Specifically, the focus lies on exploring DGMs, which encompass GANs, VAEs, and RL techniques. A systematic search was conducted within the Web of Science Core Collection database to identify relevant scholarly papers. The search query employed the keywords ("deep generative models") AND ("architecture" OR "building" OR "structural engineering") AND ("generative adversarial network" OR "variational autoencoder" OR "reinforcement learning"). The search was constrained to categories related to architecture and civil engineering, yielding 34 initial papers. Additionally, supplementation from the references of these papers led to the inclusion of 55 relevant papers for review.

Deep Generative Models (DGMs)

GD supported by AI techniques has emerged as a transformative force in reshaping the landscape of architectural and engineering practices. This synthesis of GD principles with the power of AI, particularly DGMs, holds great promise for optimizing structures and steering the construction industry towards a more sustainable future.

DGMs, encompassing models like generative adversarial networks (GANs), variational autoencoders (VAEs), and reinforcement learning (RL), play a pivotal role in the automation of GD and analysis processes. Diagrams of these methods are presented in Figure 1. They utilize the capabilities of AI algorithms to facilitate the exploration of complex design spaces, generating efficient and aesthetically pleasing solutions. Notably, DGMs streamline the analysis and design of structural members, eliminating the need for extensive manual iterations by structural engineers.

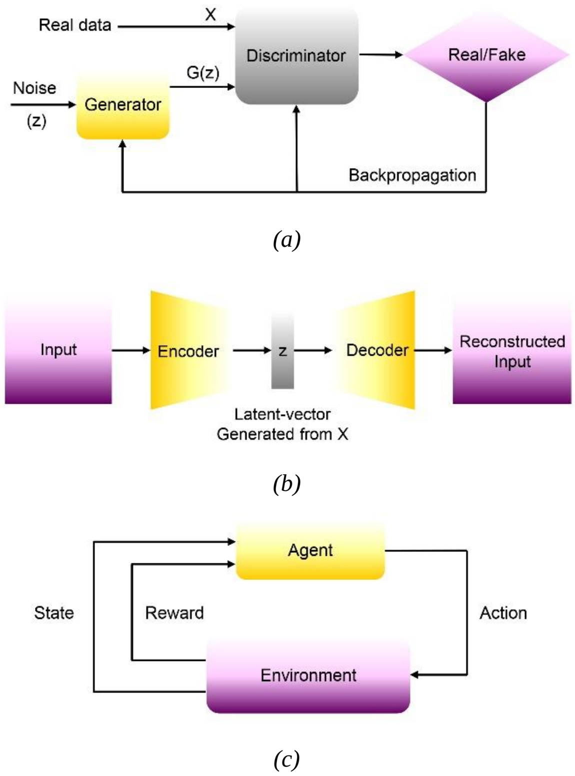


Figure 1. DGMs: (a) GAN; (b) VAE; (c) RL.

The integration of grid search accessible hyperparameters enhances the effectiveness of deep learning (DL) models. This integration enables engineers to anticipate and determine optimal performance levels with greater accuracy and efficiency, contributing to the advancement of automated GD and analysis (Torky and Aburawwash, 2018).

Several studies exemplify the application of DGMs in the field of automated GD and analysis. Yoo et al. (Yoo et al., 2021) propose a DL-driven computer-aided engineering system for conceptual design, demonstrating enhanced efficiency in structural engineering. Steuben et al. (Steuben et al., 2016) focus on AI-driven optimization of geometry partitioning for 3D printing, showcasing the potential for efficient prefabrication. Li et al. (Li et al., 2023) introduce a process for optimizing clash-free rebar design through the combination of graph neural networks (GNN) and exploratory genetic algorithms (GA).

DGMs, such as GANs, VAEs, and RL, empower architects to streamline their workflow and explore innovative design possibilities efficiently. These models find applications in indoor scene synthesis, conceptual designs, floor plan generation, architectural style identification, and architectural drawing recognition. Notable advancements include single-image 3D reconstruction and the utilization of denoising autoencoders for completing 3D shapes. Higharc, a company employing AI, utilizes DGMs for diverse architectural plans, utilizing tools like Finch that integrate machine learning (ML) and DL for conceptual designs and plan generation within the architecture field.

In summary, GANs, VAEs, and RL have improved automated GD and analysis in the building industry. GANs focus on data generation, VAEs on data compression and generation, and RL on training agents for sequential decision-making. These models offer opportunities to enhance structural analysis, optimize geometry and material distribution, and foster innovative architectural designs. The integration of DGMs into building design processes holds significant potential for advancing sustainable and visually appealing structures, aligning with global efforts to address environmental concerns, particularly in the context of smart cities and carbon emission mitigation. The continued evolution of DGMs promises to shape a future where AI-driven GD is a fundamental element of architectural, engineering, and construction practices.

Generative adversarial networks (GANs)

GANs have emerged as a powerful DL algorithm with widespread applications in generative modeling across diverse domains, encompassing image, video, and audio generation. Wu et al. (Wu et al., 2022) conducted a comprehensive review, elucidating the diverse applications of GANs in addressing complex challenges within the built environment. However, they highlighted a shared obstacle within this domain, namely the scarcity of meticulously curated datasets tailored for issues encountered in the built environment.

The GAN architecture fundamentally comprises two pivotal components: the generator and the discriminator. The generator is tasked with crafting new data instances replicating the training dataset, guided by feedback from the discriminator. On the other hand, the discriminator plays a crucial role in distinguishing real data from the training set and data synthesized by the generator. This adversarial interplay contributes to the overall learning process of GANs, with the discriminator offering critical feedback on the quality of the generated output.

In the realm of architecture, GANs have catalyzed a significant transformation, automating the generation and design of diverse architectural elements. This includes facades, interior layouts, building masses, and floor plans.

Notable examples such as ArchiGAN, House-GAN, and House-GAN++ showcase the expertise of GANs in generating fully furnished architectural plans and automated house layouts (Nauata et al., 2021, 2020), Architectural constraints are utilized, allowing architects to transform generated layouts into physical floor plans, exemplifying the practical applications of GANs in architectural design, as shown in Figure 2.

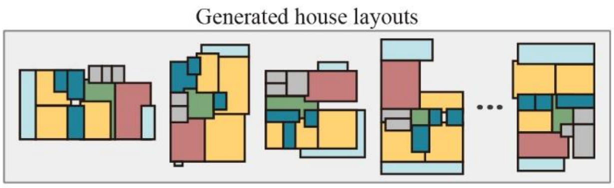


Figure 2. Automatically generating multiple house layout options by House-GAN (Nauata et al., 2020).

Subsequently, Nauata et al. (Nauata et al., 2020) illustrated a comprehensive analysis of failure and success instances, aligned with the ground-truth data, in their user study, as depicted in Figure 3. The first failure example appears unusual as a balcony is only accessible through bathrooms, and a closet is situated within a kitchen, while the second failure example seems odd due to a kitchen being divided into two disconnected spaces.

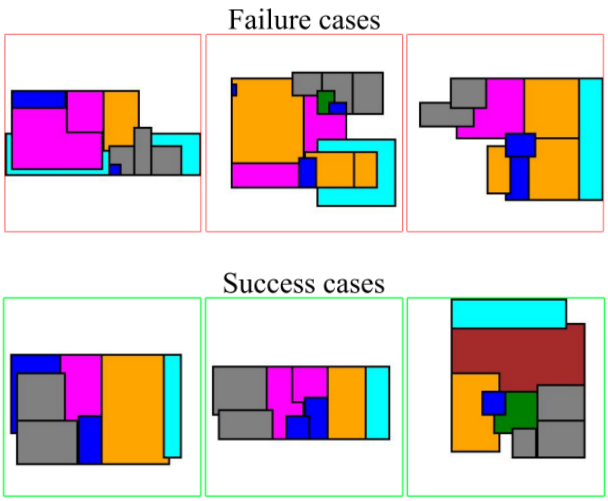


Figure 3. Failure and success examples by House-GAN from the user study (Nauata et al., 2020).

Expanding beyond conventional design realms, GANs have been employed in intelligent shear wall structure design (Zhao et al., 2023) and autonomous architectural sketch design (Qian et al., 2023, 2022). The versatility of GANs is further evidenced by their application in automating the design of shear walls, beams, and slabs in a two-dimensional plane (Liao et al., 2021; Lu et al., 2022; Zhao et al., 2022). Innovative approaches such as TxtImg2Img (Liao et al., 2022), modal expansion of reticulated shells (Zhang et al., 2023), and modular building plan creation (Ghannad et al., 2021) underscore potential of GANs to transform structural design processes, offering practical solutions in the building and construction industry.

In summary, the landscape of GANs in building and structural design is dynamic and expansive. Open-source repositories, including pix2pix, SEASAME, CycleGAN, StyleGAN, AttnGAN, and GraphGAN, further exemplify versatility of GANs in addressing specific challenges across various fields, including construction and building industry. The continued evolution of GANs and their innovative applications underscore their potential to reshape and streamline complex design processes, fostering efficiency and innovation in the built environment.

Variational autoencoders (VAEs)

VAEs have emerged as a prominent class of DGMs, garnering significant attention across various ML domains. Their success, demonstrated since their introduction in 2013, stems from their ability to address limitations faced by traditional autoencoders, particularly in sampling realistic latent vectors due to the sparsity of real data distribution in the latent space.

A fundamental characteristic of VAEs is their integration of regularization techniques to mold a well-organized latent space during training. VAEs introduce a probabilistic sampling approach in the latent space unlike conventional autoencoders, as proposed by Kingma and Welling (Kingma and Welling, 2013). The encoder of a VAE produces mean and variance values, enabling the sampling of latent vectors before decoding. The incorporation of a Kullback–Leibler divergence loss ensures a predictable latent space distribution, addressing the sparsity issue encountered by traditional autoencoders.

In the realm of deep neural networks (DNNs), VAEs have found particular utility in generative architectural designs, specifically in 2D and 3D applications. The regularization techniques employed by VAEs facilitate the creation of well-structured latent spaces, enhancing the generative process. This capability has been exemplified in the work of Wu et al. (Wu et al., 2019), who introduced an encoder-decoder network algorithm for generating floor plans, effectively enforcing room location constraints. Furthermore, Mirra and Pugnale (Mirra and Pugnale, 2021) utilized VAEs to generate topological variations of shell structure forms, highlighting the adaptability of VAEs in diverse design spaces. In addition, Danhaive and Mueller (Danhaive and Mueller, 2021) developed a methodology using conditional variational autoencoders within a performance-guided design exploration framework to construct advanced latent variable models, enabling the representation of intricate design spaces in lower-dimensional subspaces optimized for performance-driven outcomes, as shown in Figure 4. This study demonstrates that design subspace learning offers a promising solution to overcome inherent limitations in current computational design methods, providing a more intuitive and effective interface for human designers to navigate complex design spaces, thus facilitating the broader adoption of performance-informed design processes.

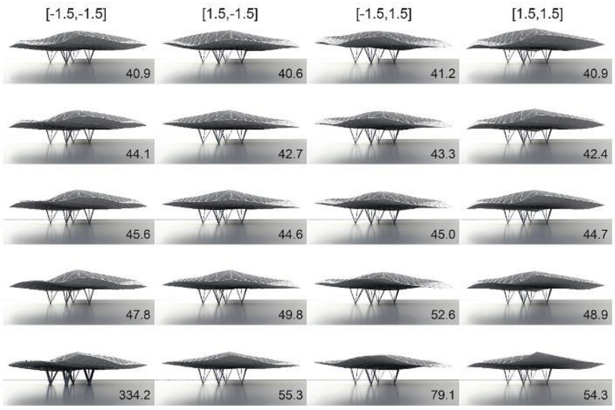


Figure 4. Morphological and performance evolution of designs in the latent space (Danhaive and Mueller, 2021).

In the context of unconditional 3D shape synthesis, VAEs exhibit significant potential (Soltani et al., 2017). Notably, they have been successfully applied to create meshes for modeling faces and human bodies, showcasing the versatility of VAEs in specific domains (Tan et al., 2017)**.** The ability to condition VAE training on design constraints or user preferences parallels the approach taken with GANs. However, VAEs possess an inherent advantage with a latent space that already exhibits some degree of structure. Researchers have developed conditional variational autoencoders (CVAEs) to enhance interpretability and control. CVAEs extend conventional VAEs by incorporating a conditioning vector as input to both the encoder and the decoder, facilitating the integration of design constraints and user preferences in the generative process (Danhaive and Mueller, 2021)**.**

In summary, VAEs have proven to be a powerful tool, offering solutions to challenges faced by traditional autoencoders. Their success in various ML applications, coupled with their adaptability to different design domains, underscores their significance in shaping the future of generative modeling.

Reinforcement learning (RL)

RL distinguishes itself from other DGMs by its unique ability to learn in an unsupervised manner, obviating the need for a pre-existing dataset. Instead, RL operates through trial-and-error interactions between an actor and an environment, with the goal of maximizing rewards accrued through decision-making and action-taking (Kaelbling et al., 1996). The central objective of the actor in RL is to optimize the total reward obtained, casting RL as a form of optimization process.

An early fusion of DL and RL was pioneered by Mnih et al. (Mnih et al., 2015), who integrated DL with Q-learning. Q-learning revolves around learning the state-action value function, known as the Q-function, estimating the potential reward associated with a specific action taken in a given state. The introduction of convolutional neural networks (CNNs) in this paradigm enabled the learning of the Q-function.

In the realm of design, RL unfolds as a sequential process, iteratively modifying or generating designs through a series of actions. The quality or performance of the resulting design serves as the reward signal, effectively acting as the environment for RL. While RL eliminates the need for a conventional dataset, it heavily depends on meaningful and reliable reward signals, often necessitating a high-fidelity simulation environment. A key advantage of RL over GANs and VAEs is its flexibility in defining the reward function based on any objective, which need not necessarily be differentiable. In contrast, GANs and VAEs rely on gradient-based optimization and require any added objectives in their loss functions to be differentiable.

In engineering design, RL has been successfully applied to address inverse design problems, seamlessly combining learning with design optimization. Dworschak et al. (Dworschak et al., 2022) proposed a comprehensive approach for design automation using RL for parametric computer-aided design (CAD) models. Cui et al. (Cui et al., 2012) employed Q-learning for design optimization, while Yonekura and Hattori (Yonekura and Hattori, 2019), and Lee et al. (Lee et al., 2019) implemented double deep Q-learning (DQN). Sun and Ma (Sun and Ma, 2020) extended four well-known exploration techniques in RL to generate multiple solutions. Shi et al. (Shi et al., 2020) introduced an innovative algorithm that combines RL with off-policy monte-carlo tree search for generative floor plan design.

Jeong and Jo (Jeong and Jo, 2021) presented a novel RL-based method for automated reinforced concrete (RC) beam design. They utilized the deep deterministic policy gradient, employing CNN function approximators. They demonstrated that the RL agent can successfully create continuous beam members with design quality comparable to optimized designs, all without requiring an iterative process. Additionally, a clearly delineated collection of reward functions is essential for the success of the suggested approach in expanding its application in structural design automation. This entails exploring a broader range of cost reward functions.

Liu et al. (Liu et al., 2020) proposed a novel approach employing Q-learning-based multi-agent RL for clash-free rebar automated designs in practical applications involving reinforced concrete structures. The outcomes of the simulation regarding success rates and average time required for rebar designs in RC members revealed that the suggested framework has the potential to substantially diminish engineering duration (up to a 90% reduction) and prevent spatial conflicts among rebars. Hayashi and Ohsaki (Hayashi and Ohsaki, 2021) introduced a novel approach combining RL and metaheuristics for automated optimal cross-sectional design of planar frame structures. It was demonstrated that the optimal simulation outcome improved after completing 100 and 1000 training sessions, successfully reducing the cross sections while adhering to the constraints. Maghami and Hosseini (Maghami and Hosseini, 2022) innovatively incorporated deep reinforcement learning (DRL) to facilitate the automated reverse engineering of layered phononic crystal beams, as depicted in Figure 5. It was demonstrated that the algorithm successfully generated appropriate geometry. Consequently, the integration of RL into engineering design research offers promising solutions for addressing complex optimization challenges.

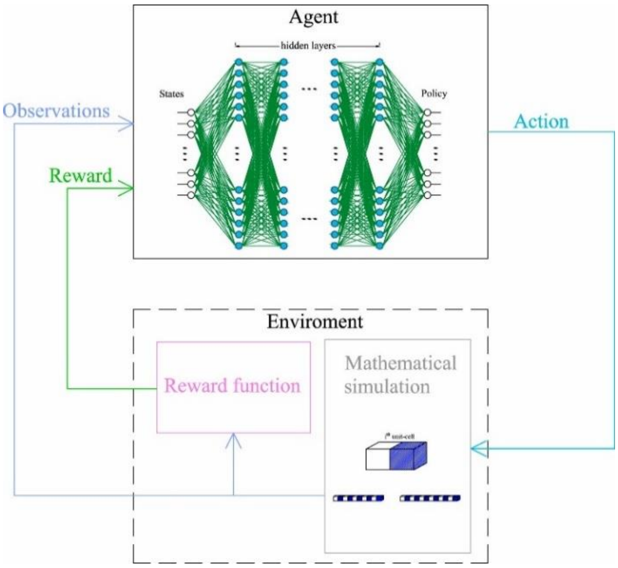


Figure 5. Framework of automated design of layered phononic crystal beams through DRL (Maghami and Hosseini, 2022).

Methods Comparison

GANs, VAEs, and RL are prominent techniques in artificial intelligence, each serving distinct purposes. GANs are primarily geared towards generating realistic synthetic data such as images, audio, and text by learning the data distribution. VAEs, on the other hand, focus on learning latent representations of data, enabling generation, reconstruction, and providing a measure of uncertainty. RL is utilized for sequential decision-making tasks, where an agent interacts with an environment to achieve specific objectives, learning through iterative trial and error processes.

In terms of implementation, GANs consist of two networks: a generator and a discriminator, trained concurrently in a minimax game setup. The generator fabricates samples while the discriminator evaluates their authenticity. VAEs comprise an encoder and a decoder, where the encoder maps input data to a latent space and the decoder reconstructs the input from sampled points. RL involves an agent interacting with an environment to learn a policy maximizing cumulative rewards through exploration and exploitation techniques such as Q-learning and policy gradients.

Applied scenarios highlight the diverse applications of each technique. GANs find utility in generating realistic images for data augmentation, style transfer, and image-to-image translation. VAEs are applied in tasks like image denoising, semi-supervised learning, and generating novel data samples. RL is widely used in robotics, game playing, recommendation systems, autonomous vehicles, design automation, and natural language processing tasks. However, each technique also faces design challenges; GANs may suffer from mode collapse and training instability, VAEs often produce blurry samples, and RL is susceptible to issues like high sample complexity and reward sparsity. Overall, the choice among GANs, VAEs, and RL depends on the specific requirements and constraints of the problem being addressed.

Challenges, Opportunities, and future work

DGMs have emerged as a transformative paradigm in the building and construction fields, offering efficient and user-friendly alternatives to traditional experience-based processes. The exploration of DGMs applications, methodologies, and optimization potential has highlighted a set of challenges and opportunities crucial for its widespread implementation and impact.

The effectiveness of DGMs heavily relies on high-quality and relevant data, posing challenges due to fragmentation, inconsistency, and proprietary nature. Addressing this necessitates efforts in establishing data standards, promoting collaboration, and developing robust data acquisition and preprocessing techniques. Utilizing sensor technology, building information modeling (BIM), and data analytics can enhance data sets, empowering GDM models to generate innovative and optimized solutions tailored to specific project requirements.

Incorporating performance evaluation is crucial for practical DGMs application. Challenges in fidelity, cost, and differentiability need attention. Balancing accuracy and computational cost are essential, and advancements in surrogate models, self-supervised data augmentation, and multi-fidelity modeling show promise in addressing these challenges. Future developments aim to create faster and more accurate evaluation methods.

DGMs involve complex algorithms and optimization processes demanding significant computational resources. Overcoming this complexity requires efficient algorithms, parallel computing techniques, and optimization strategies. High-performance computing architectures, such as tensor processing units and graphics processing units, have accelerated DGM workflows, making tools more accessible. Future advancements, like quantum computing, hold potential for even more sophisticated DGM techniques. Quantum computing holds promise for improving various fields, due to its ability to process vast amounts of data and perform complex calculations at speeds far beyond classical computers. It could potentially enhance DGM techniques by offering more efficient algorithms for data generation. This could lead to the creation of more realistic and higher-quality synthetic data, benefiting applications ranging from data augmentation to simulating complex real-world scenarios.

Current DGMs tend to imitate existing data, impeding the generation of creative or novel designs. Improvements can be achieved by enhancing models like creativeGAN and introducing new features into typical designs. Creative adversarial networks (CAN) can encourage the production of surprising and novel design elements, fostering more human-like creativity in GD.

As DGMs become more complex, the need for interpretability grows. Explainable AI techniques, visualization tools, and interactive interfaces are being developed to explain black-box models. Ensuring interpretability in DGMs enhances collaboration, decision-making, and the adoption of optimized designs.

The potential of DGMs lies in solving complex problems, necessitating collaborations between data scientists, AI experts, and construction professionals. Interdisciplinary collaborations ensure a deep understanding of domain-specific principles and constraints, leading to more sustainable and impactful designs aligned with industry standards and project requirements.

The convergence of AI and design has brought about a paradigm shift in DGMs with the advent of language models. These models offer new possibilities for exploration, diverse design generation, and innovation. However, challenges include designs that may not consider practical constraints, requiring human oversight. Responsible and ethical use of language models is crucial for ensuring designs align with safety, inclusivity, and environmental considerations in DGMs.

Conclusions

This review has delved into the expansive realm of DGMs within the architectural design sector, highlighting their promising yet intricate capabilities. While our exploration has extended the ways in which DGMs can potentially streamline design processes and foster innovation, it is crucial to underscore that the transformative potential of DGMs is derived from both established applications and forward-looking possibilities, rather than conclusive proof across the board.

The discussion on the utility and flexibility of DGMs indeed illuminates a path toward significant advancements in design methodologies, underscored by examples of efficiency gains, creative enhancements, and sustainability impacts. However, the journey from potential to proven impact is paved with challenges including data quality, performance evaluation, computational complexity, creativity limitations, interpretability concerns, and the imperative need for interdisciplinary collaboration have been identified. Recognizing these challenges is crucial for charting the course of future research endeavors. It is imperative that future investigations prioritize advancements in hardware capabilities and utilize the potential of language models to overcome existing barriers.

The industry can unlock the full potential of DGMs by addressing the challenges. The future trajectory of DGMs in the construction sector relies on a delicate balance between creativity, optimization, and human expertise. This synthesis has the potential to initiate a paradigm shift, fundamentally transforming the construction landscape and shaping a built environment that adeptly meets the intricate and evolving needs of society.

Essentially, as we navigate the evolving frontier of DGMs in the construction domain, it becomes evident that their impact extends beyond mere technological advancements. DGMs have the power to contribute meaningfully to the creation of adaptive and user-centric built environments, embodying a symbiosis between cutting-edge technology and human ingenuity. The collective efforts of researchers, practitioners, and stakeholders are instrumental in utilizing this potential, ensuring that the future of DGMs in construction is not only transformative but also socially responsible and sustainable. The path forward requires a clear-eyed focus on evidence-based advancements, collaborative innovation, and the responsible harnessing of capabilities of DGMs to meet the dynamic needs of society and the built environment.

Acknowledgments

The financial support of EPSRC via grant number EP/W018705/1 and Horizon Europe UKRI Underwrite Innovate via grant number G115676 is gratefully acknowledged.

References

Bourahla, N., Tafraout, S., Bourahla, Y., Sereir-El-Hirtsi, A., Skoudarli, A. (2022) GA based design automation and optimization of earthquake resisting CFS structures in a BIM environment. Structures 43, 1334–1341.

Cui, H., Turan, O., Sayer, P. (2012) Learning-based ship design optimization approach. Computer-Aided Design 44, 186–195.

Danhaive, R., Mueller, C.T. (2021) Design subspace learning: Structural design space exploration using performance-conditioned generative modeling. Autom Constr 127, 103664.

Dworschak, F., Dietze, S., Wittmann, M., Schleich, B., Wartzack, S. (2022) Reinforcement Learning for Engineering Design Automation. Advanced engineering informatics 52.

Ghannad, P., Lee, Y.-C., Asce, A.M., Turner, B.S. (2021) Developing an Advanced Automated Modular Housing Design System Using Deep Learning and Building Information Modeling (BIM). Computing in Civil Engineering, 587–595.

Hayashi, K., Ohsaki, M. (2021) Reinforcement learning for optimum design of a plane frame under static loads 37, 1999–2011.

Jeong, J.H., Jo, H. (2021) Deep reinforcement learning for automated design of reinforced concrete structures. Computer-Aided Civil and Infrastructure Engineering 36, 1508–1529.

Kaelbling, L.P., Littman, M.L., Moore, A.W. (1996) Reinforcement Learning: A Survey. Journal of Artificial Intelligence Research 4, 237–285.

Karan, E., Asadi, S. (2019) Intelligent designer: A computational approach to automating design of windows in buildings. Autom Constr 102, 160–169.

Kingma, D.P., Welling, M. (2013) Auto-Encoding Variational Bayes. 2nd International Conference on Learning Representations, ICLR 2014.

Kookalani, S., Cheng, B. (2022) Structural performance prediction of GFRP elastic gridshell structures by artificial neural network, in: 6th International Conference on Applied Researches in Science and Engineering. Aachen, Germany.

Kookalani, S., Cheng, B. (2021a) Structural Analysis of GFRP Elastic Gridshell Structures by Particle Swarm Optimization and Least Square Support Vector Machine Algorithms. Journal of Civil Engineering and Materials Application 5, 139–150.

Kookalani, S., Cheng, B. (2021b) Parametric-insensitive nonparallel support vector regression for structural stress prediction of GFRP elastic gridshell structures, in: International Conference on New Research and Achievements in Science, Engineering and Technologies. Seoul, South Korea.

Kookalani, S., Cheng, B., Torres, J.L.C. (2022a) Structural performance assessment of GFRP elastic gridshells by machine learning interpretability methods. Frontiers of Structural and Civil Engineering 16, 1249–1266.

Kookalani, S., Cheng, B., Xiang, S. (2021) Shape optimization of GFRP elastic gridshells by the weighted Lagrange ε-twin support vector machine and multi-objective particle swarm optimization algorithm considering structural weight. Structures 33, 2066–2084.

Kookalani, S., Nyunn, S., Xiang, S. (2022b) Form-finding of lifting self-forming GFRP elastic gridshells based on machine learning interpretability methods. Structural Engineering and Mechanics 84, 605–618.

Lee, X.Y., Balu, A., Stoecklein, D., Ganapathysubramanian B., Sarkar S. (2019) A case study of deep reinforcement learning for engineering design: Application to microfluidic devices for flow sculpting. Journal of mechanical design 141.

Li, M., Liu, Y., Wong, B.C.L., Gan, V.J.L., Cheng, J.C.P. (2023) Automated structural design optimization of steel reinforcement using graph neural network and exploratory genetic algorithms. Autom Constr 146, 104677.

Liao, W., Huang, Y., Zheng, Z., Lu, X. (2022) Intelligent generative structural design method for shear wall building based on "fused-text-image-to-image" generative adversarial networks. Expert Syst Appl 210, 118530.

Liao, W., Lu, X., Huang, Y., Zheng, Z., Lin, Y. (2021) Automated structural design of shear wall residential buildings using generative adversarial networks. Autom Constr 132, 103931.

Liu, J., Liu, P., Feng, L., Wu, W., Li, D., Chen, Y.F. (2020) Automated clash resolution for reinforcement steel design in concrete frames via Q-learning and Building Information Modeling. Autom Constr 112.

Lu, X., Liao, W., Zhang, Y., Huang, Y. (2022) Intelligent structural design of shear wall residence using physics-enhanced generative adversarial networks. Earthq Eng Struct Dyn 51, 1657–1676.

Maghami, A., Hosseini, M. (2022) Automated design of phononic crystals under thermoelastic wave propagation through deep reinforcement learning. Eng Struct 263.

Mirra, G., Pugnale, A. (2021) Comparison between human-defined and AI-generated design spaces for the optimisation of shell structures. Structures 34, 2950–2961.

Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A.A., Veness, J., Bellemare, M.G., Graves, A., Riedmiller, M., Fidjeland, A.K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S., Hassabis, D. (2015) Human-level control through deep reinforcement learning. Nature 2015 518:7540 518, 529–533.

Nauata, N., Chang, K.H., Cheng, C.Y., Mori, G., Furukawa, Y. (2020) House-GAN: Relational Generative Adversarial Networks for Graph-constrained House Layout Generation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 12346 LNCS, 162–177.

Nauata, N., Hosseini, S., Chang, K.H., Chu, H., Cheng, C.Y., Furukawa, Y. (2021) House-GaN++: Generative adversarial layout refinement network towards intelligent computational agent for professional architects. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 13627–13636.

Qian, W., Xu, Y., Li, H. (2022) A self-sparse generative adversarial network for autonomous early-stage design of architectural sketches. Computer-Aided Civil and Infrastructure Engineering 37, 612–628.

Qian, W., Yang, F., Mei, H., Li, H. (2023) Artificial intelligence-designer for high-rise building sketches with user preferences. Eng Struct 275, 115171.

Rigger, E., Shea, K., Stankovic, T. (2018) Task categorisation for identification of design automation opportunities. Journal of Engineering Design 29, 131–159.

Shi, F., Soman, R.K., Han, J., Whyte, J.K. (2020) Addressing adjacency constraints in rectangular floor plans using Monte-Carlo Tree Search. Autom Constr 115.

Soltani, A.A., Huang, H., Wu, J., Kulkarni, T.D., Tenenbaum, J.B. (2017) Synthesizing 3D shapes via modeling multi-view depth maps and silhouettes with deep generative networks. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017-January, 2511–2519.

Steuben, J.C., Iliopoulos, A.P., Michopoulos, J.G. (2016) Implicit slicing for functionally tailored additive manufacturing. Computer-Aided Design 77, 107–119.

Sun, H., Ma, L. (2020) Generative Design by Using Exploration Approaches of Reinforcement Learning in Density-Based Structural Topology Optimization. Designs 2020, Vol. 4, Page 10 4, 10.

Tan, Q., Gao, L., Lai, Y.K., Xia, S. (2017) Variational Autoencoders for Deforming 3D Mesh Models. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition 5841–5850.

Torky, A.A., Aburawwash, A.A. (2018) A Deep Learning Approach to Automated Structural Engineering of Prestressed Members. International Journal of Structural and Civil Engineering Research 347–352.

Wu, A.N., Stouffs, R., Biljecki, F. (2022) Generative Adversarial Networks in the built environment: A comprehensive review of the application of GANs across data types and scales. Build Environ 223, 109477.

Wu, W., Fu, X.M., Tang, R., Wang, Y., Qi, Y.H., Liu, L. (2019) Data-driven interior plan generation for residential buildings. ACM Transactions on Graphics (TOG) 38, 12.

Xiang, S., Cheng, B., Kookalani, S. (2020a) An analytic solution for form finding of GFRP elastic gridshells during lifting construction. Compos Struct 244.

Xiang, S., Cheng, B., Kookalani, S., Zhao, J. (2021) An analytic approach to predict the shape and internal forces of barrel vault elastic gridshells during lifting construction. Structures 29, 628–637.

Xiang, S., Cheng, B., Zou, L., Kookalani, S. (2020b) An integrated approach of form finding and construction simulation for glass fiber-reinforced polymer elastic gridshells. Structural Design of Tall and Special Buildings 29.

Yonekura, K., Hattori, H. (2019) Framework for design optimization using deep reinforcement learning. Structural and Multidisciplinary Optimization 60, 1709–1713.

Yoo, S., Lee, S., Kim, S., Hwang, K.H., Park, J.H., Kang, N. (2021) Integrating deep learning into CAD/CAE system: generative design and evaluation of 3D conceptual wheel. Structural and Multidisciplinary Optimization 64, 2725–2747.

Zhang, J., Guo, X., Zong, S. (2023) Implicit parametric modal expansion method for single-layer reticulated shells based on generative adversarial network. Structures 54, 1676–1689.

Zhao, P., Liao, W., Huang, Y., Lu, X. (2023) Intelligent design of shear wall layout based on attention-enhanced generative adversarial network. Eng Struct 274, 115170.

Zhao, P., Liao, W., Xue, H., Lu, X. (2022) Intelligent design method for beam and slab of shear wall structure based on deep learning. Journal of building engineering 57.

UNVEILING PROSPECTS THROUGH PLATFORM-DRIVEN TRANSFORMATIONS IN MODULAR CONSTRUCTION SUPPLY CHAINS

Akila Rathnasinghe[1], Niraj Thurairajah [1], and Paul Jones [1]
[1]Faculty of Engineering and Environment, Northumbria University, Newcastle-upon-Tyne, UK

Abstract

Managing intricate supply chains (SCs) in modular construction is challenging, necessitating a nuanced approach to enhance effectiveness. This study aims to explore the potential of platform thinking—a methodology that reimagines traditional, linear SCs into dynamic, interconnected networks, aligning producers' resources with user needs, through an integrative literature review method grounded in supply chain management (SCM) theoretical perspective. Identifying four prospects for platform-induced SC reconfigurations, this study offers novel perspectives on preparing the modular sector for the platform revolution, highlighting the importance of understanding its benefits and challenges.

Introduction

Change is unavoidable, regardless of how well established a business or profession is. A new type of supply chain (SC) business model has been adopted by a wide range of businesses over the last two decades. Without gratifying much about the importance of those, it would be easy to start by thinking of companies like Apple, Uber, Airbnb, and Twitter. This new generation of companies has taken a different approach to reaching scale while meeting a wide range of client demands. For centuries, businesses operated on a pipeline SC design; they would manufacture a product and push it towards a consumer in which value had been produced upstream and consumed downstream, creating a linear value model like a pipe. Choudary (2015), on the other hand, noted that the growing connectedness of entities (as part of the globalisation), decentralised manufacturing, and the embedding of artificial intelligence have driven traditional sectors to adopt an entire new SC model.

Accordingly, many firms in a wide variety of industries are acting as a platform to facilitate producers and consumers of value around the world to connect and interact with each other in a manner that was not possible in pipe design. Traditional taxi businesses, for instance, were historically established with ownership/possession of hundreds of cabs, and passengers were needed to request the cab directly from the firm or the driver without even knowing what the rate would be. Companies such as Uber, on the other hand, are now functioning as a platform to link convertible roles (no specified roles for participants; a driver in one instance might become a passenger in the next) of drivers and passengers toward an economic transaction without even owning a single taxi for the firm itself. As a result, such platform-based firms enable externally positioned producers and consumers to trade value within the SC while also allowing them to be a part of their core business function. This shattered the long-established feature of the firm being the producer of value while enabling multi-dimensional value exchanges among many SC stakeholders, which is the primary deviation from the pipeline model.

The success achieved by various industries due to the transition from pipe to platform-enabled SC models is undeniable. By the end of 1980s, there were less than ten firms who has tried at adopting platform strategy into their SC functions. However, the Forbes Insights survey (2017) found out that 31% of organisations consider themselves top performers in platform thinking, enabling both external and internal SC stakeholders to interact with the organisation and widening the availability of the organisation's digital resources to all stakeholders (McKendrick, 2017). In 2019, it was reported that seven of the world's ten most valuable firms are based on platform-enabled models with less than two decades of existence. As a result, the advantages and potential disruption to established firms from platforms seem essentially inconceivable. Parker *et al.* (2016) rephrased Marc Anderssen's famous slogan "software is eating the world" in 2011 to highlight the massive disruption caused by software-based digitalisation in every aspect of human life as "platforms are eating the world" because software alone is insufficient with current market features: increasing interconnectivity, decentralised manufacturing, and the emergence of artificial intelligence.

As a result, the potential value that can be realised by a platform opportunity in the modular industry is unmistakable. Even though the modular industry already comprises the market attributes, platform disruption has only been examined as a feasible tool for achieving mass customisation (Bertram *et al.,* 2019). However, an investigation into current platform studies in modular construction revealed that the identification of further opportunities arising from unique SC reconfigurations induced by platform applications in modular construction has received minimal attention. Also, as Thurairajah et al. (2023) reveals, even though most production procedures in manufacturing sectors are mature and ready to use, applying these approaches blindfolded without systematically thinking about or adopting them may bring additional issues in any given context. As a result, the purpose of this study is to take the initiative in presenting on how the platform thinking may provide effective SCM within the modular construction. Thereby, this study intends to demonstrate the viability of platform-enabled SC in a modular construction setting in a methodical manner.

What is platform thinking?

Thomas, Autio, and Gann (2014) establish platforms as control and monitoring mechanisms within SCs, concentrated on shared technology, standards, and assets that confer benefits to contributors. Robertson and Ulrich (1998) present a more inclusive definition of platforms as a collection of assets—encompassing components, processes, knowledge, people, and relationships—shared among a range of product SCs. In essence, platforms constitute a unification of components and rules governing SC relationships, where components encompass software, hardware, service modules, and architectural frameworks delineating their interactions (Henderson and Clark, 1990), and rules encompass standards, protocols, and regulations organising the activities of contributors within the ecosystem (Baldwin and Clark, 2000).

The application of platform thinking to restructure a firm's SC into somewhat standardised is not a recent phenomenon. In manufacturing, it is routine to leverage past designs to improve or fix shortfalls in prior activities, resulting in the emergence of platforms as subsystems and interfaces facilitating efficient development and production of product families, such as automobiles or consumer electronics devices (Muffatto and Roveda, 2002). A fundamental objective of such platform-based new product development is to augment product variety, addressing diverse customer requirements, business needs, and technological advancements, while upholding economies of scale and scope within manufacturing processes—a concept inherently linked with "mass customisation" (Pine and Davis, 1993).

Platform thinking in construction

The advent of platform thinking in the construction sector, especially through the strategic use of artificial intelligence, digital innovation, and the standardisation of processes, marks a pivotal transition towards a more synergistic, efficient, and forward-thinking modus operandi. Studies, such as of Oprach et al. (2019), elucidate the way a platform-centric approach, augmented by AI, substantially boosts data interoperability and utilisation across diverse organisational frameworks, thus engendering an ecosystem ripe for innovation and streamlined operations within the industry. Concurrently, Jansson, Johnsson, and Engström (2014), alongside Eriksson and Emilsson (2019), lay emphasis on the quintessential role of standardisation and modular design, coupled with the implementation of process platforms, as instrumental in enhancing product delivery flexibility and efficiency. Such platforms adeptly manage consumer specifications, bolstering competitive product propositions while steadfastly adhering to a platform-centric strategy. Furthermore, Lapidus (2022) broadens this discourse to encapsulate large-scale construction ventures, where organisational and technological platforms, through the leverage of digital technologies and information modelling, optimise construction undertakings. Additionally, the discourse on platforms as pivotal in knowledge management within construction firms, as posited by Styhre and Gluch (2010), illuminates their utility as boundary objects that amalgamate functions and codify procedures—a cornerstone in the efficacious management of knowledge and the propulsion of innovation.

Yet, despite the apparent advantages and the progressive integration of platform thinking within construction paradigms, a discernible lacuna persists in its application to SCM within the domain of modular construction. Characterised by its predilection for prefabrication and assembly-line proficiency, modular construction is ostensibly positioned to reap substantial benefits from platform-centric SCM methodologies. Such an amalgamation promises not only to streamline operational processes but also to foster stakeholder collaboration, thereby paving the way for more sustainable and cost-efficient construction modalities. Yet, the existing literature lacks comprehensive analysis on the effective implementation of platform thinking in SCM for modular construction. This identified gap serves as a critical research avenue for this study, aiming to explore the transformative impact of platform thinking on SCM practices, thereby ushering in a new epoch for modular construction and the construction industry at large.

Research Method

Using an integrative review methodology, this study seeks to investigate the effect of platform thinking on the effectiveness of SCM within the milieu of modular construction. The motive for using an integrative methodology was established on its ability to foster a comprehensive discourse around ongoing challenges by combining and processing existing theories or concepts (Synder, 2019). Integrative literature reviews, whether they focus on well-entrenched or emerging subjects, vary predominantly in their synthesis approach. For recognised topics, this method involves a thorough examination of existing scholarly contributions, facilitating a critical assessment to either recontextualise or expand prevailing theories. On the other hand, for emergent fields, it envisages producing innovative theoretical frameworks or models (Synder, 2019). Following this principle, this study looked for literature, employing terms such as "modular construction," "supply chain management," "platform business models," and "platform thinking," located from authoritative databases like Google Scholar, Scopus, and Web of Science. This search yielded insights into the evolution of the platform concept across various industries and the enhancements it has spurred. This phase was succeeded by an investigation into the potential of platforms to revolutionise SCs across sectors. This led to an analysis focused on the modifications within modular SCs that accompany a successful reconfiguration from product-centric to platform-centric models.

Background

Supply chain management (SCM)

The definition of "supply chain" exhibits a greater consensus compared to the definition of "supply chain management" (SCM). While La Londe and Masters (1994) offer a straightforward perspective, viewing a SC as a "set of firms that pass materials and services forward," Mentzer *et al.* (2001) present a nuanced view, outlining it as a "set of three or more functional bodies directly involved in the upstream and downstream flows of products, services, finances, and/or information from a source to a customer" (p.4). The management of these SCs serves as a mechanism for SC components, aiming to foster an innovative mindset that creates significant value for end customers, as highlighted by Ellram (1990). In line with the reviewed literature, this study takes a stance on SCM by considering the SC as a unified entity to govern the flow of inventory from inception to destination. Additionally, it emphasises placing the sub-units of the entire SC logically for enhanced integration among inter-firm functions and strategic capabilities.

Love *et al.* (2004) articulate the intricacies of supply chain management (SCM) within the construction sector, outlining it as a network involving facilities and activities contributing customer and economic value to various functions, including design development, contract management, service, material procurement, materials manufacture and delivery, and facilities management. However, the reality of SCM in the construction industry reveals a lack of cohesion, adversarial relationships, and a project-centric mindset, hindering the optimisation of value creation (Bankvall et al., 2010). In scrutinising the underlying causes for these challenges, a paramount consideration is delving into the unique attributes of the construction industry and its project environments, which are often perceived as trailing behind other sectors in terms of clear and long-term supply relationships (Thurairajah *et al.,* 2023). Notably, Vrijhoef and Koskela (2000) highlight distinctive features in construction SCM, including converging SCs directing all materials to the construction site, project-specific configurations, and make-to-order SC relationships characterised by limited recurrence. This perspective aligns with Fearne and Fowler (2006), who argue that prevailing structural features in construction SCs are suboptimal from a production standpoint, primarily due to the inherent fragmentation and transience in supply relationships.

Modular construction and its Supply Chain (SC)

Concerning modular construction, Langlois (2002, p. 19) explains modularity as "a highly general set of principles for managing complexity. By disassembling a complex system into discrete pieces—capable of communication solely through standardised interfaces within a standardised architecture—one can obviate what would otherwise be an unwieldy entanglement of systematic interconnections." Within the construction industry, modularity manifests as the production of volumetric systems, along with sizable, prefabricated sections within a supervised manufacturing setting. Volumetric systems comprise three-dimensional modules functioning either independently or multiplicatively, contributing to the building's structural integrity (Thurairajah *et al.,* 2023). These modules, produced in a factory with all finishes (i.e., fixtures and fittings), involve least on-site installation work. Essentially, the modular construction approach entails off-site SCM, diverging from traditional on-site methodologies.

Prior to delving into the SC intricacies of modular construction, a prerequisite lies in delineating the disparities between traditional construction processes and the modular method, considering these as fundamental determinants of its distinctive nature. Notably, the design of modular buildings undergoes detailed scrutiny under the Design for Manufacture, Logistics, and Assembly (DfMA) approach, with a focus on real-world execution encompassing manufacturing, logistics, and assembly (Building and Construction Authority, 2017). Another distinction surfaces in the imperative for factory production, necessitating an early freeze on the detailed design to facilitate module production—a necessity absents in traditional construction (Thurairajah *et al.,* 2023).

Regarding the SC dynamics of modular construction, in contrast to the conventional SC, modular SC can be perceived as an intricate ecosystem encompassing firms, participants, information, and resources engaged in interconnected downstream and upstream processes and activities, culminating in value delivery within modular projects. This delineation encapsulates the fluidity of information and material flow throughout the delivery chain, spanning conceptualisation, tendering, design, procurement, construction, handover, and the operation phase of construction projects (Vrijhoef and Koskela, 2000). Given the apparent distinctions between traditional construction and the modular approach, the configuration of the supply and delivery chain in the latter departs from the former. The modular SC stages are typically concretised as design, procurement, production, logistics (encompassing transportation, buffer, and storage), and on-site installation (Luo *et al*., 2020).

Discovering Prospects in Platform-induced SC Reconfigurations

Most of today's platforms (i.e., iOS, Facebook, Google, etc.) started as successful standalone products, while some began as services or one-sided platforms. The Internet, which enables SC actors from anywhere on the globe to communicate in a matter of seconds, is primarily responsible for the transformation of products into platforms (Abdelkafi et al., 2019). On the other hand, Tiwana (2014) stated that such a transformation would be the safest alternative for product owners to establish a successful platform. Because successful products have an established consumer base, from which product owners are motivated to consider a future option to construct a platform aimed at offering a more inclusive service to consumers as a collaborative effort of multiple parties. In the current world, anytime a product is successful and one side (i.e., consumers) accepts it in droves, each of these products is driven to add a second side to grow into a platform (iPhone added App Store, Facebook added advertising, Firefox added extensions). However, as Hagiu and Altman (2017) pointed out, such a transformation for successful products is not an instantaneous process, and organisations that are not born platforms are frequently left in the dark in search of a strategy to achieve this transformation.

In the domain of modular construction, Thurairajah et al. (2023) highlights a significant evolution in SCs, extending beyond the simple delivery of modular buildings to include elements that add value, such as servitisation—focusing on operational and maintenance aspects—and customisation, which empowers consumers to influence the specifications of modular products. However, the sustainability of these evolved SCs has been questioned. Persson and Lantz (2022) argued that despite the direct advantages derived from servitisation and customisation, these practices may not ensure the long-term sustainability of modular construction firms. The financial difficulties confronted by prominent modular construction companies in the UK, such as ILKE Homes and L&G, credited to the complexities of managing diverse service offerings, increased SC complexities, and the need for dynamic adaptation strategies, reinforce these concerns, causing unease among industry professionals regarding future SCM innovations.

Addressing these challenges, this study introduces four strategic opportunities, or "prospects," for modular construction companies to adapt to and leverage in the wake of SC transformations brought about by platforms. These prospects are identified through an adaptation of Tiwana's (2014) "four lenses framework," originally developed for software development, now repurposed for the construction sector. This framework uses four distinct perspectives to dissect and comprehend the extensive production processes in the context of platform thinking. The identified "prospects" offer modular construction firms a guide for strategic adaptation and enhancement in areas such as consumer interaction, operational efficiency, service diversification, and ecosystem collaboration, as depicted in Figure 01.

Prospect 1: Horizontal supply chain collaborations.

Barrat (2004) outlines horizontal SC collaboration as the collaborative inter-firm dynamics existing among entities operating at similar SC tiers (i.e., this could be component manufacturers considering the modular context). Barrat (2004) also suggests that these collaborations hold the potential for enhanced advantages, encompassing shared risks, mutual acknowledgment, and interdependence directed at shared or industry-wide goals which could not have been achieved by a single industry player.

Accordingly, a platform requires to have at least two sides (Eisenmann *et al.*, 2006), resulting the significant SC reconfiguration within any production process. Schmalensee and Evans (2007, P.38) emphasised that the unifying factor of all platforms is that they facilitate interactions between at least two or more "distinct" groups who want to engage with and require each other but cannot initiate such interaction by the current means of. As a result, Tiwana (2014) determined that to be a true platform, two distinct parties must be offered, with the platform actively engaged as the medium for the interaction (i.e., based on that claim, ERP is a one-sided platform in which the platform owner-software developer serves the needs of the enterprise/company without interfering with or governing communications between the company and its third-party suppliers/professionals). As a result, in any product environment, if there is any chance that two or more parties might enhance their values by having someone support them in finding each other and engaging cost-effectively, then there is an opportunity for a platform to be established. By saying so, it does not imply that the interaction would be impossible to occur in the absence of a platform; rather, the existence of a platform has made such interaction simpler and less expensive by linking in ways that those parties could not directly connect (Evans, Hagiu and Schmalensee, 2006). Platforms could serve as a conduit for facilitating value-infused interactions between two or more parties.

Prospect 2: Demand-responsive supply chains.

The imperative of SC Responsiveness (SCR) has escalated, presenting paths for competition in business landscape. Formerly a conceptual abstraction, SCR has become into a strategic tool, necessitating the evolution of SCs towards heightened flexibility and responsiveness (James-Moore,1996). Accordingly, SCR is seen as the capacity of the SC to adeptly address variations (i.e., niches) in general consumer demand (Holweg, 2005). It also captures the coordination of production, inventory, location, and transportation within the business firms of SC, strategically optimising responsiveness and efficiency tailored to the relevant market dynamics.

In light of this, consumers (markets) for any product SC can be identified under two categories: core and niche markets. However, satisfying the needs of the niche group of consumers (also known as long-tails in the product market) is fairly challenging within the economics of scale in manufacturing (including modular construction) since such consumers' needs are somewhat distinct from the basic/ every-day consumers' demands (which majority of the product market) (Rogers, 2004). As a result, product owners might consider a platform solution to handle the market's long tail through smaller competitors in the industry. Such a choice by product owners to form a platform may be an alluring option to: digitise and modularise (decentralise) the production activities that are currently provided in-house; provide additional goods or services to enhance the value of the core product; bring specialised collaborators who are excelled at performing a certain activity rather than manufacturers themselves; enable smaller firms to improve their work based on producers' work without replicating it (Gawer and Cusumano, 2008).

Prospect 3: Vertical supply chain collaborations.

In contrast to the previously examined horizontal collaboration, vertical SC collaboration delves into inter-firm relationships at various levels within the SC (Barrat, 2004). This entails exploring the dynamic between suppliers and manufacturers, specifically in a modular context this could be manifested as the collaboration between component manufacturers and mega manufacturers.

coordination among business entities, industries must methodically define problems, explore available options, and forge agreeable solutions by fostering collaboration across diverse functions within the industry (i.e., in the context of the modular industry, these functions span design, planning, procurement, manufacturing, logistics, and operation & maintenance).

When a platform facilitates value-added interactions among and between two or more distinct groups, cross-functional collaboration impact (i.e., in platform literature this has been termed as network effects) define the level of value impact that every new contributor (from one group) may offer to the rest of the existing platform contributors (Tiwana, 2014).

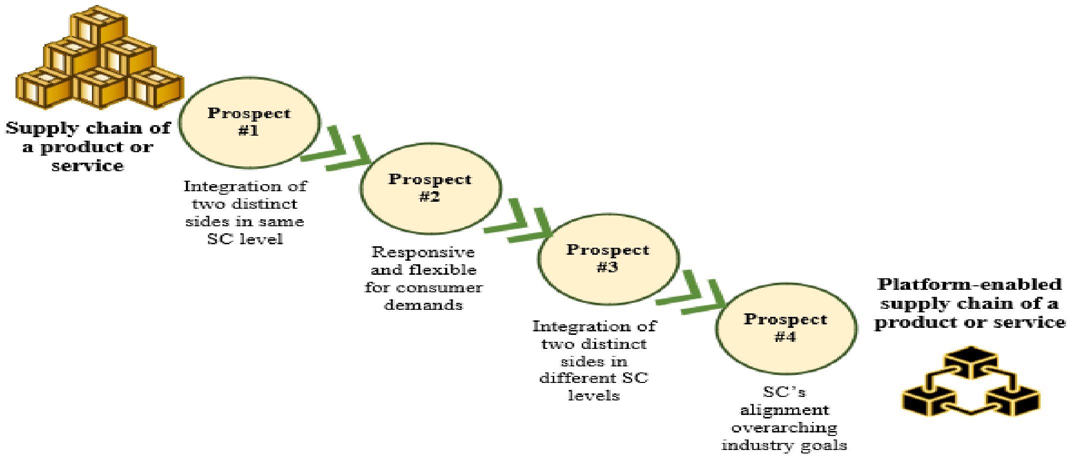


Figure 1: Four prospects due to platform-induced supply chain reconfigurations

Also, as discussed in the first prospect, a real platform should at least include two distinct parties. However, the initial inclusion of both parties by a platform owner proves unfeasible. Hence, Tiwana (2004) emphasised the importance of presenting at least one party from two distinct groupings to bring a platform to its true essence. Platform development, on the other hand, may face a "chicken-egg dilemma," in which interested parties may find a platform unappealing to join in the absence of a strong presence on the opposing side (Calliaud and Jullien, 2003). Consequently, Tiwana (2014) elucidates that if a prominent market player aligns with the platform solution, it induces other entities to join, driven by perceived prestige and the ensuing competitive advantage in their respective business domains. Once a party agrees to be corporative in the platform creation, the platform solution should include a compelling value proposition with the already recruited party to bring the other side onboard as well.

Prospect 4: Cross-functional collaboration impact.

Cross-functional collaboration within SCs facilitates a comprehensive evaluation of SC's state in alignment with industry needs. This evaluation serves as a foundation for crafting and perpetuating value through a collaborative assessment (Oliva & Watson, 2009). Beyond mere

This level of collaboration impact (network effect) might be perceived by contributors in the same group (same-side network effects) or by contributors in a different group (cross-side network effects). Tiwana (2014) emphasised the importance of a clearly established second side to produce cross-side effects rather than same-side effects because such effects would assist the platform in generating downstream sub-products (rather than just increasing the consumer base or supply rate by having more contributors on the same side) that the core market group could use to enhance their experience with the main product or service.

As a result, if any production process is transformed into a platform, the platform owner should be able to recognise that there is more than one "other side" that could inflict the most obvious cross-side network effects. Therefore, Tiwana (2014) asserted that every production or manufacturing scenario that can be observed via at least two of the impacts depicted in the four lenses (refer to Figure 01) that generate due the product's transformation into a platform.

Expected prospects due to platform-induced modular construction SC reconfigurations

Modularity in construction fundamentally transforms complex buildings into manageable, independently produced components that together constitute a unified structure. This approach, deeply embedded in the principles of mass production within controlled settings, has traditionally been recognised as a collective endeavour among various manufacturers (Langlois, 2002). Yet, recent challenges faced by modular factories, especially those related to scaling component production, necessitate a critical reassessment of these collaborative frameworks (Ethiraj & Levinthal, 2004). Such challenges not only question the efficacy of conventional collaboration in component manufacturing but also spotlight the limitations in applying modular principles in practice.

This situation highlights the necessity for a thorough reassessment of the manufacturing strategies within the modular SC, suggesting the need for a unified approach across various SC entities. However, we would like to highlight that the purpose such evaluation is not aiming to detract from the value of current alliances within the modular domain but rather to improve the collaborative efficacy of these partnerships (Brusoni & Prencipe, 2001). Within this context, the adoption of platform thinking is pivotal, promoting an integrated ecosystem of stakeholders in the modular construction SC. This approach enables direct access to a broader spectrum of component manufacturers, overcoming conventional geographic limitations (Parker, Van Alstyne, & Choudary, 2016). Such strategic unification not only streamlines procurement processes but also ensures the participation of the most appropriate contributors for each initiative.

Moreover, the modular construction industry's inherent adaptability, particularly through the use of prefabricated components, positions it uniquely to address diverse customer requirements (Baldwin & Woodard, 2009). However, the pursuit of extensive customisation within this model presents challenges, notably in reconciling the advantages of standardisation with the intricacies of individualisation. The pursuit of excessive customisation may dilute the efficiency and scalability that are hallmarks of modular construction, risking resource allocation away from broader market needs and impacting production timelines and efficiency (Thurairajah et al., 2023). Embedding modular offerings within platform-based frameworks offers a systematic method to maintain both the autonomy of individual component production and the integrity of the assembled product. This approach aligns with Tiwana's (2014) perspectives on innovation and is exemplified by Dell's successful mass customisation model, which assembles personalised products from a spectrum of supplier components (Berman, 2002).

Despite apprehensions that platforms might enable dominant entities to overshadow smaller firms, a more discerning view recognises the ambivalent nature of platforms. While platforms can introduce imbalances and governance complexities requiring vigilant regulation to ensure fair participation (Tiwana, 2014), they also present strategic benefits for all stakeholders. Platforms enable small and medium enterprises (SMEs) to engage with larger markets, gaining visibility and opportunities for growth by addressing niche demands. Apple Inc.'s platform strategy exemplifies how integration with a global network of developers can augment service offerings while supporting smaller contributors (Stratton, 2020).

Further, platforms centric SCs equipped with modern digital technologies and artificial intelligence substantially optimise design and production workflows, ensuring that supply dynamically matches demand. This enhancement not only streamlines production and inventory control but also equips the modular construction industry to swiftly adapt to changes in market dynamics and consumer preferences. Also, the analytics derived from interactions on these platforms provide invaluable insights into market trends and operational performance, enabling the refinement of products, reduction of waste, and improvement of customer satisfaction. This shift towards an analytics-driven, consumer-oriented approach signifies a notable advancement in the strategic market positioning of the modular construction sector (Kagermann, Helbig, Hellinger, & Wahlster, 2013).

Overall, the advent of platform thinking within the modular construction SC represents a significant departure from traditional methods, ushering in an era defined by enhanced integration, flexibility, and creative innovation. These platforms do more than just improve collaborative efforts and customisation capabilities; they transform the future of the industry, guiding it towards a path of increased agility, efficiency, and a stronger focus on customer needs.

Conclusion and Way Forward

The construction industry including modular, known for its historical resistance to change, finds itself on the brink of the impending platform revolution, an unstoppable force that has already reshaped various manufacturing and production industries. While acknowledging the industry's apprehensions and reluctance, it is also imperative for construction professionals to embrace this transformative wave. Accordingly, this study serves as a forewarning for construction professionals, encouraging them to embrace the imminent platform revolution. However, it is also crucial to dispel the notion that platform solution serves as a one-size-fits-all solution for the modular industry's challenges, a tendency often observed in past innovation breakthroughs. Instead, this study critically explores how platforms may reconfigure modular SCs, unveiling opportunities for effective horizontal and vertical integrations, demand responsiveness, and the ability to generate more value impacts through cross-functional collaborations. The envisioned platform transformation in the modular SC is nuanced, neither a purely positive nor negative scenario. While it promises an efficient, collaborative, and responsive SC, it concurrently poses threats of consolidating power among major industry players. While the primary objective of implementing a platform solution is to facilitate an effective industry-wide collaborative SC, the existing competitive dynamics within the market structure pose a substantial impediment to realising the identified prospects. Consequently, advocating for a transparent and equitable platform governance model is imperative to pre-empt potential adversities and ensure the success of the initiative. This governance model is anticipated to prevent exploitation of medium and small-scale market players by larger entities, fostering a mutually beneficial, win-win scenario. Moreover, acknowledging the perceived hindrance posed by data sharing, particularly within a competitive environment involving market competitors, necessitates the establishment of predetermined governance mechanisms. These mechanisms are instrumental in delineating how data is shared, with whom, and to what extent, thereby fostering a consensus among platform participants while addressing concerns related to the exposure of intellectual properties within the industry. Therefore, it is essential for the modular sector to prepare for the platform revolution, not only by understanding its potential benefits but also by being cognisant of the challenges it may pose. Prepared with this knowledge, construction professionals can realise preventive mechanisms, drawing insights from other sectors that have already navigated the complexities of the platform revolution. The call to action envisioned in this study is clear: engage, adapt, and lead the way forward for modular industry in this platform-induced transformative era.

Acknowledgements

The authors would like to acknowledge the fundings received from the Research Development Fund at Northumbria University under the grant number: RDF20/EE/ABE/THURAIRAJAHNiraj.

References

Abdelkafi, N., Raasch, C., Roth, A. *et al.* (2019). Multi-sided platforms. *Electron Markets,* 29**,** 553–559. Doi: 10.1007/s12525-019-00385-4

Baldwin, C. Y., & Clark, K. B. (2000). *Design Rules: The Power of Modularity Volume 1*. Cambridge, MA, USA: MIT Press.

Baldwin, C. Y., & Woodard, C. J. (2009). The architecture of platforms: A unified view. In Platforms, markets and innovation (pp. 19-44). Edward Elgar Publishing.

Bankvall, L., Bygballe, L. E., Dubois, A., & Jahre, M. (2010). Interdependence in supply chains and projects in construction. *Supply Chain Management*, *15*(5), 385-393.

Barrat, M. (2004). Understanding the meaning of collaboration in the supply chain, *Supply Chain Management*, *9*(1), 30-42.

Berman, B. (2022). Should your firm adopt a mass customisation strategy? Business Horizons, 45(4), pp. 51-60.

Bertram, C. A., Müller, G. O., Løkkegaard, M., Mortensen, N. H., & Hvam, L. (2019). Consolidated Challenges Regarding Execution of Portfolio Rationalization Projects in an Engineer-to-Order Context (Version 1). Technical University of Denmark.

Brusoni, S., & Prencipe, A. (2001). Unpacking the black box of modularity: technologies, products and organizations. Industrial and Corporate Change, 10(1), 179–205.

Building and Construction Authority. (2017), Overview of Design for Manufacturing and Assembly (DFMA), Singapore.

Calliaud, J., & Jullien, B. (2003). Chicken & Egg: Competition Among Intermediation Service Providers. The RAND Journal of Economics, 34(2), 309-328.

Choudary, S. (2015). *Platform Scale: How a new breed of startups is building large empires with minimum investment.* Platform Thinking Labs Pte. Ltd.

Eisenmann, T. R., Parker, G., & Van Alstyne, M. W. (2006). Strategies for two-sided markets. *Harvard Business Review*, 84(10): 92-101.

Eriksson, H., & Emilsson, E. (2019). Platforms for Enabling Flexibility at Two Construction Companies. *Proceedings of the 36th International Symposium on Automation and Robotics in Construction (ISARC)*.

Ellram, L. (1990). The Supplier Selection Decision in Strategic Partnerships. *Journal of Purchasing And Materials Management, 26*(4), pp.8-14.

Ethiraj, S. K., & Levinthal, D. (2004). Modularity and Innovation in Complex Systems. Management Science, 50(2), 159–173.

Evans, D. S., Hagiu, A., and Schmalensee, R. (2006). *Invisible Engines: How Software Platforms Drive Innovation and Transform Industries.* Cambridge, MA, USA: The MIT Press.

Fearne, A. & Fowler, N. (2006). Efficiency Versus Effectiveness in Construction Supply Chains - The Dangers of Lean Thinking in Isolation. *Supply Chain Management, 11* (4). pp. 283-287.

Gawer, A., and Cusumano, M. A. (2008). How companies become platform leaders. *MIT Sloan Management Review*, 49(2), pp.28.

Hagiu, A., & Altman, E. J. (2017). Finding the Platform in Your Product: Four Strategies That Can Reveal Hidden Value. *Harvard Business Review*.

Henderson, R. M., & Clark, K. B. (1990). Architectural innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, *35*(1): 9-30.

Holweg, M. (2005). An Investigation into Supplier Responsiveness, *International Journal of Logistics Management, 16*(1), 96-119.

James-Moore, S. M. R. (1996). Agility is Easy; But Effective Agile Manufacturing is not, *IEE Colloquium (Digest), 179*, 1- 4.

Jansson, G., Johnsson, H., & Engström, D. (2014). Platform use in systems building. *Construction Management and Economics, 32*(1-2), 70-82

Kagermann, H., Helbig, J., Hellinger, A., & Wahlster, W. (2013). Recommendations for implementing the strategic initiative INDUSTRIE 4.0. Final report of the Industrie 4.0 Working Group.

La Londe, B., & Masters, J. (1994). Emerging Logistics Strategies. *International Journal of Physical Distribution & Logistics Management, 24*(7), pp.35-47.

Langlois, R. N. (2002). Kirznerian Entrepreneurship and the Nature of the Firm. *Journal des Economistes et des Etudes Humaines, 12*, 1-8.

Lapidus, A. (2022). Organizational and technological platform of construction. *Vestnik MGSU*.

Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: a network model of category learning. *Psychological review*, *111*(2), 309–332.

Luo, H., Liu, J., Li, C., Chen, K., & Zhang, M. (2020). Ultra-rapid delivery of specialty field hospitals to combat COVID-19: Lessons learned from the Leishenshan Hospital project in Wuhan, *Automation in Construction, 119*, Article 103345

McKendrick, J. (2017). Once They Were Companies, Now They Are Platform Businesses. Forbes Insights Survey.

Mentzer, J., De Witt, W., Keebler, J., Min, S., Nix, N., Smith, C., & Zacharia, Z. (2001). Defining Supply Chain Management. *Journal Of Business Logistics, 22*(2).

Muffatto, M., & Roveda, M. (2002). Product architecture and platforms: a conceptual framework. *Int. J. Technol. Manag., 24*, 1-16.

Oliva, R. & Watson, N. (2009). Cross-Functional Alignment in Supply Chain Planning: A Case Study of Sales and Operations Planning, *Working Paper, 07-001*, Harvard Business School, USA.

Oprach, S., Bolduan, T., Steuer, D., Vössing, M., & Haghsheno, S. (2019). Building the Future of the Construction Industry through Artificial Intelligence and Platform Thinking. *Digitale Welt, 3*, 40-44.

Parker, G. G., Van Alstyne, M. W., & Choudary, S. P. (2016). *Platform Revolution: How Networked Markets are Transforming the Economy--and How to Make Them Work for You*. W. W. Norton & Company.

Persson, M. & Lantz, B. (2022). Effects of customization and product modularization on financial performance, *Journal of engineering and technology management, 65*, 101704

Pine, B. J., and Davis, S. (1999). *Mass customization: The new frontier in business competition*. Cambridge, MA, USA: Harvard Business School Press.

Robertson, D., and Ulrich, K. (1998). Planning for product platforms. *MIT Sloan Management Review*, 39, pp.19-32.

Rogers, B.J. (2004), *Market Based Management: Strategies for Growing Customer Value and Profitability*, 3rd ed. Prentice Hall.

Schmalensee, R. & Evans, D. (2007). Industrial Organization of Markets with Two-Sided Platforms. *Competition Policy International, 3*(1),

Stratton, C. (2020). Platfrom politics: software as strategy in Apple's platform ecosystem. *First Monday, 25*(2),

Styhre, A., & Gluch, P. (2010). Managing knowledge in platforms: boundary objects and stocks and flows of knowledge. *Construction Management and Economics, 28*(6), 589-599.

Snyder, H. (2019). Literature Review as a Research Methodology: An Overview and Guidelines. Journal of Business Research, 104, 333-339.

Thurairajah, N., Rathnasinghe, A. P., Ali, M., & Shashwat, S. (2023). Unexpected Challenges in the Modular Construction Implementation: Are UK Contractors Ready? *Sustainability*, *15*(10), [8105].

Thomas, L. D. W., Autio, E., and Gann, D. M. (2014). Architectural leverage: Putting platforms in context. *Academy of Management Perspectives*: *28*(2) 1-36.

Tiwana, A. (2014). *Platform ecosystems: Aligning architecture, governance, and strategy*. Waltham, Mass.: Morgan Kaufmann.

Vrijhoef, R., & Koskela, L. (2000). The four roles of supply chain management in construction. *European Journal of Purchasing and Supply Management*, *6*(3-4), 169-17

ASSESSING THE ENVIRONMENTAL IMPACT OF THE CONSTRUCTION PHASE – A SIMULATION-BASED METHODOLOGY

Danh Toan Nguyen[1] and Walter Sharmak[2]

[1] University of Kassel, Kassel, Germany
[2] Technische Hochschule Mittelhessen - University of Applied Sciences, Giessen, Germany

Abstract

The construction industry makes a significant contribution to global resource and energy consumption and the associated CO2 emissions. The environmental impacts in the execution phase compared to the planning and operating phase have not yet been comprehensively investigated in research for several reasons. This paper presents a methodology for the preliminary calculation of greenhouse gas emissions on the construction site by simulating the nonlinear behavior of the complex and dynamic construction processes. The estimated results will help to optimize the execution planning of the construction processes from an environmental perspective.

Introduction

According to a report by the International Energy Agency, the construction and operating of buildings accounted for the largest share of both global final energy consumption (36%) and energy-related greenhouse gas (GHG) emissions (39%).

However, the Intergovernmental Panel on Climate Change (IPCC) predicted that the construction sector has the greatest potential for significant reductions in greenhouse gas emissions compared to other major manufacturing sectors (IPCC, 2014). Particularly during the operating phase, the building energy requirement will decrease due to the increased requirements for new buildings and renovations of old buildings. Due to this development, the embodied energy will become more important as its share of total energy has increased significantly, based on the calculated life cycle of 50 years. For this reason, studies into sustainable materials like clay are currently gaining great importance in research and development (Tourtelot et al., 2023; Muntari and Windapo, 2020).

Compared with the operating phase, the construction phase is a short-term phase and is believed to have small environmental impacts, so it has not attracted much research (Slossharek et al., 2021). However, to exploit the full potential for reducing greenhouse gas emissions in the construction industry, the execution phase should also be optimized in this regard.

Nonetheless, the complex and dynamic parameters and their relationships on the construction site make it difficult to quantify the environmental impacts in the execution phase. In current practice, only qualitative analyses are carried out to demonstrate the sustainability of the construction site.

The use of simulations can provide a cost-effective and efficient approach to analyzing construction processes and evaluating the resource consumption of different construction methods.

Several researchers have integrated simulation methods such as discrete event simulation (DES), system dynamics (SD) and agent-based simulation (ABS) for various purposes in construction (Xie and Peng, 2012; Xia and Sun, 2013; de Assis et al., 2021). However, such a simulation-based approach is rarely used in the environmental impact assessment of the construction phase.

In this paper, a hybrid simulation model combining ABS and SD is proposed to analyze the environmental performance during the construction phase.

Autonomous agents represent construction processes, and their needed resources can interact with each other according to predefined operating rules. SD provides causal effect loops to simulate the nonlinear behavior of dynamic construction processes. A simple example is presented as well to test the developed simulation model.

Work methodology

The simulation of the processes in the construction phase requires a variety of inputs from different areas for realistic results, such as data on the work schedule, quantities, the planned construction methods and the construction site facilities.

From an IT perspective, the different data structures of the required data represent a major challenge for the simulation. Another challenge is that the construction processes are nonlinear and therefore the traditional discrete event simulation may not lead to reasonable results.

For this purpose, hybrid modelling using multiple interacting agents and system dynamics is used for the simulation in the presented approach of this paper.

Agents are defined in ABS as intelligent autonomous objects that represent real parties without global control and unified goal (Ren and Anumba, 2003).

Autonomously acting agents interact directly with their predecessors and successors. According to predefined operating mechanisms, process agents register their required resources in a central board system that is used for all agent negotiations. Depending on the availability

of resource agents, the control centre processes all information in a specific protocol and initiates resource allocation accordingly. Both construction processes and needed resources are represented by agents.

System dynamics simulation is a top-down approach based on information feedback to analyse complex system behaviour between project elements within a predefined boundary (Ding et al., 2016; Ding et al., 2018). Typically, system dynamics models are structured by stock-flow diagrams, which describe the movement of entities from start to finish in a model, and by causal loop diagrams, which capture the chain influences of a cause that can be traced back to the original cause and a set of related variables.

The concept of causal loop diagrams is explained in the simple well-known example related to the human population (Figure 1).

If the human population increases (larger birth rate), this leads in loop 1 (left) to the population continuing to grow in the future. Such a loop is called a Reinforcement Loop (R). Since people need some time after birth before they are able to have children of their own, the loop has a delay, which is represented by two pipes in figure 1. This delay is not deterministic (e.g. 18 years), but depends on many cultural, economic, and other factors. For this reason, this delay should be represented using a suitable probability distribution. Deaths cause the population to decrease. This is shown in the loop on the right (Balancing Loop B). Loop B also has a delay because most people do not die immediately after birth. This delay also has to do with many factors, such as available medical care, exercise or eating habits and cannot be used as a deterministic value (e.g. 75 years), but in this case too a suitable probability distribution should be used.

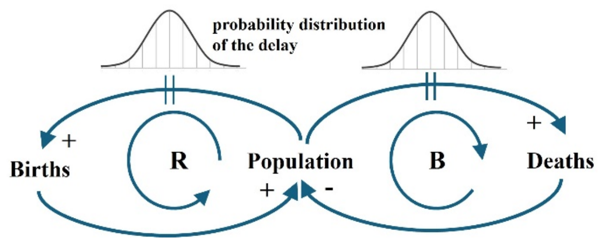


Figure 1: a general example of Causal Loop Diagrams

Another well-known example is related to using work overtime to accelerate the schedule (Figure 2).

Two feedback loops can help decision-makers become aware of both the positive and negative impacts of this management policy. In Figure 2, loop 1 is a reinforcement loop. Indeed, starting with an assumption that Overtime_hours grows and going around the loop counterclockwise (according to the direction of the links going from Overtime_hours through Fatigue, Completed_Work, Remaining_Work and come back to Overtime_hours), we get growth of Fatigue, but at the same time this leads to decrease of Completed_Work, increase of Remaining_Work and again get growth of Overtime_hours, so the same result as the initial assumption. Alternative check by counting the number of negative links: This loop has two negative links, so it is a reinforcing loop.

This reinforcing loop shows a negative impact of the overtime policy. On the other hand, loop 2 is a balancing loop, which indicates the positive effect of working overtime. The adoption of overtime means an increase in Overtime_hours. Let's go around the loop counterclockwise. The increase of Overtime_hours leads to an increase of Completed_Work, and that by-turn leads to a decrease of Remaining_Work. Consequently, the decreased Remaining_Work will result in the reduction of Overtime_hours, so the result contradicts the initial assumption. Alternative check by counting the number of negative links: This loop has one negative link, so it is a balancing loop.

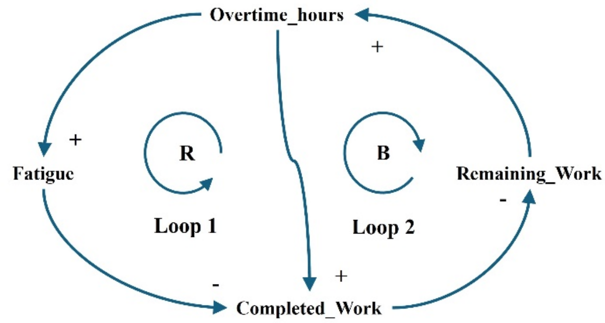


Figure 2: a general example of Causal Loop Diagrams using the Feedback processes related to overtime.

In the presented approach, SD including the schedule pressure loop and the rework loop, is embedded in the processes to capture the dynamic behaviour.

Process patterns are used to facilitate the use of the construction processes in the simulation. A relational database is developed to manage the data of the process patterns. Process patterns are generally used to describe repetitive processes in various construction projects (Dori, 2016; Nguyen, 2024). The process patterns can therefore be used when simulating various projects.

The process pattern structure includes the following elements (Figure 3):

- Description: depicts the process pattern and its application.
- Metadata: contains information about when the process pattern was created, who the author is and some tags to help find patterns.
- Object(s), one or more building object types are created by completing all steps as described in the process pattern, such like walls, columns etc.
- Operations: represent required steps of the process at the second level of granularity.
- Activities: represent required steps of the operation at the third level of granularity.

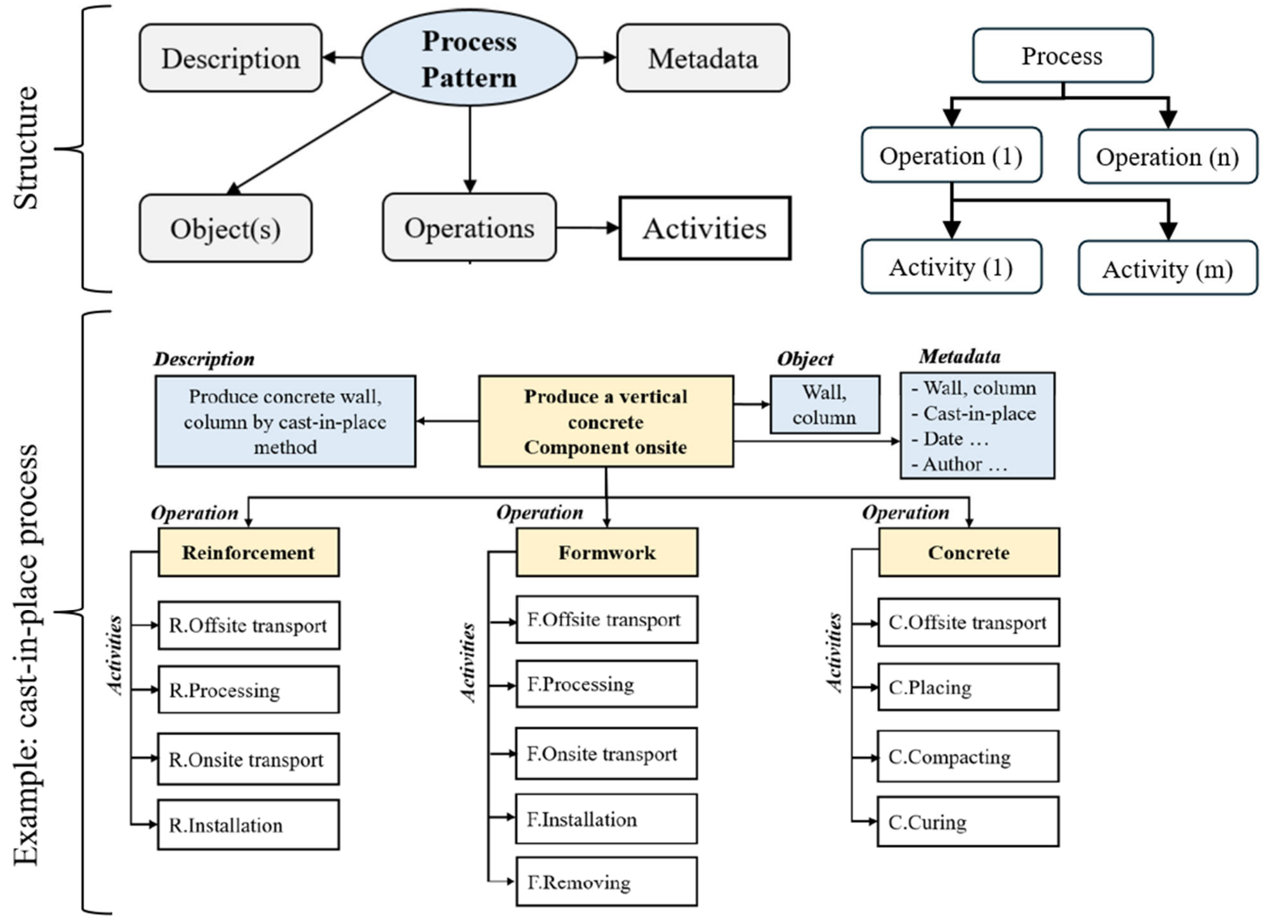


Figure 3: the structure of the process patterns with an example

The necessary information about the building components (such as classification, quantities and dependencies) is extracted from the Building Information Model (BIM). Information about processes and resources is captured by reference to appropriate process patterns. These sources of information are used when initializing agents and populating attributes. In a later step, the required LCA-datasets are queried from a suitable LCA database as a CSV file.

System Dynamics modules are created by combining stock-flow structures and effect-causal loops diagrams. Causal-effect loops are located within the operation agents and consists of reinforcing and balancing loops. A reasonable delay in each loop is considered in the simulation model and its value is represented as a probability distribution using practical experience or literature sources. Examples of the considered dynamic factors in SD Model include overtime and error-related rework. Figure 3 shows schematically the simulation framework planned in the presented methodology.

Figure 4 shows the Error Detection and Rework Loop built in the simulation Software Anylogic. Errors in construction tasks are often inevitable because of unreliable workflow in an uncertain environment (Alzraiee et al., 2015; SangHyun Lee, 2006). Errors lead to rework that consumes more resources than expected, so the environmental impact also increases significantly. However, errors and rework are often overlooked in previous studies on environmental impact assessment.

The rework cycle is affected by the workforce level, productivity, and quality of the performed work. While undiscovered errors are inherently unobservable, the final stages of a project tend to see a big increase in rework discovery. Indeed, crews are aware of their mistakes when the work comes near the end, so the time of error discovery depends on the completion rate. It is very much like finally putting the pieces of a puzzle together. At the end, it becomes quite obvious which pieces are missing or have the wrong shape.

The suggested rework cycle includes four typical loops that result from management policy. The first loop R1 shows the causal-effect relationship between schedule pressure and work with errors. High schedule pressure increases error rate and flawed work. Consequently, the remaining work or initial scope increases. This adversely impacts and increases the schedule pressure again. The second loop B1 shows that if the management policies increase the productivity rate, the work remaining will decrease. The arrow link with minus symbol connecting

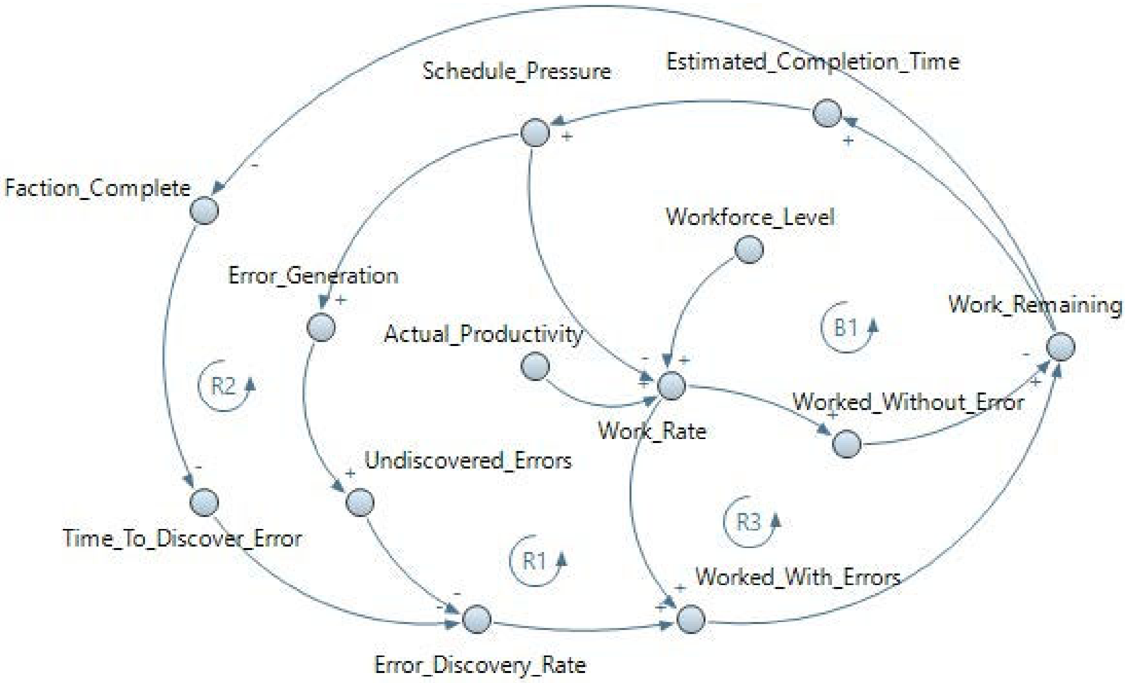


Figure 4: The implemented Error detection and rework feedback loops in AnyLogic.

node Worked_Without_Error and node Work_Remaining has negative polarity (minus symbol) indicating that if node Worked_Without_Error increases, node Work_Remaining decreases. The third loop R2 shows the fact that the errors are usually detected when the project nears completion. The fourth loop R3 indicates that the increase in work rate could increase the errors, and hence increase the work remaining and schedule pressure.

Case study

To test the applicability of the proposed hybrid simulation methodology for assessing the environmental impacts of the construction phase, a simple process scenario is introduced.

The described simulation environment is implemented in the AnyLogic software. Interfaces to various necessary data, such as the process patterns, are carried out through the SQL interface. Other data, such as LCA data sets or the construction quantities, are prepared in CSV format and imported into the AnyLogic environment.

In the case study, the construction phase of the shell structure of a virtual building is selected to apply the proposed methodology. The building has a reinforced concrete frame structure that is constructed using the in-situ concrete method.

The project delivery method is design-bid-build. The planners select the primary building materials in advance, while the construction companies select the auxiliary materials based on their construction method.

Construction companies and their contractors are responsible for transporting materials to the construction site and carrying out construction work on site. Construction works cause environmental impacts through the use of equipment and vehicles, consumption and waste of materials.

The range of environmental sources of influence lies within the decision-making area of construction companies, for example when selecting vehicles, equipment, machines and building materials during the construction phase.

Therefore, the scope of simulation in this case study includes the associated material supply chains and on-site construction operations.

All necessary data on resource consumption and greenhouse gas emissions are obtained from available construction standards and open LCA datasets or international studies.

According to the simulation framework shown in Figure 5, the simulation model is designed in three steps. First, the construction phase of the building is decomposed into units, including product units and process units. Secondly, product units and process units are analyzed to select corresponding process patterns. Third, agents are established and enriched by embedding SD models.

In this case, the building consists of two product units. These are vertical elements such as columns and walls and horizontal elements such as floors and beams.

Process entities represent the cast-in-place concrete method, which is applied to produce product entities of the building shell structure. There are two process entities to produce vertical and horizontal product entities. Each process entity includes three main construction operations: reinforcement, formwork, and concrete. Each operation contains several activities which utilize different resources. An operation can be carried out by

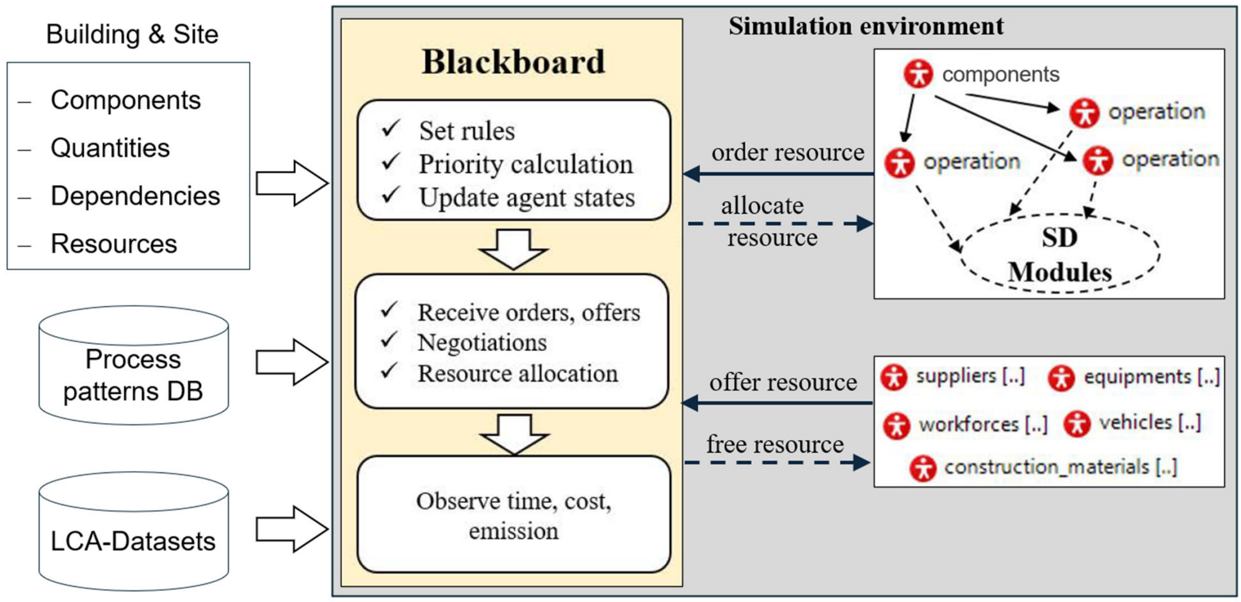


Figure 5: Schematic representation of the planned methodology

different methods. For example, several formwork options can be used to produce one type of concrete structure. When it comes to reinforcement work, reinforcing bars can be processed in the factory or on construction sites.

Depending on the variants selected, modelers can query the process pattern database to identify the specific activities of each process. In addition, the corresponding resources of these activities are determined based on the process pattern.

The abstracted simulation elements of the construction phase are shown in Table 1. For example, building components are defined as product units. Construction-related equipment and building materials that have a direct or indirect impact on the environment are defined as renewable or non-renewable resources.

Table 1: Included Concepts in the simulation model.

Real world	Simulated world
Building components	Product agents
Construction operations	Operation agents
Construction material	Resource agents (non-renewable)
Equipment, workforce	Resource agents (renewable)
Causal-effect relations	Causal-effect loops

Using the proposed simulation model, four scenarios are analyzed to quantify the impact of implemented methods on the environmental performance of the construction phase. The scenarios take into account push or pull controlled processes as well as different resources for the planned construction.

The greenhouse gas emissions (expressed as carbon dioxide equivalents CO2eq) of the four scenarios using the simulation are shown in Table 2. By applying pull-driven processes, scenarios two and four can significantly reduce emissions compared to scenarios one and three.

Using the ABS simulation without the SD models shows lower greenhouse gas emissions results than the ABS-SD simulation. Understandably, the ABS-SD model assumes that error and rework lead to a higher material consumption, so the greenhouse gas value should be higher than the ABS model results. In addition, the ABS-SD model takes into account overtime during night shifts, requiring vehicles, machines and crews to work in uncomfortable conditions, resulting in longer working hours and higher energy consumption.

The SD components take schedule pressure loops and rework loops into account in the construction processes. These loops are needed to capture the realistic behaviour of the construction processes. By effectively applying lean methods to reduce waste in the processes, the effect of these loops will be reduced, and the results of the hybrid ABS-SD simulation will, in the best case, be identical to those of the ABS simulation.

Notably, material consumption contributed the highest CO2 equivalent (61%-64%) among all construction phase impact resources. This result is consistent with results from previous studies that found that materials account for approximately 60–80% of the CO2eq ratio depending on the construction type (Feng et al., 2018; Y. Wang et al., 2017).

Table 2:Simulation results for one typical story.

Scenario	CO2-eq (kg) ABS-SD	CO2-eq (kg) ABS
1	34.809	29.239
2	31.055	29.107
3	33.633	28.124
4	29.589	27.978

Conclusions

The proposed methodology in this paper helps to assess the environmental impact of the construction site and contributes to improve the processes on future construction sites from an environmental perspective. It does not consider only the technical aspect such as the choice of the construction method but also the dynamic changes to the processes. Direct performance measurement on the construction site seems to be difficult because it requires precise documentation of many factors in real-time, which needs a lot of effort and is not justifiable for the construction companies. This makes the target-performance comparison almost impossible. In any case, the proposed simulation-based methodology in the execution planning phase offers a pragmatic solution in this context.

Future development efforts are needed regarding the definition of new patterns for different construction processes and the developing of simulation templates to simplify future use.

The method can also be extended to consider not only the environmental aspect, but also economic and social aspects, which are essential components of sustainable construction.

References

Alzraiee, H., Zayed, T., & Moselhi, O. (2015) Dynamic planning of construction activities using hybrid simulation. Automation in Construction, 49, 176–192. https://doi.org/10.1016/j.autcon.2014.08.011

de Assis, R. F., de Santa-Eulalia, L. A., Ferreira, W. de P., Armellini, F., Anholon, R., Rampasso, I. S., & Santos, J. G. C. L. dos. (2021) Translating value stream maps into system dynamics models: a practical framework. International Journal of Advanced Manufacturing Technology, 114(11–12), pp. 3537–3550. doi: 10.1007/s00170-021-07053-y.

Ding, Z., Yi, G., Tam, V., & Huang, T. (2016) 'A system dynamics-based environmental performance simulation of construction waste reduction management in China', Waste Management. Elsevier Ltd, 51, pp. 130–141. doi: 10.1016/j.wasman.2016.03.001.

Ding, Z., Gong, W., Li, S., & Wu, Z. (2018). System dynamics versus agent-based modeling: A review of complexity simulation in construction waste management. Sustainability (Switzerland), 10(7). https://doi.org/10.3390/su10072484

Dori, G. (2016). Simulation-based methods for float time determination and schedule optimization for construction projects.

IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, R.K. Pachauri, L.A. Meyer (Eds.), IPCC, Geneva, Switzerland, 2014, p. 151. 84.

Feng, K., Lu, W., Olofsson, T., Chen, S., Yan, H., & Wang, Y. (2018). A predictive environmental assessment method for construction operations: Application to a Northeast China case study. Sustainability (Switzerland), 10(11). https://doi.org/10.3390/su10113868

Muntari, M.Y. & Windapo, A.O. (2020) Clay as Sustainable Building Material and its Benefits for Protection in the Built Environment. IOP Conference Series: Materials Science and Engineering, Volume 1144, 3rd International Symposium on Civil and Environmental Engineering (ISCEE 2020) 1st-2nd December 2020, Batu Pahat,Johor, Malaysia

Nguyen, D.T. (2024) Hybrid Simulation-based Lean Management Methodology to Improve the Sustainability of the Construction Phase. Kassel University Press. ISBN: 978-3-7376-1155-8

Lee, SangHyun. (2006). Dynamic Planning and Control Methodology: Understanding and Managing Iterative Error and Change Cycles in Large-Scale Concurrent Design and Construction Projects. Massachusetts Institute of Technology.

Tourtelot, J., de Lacaillerie, J., Duc, M., Mertz, J., Bourgès, A., & Keita, E. (2023). Strengthening mechanisms of clay building materials by starch, Construction and Building Materials, Volume 405, 2023, 133215, ISSN 0950-0618, https://doi.org/10.1016/j.conbuildmat.2023.133215.

Ren, Z., & Anumba, C. J. (2003). Multi-agent systems in construction-state of the art and prospects. Automation in Construction, 13(3), 421–434. https://doi.org/10.1016/j.autcon.2003.12.002

Slosharek, B., Dlouhy, J., Schneider-Marin, P., & Lang, W. (2021) Takting the Sustainability of Construction Processes: An Environmental Assessment Method, Proc. 29th Annual Conference of the International Group for Lean Construction (IGLC), pp. 902–912. doi: 10.24928/2021/0198.

Wang, Y., Feng, K., & Lu, W. (2017). An environmental assessment and optimization method for contractors. Journal of Cleaner Production, 142, 1877–1891. https://doi.org/10.1016/j.jclepro.2016.11.097

Xia, W. and Sun, J. (2013) Simulation guided value stream mapping and lean improvement: A case study of a tubular machining facility, Journal of Industrial Engineering and Management, 6(2), pp. 456–476. doi: 10.3926/jiem.532.

Xie, Y. and Peng, Q. (2012) Integration of value stream mapping and agent-based modeling for OR improvement. Business Process Management Journal, 18(4), pp. 585–599. doi: 10.1108/14637151211253747.

FIRE SAFETY DESIGN OF BUILDINGS: A DECISION SUPPORT FRAMEWORK TO ASSESS THE SAFETY OF OCCUPANTS

Alessandro Carbonari, Massimo Lemma, Martina Tarsi
Polytechnic University of Marche, Department of Civil and Building Engineering and Architecture
60131, Ancona, Italy

Abstract

At the design phase, designers must investigate what design solution represents the best trade-off among a set of simulated fire scenarios. This choice requires a quantitative assessment, embracing the concept of performance-based design. This task can be facilitated using key performance indicators (KPIs). A KPI assessing life safety in the event of a fire is tested in this paper. Also, a framework to facilitate the mapping of KPI values through a graphical visualisation superimposed on the BIM model will be showcased. As a results, the life safety performance assessment is performed at any points of the building layout.

Introduction

The process of fire safety design of buildings entails the assessment of a variety of fire scenarios chosen among the most severe events that could reasonably occur. Out of all the solutions assessed in the various fire scenarios, that one which represents the best trade-off for the specific case will be selected. To this purpose, the adoption of quantitative assessment of the effects of fire on life safety related to the assumed design solutions would help make a grounded, non-personal decision. Quantitative assessment is permissible within the performance-based approach. This paper adopts this approach and tries to go even beyond the current concept of this method, which usually assumes that one scenario at a time is evaluated. Indeed, individual design solutions are evaluated in terms of whether or not they exceed performance thresholds, i.e., the quantitative translation of fire safety objectives. Whereas the approach proposed in this paper wishes to supply the designer with a quick and easy to handle method for choosing the best design solution for the specific case under scrutiny. The safety for occupants will be assessed at every point in the building. To this purpose, the building plan was discretized by creating a square mesh grid. This approach facilitates the comparison among alternative design solutions by making a spatial evaluation of life safety performances in case of fire. In order to make this process effective, one or more KPIs could be identified, not only to check whether a solution meets required design specifications, but also provides feedback about the tested design solution.

Therefore, the aim of this paper is to develop a framework, compatible with a BIM environment, capable of implementing performance-based procedures to assess the degree to which fire safety objectives are being met, in terms of life safety of occupants, through the use of a KPI. The KPI, in addition to assessing any design solutions, must supply a clear and immediate representation of safety conditions levels for occupants. Technically, a graphical display of the KPI has been worked out, taking advantage of the BIM-based design environment. Thanks to the approach proposed in this paper, not only do we display the results of simulations in the BIM environment, rather ASET and RSET values are combined and processed to check the positions where the KPI fails to be verified. It suggests what areas in a building are critical and what countermeasures can be adopted to improve life safety in the event of a fire. This approach enables a building-wide assessment and allows designers to compare an increasing number of design solutions until reaching a decision about the best one for the type of building under consideration. An analysis of the overall framework and workflow is performed to leverage the BIM environment as the preferred collaborative environment.

Literature review

Fire engineering has undergone significant innovations in the latest years, even introducing performance-based approaches. These approaches have proven particularly effective for those scenarios where traditional prescriptive design techniques are not sensibly applicable. Specifically in complex buildings and in buildings that combine different uses, fire safety engineering (FSE) may be the only practical way to achieve a satisfactory standard of fire safety (Fire Protection Association, 2008). Fire engineering has played a pivotal role in liberating building design, allowing for greater architectural flexibility while maintaining high levels of safety (Wilkinson et al., 2013). The safety of occupants in case of fire in buildings is one of the most crucial aspects of FSE (Kong et al., 2013). This is complex due to the high complexity of fire dynamics and the variability of human behaviour. As reported in the Italian fire regulation (D.M. 03/08/2015, 2015), the criteria underlying the performance-based approach for assessing life safety is ASET > RSET, that is occupants can feel safe if the time available to a safe escape (ASET) is greater than the time required to a safe escape (RSET). The difference between the two values is called 'safety margin'. However, this concept has never evolved over time and has not adapted to developments in numerical evaluations (Schröder et al., 2020). Thus, mainly macroscopic assessments have been provided, evaluating only selected points in the building, and lacking a comprehensive view of the safety of occupants. Some methods for a more specific assessment have been proposed. For example, a system proposes expressing tenability conditions and egress accessibility of areas of the building (Poon, 2014). This introduces the concept of available safe utility time (ASUT) that must be longer than the required safe utility time (RSUT), where ASUT represents the time whereby an area is safe to use and RSUT indicates the time for the last person to safely use the area for egress. Another approach evaluates the safety level of egress routes and compartments (Mirahadi et al., 2019). It showed blocked routes, determined the safest egress route from each compartment, and identified critical fire initiation locations. In addition, for an even increasing number of details, a visual representation to depict safety conditions across the entire space was introduced (Schröder et al., 2020). Contrary to earlier evaluations focused on particular areas, this methodology measured safety thresholds comprehensively. Consequently, they produced maps concerning ASET and RSET, and the difference map was derived by subtracting RSET from ASET, effectively displaying the safety margin in the illustrative case of one room. The process of establishing ASET and RSET at each discrete point on the maps aim to move beyond macroscopic representation in favour of a dynamic approach supported by advanced simulation models.

Another critical aspect in the application of FSE concerns the definition of fire scenarios and the choice of design fire scenarios for building structures (Del Prete et al., 2016). The number of possible fire scenarios is typically huge and can hardly be analysed singularly. Therefore, it becomes critical that those design fire scenarios that best represent the most severe cases for the structure are chosen. The US fire regulations NFPA 101 (National Fire Protection Association, 2017) provide us with eight predefined starting design fire scenarios on which to set specific scenarios for different activities. Scenarios selected as design fire scenarios should include but not be limited to those eight predefined. However, they are an excellent starting point on which to develop multiple scenarios for the specific activities.

The performance approach can be facilitated by BIM's ability to generate and return structured information. Fire safety engineers have been slower to adopt BIM as compared to other disciplines (Malagnino et al., 2022). Several research focus on the integration of BIM and fire prevention simulation tools (Wang et al., 2015; Sun and Turkan, 2020). The great limitation of one-way exchange between BIM and FSE tools was also identified. An example to overcome this limitation concerns the development of an open-source framework called 'Evac4BIM' to facilitate two-way data exchange, with a specific focus on fire evacuation (Yakhou et al., 2023). The authors yield the results of numerical evaluations available in the BIM environment which become accessible to all stakeholders but no graphical representation of the results covering every point of the building was provided in return.

Research questions

This paper suggests a methodology and develops tools to answer the following research questions (RQs):

1. can a BIM-based framework assessing fire safety performances of buildings, in terms of life safety for occupants, be defined and implemented to support the designer during the design phase?
2. can a KPI be defined to assess the safety conditions of occupants and to gives back a clear and immediate visualization in a BIM environment?

Materials and methods

In order to answer the research questions formulated in the previous section, a framework to support the designer's choices during the assessments of life safety in case of fire emergencies was developed. This evaluation was meant to overcome the current approach based on calculating ASET and RSET only at a limited number of locations in a building and working out results that do not represent the consequences of the fire on occupants. Indeed, the current numerical evaluations can tell us whether occupants were able to exit the building or were entrapped, but it does not tell us where in the building they may have been hampered by the effects of fire. Thanks to the approach proposed in this paper, we want to assess the safety of the occupants at all points of the building, discretizing the entire area using a grid. In this way, it will be possible to visualise the points where the occupants may be subject to the effects of fire and the areas that are most dangerous for every different design scenario.

The project workflow for each design fire scenario

A representation of the processes that could be carried out to handle each design fire scenario is depicted in Figure 1.

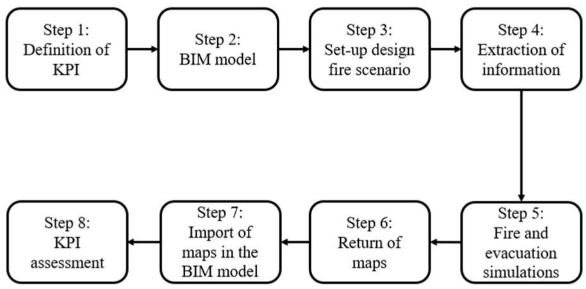


Figure 1 - Conceptual representation of the flow to handle design fire scenarios

By applying all the steps to the design fire scenarios identified, it will be possible to compare the different design solutions and decide which represent the best trade-off. According to the performance-based approach, KPI must be identified as a first step (step no. 1). Having set life safety as the aim of our work and taking the Italian fire regulations (D.M. 03/08/2015) as a reference, the safety margin (t_{safety}), i.e., the difference between ASET and RSET, was chosen as the KPI. As reported in paragraph M.3.2.2 of the reference regulation (D.M.

03/08/2015), this safety margin can either be set as a percentage value, e.g., RSET must double ASET, or as a predefined time limit. The minimum required threshold in the fire regulation is 30 seconds (Eq. 1), which is the threshold set in the test case reported in this paper. Hence, the KPI was set as:

$$t_{safety} > 30\ seconds \quad (1)$$

The BIM model of the building must be available (step no. 2). In the application proposed in this paper it was created within the authoring software Revit™. Once the design solution has been assumed for the design fire scenario under consideration (step no. 3) the required information is extracted from the BIM model (step no. 4). The latter will be needed as input to simulation tools (step no. 5). Two different simulation software tools were used for this step. The execution of fire simulations was entrusted to the software FDS (fire dynamics simulator). It is a field model, i.e., it divides the environment into elementary volumes, calculating and outputting data for interested variables in each volume. In the input list for FDS, in addition to geometry and fireplace, even the positions and quantities of the slices are specified as output, because they are necessary for the graphical visualisation of results. The use of this software becomes relevant to our method, as long as the building's space must be discretized to assess the safety of the occupants at every location. Then, the Jülich Pedestrian Simulator (JuPedSim) was used to conduct evacuation simulations. It represents an extensible framework designed for the simulation and analysis of pedestrian movement at the microscopic level (Kemloh Wagoum et al., 2015). Again, the simulation environment was discretized to assess the passage of occupants across every cell of the grid. In order to read the results of the simulations clearly and immediately, three Python scripts have been implemented to obtain a visual representation of them (step no. 6). In practice, maps will be produced containing the ASET and RSET evaluated at each point of the building and, as a result, the safety margin map, i.e., the aforementioned KPI, can be obtained. The maps created are imported into the BIM model of the building to visualise directly within it (step no. 7). At this point, the KPI must be evaluated by the designer (step no. 8). If the KPI is verified and, therefore, the threshold is not exceeded, the proposed solution can be considered and compared with other ones. Otherwise, the designer is required to investigate other solutions on the considered design fire scenario. In this way, the designer can assess the effect of different design solutions on a design fire scenario quantitatively and efficiently.

Extraction of information

Structured information needed to carry out evaluations can be extracted from the BIM model of the building. In particular, it is possible to automatically create the input list for the simulation software tools. Having implemented the building model in Revit™, Dynamo™ was chosen as the tool for data export. In addition, Dynamo™ was used to read the geometry of a building, to create a mesh of the building area composed of quadrilateral finite elements characterized by a predefined parametric value, and to export the list of mesh points along with their coordinates automatically. In the case study reported in this paper, a mesh element size, to create the maps, as big as 0.50 x 0.50 m was used. This size of the mesh was considered accurate enough to represent the path followed by people and the spread of fire. The authors' opinion is that a smaller mesh size would not add relevant information about hazardous areas in the building.

Return of maps

In order to provide a clear and immediate visualisation of the results of the occupants' life safety assessment, maps have been implemented. For this, we have taken into consideration what was proposed by Schröder et al., 2020. The authors proposed using the Python programming language to create ASET, RSET and difference maps by importing the results of fire and evacuation simulations into scripts. The Python scripts implemented in our paper are an adaptation of those ones made available by Schröder et al., 2020.
ASET maps were worked out for two environmental parameters: the temperature room and the light extinction coefficient. The thresholds for these parameters are 45°C for the temperature field and 0.23 m^{-1} for the light extinction coefficient field, assessed at the height of 2 metres above the floor level (Zehfuβ, 2020). The script to obtain the ASET map converts the slices given as output by the FDS simulation as an ASCII file. The geometry of a building, the mesh size, and the colour scale of the maps are the other input data needed. The resulting ASET map is produced in both 'png' and 'txt' formats. RSET maps were worked out according to the trajectories of the occupants. In our case, ten different trajectories for each hypothesised design solution were assumed for the occupants, arranging them randomly. The script to realise the RSET map is divided into two parts. The first part processes the results of the imported occupant trajectories. These are contained within an 'xml' format file processed from the 'txt' format file provided by the evacuation simulator. The second part of the script creates the global RSET map, interpolating the worst cases from all the earlier ones. The input set consists of the building geometry, the mesh size, and the map colour scale. The output is the RSET map in 'png' and 'txt' formats. Once the final ASET and RSET maps were created, the difference map was completed. The script creates the difference map, which is the map including safety margin values, by subtracting the RSET map from the ASET map. It accepts ASET and RSET maps in 'txt' format as inputs, as well as building geometry, mesh size, colour scale of the map and produces the difference map in 'png' and 'txt' formats.

Import of maps into the BIM environment

Once the maps have been created via Python, these were imported into the BIM environment to be assessed by the designer. To conduct this step, a flow was created in Dynamo™ that imports the text file produced by Python, containing the KPI values. In the BIM model, the discretization mesh was already created, as shown in the section 'Extraction of information'. A geometric element was placed in the centre of each mesh element. Then, a new design parameter was created, named 'T_{safety}' and associated with the above element. The text file created by Python was converted into Excel and imported into the Dynamo™ flow. Next, Dynamo™ nodes were used to transform the excel matrix into a list, so that it could be imported as a list node. The values contained within it were assigned to the 'T_{safety}' parameter according to the order of the list of elements. At this point, each element was assigned a different colour according to the value of the parameter. Values below 30 seconds, i.e., the KPI threshold (Eq. 1), were coloured red, the others green. Some points are left without a numeric value. These have been colored white and they indicate that no occupant has ever walked through that point.

Application on a case study

The feasibility of the proposed framework was assessed through the case of a pilot building. It is a building located in the campus of the Polytechnic University of Marche (Ancona, Italy), which hosts the school of Medicine (Figure 2). It is an 8-floor level building, 7 floors above ground and 1 basement. The building is used for teaching, research, a library, and other services. For the purposes of this paper, only the ground floor of the building has been considered.

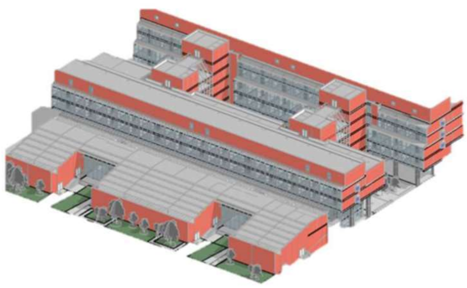


Figure 2 - 3D model of the pilot building

Fire and evacuation scenarios

Fire scenarios represent the most severe but realistic events that can occur while performing activities allowed in the building. These can be very numerous, especially in complex buildings. For the purpose of this paper, the default scenarios contained in the US fire regulations were considered. In fact, NFPA 101 covers a wide range of situations and defines fire scenario in Chapter 5.5.3. The latter provides us with eight starting scenarios on which to develop the most appropriate scenarios based on the specific case. In particular, two out of the eight predefined scenarios listed in NFPA 101 were analysed, namely:

- scenario no. 2, which describes a fire that develops through the burning of a material having an ultrafast growth curve, located along any major egress route. The interior doors at the beginning of the fire are supposed to be open. This scenario must address the simulation of the fire paying particular concern to people's egress problems. In fact, due to the hypothetical fire having a rapid spread, special attention should be paid to reducing the number of available egress routes, evaluating the availability and effectiveness of alternative egress systems and the consequences of the fire on the assets;
- scenario no. 6, involves an intense fire as a consequence of the highest possible fire load in normal operations in the building. It refers to the rapid growth in the presence of people.

In fact, designers are expected to test and compare several fire scenarios. They usually assume a first design solution and investigate the obtained performances. Then, changes could be made to the first solution to improve it and new investigations are performed until the verification process is concluded. In this paper, two scenarios have been assumed as relevant, each split into three subsets.

Referring to the scenario no. 2 of the NFPA 101 described above, the fire hearth was positioned along the corridor as this is the main escape route. The corridor serves offices, the library, the cafeteria and seven classrooms for a total capacity of 1040 occupants. The occupants inside the seven classrooms have also emergency exits directly into the classrooms. This set of scenarios is called 'scenario LS1'. Taking into consideration the scenario no. 3 of the NFPA 101, the fireplace was positioned inside the library. This set of scenarios is called 'scenario LS2'. In Figure 3 the fire compartment and the library are marked in green and blue, respectively. Each fire scenario was combined with ten different occupant evacuation trajectories.

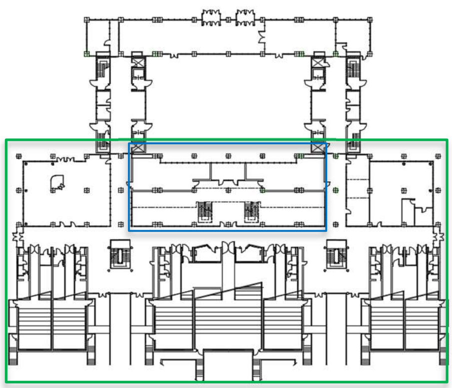


Figure 3 - Ground floor of the pilot building

Scenario LS1

For this first scenario, three alternatives were analysed. In the first case, called 'LS1_1 scenario', two emergency exits were placed along the corridor for the safe exit of the occupants. From the fire simulation, the values of the environmental parameters specified in the input list are obtained. In addition, it is possible to display them graphically with the set slices. An example is provided in Figure 4, displaying a slice taken from the FDS tool called

'Smokeview', concerning the light extinction coefficient reached at the end of the simulation (300 seconds) for the 'LS1_1 scenario'.

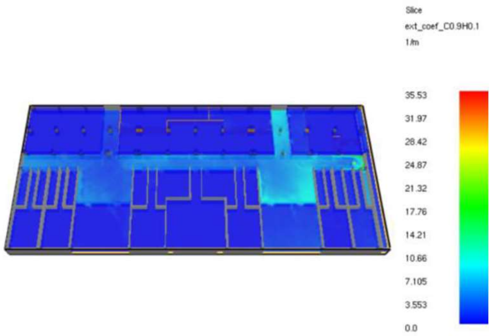


Figure 4 - Slices generated by FDS about light extinction coefficient for the 'LS1_1 scenario'

Regarding the evacuation simulation, the times required by the occupants are shown in Table 1 for the ten different trajectories (denoted by the term 'Seed_ID' from the number 1254 to 1263). In the second alternative ('LS1_2 scenario'), two emergency exits were added along the corridor. The effects of the fire remain the same as in the 'LS1_1 scenario', while the times required by the occupants change and are shown in Table 1. In the last alternative 'LS1_3 scenario', two more doors were added along the corridor for a total of six emergency exits. Again, only the RSET values change (Table 1).

Table 1 - Required safe escape time as a result of the JuPedSim simulation

	Required safe escape time [s]					
Seed_ID	LS1_1	LS1_2	LS1_3	LS2_1	LS2_2	LS2_3
1254	67.88	67.13	67.13	67.13	65.50	65.50
1255	67.63	66.50	66.50	66.50	66.62	66.62
1256	69.25	67.38	67.38	67.38	67.38	67.38
1257	68.50	68.50	68.54	68.54	68.54	68.54
1258	70.50	68.75	68.80	68.80	68.75	68.75
1259	67.38	69.00	69.00	69.00	69.00	69.00
1260	69.38	66.88	66.91	66.91	66.90	66.90
1261	68.50	68.75	68.87	68.87	68.87	68.87
1262	68.25	68.25	68.30	68.30	68.30	68.30
1263	64.63	65.63	65.72	65.72	65.73	65.73

Scenario LS2

In the 'LS2 scenario', the fire was placed inside the library, where the greatest fire load is expected. A first analysis assumed 38 occupants inside the library, which is divided into a reading room and the book depository ('LS2_1 scenario'). Also in this case, ten different simulations were conducted for the trajectories of the occupants and the times are shown in Table 1. Subsequently, an increase in the number of occupants, from 38 to 88, within the library and the elimination of subdivisions was assumed. This scenario was called 'LS2_2'. In this case, the fire simulation was run again because the building layout was changed and, as a result, the fire will spread differently. The third alternative ('LS2_3 scenario') differs from 'scenario LS2_2' only in that a library wall is replaced by a glass wall. The RSET values are contained in Table 1, but they are the same as in the 'LS2_2 scenario', while the fire simulation is re-run.

Maps generation

Once the simulations for the fire and evacuation have been carried out, the results are imported into the scripts developed with Python. The latter gives as output the graphical representations of the results in the form of a map. In particular, the values of ASET, RSET and the differences between the two are calculated. In this way, we can get a graphical representation of the safety margin, i.e., our KPI.

Scenario LS1

The ASET maps obtained for temperature and light extinction coefficient are shown in Figure 5-a and Figure 5-b, respectively. Two maps are shown because there are two environmental parameters to be assessed. The final ASET map is the one with the highest risk, i.e., with the least available time. The RSET map, on the other hand, can be seen in Figure 5-c. Finally, Figure 5-d shows the difference map obtained by subtracting RSET values from ASET values. In this map, therefore, the KPI is shown. When the value falls below the threshold of 30 seconds (Eq. 1) indicates that the point is not safe for the occupants. For the second and the third alternative of the 'LS1 scenario', since the position of the fire hearth and the layout of the building remain unchanged, the maps of ASET will be the same as those already shown in Figure 5-a and Figure 5-b. The new RSET maps have been produced and, as a result, difference maps were shown in Figure 6-a for the 'LS1_2 scenario' and Figure 6-b for the 'LS1_3 scenario'.

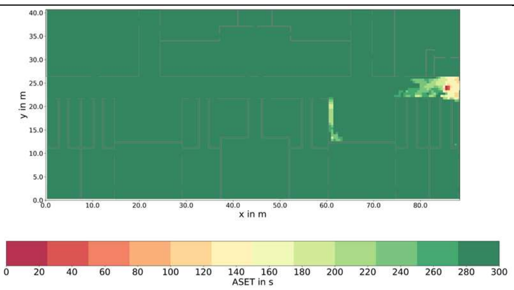
Scenario LS2

For the 'LS2 scenario', the difference maps are shown directly in Figure 7-a (scenario LS2_1), Figure 7-b (scenario LS2_2) and Figure 7-c (scenario LS2_3).

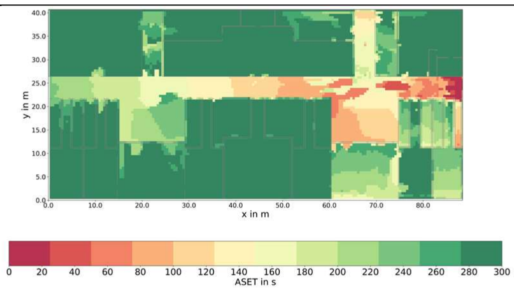
Visualization of maps in a BIM environment

Once the maps have been created via Python, these were imported into the BIM environment for a seamless evaluation by designers. To conduct this process, the Dynamo™ script created as described in 'Import of maps into the BIM environment' section was used. In this way, the designer can visualise the results directly within the BIM model and all stakeholders can collaborate in one environment. Even those who are not familiar with fire simulation tools can visualise simulation results in the BIM environment. For the sake of clarity, the Revit™ screenshots including the imported difference maps were shown in the following. For representative purposes, only the maps imported into Revit™ relating to the 'LS1_1 scenario' in Figure 8-a, 'LS1_2 scenario' in Figure 8-b and 'LS1_3 scenario' in Figure 8-c have been shown.

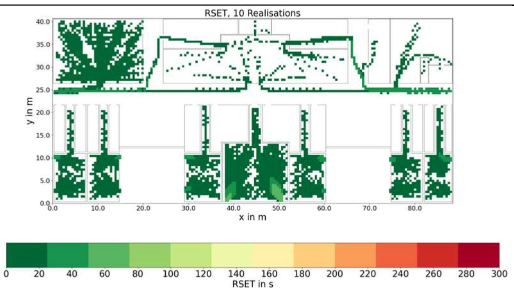
(a)



(b)



(c)



(d)

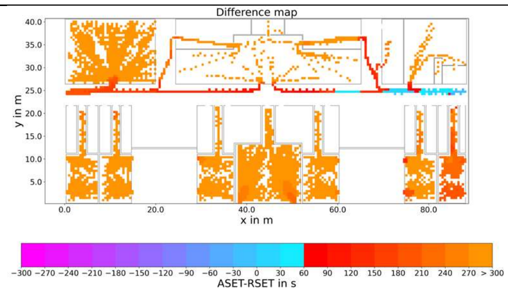
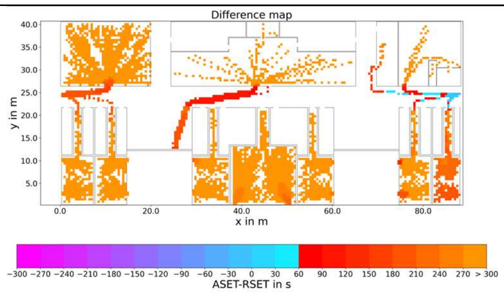


Figure 5 - ASET, RSET and difference maps for the 'LS1_1 scenario' (output of Python)

(a)



(b)

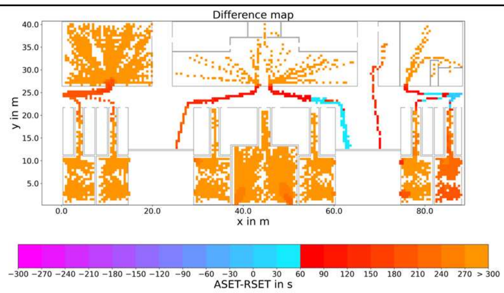
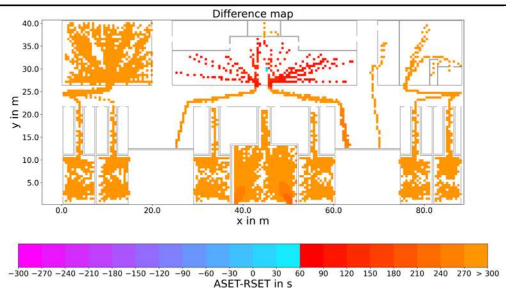
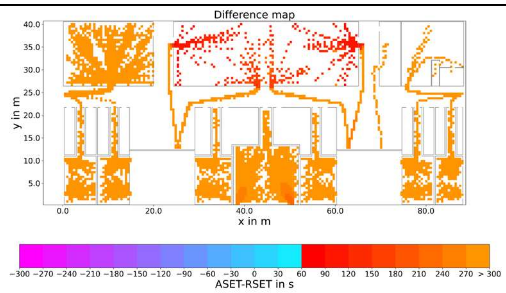


Figure 6 - Difference maps for the 'LS1_2 scenario' and 'LS1_3 scenario' (output of Python)

(a)



(b)



(c)

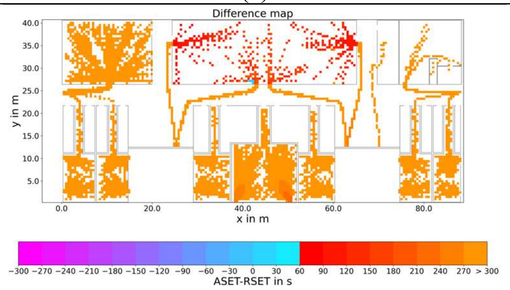


Figure 7 - Difference maps for the 'LS2 scenario' (output of Python)

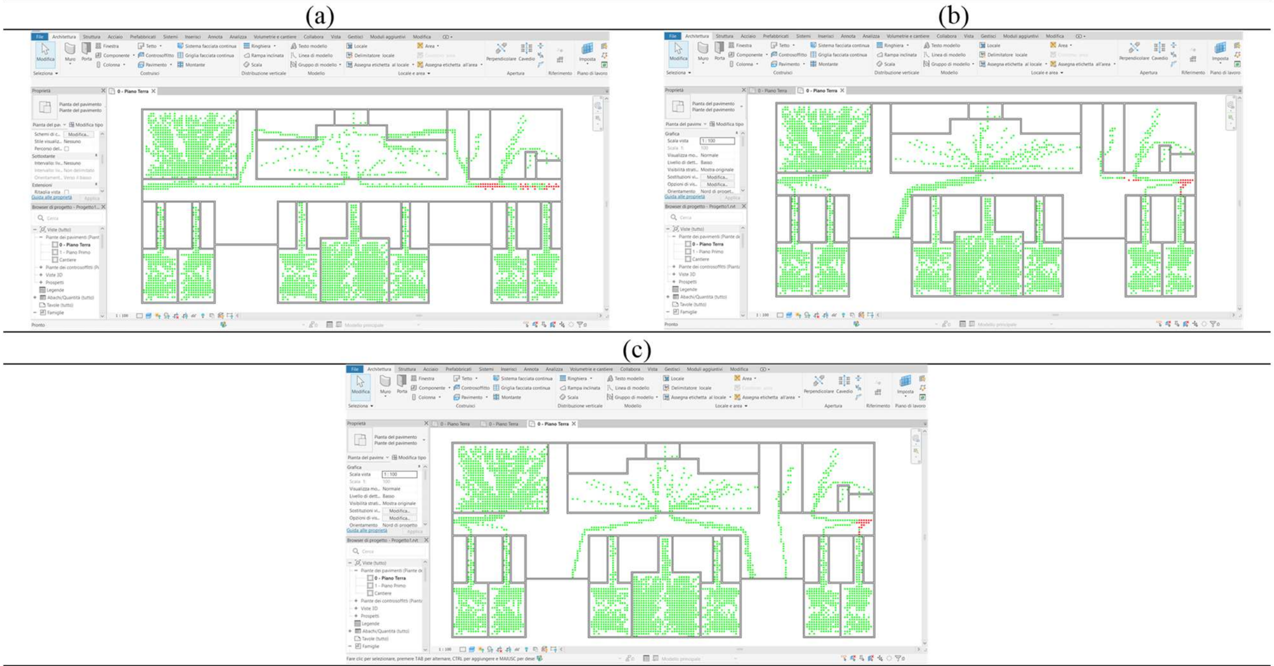


Figure 8 - Difference map for the fire compartment for the 'LS1_scenario' once it has been imported into Revit^{TM}

The pictures show the geometric elements inserted in the centre of each grid mesh. These take on a green colour when the value of the imported KPI associated with the 'T_{safety}' parameter is greater than 30 seconds. The red colour corresponds to a KPI of less than 30 seconds. When no value is associated with the parameter, it means that no occupant has passed through that grid mesh, so there is no RSET value and, so, no safety margin can be calculated. For this reason, these points have been left white.

As a result, Revit^{TM} was able to integrate the KPI information and display the points where safety conditions could not be verified.

Discussion

Thanks to the results reported in the earlier section, it can be seen that the methodology introduced gives us a clear and immediate view of the KPI values and, therefore, of the level of safety of the occupants inside the building in case of fire. Concerning the 'LS1 scenario', two emergency exits are not sufficient for a safe escape of the occupants. Indeed, in Figure 8-a includes forty-six points are coloured in red along the corridor. This means that the KPI threshold has been exceeded at those points and the occupants passing through that area are not safe. Consequently, the designer could think of adding two emergency exits along that corridor. The result is depicted in Figure 8-b and shows that conditions have been improved, despite twenty-three points close to the fire hearth still remained below the required safety level. To overcome this problem, the designer could choose to add two additional emergency exits, for a total of six. Figure 8-c shows that the life safety level has increased further, even if some but smaller areas are still unsafe, due to the position of the exits close to the fire hearth, and eighteen points are still red-coloured. This gives a hint to the designer about the number of emergency exits required and their location for a safe escape of the occupants. Comparing Figures 5-d, 6-a, and 6-b with Figures 8-a, 8-b, and 8-c, respectively, we can see the correspondence between the maps obtained through Python and the representation obtained once the maps are imported into the BIM model.

In the 'LS2 scenario', on the other hand, the maps obtained with Python show that the safety conditions are maintained even when increasing from 38 occupants inside the library (Figure 7-a), where only one point is unsafe, to 88 occupants (Figure 7-b), while leaving the number of exits unchanged. In the third hypothesis (Figure 7-c) where a glass wall is inserted, four unsafe points are seen close to the fire hearth, but there is no glass breakage due to the thermal action of the fire. This shows the designer that the fire that occurs inside the library does not spread rapidly along the escape routes. In the same way as the 'LS1 scenario', the maps imported into the BIM model were also obtained for the 'LS2 scenario'.

Thanks to the visualisation of the maps directly in the BIM model, a single working environment is created where even non-experts in fire simulation software can visualise the results clearly and immediately. In addition, the designer can compare different design solutions for each design fire scenario based on quantitative results (KPI values) that refer to the entire building. Thus, the safety of the occupants is assessed at every point, visualising the points and areas with critical conditions for them. The latter occurs when the KPI threshold is exceeded, that is in case the consequences of combustion result in environmental conditions that reduce the time available for occupants to get to a safe place.

Conclusions

In this paper, an approach has been proposed for assessing occupant safety in case of fire during the building design phase. This approach aims to go beyond the current concept of ASET/RSET evaluated only at specific points. Furthermore, current assessments do not supply insights into the zones or points where conditions become critical for occupants but instead offer a global evaluation of the safety of evacuation routes. The graphical visualisation of the KPI, in the form of maps, enables designers to immediately pinpoint those locations in the building that are critical for the occupants. This assessment was carried out by spatially discretizing the entire building, thus considering every point in the layout. Furthermore, thanks to the visualisation approach, the designer can quantitatively assess the assumed design solutions over numerous design fire scenarios. The KPI initially chosen to monitor fire safety objectives was mapped inside a BIM environment, importing maps created through Python. The KPI enriches the design model, and the entire method provides support to the designer when looking for the best trade-off, which is meant to implement the approach of performance-based design. In this way, the designer's decisions can be made quickly in a complex environment and a single collaboration environment can be created by importing the results of fire assessments directly into the BIM environment. One possible recommendation for future research involves identifying additional KPIs related to the other domains of FSE, such as structural safety. Another interesting aspect would be to implement the whole procedure within an open software and computing environment.

References

Del Prete, I., Cefarelli, G., Nigro, E., 2016. Application of criteria for selecting fire scenarios for structures within fire safety engineering approach. J. Build. Eng. 8, 208–217. https://doi.org/10.1016/j.jobe.2016.10.010

D.M. 03/08/2015, 2015. Codice di Prevenzione Incendi. https://www.vigilfuoco.it/aspx/page.aspx?IdPage=10259

Fire Protection Association, 2008. Approved Document B (Fire safety) Volume 2. https://assets.publishing.service.gov.uk/media/639ae876e90e0721839ea637/Approved_Document_B__fire_safety__volume_2_-_Buildings_other_than_dwellings__2019_edition_incorporating_2020_and_2022_amendments. pdf

Kemloh Wagoum, A.U., Chraibi, M., Zhang, J., Lämmel, G., 2015. JuPedSim: an open framework for simulating and analyzing the dynamics of pedestrians. 3rd Conf. Transp. Res. Group India 3rd CTRG.https://www.researchgate.net/publication/28937 7829

Kong, D., Lu, S., Frantzich, H., Lo, S.M., 2013. A method for linking safety factor to the target probability of failure in fire safety engineering. J. Civ. Eng. Manag. 19, S212–S221. https://doi.org/10.3846/13923730.2013.802718

Malagnino, A., Corallo, A., Lazoi, M., Zavarise, G., 2022. The Digital Transformation in Fire Safety Engineering over the Past Decade Through Building Information Modelling: A Review. Fire Technol. 58, 3317–3351. https://doi.org/10.1007/s10694-022-01313-3

Mirahadi, F., McCabe, B., Shahi, A., 2019. IFC-centric performance-based evaluation of building evacuations using fire dynamics simulation and agent-based modeling. Autom. Constr. 101, 1– 16. https://doi.org/10.1016/j.autcon.2019.01.007

National Fire Protection Association, 2017. NFPA 101 - Life Safety Code. https://www.nfpa.org/codes-and-standards/1/0/1/nfpa-101

Poon, S.L., 2014. A Dynamic Approach to ASET/RSET Assessment in Performance based Design. Procedia Eng., 2013 International Conference on Performance-based Fire and Fire Protection Engineering, Wuhan (ICPFFPE 2013) 71, 173–181.https://doi.org/10.1016/j.proeng.2014.04.025

Schröder, B., Arnold, L., Seyfried, A., 2020. A maprepresentation of the ASET-RSET concept. Fire Saf. J. 115, 103154. https://doi.org/10.1016/j.firesaf.2020.103154

Sun, Q., Turkan, Y., 2020. A BIM-based simulation framework for fire safety management and investigation of the critical factors affecting human evacuation performance. Adv. Eng. Inform. 44, 101093. https://doi.org/10.1016/j.aei.2020.101093

Wang, S.-H., Wang, W.-C., Wang, K.-C., Shih, S.-Y., 2015. Applying building information modeling to support fire safety management. Autom. Constr. 59, 158–167. https://doi.org/10.1016/j.autcon.2015.02.001

Wilkinson, P., Glockling, J., Bouchlaghem, D., Ruikar, K., 2013. Using business impact analyses to enhance resilient fire engineering building design. Archit. Eng. Des. Manag. 9, 229–249. https://doi.org/10.1080/17452007.2012.738043

Yakhou, N., Thompson, P., Siddiqui, A., Abualdenien, J., Ronchi, E., 2023. The integration of building information modelling and fire evacuation models. J. Build. Eng. 63, 105557. https://doi.org/10.1016/j.jobe.2022.105557

Zehfuβ, J., 2020. Leitfaden Ingenieurmethoden des Brandschutzes. https://www.vfdb.de/media/doc/technischeberichte/TB_04_01_Leitfaden_IngMethoden_4Auflage_2020-03-26.pdf

IFC-BASED COST ESTIMATION: APPLICATION TO A STRUCTURAL MODEL

Jacopo Cassandro, Claudio Mirarchi, Alberto Pavan
Polytechnic of Milan, Milan, IT

Abstract

Accurate cost estimation is an essential factor in the success of any construction project. Based on previous work, this study aims to use the IFC data model to support cost estimation processes and reduce human error. Through a code, developed in IfcOpenShell, a cost estimation has been made in IFC associating the new *IfcCostItems* with the IFC 3D model objects. The study was validated by applying the methodology to a real case of a structural model and defining a cost schedule for the structural project. Ultimately, this study seeks to support, verify, and improve public tenders quality of cost estimates.

Introduction

Accurate cost estimation is a crucial factor for any construction project's success. In the Architecture, Engineering, and Construction (AEC) industry, this task demands a high degree of accuracy and attention to detail. Without proper cost estimation, projects can easily exceed the budget, leading to delays, reduced profit margins, and ultimately project failure.

The construction process is characterized by its dynamic, complex, and time-consuming nature. Within the AEC industry, effective collaboration among all stakeholders is essential for sharing diverse information generated throughout the construction phase. This information includes quantities of physical building components, schedule plans, resource consumption, costs, site layout, safety management, and quality evaluation (Adeli et al., 2001). Cost estimation is a key task involving calculating quantities, and project costs, and classifying essential products within a construction project. Cost estimation is commonly defined as the process of predicting project costs at the operational level, based on detailed design drawings/documents and specific construction methods/specifications. Construction cost estimation is typically carried out separately by construction professionals due to the lack of a well-developed integrated system for both activities (Liu et al., 2014). Despite the existence of advanced technologies for quantitative take-off (QTO), scheduling, and costing, a significant gap remains in seamlessly integrating these activities into a cohesive system. The potential solution lies in a system founded on modeling technologies associated with scheduling and economic management, promising automatic and effective integration (Lu et al., 2016).

In the AEC industry, the standard format for information exchange is the Industry Foundation Classes (IFC), an open international standard developed by BuildingSMART.

Therefore, to support, verify, and improve the quality of cost estimates, in public tendering, and reduce human error-prone, this paper studies how to develop a cost estimation based on IFC.

IFC provides some entities to represent information in building management, including *IfcCostSchedule* (cost planning), *IfcCostItem* (unit cost estimation item), and *IfcCostValue* (value).

A cost estimation (*IfcCostSchedule*) for a structural project has been created including all cost classes (*IfcCostItem*) needed to define a correct estimate of costs. The *IfcCostItems* are linked to the 3D model's geometric objects to determine the quantity of individual works. They are also associated with unitary items in the public works price list (in this specific case of the Lombardy Region's Price List), to establish the unit cost values. The price list items are characterized by a new architecture based on IFC that contains a set of attributes in which information can be instantiated and created. This new architecture of cost items was validated in a previous study by the same research team (Cassandro, et al., 2023). Furthermore, the study carried out seeks to analyse, evaluate, and use the IFC data model to support the economic analysis processes of the projects.

The paper is structured as follows. Firstly, it presents an analysis of the existing literature on cost estimation, including the use of IFC classes, and the IFC standard with a specific focus on the *IfcCostItem* and *IfcCostSchedule* entities (attributes and relations). Currently, no BIM authoring software can write this entity, so it was decided to use the IfcOpenShell library to create the entity and the cost estimation through code. The final step included the validation of the data model generated through IFC viewer/reader, the analysis of the results, the limitations of the study, and future developments.

Literature Review

Cost Estimation Process in Construction

The process of cost estimation involves using available information to estimate the total cost of construction work activities of the project (Ji et al., 2019). Estimating the cost of the project plays a key role in the success of a construction project (Ji et al., 2019), (Jrade & Alkass, 2007). This involves assessing and predicting total costs over a specified period, incorporating all available project information and resources (A. Kwakye, 1994). Initial estimates, with an accuracy of -30% to +50%, are refined at the budget and conceptual level (-15% to +30%) and ultimately at the definitive level (-5% to +15%) (Rast & Peterson, 1999). Cost estimation serves the purpose of providing decision-making information (Carr, 1989).

However, despite the availability of advanced technologies, construction contractors still rely on conventional methods, leading to potential inaccuracies due to factors such as insufficient time and poor tender documentation (Famiyeh et al., 2017).

A cost estimation document contains all work items with a short description, estimated quantities, the unit cost of the repair work, the total cost of each item, and then summing all the items, the total project cost. However, other methods are also developed for estimating construction costs. These include expert systems, artificial neural networks, case-based reasoning, fuzzy logic systems, simulations, statistical regression approaches, decision trees, radial basis function neural networks, and particle swarm optimization (Tayefeh Hashemi et al., 2020).

Cost Estimation Process through IFC

Several studies have explored the applicability of IFC in project planning and cost estimation. Froese T et al. suggested that IFC can effectively represent information related to costs, construction processes, resources, products, and project documents (Froese T et al., 1999). Ma et al. established a BIM-based Construction Cost Estimating (CCE) software framework based on Chinese standards (Ma et al., 2010). Zhiliang et al. investigated the use of IFC standard for construction cost estimating in China, noting the need for extensions in the form of proxy elements and property sets (Zhiliang et al., 2011).

Additionally, Ma et al. discussed key issues for semi-automatic Tendering of Building Projects (TBP) cost estimation using IFC data (Ma et al., 2013). Liao et al. proposed a method to develop a collaborative construction prototype model on BIM software to enhance information-sharing efficiency (Liao et al., 2014). Xu et al. introduced a philosophic position for model-based cost estimation using IFC, emphasizing contextual information and pricing extension (Xu et al., 2013). Xu et al. innovatively used BIM data linked to a project to create items for a bill of quantity for cost estimation (Xu et al., 2016).

Other researchers, including Wu et al., Sacks et al., and Elghaish et al. explored the potential of BIM to enhance cost estimation (Wu et al., 2014), (Sacks et al., 2018), (Elghaish et al., 2020). Staub-French et al. and Lee et al. investigated the use of semantic web technology in BIM-based cost estimation (Staub-French et al., 2003), (Lee et al., 2013) Fürstenberg, et al. delved into how semantic web technology can support automated cost estimation linked to IFC property sets (Fürstenberg et al., 2021).

In the context of cost estimation, current practices often involve connecting codes in digital objects to matching keys with various price items. This process requires different codes for each element generating distinct articles (Pavan et al., 2017).

IFC & Construction Management

Developed by BuildingSMART, the IFC is an open and interoperable standard with a mission to facilitate interoperability across different domains in civil engineering projects. Its objective is to enable the sharing and exchange of project information among various computer applications used by different project participants. Initially released in 1997, the official version, IFC4 ADD2 TC1, became an ISO standard in 2018 (ISO 16739-1:2018). At the beginning of 2024, the new IFC version 4.3 ADD2 - 4.3.2.0 was finally approved by ISO and was published in April 2024 becoming the basis for IFC software certification.

Structured in EXPRESS data specification language or XML, the IFC standard describes actors, processes, controls, resources, and products within the construction domain. The IFC data model is organized hierarchically into four conceptual layers: Resource layer, Core layer, Interoperability layer, and Domain layer. Entities within this model can be related to each other and characterized by a set of attributes, facilitated by the *IfcRelationship* entity. This allows related information to reside either inside or outside the project data.

The IFC standard plays a crucial role in information exchange throughout the project lifecycle in construction or building management projects, serving as a prominent vendor-neutral data schema in the AEC industry. IFC enables the exchange of both geometric and semantic information among different stakeholders and software (Fürstenberg et al., 2021).

In the field of construction management, IFC provides entities like *IfcWorkPlan* (schedule planning), *IfcTask* (construction tasks), *IfcScheduleTimeControl* (task time information), *IfcCostSchedule* (cost planning), *IfcCostItem* (unit items of cost estimation), and *IfcResource* (construction resources such as material, product, labor, and equipment resources).

Methodology

This research aims to define a new method for cost estimation using IFC. The method is based on the use of entities within the IFC standard and the relationships that these can create to develop a cost estimate. Specifically, the method used provides for:

- Definition and structuring of the new architectures of the unit cost of the construction works, contained in the Price List of the Lombardy Region, based on what was validated in a previous study by the same research team (Cassandro, et al., 2023);

- Identification of a case study relating to a structural model in IFC format;
- Development of a code for the implementation of the new cost estimate based on the relationship between IFC classes;
- Identification of the measurement criteria of the unit cost items useful for the correct quantification of the cost items (no manual entry of the quantification rules already defined in the price list).

This methodology will also allow the validation of the IFC model obtained to assess the percentage of correctness in the association of identified cost items and estimated costs. This has already been tested and validated in a previous work of the same research group (Cassandro, et al., 2023). Figure 1 shows the workflow performed in this research.

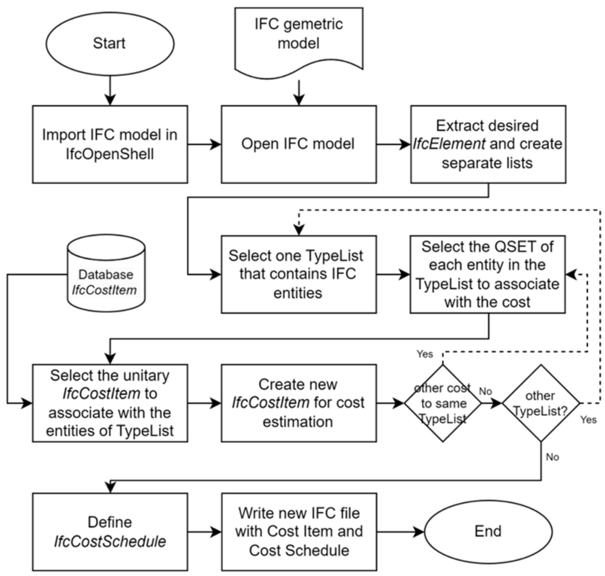


Figure 1: workflow performed in the study

Key Concept

Cost Estimation

According to Samphaongoen, cost estimating is essential for budgeting and tendering in any construction project (Samphaongoen, 2010). It reflects the inherent risks, and direct costs of a project (Olatunji et al., 2010).

The unitary cost of construction works is the sum of the unit costs of resources (materials, labor and equipment), overheads and profits. This sum is calculated through the logic of price analysis and the amount of resources used in the price analysis differs for each construction work. The cost estimate is obtained by multiplying the unitary cost of construction work with corresponding product quantities. The sum of the costs for each construction work provides the total amount of the cost estimate, as shown in equation (1):

$$Cost_i = \sum_{j=1}^{n} (R_j \times Q_j) \qquad (1)$$

Where i denotes the category of construction work, j represents the index of the cost item, n is the number of cost items, R is the unit price of resources from the Price List, and Q represents the quantity for each construction work.

Cost estimation helps stakeholders determine project resources and enables better cost and technology analysis from the start of the design process. Furthermore, it is possible to simplify later phases such as construction and operation by clearly defining cost management (Bryde et al., 2013). So, anticipating economic management during the design phase provides significant advantages for project management (Plebankiewicz et al., 2015).

Generally, the cost estimate can be divided into two distinct phases. These are the physical quantities of design components (volume of concrete columns), that have substantial impacts on the outcome of the cost estimation (Monteiro & Poças Martins, 2013), and the process quantities that are related to specific construction processes (labor hours for concrete casting) (Marzouk et al., 2018). The cost estimation cannot be separated from the extraction of quantities (Aram et al., 2014), (Khosakitchalert et al., 2019). Literature shows that the most used process is to connect the code inserted in the digital objects as matching keys with various price items (Pavan et al., 2017).

The comparison between the traditional estimating process and the 3D BIM-based estimating process is visible in Figure 2 (Sacks et al., 2018).

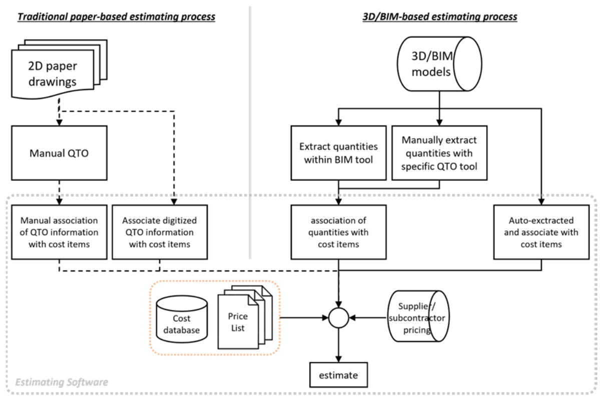


Figure 2: Conceptual diagram of an estimating process

BIM technology has had a revolutionary impact on the conventional QTO process (Smith, 2016). The use of BIM methodology automates and improves QTO processes compared to conventional 2D drawings, resulting in reduced time and errors (Smith, 2016), (Sacks et al., 2018). Nowadays some leading software in the field are Navisworks, CostX, Innovaya, iTWO, Vico, and so on (Abanda et al., 2017).

However, the accuracy of the output quantities from BIM is another major concern (Smith, 2014). The limitations of BIM data hinder a smooth QTO. The material quantities calculated from BIM authoring software rely heavily on the geometry of 3D objects and sometimes may not be sufficient for QTO (Azhar, 2011).

IfcCostSchedule and IfcCostItem

An *IfcCostSchedule* is used to consolidate instances of *IfcCostItem*, for explicit identification of cost information, for example in cost estimation of construction projects.
A Cost Schedule, like all entities in the standard, is also described through a set of attributes (Name, Description, PredefinedType, Status, etc.).
IfcCostItem is the entity that represents the cost of activities and services, the execution of works through a process, life cycle costs, cost estimates, budgets, and other financial aspects within the IFC standard. It functions as a non-geometric entity and is a subclass of *IfcControl*. The *IfcCostItem* abstract entity describes a financial cost or value, including contextual information, in a format that allows its use within a cost schedule.
IfcCostItem is also described through a set of attributes (Name, Description, CostValues, CostQuantities, etc.).
An *IfcCostItem* can link one or many *IfcCostValue's* representing a unit cost, total cost, or a unit cost with one or many quantities used to generate the total cost.

Case Study

The research, to validate the proposed methodology, focused on the implementation of a cost estimate based on the new IFC cost items and an IFC model of a structural project. This model contains 82 objects: 7 *IfcSlab*, 45 *IfcColumn*, 27 *IfcWall,* and 3 *IfcBeam* (Figure 3).

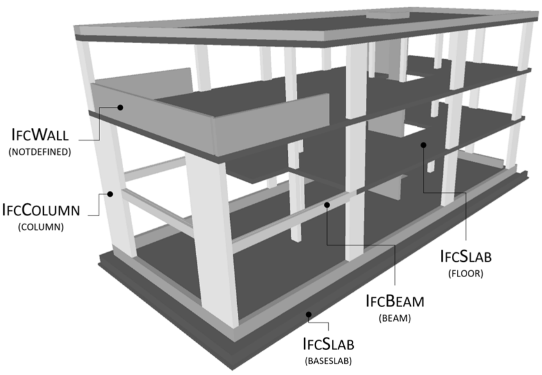


Figure 3: IFC model and the different entities contained

The developed method was implemented using the IFC4_ADD2_TC1 - 4.0.2.1, IfcOpenShell v0.7.0 and Python 3.10.
As a first step, the geometric model was interrogated to identify the IFC classes. The lists of geometric objects based on IFC classes have been identified (*IfcWall*, *IfcSlab*, *IfcColumn*, *IfcBeam*). Subsequently, these were divided into additional sub-lists according to the PredefinedType attribute in the model (for example, all *IfcSlab* were divided into BASESLAB and FLOOR). After that, each sub-list was divided into additional sub-lists according to the Object Type of entities. Finally, a query of these Object Type lists was performed to analyse the entities, verifying the LOADBEARING attribute; this allowed to verify if all objects were characterized by structural properties (LOADBEARING: True) or by non-structural properties (LOADBEARING: False). This analysis was conducted by questioning the "general" PropertySet (Pset) related to each entity, useful for a first verification of the structural model. In Figure 4 an example regarding an IfcSlab.BASESLAB (GUID 1VMFjRH1zBGg8s6Mq2oKI5) with Pset_SlabCommon and property "LOADBEARING: True".

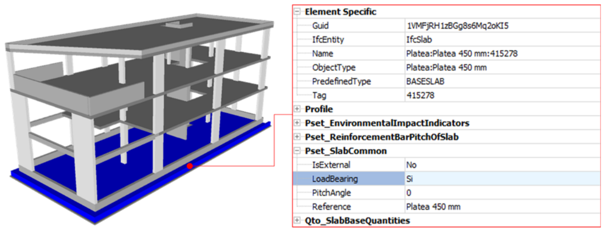


Figure 4: Example of a structural IfcSlab.BASESLAB entity

At this point, the structural objects identified previously were analyzed and through the interrogation of the individual entities, it was possible to extract the relative quantities (*IfcElementQuantity*). However as highlighted in the literature review chapter, quantification is one of the fundamental processes and cannot be based solely on the quantities contained in the 3D model. Therefore, different measurement rules are needed depending on the type of processing that needs to be economically quantified. For this reason, the code developed contains the measurement rules for the correct quantification of cost items according to the type of processing not modifiable by the user (the volume of the concrete casting, the wet surface for laying the formwork, or the weight of the steel for laying the reinforcement bars); in the specific case, these rules are in accord with how much defined in the list prices of the Lombardy Region.
At this point, the analysis performed by the code provides the first of the two manual inputs that the users have to perform. This input is the choice of the construction work (concrete casting, laying the formwork, laying the reinforcement bars, etc.) for which it is necessary to extract the quantity from the IFC model. The user does not need to define measurement rules or perform calculations as the code will perform these operations automatically based on the chosen construction work.
The formulas used by the code, in this case study, to calculate the final quantities contained in the IFC model are summarized in Table 1. These follow the rules defined in the price list of the Lombardy Region for the correct quantification of cost items.
After that, once the quantities were defined (Volume, Wet Area, Weight), the price list was queried. However, this is not a standard price list but a new cost database with cost entities based on IFC. Each of these cost items is characterized by an architecture defined and validated in another study carried out by the research group (Cassandro, Donatiello, et al., 2023). Once questioned, it is possible to identify the unit cost item (*IfcCostItem*) from which to extract the cost value (*IfcCostValue*) useful for the cost estimation and to associate with the geometric object.

Then after the selection of the unit cost item (the second and last manual input), a new cost item (*IfcCostItem*) is automatically created within the new IFC model that updates the previous geometric model and relates to the geometric object. It was possible to define "n" relationships based on how many cost items are necessary for the cost estimation of the specific workings. This new cost item contains the quantity (*IfcPhysicalQuantity*) defined previously according to the rules of measurement of the machining automatically applied by the code from the quantities extracted from the geometric objects and the unit cost value (*IfcCostValue*) extracted from the unitary cost items in the cost database (Figure 7).
Finally, the last step involves the creation of the *IfcCostSchedule* (cost sheet) in which all the cost items previously created have been collected. It will be possible to define an "intermediate" *IfcCostSchedule* to obtain the cost sheet of each set of similar construction work even if related to different objects (for example all C20/25 concrete castings for foundations, beams, slabs, and masonry) and then the final *IfcCostSchedule* corresponding to the entire cost estimate.
This will then be linked to the *IfcProject* and will allow to define the total cost estimate for the structural model.
In the end, a new IFC model was obtained which contains the *IfcCostItem* and *IfcCostSchedule* entities with the cost value and the cost quantity of the construction works for the definition of cost estimation. This new IFC file was validated through the FZK Viewer (Figure 5 and Figure 6).

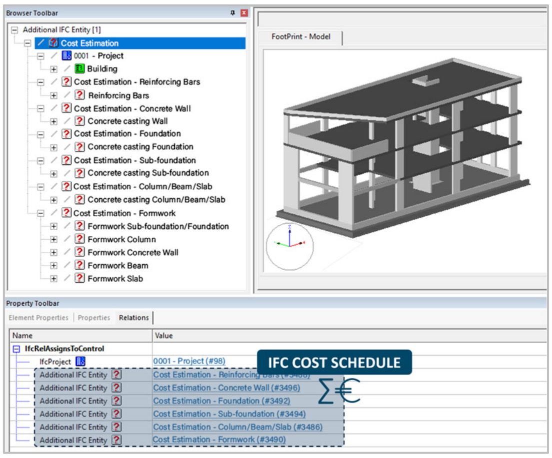


Figure 5: IFC Validation of the cost schedule on FZK Viewer

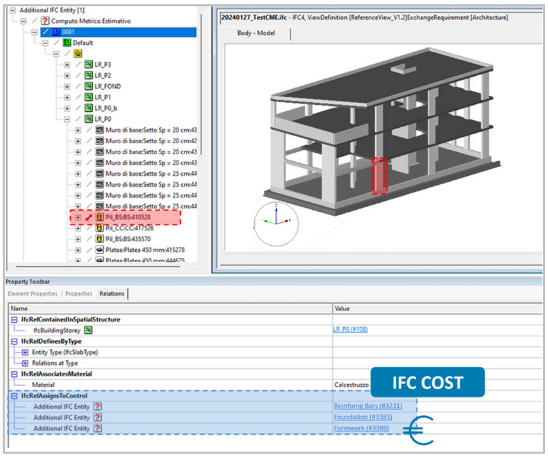


Figure 6: IFC Validation of the cost item on FZK Viewer

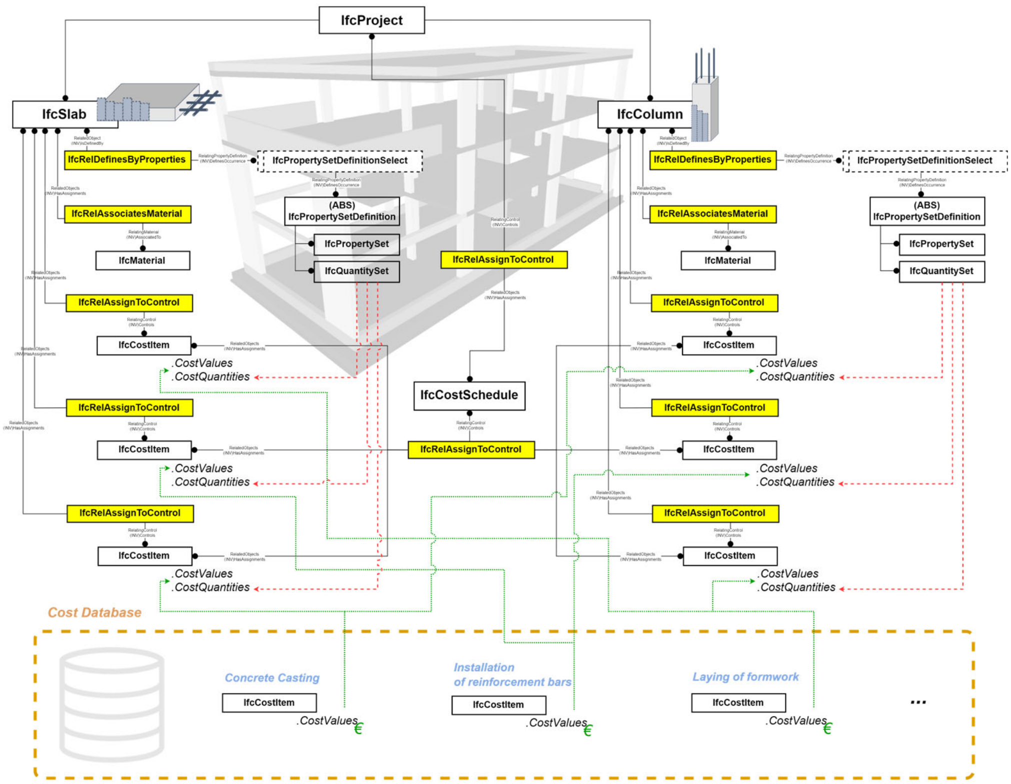


Figure 7: Example of relationships between IFC entities for estimating the costs in IFC of a slab and column

Table 1: Formulas used in the cost estimation process

Working		*Quantity*	*Formula*
Formwork	Slab (foundation)	Wet area[1] [m^2]	Perimeter * Thickness
	Column	Wet area[1] [m^2]	Perimeter * Height
	Wall	Wet area[1] [m^2]	(2*(Length + Width)) * Height
	Slab (floor)	Wet area[1] [m^2]	(Perimeter * Thickness) + (Length * Width)
Reinforcement	General	Reinforcement per cubic metre[2] [kg]	Reinforcement at m^3 * Volume
Casting	Concrete	Concrete casting [m^3]	Volume*1,1

[1] The area is the area of the single face (the height for the length of the septum).
[2] The values used for reinforcement mass per volume unit for each structural element are (Table 2):

Table 2: Theoretical values used for reinforcement mass

Element	*Reinforcement mass per unit volume*
Wall/Floor	200 *kg/m³*
Column/Beam	250 *kg/m³*
Foundation	100 *kg/m³*

Discussion

The research group conducted a study based on the results of ongoing work on a project with the Lombardy Region. The study demonstrates the feasibility of estimating costs based on IFC classes, by relating new cost items structured according to the IFC data model and their geometric objects. The study also validates the use of openBIM methods for cost estimation. This has been made possible through the implementation in IfcOpenShell of the relationship between the new cost items (*IfcCostItem)* with geometric objects (*IfcElement)*, and the creation of the cost schedule (*IfcCostSchedule)* for the identification of the final cost estimate of the entire project (*IfcProject)*.
The methodological and scientific research conducted has led to technological advancements using IfcOpenShell, achieving efficient and scalable outcomes. The proposed methodology can also be applied to estimate the costs of different projects. Currently, these results can be obtained through code, as existing tools do not support user-friendly implementations.
The studied approach, which is based on entity relationships, is effective and flexible compared to the standard approach that uses cost allocation and cost estimation through Excel spreadsheets and associating costs as simple attributes. A reliable and interoperable cost management information base is essential. This process aims to implement and encourage greater data interoperability, not only within the 3D dimension but also extending to the dimension of costs (5D BIM).
The power of the proposed methodology will ensure the possibility of developing cost estimates through procedures that allow the selection in a semi-automatic way of new cost items characterized by a standardized architecture. In addition, it is possible to query this information, verify the correctness of the associations, and relate the cost items to the object automatically. It is no longer necessary to define an attribute for each geometric entity in the geometric model to include cost information.
The cost item will be defined as a single entity related to a specific number of geometric entities within the IFC model by updating it (Figure 6). This avoids duplication of information and ensures respect for the area of knowledge of the different domains.
Furthermore, the *IfcCostItems* are grouped in the *IfcCostSchedule* to create a cost estimate. It is possible to generate different "cost views" by creating multiple cost schedules (Figure 5). This allows the extraction of complete or partial cost estimates from the same model based on specific requirements.
In addition, some problems were identified at the start of the procedure after analyzing and examining the geometric model:

- not all IFC entities had the correct structural property, despite being part of the structural model (LOADBEARING = True);
- some quantities are not stored within the *IfcElementQuantity* entity (volume, perimeter, etc.). This can lead to difficulties in creating standardized automated rules and allocating quantities to cost items. Only standard and hardcoded IFC schema data are evaluated in the research and not proprietary data that may differ between projects.

Therefore, it can be deduced that the issues with cost estimation are not only due to an incorrect cost item selection or association but also due to inaccurate information in the geometric model.

Conclusion

The cost estimation process discussed in this study is based on the IFC standard. It establishes logical connections between products, resources, cost items, and property sets in the IFC data model ensuring greater interoperability of data within the AEC sector through openBIM languages. The possibility of using cost classes (*IfcCostItem* and *IfcCostSchedule*) as independent entities, that are characterized by granular architecture and logical relationships between entities, ensures the possibility of querying, verifying, and validating this relationship to reduce cost estimation errors. It also allows for automatic cost estimation updates by modifying the IFC file of input and identifying changes to cost items or objects without associated cost entities, tracking changes and additions to the project.

During the study, limits were found in the implementation of a standardized methodology for costs. The first is the market barrier with the lack of a user-friendly tool for creating relationships, currently possible through code. A second limitation is the need to have correct information for the realization of a correct estimate of costs; not always the geometric model contains all the information useful for the realization of an estimate of costs, for example, the lack of certain dimensions that do not allow the extraction of certain quantities or the absence of fundamental parameters for the choice of cost items to be reported.

However, this approach must allow for exceptions, as it will not be possible to make all possible cases of cost estimation.

As future work, and already in progress, there is a need to prototype what is studied on a larger scale and with different and more complex geometric patterns, demonstrating how a set of typical costing strategies and economic quantification of projects can easily be implemented in the proposed approach. Develop a user-friendly application with an intuitive interface for a better definition and visualization of this data. Verify the correctness of all the cost items linked to the geometric element. Automate the choice of cost item to reduce the manual input (second manual input in this method).

Data availability

The data presented in this study are openly available on GitHub at https://github.com/Cassa97/IFC-based-Cost-Estimation.git under a Creative Commons Attribution 4.0 International License.

References

A. Kwakye. (1994). Understanding Tendering and Estimating. *Taylor & Francis Ltd*.

Abanda, F. H., Kamsu-Foguem, B., & Tah, J. H. M. (2017). BIM – New rules of measurement ontology for construction cost estimation. *Engineering Science and Technology, an International Journal*, *20*(2), 443–459.

Adeli, H., Karim, A., & York, N. (2001). Construction Scheduling, Cost Optimization and Management. In *Construction Scheduling, Cost Optimization and Management*.

Aram, S., Eastman, C., & Sacks, R. (2014). A knowledge-based framework for quantity takeoff and cost estimation in the AEC industry using BIM. *31st International Symposium on Automation and Robotics in Construction and Mining, ISARC 2014 - Proceedings*, 434–442.

Azhar, S. (2011). Building information modeling (BIM): Trends, benefits, risks, and challenges for the AEC industry. *Leadership and Management in Engineering*, *11*(3), 241–252.

Bryde, D., Broquetas, M., & Volm, J. M. (2013). The project benefits of building information modelling (BIM). *International Journal of Project Management*, *31*(7), 971–980.

Carr, R. I. (1989). Cost Estimating Principles. *Journal of Construction Engineering and Management*, *115*(4), 545–551.

Cassandro, J., Donatiello, M. G., Mirarchi, C., Zanchetta, C., & Pavan, A. (2023, July 10). *Reliability of IFC classes in ontology definition and cost estimation of public procurement*.

Cassandro, J., Mirarchi, C., Pavan, A., Donatiello, M. G., & ZANCHETTA, C. (2023). *Consistency Verification Between Cost and Geometric Information Based on IFC: Application on Structural Elements*. 801–812.

Elghaish, F., Abrishami, S., Hosseini, M. R., & Abu-Samra, S. (2020). Revolutionising cost structure for integrated project delivery: a BIM-based solution. *Engineering, Construction and Architectural Management*, *28*(4), 1214–1240.

Famiyeh, S., Amoatey, C. T., Adaku, E., & Agbenohevi, C. S. (2017). Major causes of construction time and cost overruns: A case of selected educational sector projects in Ghana. *Journal of Engineering, Design and Technology*, *15*(2), 181–198.

Froese T, Grobler F, Ritzenthaler J, Yu K, & Akinci B. (1999). Industry foundation classes for project management - a trial implementation. *Electronic Journal of Information Technology in Construction*.

Fürstenberg, D., Hjelseth, E., Laedre, O., & Wikström, L. (2021). Enabling automation of BIM-based cost estimation by semantic web technology. *Conference CIB W782021*.

ISO 16739-1:2018 - Industry Foundation Classes (IFC) for Data Sharing in the Construction and Facility Management Industries — Part 1: Data Schema (2018). https://www.iso.org/standard/70303.html

Ji, S. H., Ahn, J., Lee, H. S., & Han, K. (2019). Cost Estimation Model Using Modified Parameters for Construction Projects. *Advances in Civil Engineering*, *2019*.

Jrade, A., & Alkass, S. (2007). Computer-integrated system for estimating the costs of building projects. *Journal of Architectural Engineering*, *13*(4).

Khosakitchalert, C., Yabuki, N., & Fukuda, T. (2019). Improving the accuracy of BIM-based quantity takeoff for compound elements. *Automation in Construction*, *106*.

Lee, S. K., Kim, K. R., & Yu, J. H. (2013). BIM and ontology-based approach for building cost estimation. *Automation in Construction*, *41*.

Liao, L., Man, Q., Teo, E. A. L., Li, L., & Li, X. (2014). Improving construction schedule and cost information feedback in building information modelling. *Proceedings of Institution of Civil Engineers: Management, Procurement and Law*.

Liu, H., Lu, M., & Al-Hussein, M. (2014). BIM-based integrated framework for detailed cost estimation and schedule planning of construction projects. *31st International Symposium on Automation and Robotics in Construction and Mining, ISARC 2014 - Proceedings*, 286–294.

Lu, Q., Won, J., & Cheng, J. C. P. (2016). A financial decision making framework for construction projects based on 5D Building Information Modeling (BIM). *International Journal of Project Management*, *34*(1), 3–21.

Ma, Z., Wei, Z., & Zhang, X. (2013). Semi-automatic and specification-compliant cost estimation for tendering of building projects based on IFC data of design model. *Automation in Construction*, *30*.

Ma, Z., Zhang, X., Wu, S., Wei, Z., & Lou, Z. (2010). Framework Design for Bim-Based Construction Cost Estimating Software. *Proceedings of CIB W78: 27th International Conference*.

Marzouk, M., Azab, S., & Metawie, M. (2018). BIM-based approach for optimizing life cycle costs of sustainable buildings. *Journal of Cleaner Production*, *188*, 217–226.

Monteiro, A., & Poças Martins, J. (2013). A survey on modeling guidelines for quantity takeoff-oriented BIM-based design. *Automation in Construction*, *35*.

Olatunji, O., Sher, W., & Ogunsemi, D. (2010). The impact of building information modelling on construction cost estimation. *Engineering*.

Pavan, A., Giani, Matteo., & Mirarchi, Claudio. (2017). *BIM: metodi e strumenti : progettare, costruire e gestire nell'era digitale*. Tecniche nuove.

Plebankiewicz, E., Zima, K., & Skibniewski, M. (2015). Analysis of the First Polish BIM-Based Cost Estimation Application. *Procedia Engineering*.

Rast, J., & Peterson, K. (1999). Parametric cost estimating for environmental remediation projects. *AACE International Transactions*, *9*.

Sacks, R., Eastman, C., Lee, G., & Teicholz, P. (2018). BIM Handbook: a guide to building information modeling for owners, designers, engineers, contractors, and facility managers. In *BIM Handbook* (3rd ed). John Wiley & Sons, Inc.

Samphaongoen, P. (2010). *A Visual Approach to Construction Cost Estimating*.

Smith, P. (2014). BIM & the 5D Project Cost Manager. *Procedia - Social and Behavioral Sciences*, *119*.

Smith, P. (2016). Project Cost Management with 5D BIM. Procedia - Social and Behavioral Sciences, 226.

Staub-French, S., Fischer, M., Kunz, J., Ishii, K., & Paulson, B. (2003). A feature ontology to support construction cost estimating. Artificial Intelligence for Engineering Design, Analysis and Manufacturing: AIEDAM, 17(2), 133–154.

Tayefeh Hashemi, S., Ebadati, O. M., & Kaur, H. (2020). Cost estimation and prediction in construction projects: a systematic review on machine learning techniques. *SN Applied Sciences*, *2*(10), 1–27.

Wu, S., Wood, G., Ginige, K., & Jong, S. W. (2014). A technical review of BIM based cost estimating in UK quantity surveying practice, standards and tools. *J. Inf. Technol. Constr.*

Xu, S., Liu, K., & Tang, L. C. M. (2013). Cost estimation in building information models. *ICCREM 2013: Construction and Operation in the Context of Sustainability - Proceedings of the 2013 International Conference on Construction and Real Estate Management*, 555–566.

Xu, S., Liu, K., Tang, L. C. M., & Li, W. (2016). A framework for integrating syntax, semantics and pragmatics for computer-aided professional practice: With application of costing in construction industry. *Computers in Industry*, *83*, 28–45.

Zhiliang, M., Zhenhua, W., Wu, S., & Zhe, L. (2011). Application and extension of the IFC standard in construction cost estimating for tendering in China. *Automation in Construction*, *20*(2), 196–204.

CHATTWIN: ENABLING NATURAL LANGUAGE INTERACTIONS WITH INFRASTRUCTURE DIGITAL TWINS

Peihang Luo, Erika Parn, and Ioannis Brilakis
University of Cambridge, Cambridge, UK

Abstract

Infrastructure lifecycle management requires interactions with dynamic datasets. Traditional interfaces often hinder users' ability to rapidly locate lifecycle-specific information they need. Our work proposes ChatTwin, a system that employs Large Language Models (LLMs) to enable natural language queries related to various lifecycle stages of infrastructure, with a focus on operations and maintenance. Simulated scenarios were constructed to test the system. The results demonstrate that the system can effectively categorise interactions, fetch relevant information, and produce human-friendly outputs. With this LLM-based approach, we present an improvement in user-centricity in infrastructure lifecycle management, streamlining interactions and decision-making throughout the entire infrastructure lifecycle.

Introduction

Modern infrastructure systems are becoming increasingly complex and interconnected, posing significant challenges in their lifecycle management (Grafius et al., 2020). Decisions made throughout their lifecycles require consideration of large amounts of data from various sources. Digital twins have emerged as promising tools to overcome these challenges. They are virtual replicas of physical entities or systems, offering integrated data platforms that enable stakeholders to gain an enriched understanding of infrastructure's behaviour, to optimise maintenance and operation activities, and more (Broo et al., 2022; Rudskoy et al., 2021). Digital twins typically consist of three elements: a physical space where the actual object resides, a digital space that hosts the virtual model, and a connection that facilitates synchronisation and real-time feedback between the two spaces (Barricelli et al., 2019). However, the current human-information interaction paradigms with digital twins remain unintuitive, leading to a steep learning curve and significant investment in training. UK employers invest around £42 billion in training their employees each year, equivalent to an average spend of £2,540 per trainee (Winterbotham et al., 2020). However, even after extensive training, employees can still find it difficult and time-consuming to find the information they need, with employees spending approximately 9.3 hours weekly just sourcing information (Chui et al., 2012). Therefore, the need for more accessible, intuitive interaction methods is clear.

Background

Human-information interaction with digital twins

The interaction between users and digital twins plays a crucial role in leveraging the full potential of these complex and dynamic digital twins. Despite their potential, the current interaction modes between users and digital twins, particularly the mode of natural language interaction, remain underdeveloped. For example, Pairet et al. (2019) present a natural language interface designed for interactions with a digital twin of an offshore platform, facilitating user training across various human-robot collaboration scenarios, but without any results or implementation details. Dingli and Haddod (2019) propose a human interface of an Intelligent Digital Twin system for a semiconductor manufacturing plant, utilising three different modalities: hand gestures, a VR controller, and a voice interaction system. However, no results have been obtained for the performance of the system. Siyaev et al. (2023) introduce a neuro-symbolic reasoning mechanism to interact with a virtual aircraft maintenance digital twin constructed from structured manuals using natural language, which requires a special symbolic vocabulary and a neuro-symbolic dataset of queries to work. Furthermore, to the best of the authors' knowledge, there is no work on the natural language interactions with infrastructure digital twins.

Prompting pre-trained large language models

By enabling users to interact with computer systems using their native language, natural language interactions offer a seamless and accessible interaction experience, particularly for users with limited technical expertise or background knowledge in computing (Pazos R. et al., 2013). Recent advancements in natural language processing, especially the emergence of LLMs, have made it possible to develop more sophisticated and context-aware natural language interaction systems (Wei et al., 2022). Unlike traditional language models, which need to be fine-tuned or even re-trained on specific datasets tailored for individual tasks, LLMs have demonstrated strong in-context learning abilities. This means that they can effectively adapt to new scenarios and perform different tasks by only including a few task-specific input-output examples in the input prompt. This innovative technique is known as "few-shot prompting" (Brown et al., 2020). With few-shot prompting, LLMs can potentially be used for diverse natural language interaction tasks without the need for task-specific models or training datasets. Moreover, LLMs have also demonstrated the ability to perform tasks through "one-shot learning" or even "zero-shot learning" (Kojima et al., 2023). The former involves the LLM accomplishing tasks with just a single input-output example. The latter refers to LLMs handling specific tasks with no examples provided, relying solely on task descriptions in natural language.

While few-shot prompting, one-shot prompting, and even zero-shot prompting have demonstrated competitive performance on various NLP tasks like question-answering or translation (Brown et al., 2020), it has been observed that for complex reasoning and numerical tasks, LLMs may not always provide accurate answers. A widely adopted approach to address this challenge is "chain-of-thought" prompting. This approach provides chain-of-thought demonstrations in prompts to guide the LLM in generating a series of intermediate reasoning steps for a given task. This technique has proven to significantly improve LLM performance on tasks involving intricate reasoning or numerical skills (Wei et al., 2023).

To further enhance the performance of LLMs, some more advanced prompting techniques have also been proposed. For example, "Program-Aided Language models" (PAL) improve LLM performance in arithmetic and symbolic reasoning tasks (Gao et al., 2023). PAL employs the LLM to read and understand natural language problems, generating programs as intermediate reasoning steps based on the problem description. Unlike other methods, PAL offloads the solution step to a runtime, such as a Python interpreter. This allows the LLM to focus solely on decomposing the problem into executable steps.

Gaps in knowledge, objectives & research questions

The gaps in knowledge can be summarised as follows:

1. Domain-Specific Knowledge: LLMs lack specialised domain knowledge related to infrastructure digital twins and infrastructure lifecycle management. This includes domain-specific processes, the specialised knowledge and contextual understanding required to effectively communicate with and interpret information from digital twins, and more.
2. Contextual Information Retrieval: Integrating LLMs with digital twins requires retrieval of relevant information from the digital twin itself. LLMs must be equipped to comprehend the underlying data structure and extract meaningful information and insights from the digital twin's vast repository of information.
3. Effective Prompt Structures: The structuring of prompts plays a pivotal role in guiding LLMs to generate accurate and useful outputs. Understanding how to design prompts that effectively convey the context of a specific query is still an open question to be answered.
4. Usability and Interpretability: Ensuring that LLM-driven interactions with digital twins are not only accurate but also interpretable is critical. Gaining insights into how to post-process LLM outputs to extract valuable information in a user-friendly format is another key aspect to explore.

The primary objective of this project is to enhance human-information interaction with infrastructure digital twins by leveraging LLMs for natural language interactions, contributing to the development of more efficient, accurate, and domain-aware LLM-driven interactions with infrastructure digital twins. With this primary objective in mind, the subsequent research questions have been formulated:

1. What specific techniques can be employed to retrieve relevant information from the digital twin?
2. How can prompts be optimised to ensure that LLMs generate outputs that not only align with the specific requirements of the stakeholder but also maintain a consistent level of accuracy? Moreover, how can the retrieved data be presented to the LLMs within the prompts, ensuring that the LLM understands the broader scenario, background, or situation in which the information is embedded, leading to accurate and context-aware responses?
3. What specific procedures or tools can be employed to most effectively present data or insights to generate tangible value for infrastructure management stakeholders?

Proposed Solution

Scope

This paper focuses on investigating how LLMs can enhance interactions with infrastructure digital twins, enabling them to offer information and insights, and perform certain actions on behalf of users via natural language interactions. This paper prioritises the following five common interaction tasks, using a road digital twin as a case study. Road digital twins are comprehensive virtual models of road systems that replicate their physical counterparts in real-time. These models integrate various data sources to create a dynamic simulation of the road network. This allows for enhanced real-time monitoring, predictive maintenance, and strategic planning (Marai et al., 2021). Although our solution is demonstrated on road networks, it is designed to be adaptable to other infrastructure systems with minimal modifications.

Task 1: Data visualisation

One of the primary functions of road digital twins is integrating data from multiple sources, such as time-series data from meteorological and traffic sensors. Our first goal is to enable users to generate and view visualisations of such time-series data through natural language interactions, enabling them to gain a more intuitive view of the complex information patterns within such data.

Task 2: Information summarisation

Navigating through the vast datasets of a road digital twin to extract pertinent information is often cumbersome and time-consuming. Our system aims to simplify this by summarising key data, like road defect details, in response to user queries in natural language.

Task 3: Performing changes

A frequent and critical interaction with road digital twins is the implementation of changes to the properties of specific instances. This includes, for example, changing the status of a defect from "scheduled to fix" to "fixed" after a repair with a natural language prompt.

Task 4: Model visualisation

Accessing and visualising models, such as point clouds, is another routine task for road digital twin users. This typically involves locating specific files and using dedicated tools. We propose to simplify this process using natural language interaction, enabling users to request visualisations of model sections through simple prompts, enhancing user experience and information access efficiency.

Task 5: Work schedule enquiry

Another commonly performed task for infrastructure digital twins is generating and visualising work schedules. Our work aims to automate these tasks, traditionally a manual and laborious task, through natural language interactions. This automation will facilitate more frequent and accurate schedule adjustments and updates, improving productivity and reducing unnecessary costs associated with outdated schedules.

However, due to the data availability of the demonstrative digital twin, only summarisations of defect information are included in the scope of *Task 2*, and the scope of *Task 5* is only limited to inquiries concerning work schedules of defect rectifications. In addition, this work limits its scope of *Task 3* only to property modifications due to their regularity and natural language compatibility. More complex modifications, such as geometrical alterations or adding new instances, often require assistance from specialised equipment like laser scanners and are thus excluded.

Details

This section outlines the step-by-step process of the proposed solution that enables the user to interact with the infrastructure digital twin using natural language prompts.

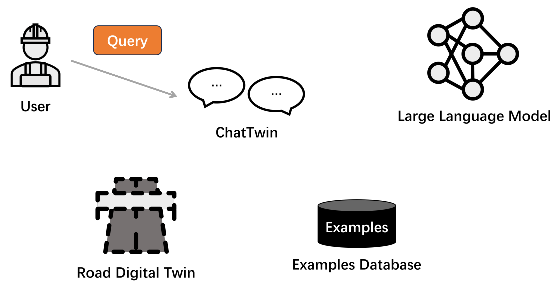


Figure 1: Step 1 - User initiating request

In step 1, as illustrated in Figure 1 above, the user initiates the interaction process by providing a natural language prompt that describes their query or request. Such prompts could either instruct the system to perform specific actions on the digital twin, or request specific data visualisations or summaries.

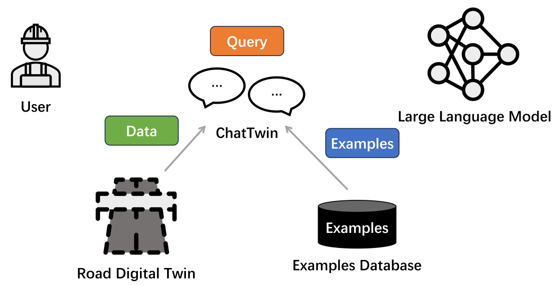


Figure 2: Step 2 - Requesting and receiving relevant data from the Road Digital Twin & Examples Database

In step 2, upon receiving the user's input, the system determines which task category this user prompt belongs to. This categorisation result will determine the subsequent actions of the system. This is done by sending a pre-structured prompt to the LLM. The structure of this prompt can be found in appendix "Prompt for categorising tasks". Depending on the task category, the system fetches relevant data from the digital twin and the examples database (as shown in Figure 2 above). Each task category dictates a unique retrieval process. Further details on the data retrieval process can be found in the "Results & findings summary" section. This aims to solve the first research question, so that relevant data can be correctly retrieved from the digital twin for each of the task categories.

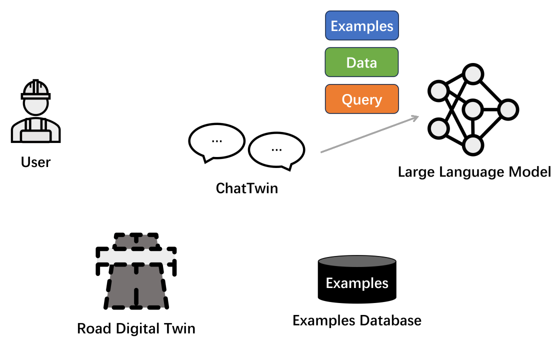


Figure 3: Step 3 - Pre-processing and sending modified prompt to LLM

In step 3 (as shown in Figure 3 above), the relevant data and examples retrieved from the digital twin and the examples database are combined with the user prompt to generate a modified prompt, which is then sent to the LLM. The modified prompt is generated by collating the user's original prompt, the relevant data from the digital twin (if required), and the relevant few-shot training examples from the examples database into a structured prompt. An illustrative example is provided in Figure 8 in the appendix. The prompt starts with one or a few few-shot training example(s) retrieved from the examples database. Each example consists of the relevant data retrieved from the digital twin (if applicable), the user prompt, and the expected task-specific output from the LLM. Additional examples are appended to the end of the previous ones. The second part of the prompt is the relevant data retrieved from the digital twin (if required), and the final part of the prompt is the user's original prompt. The modified prompt is then sent as input into the LLM. This aims to solve the second research question. The examples from the examples database provide the LLM with essential information about the data it works with. This includes but is not limited to the nature of the data, its structure, and the methods to access it. By illustrating this with sample outputs for typical user query inputs, the LLM gains a better understanding of both the data and its context, enabling it to produce more accurate results. In addition, the examples also act as guides for the LLM to ensure that its responses meet the specific needs and standards of different users and situations. This helps in maintaining the relevance and accuracy of the LLM's outputs.

One exception is that for model visualisation tasks, a distinct prompt structure is used to determine whether the requested point cloud data is available within the digital twin, and retrieve the requested point cloud data file if it is available. This prompt structure can be found in appendix "Prompt to retrieve point cloud data".

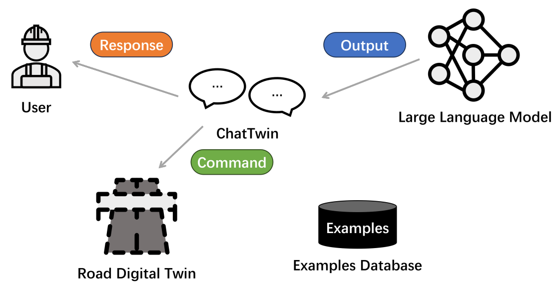


Figure 4: Step 4 - Receiving and post-processing output from LLM

In step 4, as shown in Figure 4 above, an output is received from the LLM, which contains either actionable commands for the digital twin, machine-readable information on the requested information, or a natural language summary of the information requested by the user.

The commands are generated in the form of Python scripts with embedded comments that contain chain-of-thought information. The inclusion of comments helps improve the quality and readability of the generated commands. For tasks where the system is asked to perform specific changes to the digital twin, the LLM will also generate natural language feedback to the user based on the result of the execution of the Python script. For example, if the Python script is successfully executed without errors, the LLM will then provide the user with a "success" message, confirming that the requested actions have been successfully applied to the digital twin.

Otherwise, for example, for information summarisation tasks, the summary generated by the LLM is directly sent to the user, or for model visualisation tasks, the output from the LLM is first analysed - if the output is "Not available", the system will output a message indicating that the requested data is not available within the digital twin, or if the output is the name of the requested data file, the name of the file will be passed into a pre-structured Python script to visualise the requested data. This pre-structured script can be found in appendix "Python script to visualise point cloud data".

This step aims to solve the third research question, so that the data retrieved from the digital twin can be presented to the user in an appropriate form using different tools and procedures, enabling users to quickly gain more value and insights from the data.

Research Methodology

Our work constructs a demonstrative road digital twin to evaluate the proposed interaction pipeline. This twin is populated with four data types: geometric data in point cloud format, time-series sensor data, defect information, and recommended maintenance actions.

Data collection and preparation

The geometric point cloud data, representing a specific road junction in Tallinn, is sourced from the Tallinn City Digital Twin. Time-series sensor data, simulating the input from infrastructure sensors, is obtained from the UK Environment Agency's Real Time flood-monitoring API. This API provides real-time and historical water level and flow measurement data at each of the monitoring stations across the UK. This data is indicative of the real-time sensor data one would expect in a fully operational digital twin. These sensor data have been integrated with the point cloud data, maintaining their properties such as names, IDs, and typical ranges, but changing their location-related properties to correspond with the geospatial location of the point cloud data, so that the sensors are essentially "embedded" into the point cloud at certain locations.

The defect information and suggested actions are stored in text format in a `.txt` file. The information stored within the digital twin about each registered defect includes defect type (e.g. pothole, crack, etc.), defect location (shown as a kilometre number along a specified road), severity level (either severe, moderate, or low), and suggested action based on the severity level (immediate fix for severe defects, schedule fix for moderate severity defects, and monitor progress for low severity defects).

Another component of the system is the examples database. The examples database contains few-shot training examples for each task category. However, not every example contains all three parts (relevant data, user query, and expected task-specific output from LLM). In some cases, no relevant data is retrieved from the digital twin. Instead, time-series data or geometry data are retrieved from the digital twin via Python scripts. Such an example can be found in appendix "An example for *Task 1: Data visualisation* from the examples database", where the system is asked to retrieve and visualise time-series data from a specified monitoring station.

Testing and validation

In order to test and validate the solution proposed in this work, test cases have been constructed for each type of task introduced in section "Scope". One test case has been proposed for each type of task introduced, and the proposed test cases are summarised below:

1. What is the historical water level at the station "Bourton Dickler" for the last 5 days?
2. Are there currently any defects on the road network?
3. Could you please change the severity level of the defect located at M25 28km to "severe"?
4. Could you please visualise the junction of Vabaduse väljak and Pärnu maantee for me?
5. Could you please generate a work schedule based on the current status of the road network?

Implementation details

These test cases are first categorised using a structured prompt (shown in appendix "Prompt for categorising tasks"). Then, they are processed according to their categorisation results. More details can be found in the next section.

The LLM used in this project to generate the results in the next section is the `gpt-3.5-turbo` model developed by OpenAI, and the model is accessed via the OpenAI API with the `openai` Python library.

Discussion

Results & findings summary

This section presents a comprehensive overview of the outcomes and discoveries obtained from the implementation and evaluation of our proposed natural language interaction system with infrastructure digital twin.

First, the task categorisation results are presented in Table 1 below.

Table 1: Expected categories vs actual categorisation results for the test cases

Test Case	Expected Category	Categorisation Result
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5

It can be seen from Table 1 above that the LLM is able to categorise all the test cases correctly.

Then, task-specific results are generated for each of the test cases. For test case 1, the test prompt, together with the retrieved example, are compiled according to the prompt structure introduced in step 3 in section "Details" into an input into the LLM. Then, the output from the LLM, which is a piece of Python script with chain-of-thought comments, is executed in a Python interpreter to produce the visualisation result shown below in Figure 5.

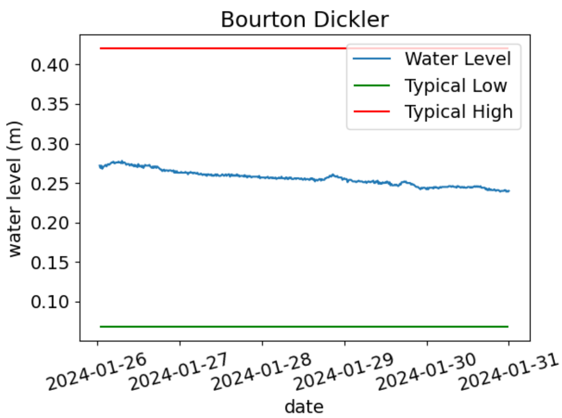


Figure 5: Visualisation result generated by the system

It can be seen from Figure 5 above that a visualisation of the requested time-series data has been successfully and correctly generated.

Then, for test case 2, the test prompt, together with the retrieved example and the defect data retrieved from the digital twin, are concatenated according to the prompt structure introduced in step 3 in section "Details" into an input into the LLM. Then, the output from the LLM is presented below:

```
Yes, I have found a total of 8 defects
```

```
on the road network. Here is a summary
of the defects:

High Severity Defects:
1. Pothole on M25 at 112km

...

3. Crack on A14 at 21km

It is suggested that you prioritize
fixing the high severity defects
immediately, schedule the moderate
severity defects for the next round of
fixing, and monitor the low severity
defects for progress.
```

The full result is available in appendix "Full result for test case 2". It can be seen that the system can retrieve the relevant data from the digital twin and generate a natural language summary. Upon cross-checking with the actual data stored within the digital twin, it was confirmed that the system retrieved the correct information. This verification process involved a comparison between the output generated by the system and the actual defect records within the digital twin. This consistency validates the reliability of the data retrieval process implemented by the system.

Next, for test case 3, the test prompt, and the retrieved example, are combined according to the prompt structure detailed in step 3 in section "Details" into an input into the LLM. The output from the LLM is then executed within a Python interpreter to apply the changes to the digital twin. The system returned a "success" message, and it was verified that the requested change had been successfully applied to the digital twin.

Then, for test case 4, the test prompt, and the point cloud file(s) retrieved from the digital twin, are collated according to the prompt structure in step 3 in section "Details" into an input into the LLM. Then, the output from the LLM, which is the name of the requested point cloud file, is passed into a pre-structured Python script (shown in appendix "Python script to visualise point cloud data") to generate the visualisation shown in Figure 6 below.

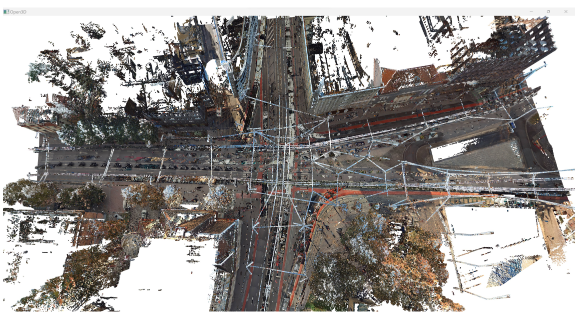


Figure 6: Model visualisation generated by the system

It can be seen from Figure 6 that the system is able to correctly retrieve and visualise the relevant point cloud data requested by the user.

Finally, for test case 5, the test prompt, together with the suggested actions retrieved from the digital twin and the relevant examples, are combined according to the prompt structure introduced in step 3 of section "Details" into an input into the LLM. The output from the LLM is then executed within a Python interpreter to generate the visualisation of the work schedule as a Gantt chart, which is shown in Figure 7 below.

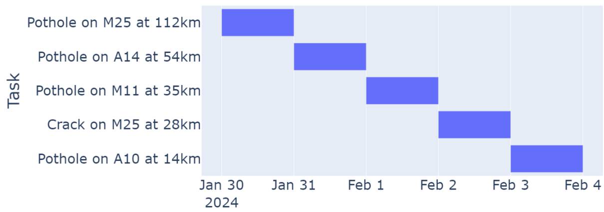


Figure 7: Work schedule visualisation generated by the system

The assumptions and criteria used here were:

1. Only "severe" and "moderate" defects are included in the schedule as the suggested action for low-severity defects is "monitor progress".
2. Prioritise "severe" defects over "moderate" defects.
3. Within "severe" and "moderate" defects, prioritise "M" roads over "A" roads.
4. Within "M" roads and "A" roads, prioritise roads with a smaller road number assigned to them.
5. Within each road, schedule the tasks according to the location of the defects, and work along each road from the defect with the smallest km number to the defect with the largest km number.
6. The time for fixing each defect is one day.

These parameters and criteria can be changed easily within the digital twin and ChatTwin to reflect the real situation. The assumptions are made purely for demonstration and to validate the results generated by the system, and do not reflect any real circumstances in the real world. It can be seen from Figure 7 above that the system is able to generate a correct work schedule according to the pre-set criteria and visualise it as a Gantt chart.

Contributions

Our work presents a novel approach to enhance the human-information interaction with infrastructure digital twins by leveraging the capabilities of LLMs, filling gaps in both the fields of digital twin research and human-computer interaction (HCI), as detailed in the following two paragraphs.

First, this work fills the gap in digital twin research by proposing a systematic approach to structure prompts that effectively leverage LLMs for generating informative and contextually relevant outputs with infrastructure digital twins. This enables the LLMs to perform domain-specific tasks without the need for extensive fine-tuning. Our work contributes to the growing field of prompting pre-trained LLMs, providing some structured templates for structuring prompts that effectively guide the responses of LLMs within the context of infrastructure digital twins. This work also presents an empirical evaluation of our LLM-based interaction system through scenario-based simulations, demonstrating its strengths, limitations, and potential areas of improvement, offering valuable insights for researchers and practitioners interested in advancing natural language interaction with digital twins in the domain of infrastructure management.

In addition, current infrastructure digital twin research often emphasises technical aspects. This work primarily enhances the interaction layer, bridging the gap between intricate domain-specific data in digital twins and straightforward, user-friendly interactions, adding a user-centred design perspective, and aligning the infrastructure management domain with foundational HCI principles by facilitating more intuitive communication with digital twin systems, further enhancing the capabilities of conventional digital twin systems. This paves the way for stakeholders to more effectively engage with digital twin systems for infrastructure monitoring, control, and decision-making.

Conclusions

The integration of LLMs with infrastructure digital twins has far-reaching implications for the infrastructure management domain and beyond. By providing a more accessible, intuitive and user-friendly interface, LLM-driven interactions reduce the learning curve required to interact with intricate infrastructure digital twin systems, empowering decision-makers at different levels, from engineers to policymakers, with timely and contextually relevant insights. This improved decision-making capability, supported by real-time data analysis and predictions provided by the capabilities of digital twins, can lead to more informed choices regarding infrastructure management, maintenance, and resource allocation.

Recommendations for future

While our research provides valuable insights, some limitations remain. This section lists some potential areas for further investigation.

Enhancing interaction feedback

While the current system can successfully generate and execute Python scripts for requested actions on the digital twin, future improvements can focus on providing more comprehensive feedback to users throughout the process to enhance user engagement and make the interface more informative. This can involve monitoring the digital twin's status during and post-execution, enabling the system to confirm the successful implementation of user requests or flag potential issues, thus ensuring the accuracy of performed actions. Users might receive a visual cue for successful operations, or a notification detailing any encountered challenges. This could establish a closed feedback loop and eliminate ambiguity about action outcomes. Potential research questions could include: How can real-time feedback during task execution enhance user trust and system transparency for users of different expertise levels? What impact does visual or textual feedback have on user comprehension and satisfaction?

Expanding task categories

While this project addresses the five most common tasks associated with infrastructure digital twins, there exists substantial potential to broaden the range of tasks covered by this approach. Expanding beyond these fundamental tasks can further improve system applicability. Furthermore, with the ReAct (Yao et al., 2023) approach, there is also potential to develop a more generally applicable natural language interaction system. Such a system could leverage the advanced reasoning capabilities of LLMs to autonomously deduce steps for various user-initiated actions beyond pre-defined task categories.

Multi-turn interactions and task clarification

The current focus of the project on single-turn interactions provides a solid foundation for future research exploring more complex multi-turn conversations. This expansion will enable the system to engage in more intricate and dynamic interactions with users, improving the adaptability of the system. This extension is particularly crucial for addressing scenarios that involve ambiguity and might require additional clarification. In such cases, the system can proactively seek clarification from users when faced with potential misunderstandings to help disambiguate potentially confusing queries. Moreover, multi-turn interactions could also be leveraged to handle situations where users require assistance in accurately conveying their intent. With multi-turn interactions, the system's ability to interact, interpret, and execute user intents can be significantly enhanced, ultimately contributing to a more effective and user-centric infrastructure management system. Potential research questions could include: How do multi-turn interactions influence user experience and system accuracy in complex query scenarios? What methodologies can effectively disambiguate user queries?

User-centric design evaluation

Future enhancements can be rigorously user-tested in simulated and controlled real-world scenarios to validate their effectiveness and user-centricity. This approach will ensure that proposed improvements align with actual user needs and preferences, thereby enhancing the practical applicability of our findings.

Evaluating performances with different LLMs

OpenAI's `gpt-3.5-turbo` model has been used in this work. In the future, tests can be conducted on other LLMs (e.g. PaLM, BERT, and Llama) to evaluate and compare the effectiveness of the system with different LLMs.

Acknowledgements

The authors gratefully acknowledge the EPSRC, Bentley Systems and the University of Cambridge for funding this research through the EPSRC Centre for Doctoral Training in Future Infrastructure and Built Environment: Resilience in a Changing World (EPSRC grant reference number EP/S02302X/1); and Christopher-Robin Raitviir and Andres Maremäe from Tallinn City Council for generously providing data from the Tallinn City Digital Twin. This work also uses Environment Agency flood and river level data from the real-time data API (Beta). For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission.

References

Barricelli, B. R., Casiraghi, E., and Fogli, D. (2019). A survey on digital twin: Definitions, characteristics, applications, and design implications. IEEE Access, 7:167653–167671.

Broo, D. G., Bravo-Haro, M., and Schooling, J. (2022). Design and implementation of a smart infrastructure digital twin. Automation in construction, 136:104171.

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., and Amodei, D. (2020). Language models are few-shot learners.

Chui, M., Manyika, J., Bughin, J., Dobbs, R., Roxburgh, C., Sarrazin, H., Sands, G., and Westergren, M. (2012). The social economy: Unlocking value and productivity through social technologies | McKinsey. `https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-social-economy`.

Dingli, A. and Haddod, F. (2019). Interacting with intelligent digital twins. In Marcus, A. and Wang, W., editors, Design, User Experience, and Usability. User Experience in Advanced Technological Environments, pages 3–15, Cham. Springer International Publishing.

Gao, L., Madaan, A., Zhou, S., Alon, U., Liu, P., Yang, Y., Callan, J., and Neubig, G. (2023). PAL: Program-aided language models.

Grafius, D. R., Varga, L., and Jude, S. (2020). Infrastructure interdependencies: Opportunities from complexity. Journal of Infrastructure Systems, 26(4):04020036.

Kojima, T., Gu, S. S., Reid, M., Matsuo, Y., and Iwasawa, Y. (2023). Large language models are zero-shot reasoners.

Marai, O. E., Taleb, T., and Song, J. (2021). Roads infrastructure digital twin: A step toward smarter cities realization. IEEE Network, 35(2):136–143.

Pairet, È., Ardón, P., Liu, X., Lopes, J., Hastie, H., and Lohan, K. S. (2019). A digital twin for human-robot interaction. In 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pages 372–372.

Pazos R., R. A., González B., J. J., Aguirre L., M. A., Martínez F., J. A., and Fraire H., H. J. (2013). Natural Language Interfaces to Databases: An Analysis of the State of the Art, pages 463–480. Springer Berlin Heidelberg, Berlin, Heidelberg.

Rudskoy, A., Ilin, I., and Prokhorov, A. (2021). Digital twins in the intelligent transport systems. Transportation Research Procedia, 54:927–935. International Scientific Siberian Transport Forum - TransSiberia 2020.

Siyaev, A., Valiev, D., and Jo, G.-S. (2023). Interaction with industrial digital twin using neuro-symbolic reasoning. Sensors, 23(3).

Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E. H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., and Fedus, W. (2022). Emergent abilities of large language models.

Wei, J., Wang, X., Schuurmans, D., Bosma, M., Ichter, B., Xia, F., Chi, E., Le, Q., and Zhou, D. (2023). Chain-of-thought prompting elicits reasoning in large language models.

Winterbotham, M., Kik, G., Selner, S., Menys, R., Stroud, S., and Whittaker, S. (2020). Employer skills survey 2019: Training and workforce development. `https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/936487/ESS_2019_Training_and_Workforce_Development_Report_Nov20.pdf`.

Yao, S., Zhao, J., Yu, D., Du, N., Shafran, I., Narasimhan, K., and Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models.

Appendices

Prompt for categorising tasks

```
Please determine which task category this user prompt belongs to: [user prompt]

-----

Candidate task categories:

Task category 1: Data visualisation
Task description: Tasks where the user asks the system to visualise specific
time-series data
Example user prompt: What is the water level data at the station "Cam" over the past 2
days?

Task category 2: Information summarisation
Task description: Tasks where the user asks the system to find and summarise certain
data or information
Example user prompt: Are there currently any defects on the road network?

Task category 3: Performing changes to digital twin
Task description: Tasks where the user asks the system to change the property value of
a specific instance within the digital twin
Example user prompt: Can you change the status of the defect at M2 112km to fixed?

Task category 4: Model visualisation
Task description: Tasks where the user asks the system to visualise a certain section
of the digital twin model
Example user prompt: Can you show me a visualisation of Regent Street between 5km and
8km?

Task category 5: Work schedule enquiry
Task description: Tasks where the user asks the system to generate/visualise a work
schedule
Example user prompt: Could you generate a work schedule for defect rectification?

-----

Please only respond with the task category number in Arabic number.
```

Prompt to retrieve point cloud data

```
The available point cloud data files are:
[the available point cloud data retrieved from the digital twin]
The user's query is: [user query]
Is the point cloud data requested by the user available? If it is available, please
only respond with the name of the corresponding point cloud data file. If it is not
available, please respond with "Not available".
```

Prompt structure

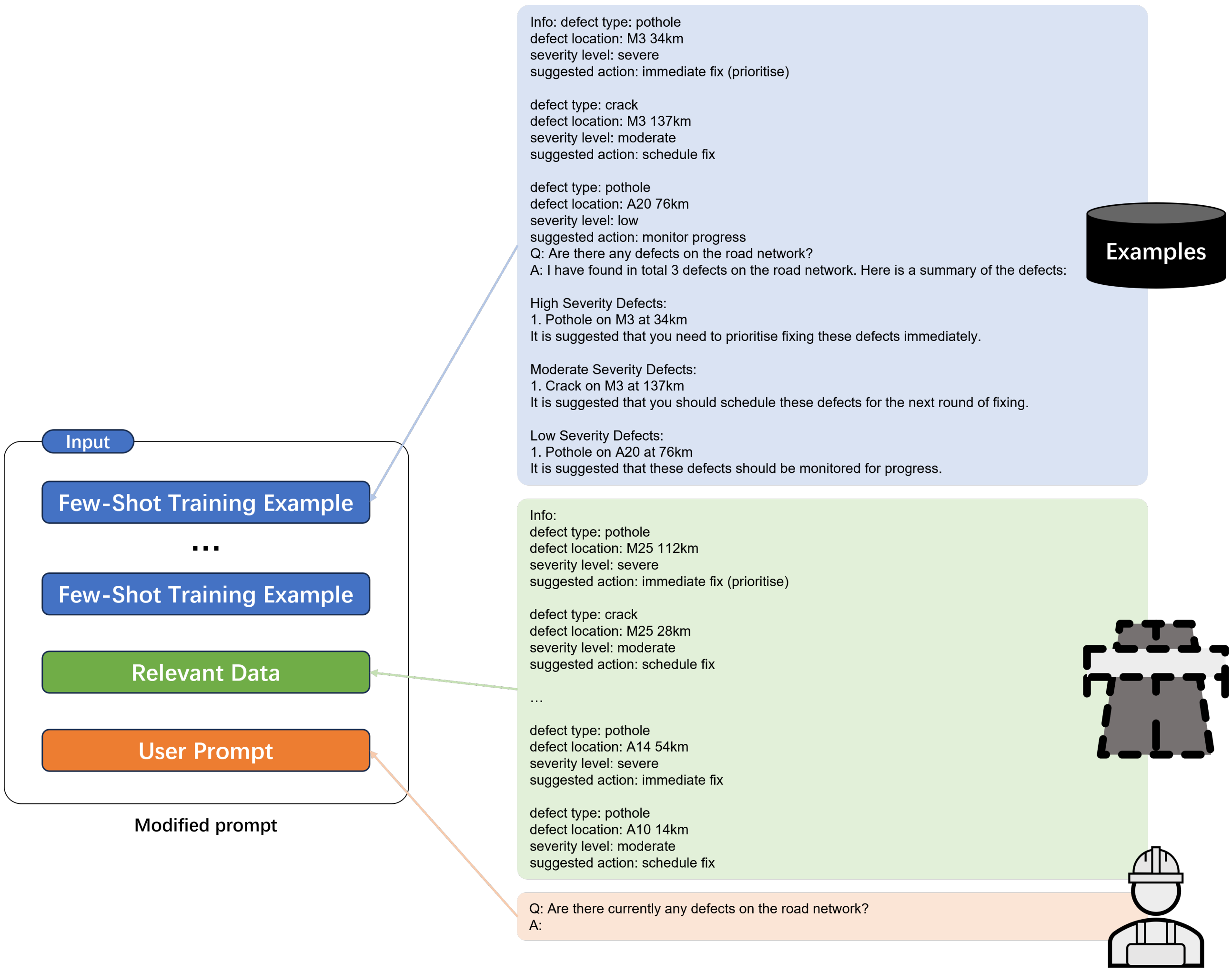


Figure 8: Prompt structure

Python script to visualise point cloud data

```python
import open3d as o3d

point_cloud_file = response["choices"][0]["message"]["content"]
top_dir = "[path to the top directory of point cloud data files]/{}"
path = top_dir.format(point_cloud_file)
pcd = o3d.io.read_point_cloud(path)
o3d.visualization.draw_geometries([pcd],
                                  zoom=0.4012,
                                  front=[0.0902, 0.2108, 0.9734],
                                  lookat=[130.1986, 20.4945, -13.5578],
                                  up=[-0.1271, -0.9669, 0.2211])
```

An example for *Task 1: Data visualisation* from the examples database

Q: What is the water level data at the station "Cam" over the past 2 days?
A:
```
# To get water level data, we first need to find the list of monitoring stations using
the function build_station_list()
stations = build_station_list()
# Then we need to find station "Cam" from the list and store it in the variable
station_cam
station_name = "Cam"
for station in stations:
  if station.name == station_name:
    station_cam = station
    break
# We then need to fetch the water level data over the past 2 days using the function
fetch_measure_levels
dt = 2
dates, levels = fetch_measure_levels(station_cam.measure_id,
dt=datetime.timedelta(days=dt))
# We can then finally plot the water level data using the function plot_water_levels
plot_water_levels(station, dates, levels)
```

Full result for test case 2

```
Yes, I have found a total of 8 defects on the road network. Here is a summary of the
defects:

High Severity Defects:
1. Pothole on M25 at 112km
2. Pothole on A14 at 54km

Moderate Severity Defects:
1. Crack on M25 at 28km
2. Pothole on M11 at 35km
3. Pothole on A10 at 14km

Low Severity Defects:
1. Pothole on M11 at 87km
2. Pothole on M11 at 134km
3. Crack on A14 at 21km

It is suggested that you prioritize fixing the high severity defects immediately,
schedule the moderate severity defects for the next round of fixing, and monitor the
low severity defects for progress.
```

RECOMMENDATIONS FOR PRIVATE AND SECURE CDE BASED BIM COLLABORATION

Anja Brelih and Robert Klinc
University of Ljubljana, Faculty of Civil and Geodetic Engineering

Abstract

The digitalization of the construction industry through BIM and CDE raises security and privacy concerns as sensitive data is stored in cloud environments. It is critical to address these concerns related to CDE and provide guidelines and recommendations for users. The paper provides an overview of security and privacy mechanisms at multiple levels that should be implemented in BIM, including data and network security, identity and access management, physical security, and compliance with standards. We examine ISO 19650-5 and recommend a direction to improve BIM related guidelines in order to help users identify potential concerns in their chosen cloud environment.

Introduction

A Common Data Environment (CDE) is a cloud computing infrastructure that serves as a shared digital workspace for project teams to collaborate on Building Information Modelling (BIM). CDE is becoming increasingly important in the construction industry as it provides a centralised repository for managing project information and improving communication between stakeholders (see Figure 1).

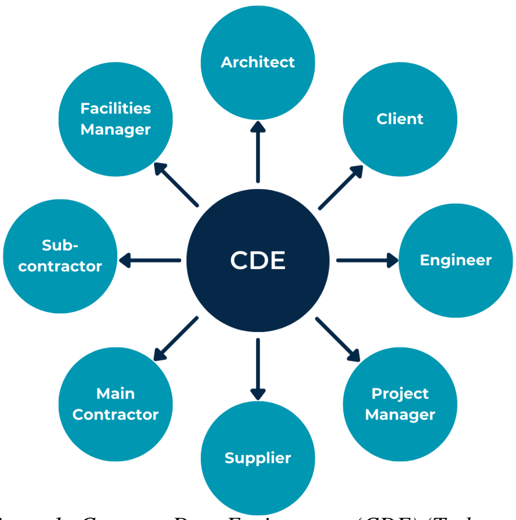


Figure 1: Common Data Environment (CDE) (Turk et al., 2022).

CDEs allow project teams to access and share BIM models and other project data in real time. This can help to improve coordination, reduce errors and shorten project timelines.

Cloud computing is a ubiquitous and rapidly growing computing paradigm that provides users with access to data and applications. Two main models are private and public clouds. In the private cloud, ownership and management of the cloud is in the hands of the organisation providing the applications and access to the resources is not open to everyone as in the public cloud. Cloud computing offers different service models, which are a specific, pre-prepared combination of information technology (IT) resources offered by a cloud service provider and relate to the way in which it offers a service to users. The three main service models are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) (Manvi and Shyam, 2021). Shared digital workspaces used in the construction industry are typically offered as a SaaS model, where users do not manage and control the cloud infrastructure, but simply utilise cloud software solutions. In commercial CDE environments, providers are responsible for managing the SaaS model and providing upgrades, maintenance and security.

The exponential growth of cloud computing brings numerous benefits, but also a number of security concerns that must be properly understood and addressed in order to successfully deploy solutions in the cloud environment (Sun, 2018). Computing and data in the cloud are associated with various security risks, including loss of governance, isolation failures, data protection, service availability, compliance and legal risks, authentication and authorisation, etc. Cloud computing security refers to maintaining the confidentiality, integrity and availability of data stored in the cloud. Cloud security requirements include robust security, trust, safety, monitoring and governance (Pavithra et al., 2019). These requirements can be directly applied to the security and privacy requirements for CDE environments in the cloud.

This paper provides recommendations for secure and private CDE based BIM collaboration environments. The second chapter provides an overview of the levels of security and privacy mechanisms that need to be implemented in these environments. The third chapter provides an overview of the ISO 19650-5 standard for Security-minded approach to information management. The fourth chapter gives an overview of the BIM Execution Plan with guidelines for users. Chapter five proposes guidelines for users of BIM and CDE environments. The last two chapters provide a discussion and a conclusion.

Cloud Computing Security and Privacy Mechanisms

As the use of BIM and CDE environments increases, so does the number of security risks and threats that need to be addressed and successfully mitigated. These mechanisms are crucial to ensure the aspects of data security, i.e. confidentiality, integrity and availability of data in the cloud, and provide protection against unauthorised access, data loss and cybersecurity threats (Manvi and Shyam, 2021).

The security and privacy mechanisms of cloud computing must be considered on multiple levels in order to provide users with a comprehensive and trustworthy service. In this paper, the mechanisms are categorised into data security, network security, identity and access management, physical security and standards compliance (see Figure 2).

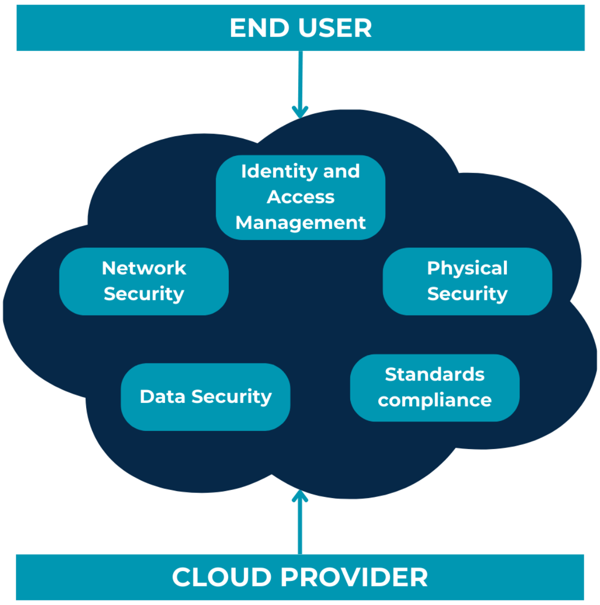


Figure 2: Cloud computing security levels.

Data security

Data security in the cloud differs from traditional data security due to its special storage architecture. Cloud computing generally uses distributed storage. Ownership and control of the data are separated. Users have ownership, but control lies with the cloud service providers. Current data protection in the cloud mainly focuses on data storage, transmission and access authentication and is mostly implemented using cryptography (Xu et al., 2019). By default, data is written in a human-readable format known as plain text, which is vulnerable to unauthorised and potentially malicious access when transmitted over a network. Cloud encryption converts the data into a unreadable format to ensure confidentiality and even data integrity. It is a service offered by cloud storage providers that uses encryption algorithms (Manvi and Shyam, 2021) to convert data into a protected and unreadable form, rendering the encrypted data unusable without knowledge of the key. The encryption process can be carried out using various techniques such as symmetric and asymmetric encryption. In symmetric cryptography, the same key is used for encryption and decryption, as opposed to asymmetric cryptography, where a pair of public and private keys is used for encryption and decryption. An important aspect of data security during transmission or remote storage is maintaining the integrity of the data. To this end, hash functions can be used to ensure that the data received is identical to the original.

Network security

Network security includes the protection of data transmitted over networks such as the internet, the protection of systems and data from network attacks and the protection of the network components themselves. A cloud service provider can ensure the security of the network environment for users by offering cloud services that ensure encryption of data traffic, network monitoring, analysis and control of data traffic, virtual private networks (VPN), firewalls and secure network services (CSA, 2012).

Identity and access management

Cloud computing has had a significant impact on identity and access management (IAM). In both public and private clouds, two parties must work together to manage IAM without compromising security. Cloud computing requires significant changes to the methods used to manage IAM for internal systems. The most important difference is the relationship between the cloud provider and the cloud user. IAM cannot be managed by just one or the other, so a mutually agreed understanding of responsibilities and a secure relationship for managing identity, authorisation and access (CSA, 2021a).

Physical security

Physical security mechanisms include various levels of control and measures to control access and activities. The cloud service provider, who has control over the physical infrastructure, is responsible for ensuring this level of security. It must ensure security against physical attacks as well as natural disasters and power outages. Every aspect of the data center, from location and accessibility to power density and redundancy, must be designed to ensure its security, resilience and efficiency. However, incidents can still occur, whether through a natural disaster, a physical attack or a cyber-attack. It is therefore important that cloud providers are prepared for such situations and have disaster recovery mechanisms in place to ensure business continuity.

Shared responsibility model

The Cloud Security Alliance (CSA, 2021b) identifies the so-called shared responsibility model, which is directly related to two recommendations for security in cloud applications:

1. Cloud providers should clearly document their inter-

nal security controls and characteristics of customers so that the cloud user can make an informed decision. Providers should also properly design and implement these controls.

2. Cloud users should create a responsibility matrix for each individual project to document who implements which controls and how.

The shared responsibility model states that when dealing with cloud computing technology, there are two parties involved who are responsible for implementing and managing different parts of the stack. Building and maintaining trust is critical to this relationship.
Cerić et al. (2021) provide a detailed analysis of the existing literature on trust in megaprojects. The authors identify three different approaches to the study of trust. Namely the psychological, the sociological and the economic. The psychological approach focuses on the dynamics of trust dynamics, while the sociological approach examines trust within groups and organisations. The economic approach, on the other hand, analyses how trust influences economic interactions. The key factors that contribute to trust in such projects are effective communication, transparency, reputation and strong relationships.
Trust is an important aspect of construction projects and it also applies to the collaboration of all parties using or providing BIM and CDE cloud environments. A trusting relationship is important for the successful implementation of these solutions so that the parties involved regulate and respect the activities of the shared responsibility model.
The basis for establishing and maintaining trust in BIM and CDE cloud computing environments between the parties involved in the projects must be clearly defined in the standards and project documentation. The ISO 19650 series of standards focuses on the management of information throughout the lifecycle of a built asset using BIM (BSI, 2023). The clear instructions for managing the security and privacy of the information should also be part of the project documentation, e.g. the BIM Execution Plan.

Standards compliance

In addition to the ISO 19650 series of standards that apply to information management using building information modelling, CDE environments must also comply with the standards that apply to cloud computing. Cloud computing has attracted increasing attention due to security concerns and standardised validation and certification is required to assess cloud security. IT security standards are a structured approach to IT security based on measurable indicators represented by controls (e.g. a checklist) or general but clear requirements (e.g. clauses or principles) Di Giulio et al. (2017). The ISO/IEC 27000 family is the most widely used standard for information security management systems (ISO, 2023). It contains guidelines for the establishment, implementation, maintenance and continuous improvement of an information security management system. The ISO/IEC 27001 standard provides instructions for the establishment, implementation, maintenance, and continuous improvement of an information security management system. Compliance with the ISO/IEC 27001 standard means that the organization or company has established a system to manage risks associated with the security of data held or processed by the company, and that this system complies with all the best practices and principles described in this international standard (ISO, 2022a). In addition to compliance with the ISO/IEC 27001 standard, consideration should also be given to compliance with the ISO/IEC 27017 standard, additional controls with implementation guidance that specifically relate to cloud services (ISO, 2015). It is intended to complement the recommendations of the ISO/IEC 27002 (ISO, 2022b) standard and various other standards in the ISO/IEC 27000 family, such as ISO/IEC 27018 on the privacy implications of cloud computing (ISO, 2019) and ISO/IEC 27031 on business continuity (ISO, 2011). The ISO/IEC 27018 standard focuses on the protection of personally identifiable information (PII) in public clouds that serve as PII processors (ISO, 2022a). Compliance with these standards demonstrates that an organisation has implemented a robust security management system and is essential for ensuring the security of data and information in CDE environments.

ISO 19650-5 standard

The fifth part of the ISO 19650 *"Organization and digitization of information about buildings and civil engineering works, including building information modelling - Information management using building information modelling"* series of standards refers to *"Security-minded approach to information management"*. This part of the standard outlines the requirements for the management of information security in BIM projects (ISO, 2020).

ISO 19650-5 provides a framework to help organisations assess security vulnerabilities in BIM and implement controls to mitigate risk. It recognises that the use of BIM and information management technologies introduces new vulnerabilities that need to be proactively addressed. The standard promotes a risk-based approach to security that can be applied across the organisation. It is designed to help organisations protect sensitive information and ensure the security and resilience of assets, products and services in the built environment (ISO, 2020).

The main part of the standard is divided into six parts, which establish the need for a security-minded approach and provide guidance on developing a security strategy, a management plan with a plan for dealing with security breaches and outline guidance for working with appointed parties. In addition, the standard also contains annexes with further information on (1) the security context, (2) the types of personnel, physical and technical security controls and the management of information security, (3) the assessment in relation to the provision of information to third parties and (4) information sharing agreements.

BIM Execution Plan

The BIM Execution Plan (BEP) is a comprehensive document that helps project stakeholders move forward with clear roles and expectations. It is an essential element that must be created before a construction project begins, and it is a powerful tool for project ownership that drives work through the various design and construction phases. Information that should be included in the BEP includes (1) how the data in the BIM files will be created, managed, documented and shared, (2) elements such as agreed roles and responsibilities within the BIM process, (3) a strategy for key deliverables and a guide to key project milestones, and (4) practical working procedure details. The BEP is a guiding document that helps the different team members to identify and execute the various phases of the project. It can help to present a clear plan with goals and objectives for each step (Ramage, 2022).

Ramage (2022) identifies seven elements of a good BEP: (1) clearly defined roles and responsibilities of each team and organisation, (2) strategic planning, definition of BIM scope and defined key deliverables, (3) project milestones and a realistic timeline, (4) project objectives, (5) model quality control procedures, (6) project reference information with key project contacts, and (7) working procedures that include file naming conventions, construction tolerance expectations, the project approach to annotation, technology infrastructure needs (including hardware and software used), BIM iteration management and data transfer management.

In the third version of the BIM Project Execution Planning Guide (Messner et al., 2021), the authors offer a structured procedure for the creation and implementation of the BEP. The five steps within the procedure include:

1. defining the goals for the implementation of BIM,
2. identifying high value model uses during the project planning, design, construction and operational phases,
3. designing the BIM execution process through the creation of process maps
4. defining the information deliverables, and
5. the development of infrastructure in the form of contracts, communication procedures, technology and quality control to support implementation.

These steps describe the general application of BEP in a project. While steps 1 to 4 mainly contain guidelines for BIM implementation and project-specific organisation, step 5 refers to the technology used. It is essential that this step includes the definition of the required level of security and privacy required, the technologies used and the responsibilities. The requirements for the technological aspect must also be the subject of the contract between the provider and the customer.

The BEP also provides guidelines for defining the infrastructure needs, including hardware, software, networks and modelling content for the project that will be used for BIM. It identifies fourteen specific categories to support the BIM project execution process.

1. **BIM Project Execution Plan Overiview**: review of the plan based on the categories developed after analysing the documents listed below, reviewing current execution plans, discussing the issues with industry experts and revising through a comprehensive review by various industry organisations.
2. **Project Information**: contains basic project information that can be useful for current and future projects. It can be used to introduce new members to the project and help others reviewing the plan to understand the project.
3. **Key Project Contacts**: contains the contacts of the owner, the planners, the consultants, the main contractors, the subcontractors, the manufacturers and the suppliers of the project.
4. **Project Goals / BM Uses**: the plan should include a clear list of BIM goals, the BIM Use Analysis Worksheet and specific information on the selected BIM Uses selected.
5. **Organisational roles / Staffing**: for each selected BIM Use, the team must specify which organisation(s) will staff and perform this Use. This includes the number of staff by job title required to perform the BIM Use, the estimated working hours, the main location where the use will be performed and the lead organisational contact for this Use.
6. **BIM Process Design**: the plan should include the overview map of BIM Uses, a detailed map of each BIM Use and a description of the elements on each map.
7. **BIM Information Exchanges**: the team should document the information exchange created as part of the planning process in the BIM Project Execution Plan. The information exchange illustrates the model elements by discipline, level of detail and any specific attributes that are important to the project.
8. **BIM and Facility Data Requirements**: project owners can have very specific BIM requirements. It is important that the plan documents the BIM requirements in the format specified by the owner.
9. **Collaboration Procedures**: the team must develop its procedures for electronic and active collaboration. This includes the management of models and standard meeting actions and agendas.

10. **Quality Control**: procedures must be defined and implemented to ensure model quality at each project stage and prior to information exchange. Each BIM created during the life cycle of the project must be pre-planned taking into account the model content, the level of detail, the format and the party responsible for the updates and distribution of the model and data to various parties.
11. **Technological Infrastructure Needs**: the team should determine the requirements for hardware, software platforms, software licences, networks and modelling content for the project.
12. **Model Structure**: the team must determine the methods that will ensure the accuracy and scope of the model. Once the planning team has agreed on the collaboration methods and technology infrastructure needs, it should reach a consensus on how the model will be created, organised, communicated and controlled.
13. **Project Deliverables**: the project team should consider what deliverables are required by the project owner. Deliverables should take into account the project phase, the format of the due date and any other specific information about the deliverable.
14. **Delivery Strategy / Contract**: When implementing BIM in a project, attention should be paid to the delivery and contract methods before the project begins.

Guidelines for users of BIM and CDE environments

As BIM and CDE environments contain sensitive information about construction projects, especially financial information, intellectual property and personal data. It is therefore important for providers and users of these environments to take measures to protect the security and privacy of this data.

Users should be aware that BIM and CDE environments are based on new cloud technologies. It is important to verify the provider's compliance with industry standards. By demanding compliance with the standards, users can help to ensure that providers adhere to the rules and standards and clearly communicate these to their users.

Users need clear guidelines on how to use cloud solutions correctly in their projects. Based on current standards and guidelines, users should pay particular attention to the following list that can help protect the security and privacy of sensitive data in BIM and CDE environments:

- **Make a conscious decision about the data you share**. Only share the data that is necessary for the project. Ask questions about the security of the environment snd consult security experts.
- **Assign a data security officer for the project**. This person should be responsible for monitoring the security of the data.
- **Research the provider's security and privacy mechanisms**. Before deciding on a cloud-based CDE, find out about the provider's security practises. Look for providers that are certified to industry standards.
- **Request documentation**. Ask for documentation from the provider outlining security practises.
- **Negotiate the terms**. Include the security requirements and consequences in your contract with the provider.

Discussion

ISO 19650-5 is a valuable standard that provides a framework for information security in information management with BIM. However, the standard could be improved by focusing more on cybersecurity, providing more clarity and consistency, and offering more detailed guidance on how to implement security measures. In particular, the standard should address cybersecurity risks associated with cloud-based technologies, provide clearer and more specific guidance, and offer more specific recommendations for policies, procedures and tools. By improving the standard, organisations using BIM and CDE and providers can better protect sensitive information.

The BIM Execution Plan outlines an important procedure for BIM implementation in construction projects. However, it only focuses on the specifics of construction projects and does not take into account the information security and privacy required when using CDE cloud environments. An important aspect that is ignored is the use of suitable providers that can guarantee an appropriate level of security and meet the required standards for cloud environments.

The shared responsibility model used in cloud environments should be included in the BEP, as it clearly defines which aspects both the provider and the user must guarantee and take into account. It would also be necessary to include in the BEP specific cloud computing technological aspects that are necessary to ensure security and privacy. Experts in the construction field do not focus on information security aspects. Therefore, collaboration with relevant experts in the field of information security and privacy is crucial, as construction projects involve a large amount of often sensitive information.

Conclusion

The use of BIM and CDE environments in construction projects is becoming increasingly common and serves as an important tool for more efficient project management and resource management. With the use of these cloud technologies, it is also of paramount importance to ensure robust security and privacy mechanisms to ensure the confidentiality, integrity and availability of data.

In this paper, we have explored the key aspects of ensuring security and privacy in cloud environments where the user does not have full control over their data in public clouds.

A review of the ISO 19650-5 standard, which focuses on the security-minded approach to information management, has shown that it does not provide adequate guidelines for users and providers of CDE environments to follow when using them. Even the BEP, which is otherwise a powerful tool for planning the specifics of a construction project, does not provide security guidelines.
For safe implementation of BIM and CDE cloud environments in construction projects, it is critical that all stakeholders are aware of the dangers that can result from inadequate security controls. A shared responsibility model should be included in the BEP guidelines, clearly stating which aspects of security must be ensured by both the provider and the user. Users must be aware of the security threats and select a suitable provider of a CDE environment also on the basis of security aspects. In the guidelines for users of CDE and BIM environments, we provide advice that can help users choose the right provider for their security requirements.
A transparent and confidential relationship between the provider and the user of BIM and CDE environments is central to the implementation of these environments. Providers must provide secure solutions, but there is an incentive for users to ensure that appropriate standards are maintained and security aspects are properly disclosed to users so that the services can be trusted.

Acknowledgments

This research was supported by the Slovenian Research Agency under the Young Researcher funding program and research program E-construction (E-Gradbeništvo: P2-0210).

References

BSI (2023). ISO 19650 Building Information Modelling (BIM). https://www.bsigroup.com/en-GB/iso-19650-BIM/ [Accessed: 2023-11-21].

Cerić, A., Vukomanović, M., Ivić, I., and Kolarić, S. (2021). Trust in megaprojects: A comprehensive literature review of research trends. International Journal of Project Management, 39(4):325–338.

CSA (2012). SecaaS Category 10 // Network Security Implementation Guidance. https://cloudsecurityalliance.org/artifacts/secaas-category-10-network-security-implementation-guidance/ [Accessed: 2023-11-21].

CSA (2021a). Security Guidance for Cloud Computing. https://cloudsecurityalliance.org/research/guidance/ [Accessed: 2023-11-20].

CSA (2021b). Security Guidance for Cloud Computing. https://cloudsecurityalliance.org/research/guidance/ [Accessed: 2023-11-20].

Di Giulio, C., Sprabery, R., Kamhoua, C., Kwiat, K., Campbell, R. H., and Bashir, M. N. (2017). Cloud standards in comparison: Are new security frameworks improving cloud security? In 2017 IEEE 10th International Conference on Cloud Computing (CLOUD), pages 50–57.

ISO (2011). ISO/IEC 27031:2011 information technology, security techniques, guidelines for information and communication technology readiness for business continuity. https://www.iso.org/standard/44374.html [Accessed: 2024-03-24].

ISO (2015). ISO/IEC 27017:2015 information technology, security techniques, code of practice for information security controls based on iso/iec 27002 for cloud services. https://www.iso.org/standard/43757.html [Accessed: 2024-03-24].

ISO (2019). ISO/IEC 27018:2019 information technology, security techniques, code of practice for protection of personally identifiable information (pii) in public clouds acting as pii processors. https://www.iso.org/standard/76559.html [Accessed: 2024-03-24].

ISO (2020). ISO 19650-5:2020. https://www.iso.org/standard/74206.html [Accessed: 2023-12-06].

ISO (2022a). ISO/IEC 27001:2022 information security, cybersecurity and privacy protection, information security management systems, requirements. https://www.iso.org/standard/27001 [Accessed: 2024-03-24].

ISO (2022b). ISO/IEC 27002:2022 information security, cybersecurity and privacy protection, information security controls. https://www.iso.org/standard/75652.html [Accessed: 2024-03-24].

ISO (2023). ISO/IEC 27000 family. https://www.iso.org/standard/iso-iec-27000-family/ [Accessed: 2023-12-04].

Manvi, S. S. and Shyam, G. K. (2021). Cloud computing: concepts and technologies. CRC Press, Taylor Francis, 1st ed. edition.

Messner, J., Anumba, C., Dubler, C., Goodman, S., Kasprzak, C., Kreider, R., Leicht, R., Saluja, C., Zikic, N., and Bhawani, S. (2021). BIM Project Execution Planning Guide, volume 3.0. Computer Integrated Construction Program, Penn State.

Pavithra, S., Ramya, S., and Prathibha, S. (2019). A survey on cloud security issues and blockchain. In 2019 3rd International Conference on Computing and Communications Technologies (ICCCT), pages 136–140.

Ramage, M. (2022). Trimble Contruction: What is a BIM Execution Plan and what should it include? https://constructible.trimble.com/construction-industry/what-is-a-bim-execution-plan-and-what-should-it-include [Accessed: 2024-01-09].

Sun, X. (2018). Critical security issues in cloud computing: A survey. In 2018 IEEE 4th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), pages 216–221.

Turk, Ž., Sonkor, M. S., and Klinc, R. (2022). Cybersecurity assessment of BIM/CDE design environment using cyber assessment framework. Journal of Civil Engineering and Management, 28(5):349–364.

Xu, H., Cao, J., Zhang, J., Gong, L., and Gu, Z. (2019). A survey: Cloud data security based on blockchain technology. In 2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC), pages 618–624.

THE INTEGRATION OF EARLY STAGE LIFE CYCLE ANALYSIS INTO ARCHITECTURAL PRACTICE THROUGH AN IT SUPPORTED PROCESS

Catherine Sinclair[1,2], Zaid Alwan[2]
[1]Space Architects, Newcastle Upon Tyne, UK
[2]Northumbria University, Newcastle Upon Tyne, UK

Abstract

Within the construction industry, a variety of tools undertake Life Cycle Analysis (LCA), including OneClick LCA, Tally and others recently developed. Early design decisions provide the biggest opportunity to reduce the whole life carbon (WLC) of a design. A mixed methodology approach including software appraisals, questionnaires and literature reviews found that current LCA tools do not encourage a free flowing design process and cannot analyse early stage decisions for their carbon impact. Integrating BIM and LCA at early design stages, reviewing decisions, using modern computing techniques and automated analysis is essential to addressing and reducing impacts through the construction phases.

Introduction

There is a growing base of evidence to suggest the impact of climate change is having an adverse effect on the natural and built environment. One example of addressing this is the UN adopting 17 sustainable development goals which were adopted by all United Nation Member States. These goals stated that ending poverty and other deprivations must work alongside improving health and education, reducing inequality and spurring economic growth, whilst tackling the threat of climate change (United Nations, 2015). Within the UK, the built environment is responsible for 25% of greenhouse gas emissions produced (Environmental Audit Committee, 2022). If the UK is to reach their Net Zero by 2050 goal, these figures need to be significantly reduced. Worldwide, the construction industry is currently responsible for 39% of energy related carbon emissions with 28% being from operational carbon and 11% being from embodied carbon (World Green Building Council, 2022). The construction industry has a vital part to play in addressing the climate emergency. Furthermore, the focus of WLC reduction within the construction industry is shifting from focussing on operational carbon reduction to reducing embodied carbon, this is due to the major improvements in the operational energy efficiency of the construction sector (Victoria and Perera, 2018). Operational carbon emissions are also more regulated than embodied carbon emissions with policies such as Approved Document L in the UK, which regulates operational carbon emissions. Currently there are no Government policies within the UK that require the assessment or control of embodied carbon emissions (Building to net zero: costing carbon in construction: Government Response to the Committee's First Report, 2022). Voluntary targets have been set in the UK by construction professionals for the reduction of embodied and operational carbon, this includes targets set by the Low Energy Transformation Initiative (LETI) and the Royal Institute of British Architects. LETI is a UK based network of built environment professionals that are collaboratively working towards a zero carbon future for both the planet and the UK (leti, n.d.).

The RIBA has developed the 2030 Climate Challenge, this sets a series of targets for Architectural practices to meet with the aim to reduce embodied and operational carbon (RIBA, 2019). LETI have also published a series of guidance documents to be used by construction professionals in order to reduce the WLC of the construction industry. Architects have a key role in reducing emissions within the construction industry, with consideration and analysis across all stages of the RIBA Plan of Work. Within the design process the key design decisions that affect both operational and embodied emissions cannot currently be automatically analysed and there is no automation within the design process. The software currently available, such as Autodesk Revit and Autocad, act as an essential tool within the design process but do not encourage the reduction of emissions. Low carbon design is essential within the construction industry in order to address the climate crisis.

The objective of this paper will be to evaluate the current processes and software used within the architectural design process and use for the calculation and reduction of WLC based on a large architectural practice in the north east of England. It will also be reviewed how embodied carbon is currently calculated within architectural practice and whether it is currently integrated within the process. Consideration will be taken regarding the perception Architects and designers have regarding embodied and operational carbon analysis across the RIBA plan of work. This will lead to discussions on how the process can be automated using IT in order to be integrated into architectural practice. In turn this will help draw conclusions on how to reduce embodied carbon within the construction industry through automation in design and integration of analysis.

The use of Building Information Modelling (BIM) within architectural practice is still developing. There are still barriers to the implementation of BIM in practice, these include lack of awareness, lack of training, training expenses and social resistance to change (Ahmed, 2018). Research has shown that the construction industry has been slow in adopting advancements and technological innovations that have enhanced workflow and productivity within other industries (Ojo and Pye, 2020). There are a variety of tools that are available that both support the utilisation of BIM and tools that incorporate LCA. These tools have not been widely implemented in practice or across the RIBA plan of work, creating a fragmented process. The RIBA plan of work organises the process of briefing, designing, constructing, maintain and operating into 8 stages (RIBA, 2024). The tools need to be seamlessly integrated into the RIBA plan of work allowing use across the lifecycle of a project. The current fragmented approach has not encouraged use of the tools during the design process where major design decision affecting embodied carbon are made.

Literature Review

Through studying relevant literature, from both academic and industry sources. The aim was to find out information regarding current software, current LCA tools, importance of early stage analysis, automation within design and the ability to interlink BIM and LCA.

Current Software and analysis tools

It is imperative that computational tools are utilised within architectural practice, both throughout the design process and for LCA. Within the construction industry there are a variety of tools that can conduct LCA of a design. These include but are not limited to Tally, One Click LCA, Eccolab, HBERT and FCBS Carbon. From this list, One Click LCA and Tally are the only tools that can be fully integrated into the BIM environment (Al-Ghamdi and Bilec, 2017). At the early design stages, these tools are not able to be utilised as there is not enough information within the design to undertake a LCA. The majority of tools that have been developed are designed for either post-construction evaluations or evaluations at the detail design stage, these tools are then used as part of scientific research as opposed to as a design tool (Meex et al., 2018).

LCA are currently a data intensive process, requiring a high level of information and computing ability by the user to calculate the embodied carbon of a design. LCA are currently used to give feedback as opposed to being able to be used as a tool to improve the design (Röck et al., 2018). Currently LCA's have not been widely adopted into architectural practice due to the complexity of analysis, the tools are very time consuming to use and there are challenges in gathering the required data (Kamari, Kotula and Schultz, 2022).

There is a high demand for a simplified approach to LCA that Architects can use within practice, and integrate into their current workflow. This will have to have an efficient and simple approach that does not require the detailed knowledge of LCA (Hollberg and Ruth, 2016). In order for LCA to be used and accepted into practice, the tools must be seamlessly integrated into the workflow and also become embedded in the culture of the practice (Means and Guggemos, 2015).

The Royal Institute of Chartered Surveyors (RICS) have released whole life carbon assessment (WLCA) for the built environment. This document will provide guidance and become the standard for consistent and accurate carbon measurement in the built environment (RICS, 2023). Previously, there was a need to standardise embodied carbon measurement as there was a huge variation in embodied carbon figures reported, which was attributed to the variability of the assumptions made in the measurements (Victoria and Perera, 2018). The results of assessments are still dependent on a series of factors such as system boundaries which are specified in BS EN 15978. The boundary choices can make a significant difference in embodied energy calculations (Dixit et al., 2010). Another variable could include choice of material library or data sources, within the UK that could include but is not limited to The Inventory of Carbon and Energy (ICE) and the Built Environment Carbon Database (BECD).

Building performance Assessment and automation within design

Worldwide there are a variety of building performance assessments including BREEAM, LEED, DGNB, Estidama and more. The schemes have some key differences which include the weighting that they give to different environmental categories. The weighting systems take into account climate and culture (BSRIA, 2011). These systems assist in creating goals and benchmarks and assist comparing alternative methods, however these methods are rarely incorporated at the early stages of design development (Häkkinen et al., 2015).

Currently there is no continuity throughout the RIBA stages in order to assess the WLC of a building (O'Sullivan et al., 2004). As it is a fragmented approach, current tools and systems have been described as too time consuming and complicated to effectively integrate into the design process. It was stated that it is important to develop the usability of the tools, integrating methods for WLCA into the design process, utilising a semi-automated tool (Häkkinen and Belloni, 2011).

As IT use within architectural practices is still growing, especially the use of BIM, design automation needs to be utilised. Design automation applications can show the environmental impact schemes can have, providing the opportunity to evaluate early design decisions against the WLC of a project. This can be done using a set of rules or parameters that evaluate the design (Gerth et al., 2016). Automation can also be utilised in order to reuse solutions that have been used on other projects, removing some of the routine work from the workflow of a project, allowing the designer to focus on other areas of the project (Gerth et al., 2016).

Digital tools provide the opportunity to decrease the additional work that is required to accurately undertake a life cycle assessment. Workflows have been developed to connect an LCA database such as the Inventory of Carbon and Energy (ICE) to a BIM software. For example there are cases where Dynamo for Autodesk Revit has been utilised (Hollberg, Genova and Habert, 2020). BIM and LCA have the potential to overcome the time consuming nature of undertaking a LCA, which will in turn facilitate the integration of environmental analysis, obtaining quick results about the impact of a building from an early design stage (Kamari, Kotula and Schultz, 2022). Soust-Verdauger et al. (2017) highlight significant challenges in the interoperability among the various applications, software integration of BIM and LCA and the amount of data that is required for LCA. It was also discussed that BIM-LCA integration will only be useful if the results and amount of data provided are user friendly (Soust-Verdaguer, Llatas and García-Martínez, 2017). These tools are able to be used when there is a high LOD in the design.

Early Stage Analysis

The majority of Architectural practices currently rely on in-house knowledge to make design decisions regarding WLC due to lack of access to specialist knowledge or simple design tools (Meex et al., 2018). Meex et al. (2018) also stated that in the earliest design stages Architects define the WLC of a building over its whole life cycle, often unknowingly, due to the lack of availability of analytical tools. Dunant et al stated that good early stage design decisions can halve embodied carbon (Dunant et al., 2021). To achieve this, it has been stated that designers have to have a greater understanding of which material and dimensioning decisions most significantly determine a building's environmental impact and which decisions are less important (Basbagill et al., 2013). It has been discussed that the building design process does not have the tools to analyse and consider WLC in the early stages of design. The most fundamental decisions which will affect the WLC of a building are conducted at an early stage. WLC reduction is not limited to material selection at an early stage design stage. It is widely discussed that a better form factor can make a huge difference in the reduction of embodied and operational carbon (make, 2021). Other considerations that should be analysed at an early design stage include window to wall ratio, orientation and structural grid spacings (Dunant et al., 2021). Basbagill et al. (2013) stated that considering shape and orientation during the early design stages, it is possible to reduce a designs environmental impact by 40%. According to Dunant et al. (2021), the typical building frame could have 40-60% less embodied carbon if a simple structural grid was chosen over a complex frame.

Window to wall ratio has a great impact on the operational energy of a building. Higher window to wall ratio is expected to be associated with increased heating energy due to more conductive loss through larger windows, but decreased lighting energy because bigger windows let in more daylight (Troup et al., 2019). An optimal window to wall ratio needs to be considered, this is one that minimises the sum of the energy use for heating, cooling and lighting (Goia, 2016). The optimum figure for different building typologies needs to be taken into account at early design stages. A building's orientation combined with its glazing ratio is key to minimising energy demand (LETI, 2020). LETI conducted a study that proved difference orientations of a building can almost double the space heating demand of a building (LETI, 2020).

Currently analysis is usually carried out after the detailed design stage has taken place. This is due to earlier stages of the BIM model having a low Level of Development (LOD), where generic building elements are used and accurate whole life carbon analysis cannot take place (Röck et al., 2018). Whilst analysis at a later stage gives more accurate data with the higher LOD, the ability to influence change is significantly lower. LCA should be able to take place at the earliest BIM LOD (LOD 100) in order to have the highest impact throughout the design. Changes to the developed design following later stage analysis will have larger cost implications than at an earlier stage, and the ability to influence change is lower (Gervásio et al., 2014). Gervásio et al. (2014) suggest an approach that involves a set of pre-defined construction components which incorporate embodied carbon impacts. This will provide the designer with the ability to compare options. Similar analysis is available within the H\B:ERT tool, an open source Revit tool that allows the designer to analyse building components (Hawkins Brown n.d.).

Methodology and Approach

Questionnaire

In order to gather data from a wide range of participants across a variety of stages of their careers in the field of Architecture, a questionnaire was undertaken. This included but was not limited to Architectural Technologists, Architects and Architectural Directors. Being able to gain information on the level of the participants career pinpointed the gaps in knowledge and skills within architectural practice. The questionnaire was systematically constructed following the process of the RIBA plan of work. It was structured to review strategies for assessing carbon impacts and gaps in knowledge/software across the RIBA plan of work. The questionnaire discusses levels of knowledge of WLC, analysis conducted within practice and software used at different RIBA stages.

Review of Current Software and workflow

There has been a variety of LCA tools, utilising different carbon databases. In order to gain an understanding of the opportunities and constraints of existing tools and software, a range of LCA tools will be reviewed. It will be considered how they will be able to be implemented into practice and into the workflow of a project.

Through the sources mentioned above, a workflow within practice will be identified. The workflow will be analysed and areas will be identified that would benefit from automation and LCA using IT within the design process. This will lead to a process or framework being developed that will assist in the reduction of embodied and operational carbon over the all stages of the RIBA plan of work, utilising appropriate tools.

Results

Questionnaire

The results of the questionnaire covered a wide spread of people across various stages of their career.

The results provided an insight of when the design process progressed from 2D on paper to 3D utilising Computer Aided Design (CAD). CAD was used from RIBA Stage 2, Concept Design. The main software used according to the participants was Autodesk Revit, which is a software that supports BIM management (breakwithanarchitect, 2018).

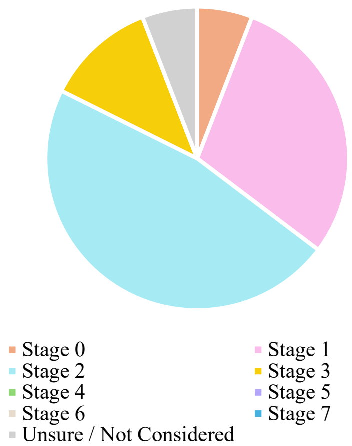


Figure 1: At what RIBA stage do you begin to use Revit

50% of participants thought that the key decisions regarding embodied carbon in a project were made at RIBA Stage 3 (Spatial Coordination). 22% stated that they considered the key decisions to be made at RIBA Stage 4 (Detailed Design).

Only one participant had conducted LCA on a project, with 78% never having conducted or outsourced an assessment. The one participant that had conducted LCA stated that they completed it at RIBA Stage 4 as it was a requirement of the project brief.

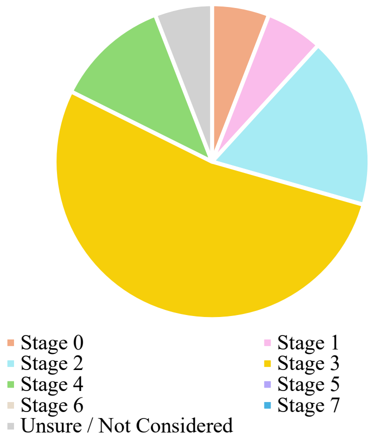


Figure 2: At what stage do you think the key decisions regarding embodied carbon are currently made on our projects

All of the participants stated that more guidance is needed within architectural practices to reduce embodied carbon at early design stages and throughout the project.

The most common response for lack of integration of embodied carbon analysis is the lack of knowledge and time. Alongside this there were comments about the lack of suitable software in order to automate the process. It was discussed that the more automated the process of analysis, the more likely it will be to be integrated into the design process. It was also commented that the main tools that are currently used do not produce live feedback and make reviewing design changes possible. Furthermore they are more mathematical as opposed to using design data, with a data heavy interface, making them unappealing to the design based users. They do not provide results that are easy to understand and are not visual.

Review of Current Software and workflow

A review of a range of tools and software was undertaken. The tools chosen are primarily aimed for use within Architectural practice, reducing the WLC of a project.

The first tool that was reviewed was OneClick LCA, this tool has the ability to extract information from a Revit model. This tool can only be utilised if the design is fully developed and the materials have been correctly set. Given the detail required, this will be used at RIBA Stage 4. It is also reliant on the Revit model being accurately and correctly modelled. Once the analysis has been conducted on the OneClick LCA web platform, it is necessary to manually check the analysis to confirm accuracy, this is a very time consuming process. OneClick LCA can compare a variety of EPD's, and states that it has the worlds largest construction LCA database (One Click LCA® software, n.d.).

FCBS Carbon is an excel based tool and requires all information to be manually input. The tool uses data from the Inventory of Carbon and Energy (ICE) and product EPD's (portal.fcbstudios.com, n.d.). The tool can be utilised at an earlier stage of design, but quantities and material choices are still required.

Within Architectural Practice, Computer Aided Design (CAD) is utilised from RIBA Stage 2. No evidence was gathered showing that embodied carbon was analysed at this stage, or further through the project. This was supported by the literature that the sector is resistance to new innovation technologies. Work by Alwan and Ilhan Jones, 2022 claimed that within the industry, one of the main barriers is the resistance to change from the traditional working practices alongside the time required to adapt to new innovative technologies.

Discussion

The results show on the whole that there is a requirement for both early stage LCA analysis and also simpler and more accessible design tools to be put in place to implement this throughout architectural practice. This is an important development to pursue, with the goal to reduce the environmental impact of the construction industry. This study has reinforced the need for development within architectural practice focussing on linking LCA and BIM. The literature has shown that not possible to conduct an accurate LCA at early design stages. Even though the current tools available cannot accurately measure WLC at early design stages, decisions can be made that affect the WLC of a project. Current literature shows that early stage design decisions can significantly reduce the WLC of a project. These early stage design decisions need to be analysed in order to set the design on the right trajectory to lower WLC. The comments received in the questionnaire also stated that more guidance is needed both with regard to design decisions and material choices.

The results of the questionnaire suggest that the key decisions regarding WLC are made at RIBA Stage 3. Existing literature suggests that these decisions are in fact made at RIBA Stage 2. As both sources state that the most important decisions regarding WLC are made in the early stages of the design process, it is evident that further tools and guidance is needed within the industry to assist in decision making. The early stage design decisions need to be analysed in order to provide guidance in low carbon design. The analysis needs to be undertaken as an automated process so that it can run hand in hand with the standard design process. This will encourage the new system to be widely adopted, especially if the process is neither complex or time consuming. Having design analysis from RIBA Stage 2 will assist in the reduction of WLC, before material selection takes place. The results discuss that key early design decisions effect the WLC of a project before material selection takes place.

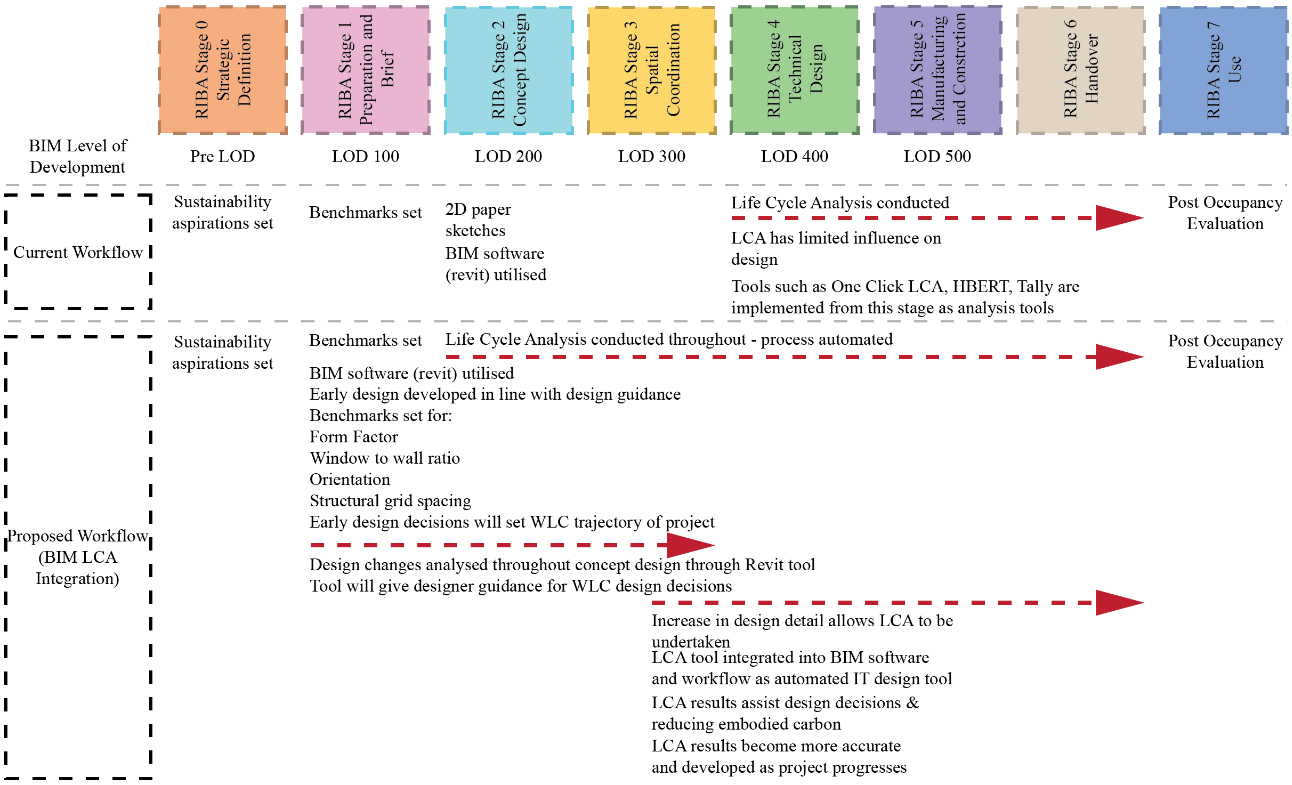


Figure 3: Existing and proposed workflow

With Revit being used from RIBA Stage 2, automation of analysis using IT tools within the design process would be able to be implemented from an early stage. Alongside this, using computer modelling from concept design stage should allow embodied carbon analysis to take place throughout the design process and through the RIBA work stages. Analysing the project from early stages will allow the design process to be influenced by constantly reviewing the information. Integrating analysis through an automated process at an early stage will assist in this decision making, increasing the ability to reduce embodied carbon from an early stage. With the industry being resistant to change from the traditional process, the tool will have to be simple to use and analyse the results, as well as not being a time consuming process.

Following analysis of both the questionnaire and existing software available it is obvious that there is a gap within Architectural Practice regarding automated LCA within the project workflow. The tools that are currently available either require a higher skill level than is available in practice or are time consuming to undertake. This is evident in the fact that only one participant had made use of a LCA tool within their projects. Furthermore, as the analysis was only conducted at RIBA Stage 4, we have to consider whether this had an impact on the final design, or whether the tool was used to only report the impact.

Conclusion

Analysis has confirmed that there is a gap in WLC design and analysis during the concept design stage or RIBA Stage 2. The tools currently available can analyse projects with a high LOD and the tools require a high skill level and are time consuming. Creating an IT based tool that is able to be utilised at an early concept design stage where we have the ability to have the biggest impact on the WLC of a project is necessary.

After reviewing relevant literature it is apparent that it is not possible to accurately undertake LCA at early stages of a design. The literature did provide an insight into the fact that early stage design decisions can have a large impact into the WLC of a project. These early stage decisions need to be analysed, this will provide guidance to the designer that the decisions made will set the project on the right trajectory to reduce the WLC. As the survey states, key decisions regarding WLC are made at RIBA 2 & 3, as this is where the major design decisions are made, setting the trajectory of the project.

A tool needs to be developed that can analyse early stage design decisions and provide the designer with feedback about the decisions they make. With the lack of detail at early stages, properties analysed will include form factor, window to wall ratio, structural grid and orientation. In line with the tool, a series of benchmarks need to be developed so that the designer understands the decisions that are being made. Following the analysis, it is not possible to conduct an accurate LCA at early design stages.

Following the feedback, this tool will have to be automated and take information from the model and provide the designer with a graphical representation of any carbon hotspots and room for improvement. The tool will have to be digitally integrated into the workflow of a construction project, giving design guidance throughout the process. This research shows that there is a need for future tools and gives the foundations for a future tool to be developed. The analytical process conducted within architectural practice shows the need for a design based tool that utilises IT tools that are integrated into the design and construction process.

In conclusion it is apparent that there is a gap in early stage analysis. Following this initial research a tool will be developed that can be utilised from RIBA Stage 2. The tool will assess early stage design decisions such as form factor, window to wall ratio, structural grid and orientation automatically through a Revit based plugin, assisting the designer in making decisions that will affect the WLC of a project before being able to conduct LCA.

This research is being developed as part of a two year Knowledge Transfer Partnership funded by Innovate UK and UK Research and Innovation. This research will continue to be developed as part of the Knowledge Transfer Partnership.

References

Ahmed, S. (2018). Barriers to Implementation of Building Information Modeling (BIM) to the Construction Industry: A Review. *Journal of Civil Engineering and Construction*, 7(2), p.107. doi:https://doi.org/10.32732/jcec.2018.7.2.107.

Al-Ghamdi, S.G. and Bilec, M.M. (2017). Green Building Rating Systems and Whole-Building Life Cycle Assessment: Comparative Study of the Existing Assessment Tools. *Journal of Architectural Engineering*, 23(1), p.04016015. doi:https://doi.org/10.1061/(asce)ae.1943-5568.0000222.

Alwan, Z. and Ilhan Jones, B. (2022). IFC-based embodied carbon benchmarking for early design analysis. *Automation in Construction*, 142, p.104505. doi:https://doi.org/10.1016/j.autcon.2022.104505.

Basbagill, J., Flager, F., Lepech, M. and Fischer, M. (2013). Application of life-cycle assessment to early stage building design for reduced embodied environmental impacts. *Building and Environment*, 60, pp.81–92. doi:https://doi.org/10.1016/j.buildenv.2012.11.009.

breakwithanarchitect (2018). *The difference between BIM and Revit*. [online] breakwithanarchitect. Available at: https://www.breakwithanarchitect.com/post/the-difference-between-bim-and-revit [Accessed 23 Jan. 2024].

Bsria.com. (2011). *Review of global environmental assessment methods*. [online] Available at: https://www.bsria.com/uk/news/article/review-of-global-environmental-assessment-methods/ [Accessed 23 Jan. 2024].

Building to net zero: costing carbon in construction: Government Response to the Committee’s First Report. (2022). [online] House of Commons Environmental Audit Committee. Available at: https://committees.parliament.uk/publications/30124/documents/174271/default/ [Accessed 24 Jan. 2024].

Dixit, M.K., Fernández-Solís, J.L., Lavy, S. and Culp, C.H. (2010). Identification of parameters for embodied energy measurement: A literature review. *Energy and Buildings*, [online] 42(8), pp.1238–1247. doi:https://doi.org/10.1016/j.enbuild.2010.02.016.

Dunant, C.F., Drewniok, M.P., Orr, J.J. and Allwood, J.M. (2021). Good early stage design decisions can halve embodied CO2 and lower structural frames’ cost. *Structures*, 33, pp.343–354. doi:https://doi.org/10.1016/j.istruc.2021.04.033.

Gauch, H.L., Hawkins, W., Ibell, T., Allwood, J.M. and Dunant, C.F. (2022). Carbon vs. cost option mapping: A tool for improving early-stage design decisions. *Automation in Construction*, 136, p.104178. doi:https://doi.org/10.1016/j.autcon.2022.104178.

Gerth, R., Sandberg, M., Lu, W. and Jansson, G. (n.d.). Design automation in construction: An overview. In: *Proceedings of the 33rd CIB W78 Conference*. Proceedings of the 33rd CIB W78 Conference.

Gervásio, H., Santos, P., Martins, R. and Simões da Silva, L. (2014). A macro-component approach for the assessment of building sustainability in early stages of design. *Building and Environment*, 73, pp.256–270. doi:https://doi.org/10.1016/j.buildenv.2013.12.015.

Häkkinen, T. and Belloni, K. (2011). Barriers and drivers for sustainable building. *Building Research & Information*, 39(3), pp.239–255. doi:https://doi.org/10.1080/09613218.2011.561948.

Häkkinen, T., Kuittinen, M., Ruuska, A. and Jung, N. (2015). Reducing embodied carbon during the design process of buildings. *Journal of Building Engineering*, 4, pp.1–13. doi:https://doi.org/10.1016/j.jobe.2015.06.005.

Hollberg, A., Genova, G. and Habert, G. (2020). Evaluation of BIM-based LCA results for building design. *Automation in Construction*, [online] 109, p.102972. doi:https://doi.org/10.1016/j.autcon.2019.102972.

Hollberg, A. and Ruth, J. (2016). LCA in architectural design—a parametric approach. *The International Journal of Life Cycle Assessment*, 21(7), pp.943–960. doi:https://doi.org/10.1007/s11367-016-1065-1.

Kamari, A., Kotula, B.M. and Schultz, C.P.L. (2022). A BIM-based LCA tool for sustainable building design during the early design stage. *Smart and Sustainable Built Environment*, [online] 11(2). doi:https://doi.org/10.1108/sasbe-09-2021-0157.

LETI (2020). *Climate Emergency Design Guide*. [online] leti. Available at: https://www.leti.uk/cedg.

leti. (n.d.). *LETI*. [online] Available at: https://www.leti.uk/about [Accessed 26 Jan. 2024].

make (2021). *Reducing embodied carbon isn’t all about materials*. [online] Make Architects. Available at: https://www.makearchitects.com/thinking/reducing-embodied-carbon-isnt-all-about-materials/ [Accessed 24 Jan. 2024].

Means, P. and Guggemos, A. (2015). Framework for Life Cycle Assessment (LCA) Based Environmental Decision Making During the Conceptual Design Phase for Commercial Buildings. *Procedia Engineering*, 118, pp.802–812. doi:https://doi.org/10.1016/j.proeng.2015.08.517.

Meex, E., Hollberg, A., Knapen, E., Hildebrand, L. and Verbeeck, G. (2018). Requirements for applying LCA-based environmental impact assessment tools in the early stages of building design. *Building and Environment*, 133, pp.228–236. doi:https://doi.org/10.1016/j.buildenv.2018.02.016.

O'Sullivan, D.T.J., Keane, M.M., Kelliher, D. and Hitchcock, R.J. (2004). Improving building operation by tracking performance metrics throughout the building lifecycle (BLC). *Energy and Buildings*, 36(11), pp.1075–1090. doi:https://doi.org/10.1016/j.enbuild.2004.03.003.

Ojo, A. and Pye, C. (n.d.). BIM Implementation Practices of Construction Organizations in the UK AEC Industry. *PM World Journal*, IX(X).

One Click LCA® software. (n.d.). *World's fastest Building Life Cycle Assessment software - One Click LCA*. [online] Available at: https://www.oneclicklca.com/ [Accessed 23 Jan. 2024].

portal.fcbstudios.com. (n.d.). *FCBS Carbon*. [online] Available at: https://portal.fcbstudios.com/fcbscarbon [Accessed 23 Jan. 2024].

RIBA (2019). *2030 Climate Challenge*. [online] Architecture.com. Available at: https://www.architecture.com/about/policy/climate-action/2030-climate-challenge [Accessed 23 Jan. 2024].

RIBA (2024). *RIBA Plan of Work*. [online] Architecture.com. Available at: https://www.architecture.com/knowledge-and-resources/resources-landing-page/riba-plan-of-work [Accessed 30 Jan. 2024].

RICS (2023). *Whole life carbon assessment (WLCA) for the built environment*. [online] Rics.org. Available at: https://www.rics.org/profession-standards/rics-standards-and-guidance/sector-standards/construction-standards/whole-life-carbon-assessment [Accessed 24 Jan. 2024].

Röck, M., Hollberg, A., Habert, G. and Passer, A. (2018). LCA and BIM: Visualization of environmental potentials in building construction at early design stages. *Building and Environment*, 140, pp.153–161. doi:https://doi.org/10.1016/j.buildenv.2018.05.006.

Soust-Verdaguer, B., Llatas, C. and García-Martínez, A. (2017). Critical review of bim-based LCA method to buildings. *Energy and Buildings*, 136, pp.110–120. doi:https://doi.org/10.1016/j.enbuild.2016.12.009.

Troup, L., Phillips, R., Eckelman, M.J. and Fannon, D. (2019). Effect of window-to-wall ratio on measured energy consumption in US office buildings. *Energy and Buildings*, 203, p.109434. doi:https://doi.org/10.1016/j.enbuild.2019.109434.

UK Parliament Committees. (2022). *Emissions must be reduced in the construction of buildings if the UK is to meet net zero, MPs warn*. [online] Available at: https://committees.parliament.uk/committee/62/environmental-audit-committee/news/171103/emissions-must-be-reduced-in-the-construction-of-buildings-if-the-uk-is-to-meet-net-zero-mps-warn/#:~:text=From%20residential%20to%20commercial%20buildings,assess%20and%20reduce%20these%20emissions. [Accessed 23 Jan. 2024].

United Nations (2015). *The 17 Sustainable Development Goals*. [online] United Nations. Available at: https://sdgs.un.org/goals.

Victoria, M.F. and Perera, S. (2018). Parametric embodied carbon prediction model for early stage estimating. *Energy and Buildings*, 168, pp.106–119. doi:https://doi.org/10.1016/j.enbuild.2018.02.044.

World green building council (2019). *Bringing embodied carbon upfront*. [online] World Green Building Council. Available at: https://worldgbc.org/advancing-net-zero/embodied-carbon/ [Accessed 23 Jan. 2024].

www.hawkinsbrown.com. (n.d.). *H\B:ERT Emissions reduction tool* |. [online] Available at: https://www.hawkinsbrown.com/sub-services/hbert-emissions-reduction-tool/ [Accessed 24 Jan. 2024].

OPENBIM DIGITIZATION OF THE ITALIAN FIRE PREVENTION CODE: AN IMPLEMENTATION ON REACTION TO FIRE

Nicola Iraci Sareri, Carlo Zanchetta, Maria Grazia Donatiello and Andrea Zamborlini
Department of Civil, Environmental and Architectural Engineering, University of Padova, Padova, Italy

Abstract

The study explores the topic of digital transition applied to fire prevention. It investigates the potential of Industry Foundation Classes (IFC) implementation in Fire Safety Engineering. Literature review highlights that, at this moment, IFC classes and attributes related to fire prevention are limited and poorly defined. In particular, the study analyses the reaction to fire of building elements and develops an algorithm that automates the code checking process through an experimental application based on a simplified building model. The results prove that the IFC data model is promising for the Fire Prevention Code's digitization.

Introduction

In recent years, the AEC industry has undergone many changes due to the progressive implementation of Building Information Modeling (BIM). Globally, based on the ISO 19650 standard, the use of open formats has become mandatory to ensure interoperability among various actors and to protect data transmission throughout the building lifecycle. In this sense, ISO 16739 defines the Industry Foundation Classes (IFC) format as an international standard for information exchange within BIM processes. Despite the ambitions, the standard is not fully usable. To date, some engineering topics, such as energetics and its simulations, acoustics, fire protection or maintenance scheduling, are not yet fully manageable through the IFC standard. The main problem concerns the lack of a common international standard for engineering domains. Each country, having its own legislation, imposes different approaches that prevent a unified vision. This is notably typical in the field of fire prevention. An initial harmonization effort among European states concerns the reaction to fire of building materials. Many countries use Euro-classes to classify the reaction to fire of materials (Zaccarelli, 2020). Even Italy, while having its own system, has made mandatory for manufacturers to use products classified only according to the Euro-classes since October 2023.

To encourage digitization, in the 2020 the National Fire Service promoted the Fire Digital Check (FDC) project with the circular STAFFCNVVF prot. n. 18332, which aimed at the acceptance of the requirements of the Ministerial Decree 560/17. The final goal of this project is to automate the verification of the authorization process, thus reducing the time to verify projects compliancy, optimizing resources and at the same time ensuring greater control of projects, which results in high levels of safety and quality (Negri, 2022).

This initiative is aligned with the activity of the Regulatory Room of BuildingSMART, which points at the automation and digitization of building approvals (Moreno, 2020).

To achieve it, the FDC project identified three steps:

1. Digitization of the code: digitize and standardize the language for compiling a fire prevention project;
2. Project validation/verification: digitize and standardize the verification process of a fire prevention project to automate the check;
3. Performance management: digitize and standardize the performance management language for the entire life of the building also in terms of fire prevention.

Currently, only the first of the three steps outlined by the FDC has been analysed and lead to the creation of proprietary parameters (Lala et al., 2022). However, the work does not consider the actual development of the IFC standard; actually, proprietary parameters were created to describe information which was already available in IFC data model. This study, first, focuses on the reaction to fire to demonstrate the feasibility of digitizing the Italian Fire Prevention Code predominantly through the use of native IFC parameters; second, it wants to demonstrate how it is possible to pursue, the automation of project approval based on the ISO 16739 standard, in accordance with BuildingSMART's Regulatory Room and the principles behind the FDC project,. A small project was implemented with the identified parameters, and finally, using a Python script, it was possible to check the compliancy of the information model with the national standard.

Background

Italian Fire Prevention Code and reaction to fire

According to a study conducted by a European Commission of 34 countries, only 30 of them have a legislation based on a performance approach (Athanasopoulou et al., 2023). The change for most states came after 2015. In Italy, just in 2015, with the approval of the new Fire Prevention Code, there was a shift from a prescriptive approach to a performance based one.

This transition requires a technical understanding of fire safety principles, followed by the application of proven engineering methods and calculations (Puchovsky, 1996). Through a performance regulation, it is possible to identify, from the actual fire risk, a precise strategy targeted to the needs of the specific project.

The Italian Fire Prevention Code is divided into four parts, two of them were analyzed: sections G and S. In the first section, the risk profiles of activities and the methods for their determination are defined. In the second section, ten fire prevention measures applicable to all activities are described. The first measure is reaction to fire.

Reaction to fire expresses the "behaviour of a material that, with its decomposition, participates in a fire to which it has been subjected under specific conditions"(Ponziani, 2019). It is a passive measure of fire protection that effects the fire prevention strategy in the initial phase (pre-flashover phase). This measure delays the achievement of the flashover condition because it limits the ignition and participation of materials in combustion, preventing or delaying the spread of fire and smoke.

On a regulatory level, the degree of material participation is assigned through the determination of certain classes of reaction to fire. These must be assigned to all building materials and products, including coatings, insulation, mobile fittings, and curtains.

This performance specification is assigned by laboratory tests performed by technical reference standards, through which specific classes of reaction to fire are determined. In the Italian Fire Prevention Code minimum reaction to fire classifications are expressed, referring both to Italian and European classification systems. These are given for each performance level. From 2022 with Ministerial Decree No.251 of 14/10/2022, the use of Italian classes for the classification of reaction to fire of construction products is no longer allowed. These products become classifiable only with the European classes according to EN 13501-1.

The Italian reaction to fire classifications, indicated with [Ita], refer to the DM 26/06/1984 which identifies six classes: from class 0, relative to non-combustible materials, to class 5 with the increasing of the material's participation in combustion. These classes refer to furnishing materials and products, such as furniture, moquettes, curtains, blankets and parquet floors, but also to scenic elements and tent covers. The reaction to fire class for upholstered furniture (armchairs, sofas, mattresses, upholstered chairs) are 1IM, 2IM, 3IM, this increases as the fire participation rate increases. The reaction to fire class assigned must refer to the whole composition of the covering, the upholstery and any interposed element.

European reaction to fire classes is only assigned to products compliant to Regulation 305/2011/EU concerning construction products, with reference to DM 10/3/2005. In Code's tables, they are indicated with the abbreviation [EU]. According to the EN 13501-1: 2018, the European classification is carried out by an alphanumeric code made up of the combination of main classes (A1, A2, B, C, D, E, F) with complementary classes relating to smoke production (s1, s2, s3), flaming droplets (d0, d1, d2) and acid smoke production (a1, a2, a3).

Table 1: Additional classification for smoke production

Acronym	Description
s3	No limitation of smoke production required
s2	The total smoke production as well as the ratio of increase in smoke production are limited
s1	More stringent criteria than s2 are satisfied

Table 2: Additional classification for flaming droplets

Acronym	Description
d2	No limitation
d1	No flaming droplets/particles persisting longer than a given time occurred
d0	No flaming droplets/particles occurred

Products that, although not intrinsically hazardous, have not yet been classified (Sabatino et al., 2021) belong to class F. Moreover, subscript abbreviations are sometimes added to the main classes in the following cases:

- fl in the case of floor products;
- L in the case of linear thermal insulation;
- ca in the case of electrical cables.

The compliance check of materials' behaviour in terms of reaction to fire depends on the performance levels assigned. Thus, it is necessary to establish the performance levels for each area of the activity through the attribution criteria. These are indicative and in the case of reaction to fire, a distinction is made between an escape route and a room. This distinction is necessary because of the importance of escape routes in saving lives, which require a more demanding performance. Consequently, the accepted performance level for rooms is generally immediately lower than the one indicated for escape routes of the corresponding risk profile.

As shown in Table 3, four performance levels were identified, each with its own description in qualitative terms. It emerges that only those spaces falling within performance levels II, III and IV are those in which the participation of materials in combustion is to be limited (Dattilo & Cavriani, 2019).

Table 3: Reaction to fire's performance levels

Performance level	Description
I	The contribution of materials to the fire is not assessed
II	The contribution of materials to the fire is significant
III	The contribution of materials to the fire is moderate
IV	The contribution of materials to the fire is almost negligible

For each performance level there is a corresponding compliant solution which consists of the use of materials belonging to certain groups (GM) identified by a number varying from 0 to 4 as their contribution to the fire increases.

The GM0 materials group does not make any contribution to fire and consist of all those materials with Italian reaction to fire class 0 or European reaction to fire class A1. It includes all the, so-called, incombustible materials. Materials group GM4, on the other hand, consist of all the ones that are not included in groups GM0, GM1, GM2, GM3. Thus, it includes unclassified materials and class F construction products. Since there is no requirement at the first performance level, compliant solutions are only defined for performance levels II, III and IV as shown in Table 4.

Table 4: Compliant solutions for each performance level

Performance level	Compliant solutions
I	No requirement
II	GM3
III	GM2
IV	GM1

The reaction to fire classes of materials included in materials groups GM1, GM2, GM3 are specified from table S.1-5 to table S.1-8 of the Italian Fire Prevention Code.

Reliability of IFC data model to support BIM based Fire Safety code checking

IFC is an international standard, organized into hierarchical classes through which the building system can be described. IFC should describe all aspects of all engineering disciplines through entity, types and relationships.

All subjects, from different disciplines, can use the IFC standard to exchange information. To ensure fast and lightweight exchange, IFC can be filtered with reference to the Model View Definitions (MVD). BuildingSMART has created many official MVDs. Currently, a sponsor search is underway to build a new MVD dedicated to FireSafety Engineering. The main limitation always concerns the lack of single legislation, in fact this new MVD is a consequence of a proposal of the International Fire Safety Standards (IFSS) Coalition. To reach "the first agreement on fire safety actions on an international scale", the consortium agreed to a 10-year commitment until 2032 (buildingSMART International, 2022). Without the IFC standard the exchange of information between the various subjects would be much longer and complicated because the various software would have to find a way to talk to each other, instead by using the IFC standard all software would be able to speak the same language (Laakso & Student, 2012). To ensure a future in the AEC industry based on interoperability among various players, BuildingSMART continues to rework and improve the IFC standard. With the roadmap developed in 2020, the organization wanted to emphasize how the future of the construction industry must move away from adopting solutions that do not work "out of the box" and embrace generic solutions (buildingSMART, 2020).

However, the use of IFC encounters difficulties related to incorrect use by actors and inappropriate reading and writing by some software. In the field of fire prevention, the prime example of a lack of standardized communication between different software is with fire dynamic simulation (FDS). To be able to simulate the fire, inside the building, simulation software first needs geometric information. Although within the IFC data model the geometric and spatial structure is well defined many studies have shown that it is not so easy for dynamic simulation software to receive information from the standard (Dimyadi et al., 2008) (Shi et al., 2019). Most software can only read CAD and DXF format so all relational information cannot be exported and read, or software only processes a limited set of data from IFC, sufficient to obtain basic building geometry and properties (Spearpoint & Dimyadi, 2007). Another research (Siddiqui et al., 2021) highlights the difficulty of operators in exchanging performance information. Most fire and evacuation modelling tools do not export results via open formats, while those that are able to do so, limit the export to geometric data, and not to the information that must be provided. The most obvious problem relates to the gaps that IFC has in relation to fire prevention regulations.

Previous research (Zanchetta et al., 2023) highlights precisely the limitations of the IFC data model with reference to some aspects of fire prevention. This research shows how to include information related to fire resistance, compartmentalization, and evacuation; it also introduces the discussion related to the reaction to fire of the materials highlighting how a unified reference is lacking. This lack leads different parties to freely create proprietary parameters, although they are already available in the standard, clashing with the concept of interoperability. One draft project, developed by a team related to the Italian Fire Digital Check project followed this second approach. Although the IFC standard was analysed, the team merely transcribed the Fire Prevention Code through a set of proprietary parameters. Regarding the reaction to fire class, the research team created a proprietary PropertySet named "VVF_Reazione al fuoco" within which there is the property to express the fire reaction class "Classe di reazione al fuoco".

BIM and automated code checking

Regulatory check is one of the aspects most prone to human error, so in recent years there has been an increasing interest toward automation. Automated rule checking (Eastman et al., 2009) consists of verifying that the project meets certain requirements. Rules should have the outcome “passed”, “failed”, “warning”, and it can be structured into four stages: (1) rule interpretation and logical structuring of rules for their application; (2) building model preparation, where the necessary information required for checking is prepared; (3) the rule execution phase, which carries out the checking, and (4) the reporting of the checking results.

There are several approaches available in literature related to automated rule checking. These are divided into systems known as “black box” and “white box” (Nidhra, 2012). The first ones involve the use of specific software generally difficult to implement and built through hard coding, while the second ones are transparent and involve the use of programming with codes or through visual programming tools (Ghannad et al., 2019). The latter approach is widely used mainly for the easy accessibility.

The software-based code is proving to be a popular choice for public administrations, which are thus able to combine a simplification of project approval work and with a potential saving of time and costs (Eastman et al., 2009). The most famous project is undoubtedly Singapore's CORENET (COnstruction and Real Estate NETwork), founded by the Minister of National Development (available at https://www.corenet.gov.sg/general/e-info.aspx). This project consists of three modules: CORENET e-Submission, CORENET e-PlanCheck and CORENET e-Info. With the e-PlanCheck module and FORENAX, it is possible to conduct a code check that includes different standards including those related to the fire Code (Greenwood et al., 2010). This project is a technology initiative that aimed to bring Singapore to the full integration of the phases of a building's lifecycle using IFC building model data.

This project was emulated in Norway with ByggSok project. The pilot project HITOS was checked with Solibri Model Checker (SMC) an IFC based commercial solution used for code checking. This software reads IFC model and is based on the Norwegian Stats'bugg handbook.

SMC and CORENET like many other applications, however, are “black box” because they present a set of rules that can be modified by the user and adapted to the specific project. On behalf they do not maintain the same freedom in creating new rules; the user must work on the product API to create them, which, however, are not public but only the developers themselves have access to them. (Simone et al., 2017).

Another approach to code checking is offered by visual programming solutions, which referring plug-ins such as Dynamo for Revit and Grasshopper for Rhino that support visual programming. A method, based on Dynamo, combines a semantic translation of the Malaysian fire safety regulations and the corresponding properties in Revit (Ismail et al., 2023). The regulations phrases are marked-up with four types of markers, using two of them (“requirements” and “exceptions”) to identify the necessary parameters in Revit. Eight scripts in Dynamo are created to check the compliance of the BIM modelled elements with the regulation’s provisions. Although this approach is very interesting, especially for the semantic translation implemented on regulatory aspects, it is a method focused on Revit's proprietary format and Dynamo, leaving no room for implementations in open formats.

Visual programming offers ease of programming and access to information, but it is based on solutions tied to proprietary formats as opposed to solutions implemented through scripts in python or java, which offer a wide range of completely open-source tools. IfcOpenShell is a python module which allows access to the elements contained in an IFC file and their properties, and its use is widespread because of its ease in extracting, manipulating, and creating information in an IFC data model. Several examples showing the potential of IfcOpenShell and its ability to extract information are available in literature (Narinder Singh et al., 2020; Singh et al., 2022). An interesting example of this Python module is used in an automated implementation of escape route information in IFC files (Haselberger, 2023).

The last relevant approach to code checking lies in the semantic web and graph databases. An interesting example (Peng & Liu, 2023) proposes an automatic code compliance checking based on BIM and Web technology. This study added the technique of knowledge graph to the structured processing of the rules. Its power lies in the application of natural language processing techniques to transform requirements into checking rules in the form of knowledge graphs.

Experimentation

In this section, it is illustrated the experimentation concerning the digitization of the reaction to fire measure. The purpose of this experimentation is to demonstrate that it is possible, through painstaking research, analysis, and definition, to achieve the objectives of the Fire Digital Check using exclusively ISO 16739. This turns out to be much more correct for the collaboration, interoperability and knowledge reuse because it does not consider parameters defined arbitrarily but relies on parameters already defined within an international and recognized standard, embracing the OpenBIM methodology. The following sub-sections explain the conceptual workflow and the translation in OpenBIM protocols.

Model construction

To test the experimental application, it was decided to consider a prototype nursing home. Therefore, the risk profile concerning human life safety (R_{life}), determined by the prevailing characteristics of the occupants, would be either D1 or D2. This assumption will then be used to determine the performance levels to be associated to the spaces. The digital model was developed using Autodesk's Revit. The modelled building, rectangular in shape, consists of two floors above ground, identically spaced. Each of the two floors, connected by a stairwell, consists of three rooms of equal size and a hallway (Figure 1, Figure 2, Figure 3).

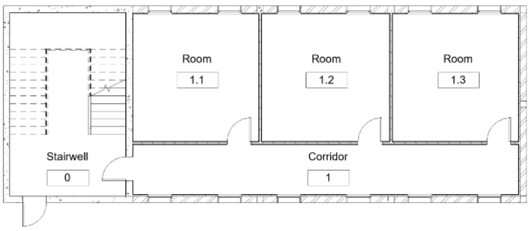


Figure 1: First-floor rooms

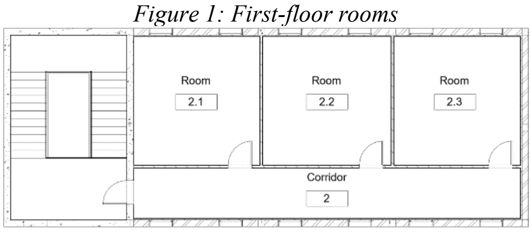


Figure 2: Second-floor rooms

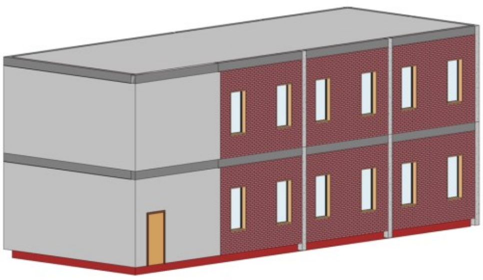


Figure 3: Building's Revit model

Check procedure

Before proceeding with the experiments, a verification procedure was devised. Starting with a room, which defines the required performance level, it is possible to identify all the boundary elements and, for each of them, check if their reaction to fire classes is compliant or not.

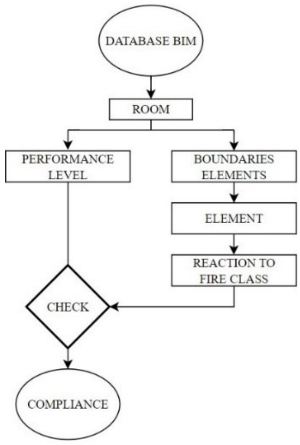


Figure 4: The assumed check procedure

Analysis of IFC attributes

For the study, the IFC version used is 4.3.2.0, made official in January 2024 (buildingSMART International, 2024).

Before compiling in the model the reaction to fire properties, it was necessary to understand where the information had to be placed according to IFC.

Reaction to fire is essentially a material property. In IFC materials are described through the class *IfcMaterial*. This entity is part of the resource domain, so it is just used to describe objects, it cannot have an instance of its own but must be related through *IfcRelAssociatesMaterial* to an entity of a higher layer. The entities in the resource layer do not have a unique identifier (Laakso & Student, 2012). For this reason, to clearly identify the uncompliant elements of the building, the reaction to fire property must be related to an object itself and not be attributed trough the material. This property can be attributed to objects such as walls, covering, roofs, siding, floors. Entities hierarchically lower than *IfcBuiltElement* and their properties were analysed. Specifically, the *IfcWall*, *IfcCovering*, *IfcSlab* and *IfcRoof* entities were analysed.

Regarding the reaction to fire performance of the elements, it was seen that *IfcWall*, *IfcSlab*, *IfcCovering* have two interesting properties: *Combustible* and *SurfaceSpreadOfFlame.* As highlighted in the previous research (Zanchetta et al., 2023) *IfcCovering* has a further possible property: *FlammabilityRating*. *IfcRoof*, instead, has none of these properties, as shown in Table 5.

Table 5: Entities of the IFC data model and the highlighted properties

	Property		
	FlammabilityRating	SurfaceSpreadOfFlame	Combustible
IfcCovering	✓	✓	✓
IfcRoof			
IfcSlab		✓	✓
IfcWall		✓	✓

Firstly, the properties *FlammabilityRating* and *SurfaceSpreadOfFlame* can be filled in with a string of text or numbers up to a maximum of 255 characters, while the property *Combustible* can assume a boolean value. The boolean value TRUE indicates that the object is made up of combustible material. If only this property is compiled, it could be assumed whether the material is combustible or not and thus whether it belongs to the GM0 material group (European class A1 and Italian class 0). However, this information is not relevant for the purpose of this research, because it does not provide any information about the specific reaction to fire class of the material and is therefore not useful for automatic verification of regulatory compliance. To best specify the fire behaviour of a material, it is necessary to assign to it a specific reaction to fire class, so the study proposes the inclusion of the specific class within the property.

FlammabilityRating and *SurfaceSpreadOfFlame* are very similar to each other (Table 6). The first one refers to the flammability of materials and the second to the way flames spread on the surface. The *SurfaceSpreadOfFlame* property seems to refer to the British standard, BS 476: Part 7: 1997, which defines a test method to measure lateral flame spread along wall and ceiling surfaces. The standard identifies four classes, from 1 to 4, of which Class 1 is the best in terms of flame propagation. However, this test doesn't specify the combustibility of a material. Instead, the European system based on Euroclasses (*EN 13501-1*, 2018), introduced in 2000 to eliminate trade barriers between member states and ensure consistent quality levels, refers to the reaction to fire of materials by considering not only the spread of flames but also many aspects, such as combustibility, the rate of participation in the fire and the ability to produce smoke. For this reason, the UK is still going through a transitional period in which it is still possible to find references to the national standard rather than the mandatory harmonized European classification, sometimes causing confusion about the performance of a given material as the two classifications refer to different aspects.

Table 6: IFC properties of interest

Property	Description
Flammability Rating	Flammability Rating for this object. It is given according to the national building code that governs the rating of flammability for materials
SurfaceSpread OfFlame	Indication on how the flames spread around the surface. It is given according to the national building code that governs the fire behaviour for materials
Combustible	Indication whether the object is made from combustible material (TRUE) or not (FALSE)

At the level of definition, it would be more correct to use the *FlammabilityRating* property because the concept of reaction to fire does not only refer to how the flame spreads on the wall, at the same time, this property it is attributable only to covering (*IfcCovering*). When the designer evaluates the reaction to fire of a wall, he must consider the whole stratigraphy not only the outermost layer although this one is more significant. So, it was decided to use the property *SurfaceSpreadOfFlame*. For the code checking, the reaction to fire within a space must be verified. The *IfcSpace* properties were analysed. As shown in Table 7, within the PropertySet *Pset_Risk*, there are four useful properties: *RiskType, RiskAssesmentMethodology, NatureOfRisk* and *RiskName*. The risk type Fire_Explosion was selected within the *RiskType* property, the regulatory reference (Fire Prevention Code) within the *RiskAssessmentMethodology* property, and the fire measure (reaction to fire) within the *NatureOfRisk* property. Finally, the performance level required for the fire reaction measure was associated with the *RiskName* property.

Table 7: IfcSpace Pset_Risk's properties

Property	Description
RiskType	Identifies the predefined types of risk from which the type required may be set
RiskAssessment Methodology	An indication or link to the chosen risk assessment methodology
NatureOfRisk	A description of the generic nature of the context or hazard that might be encountered
RiskName	A locally unique identifier for the risk entry that can be used to track the development and mitigation of the risk throughout the project life cycle

Selection of performance levels

Based on the use of the building, the model was implemented with the performance levels assumed within the *RiskName* property attributed to all rooms. According to the Fire Prevention Code, rooms and escape routes have different performance levels. Performance level 3 was assigned to rooms while performance level 4 (highest) was assigned to escape routes.

<Settings>

A	B	C	D	E	F
Type of room	Room number	Pset_Risk.RiskType	Pset_Risk.RiskAssessmentMethodology	Pset_Risk.NatureOfRisk	Pset_Risk.RiskName
Room	1.1	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Room	1.2	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Room	1.3	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Corridor	1	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	4
Stairwell	0	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	4
Room	2.1	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Room	2.2	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Room	2.3	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	3
Corridor	2	FIRE_EXPLOSION	D.M. 3/8/2015 e s.m.i.	Reaction to fire	4

Figure 5: Properties of each room and evacuation route

Material selection

Various commercial materials were analysed and some with different reaction to fire classes were identified. A material with a reaction to fire class higher than the regulatory minimum was assigned to all walls except for one wall in rooms 1.1 on the first floor and 2.1 on the second one (Figure 6). For these two, wallpaper with a reaction to fire class lower than the regulatory minimum was installed. This distinction was necessary to validate the model-checking algorithm.

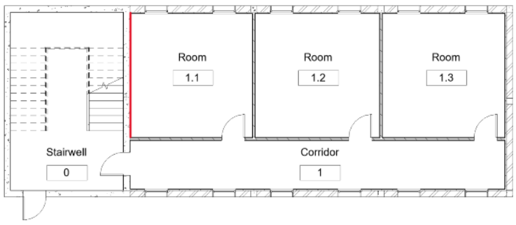


Figure 6: The wall where non-compliant wallpaper has been applied is highlighted in red. The same for the second floor

Algorithm development

The code checking algorithm was developed in open code using the Python programming language. Thanks to the IfcOpenShell library, it is possible to search all entities hierarchically below the IfcBuiltElement class previously identified properties for reaction to fire. This made it possible to implement computationally based compliance check. The script searches for every *IfcSpace* in the project the *IfcSpaceBoundaries* which enclose the room. The *IfcSpace* present the *Pset_Risk* which contains the *RiskName* property that express the reaction to fire's performance level. The algorithm reads the *RiskName* and compares the *SurfaceSpreadOfFlame* of each building element enclosing the *IfcSpace*.

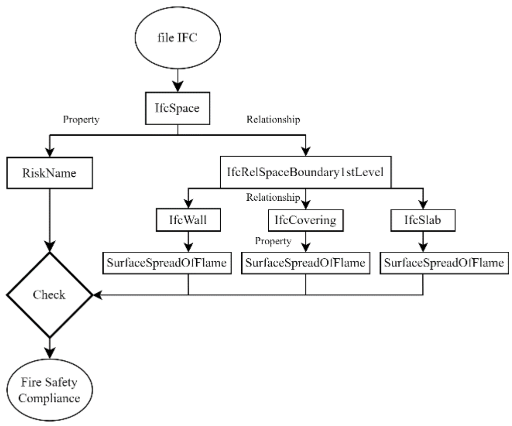
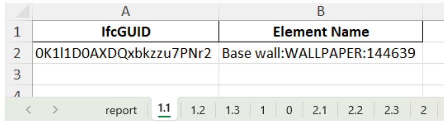


Figure 7: Script procedure

Results analysis

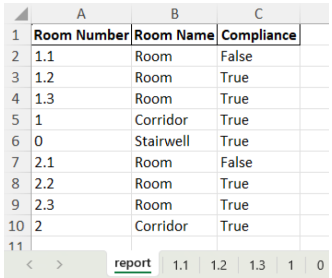
The algorithm was also set up to write an Excel file made up of several sheets, one for each room of the building, and one reporting every room and its compliance, which automatically is filled with True or False values. In the case of non-conforming elements, the global unique identifier (IfcGUID) and the name of the modelled element are reported. As can be seen from the Excel spreadsheets obtained, the only elements reported were the wallpapers installed in rooms1.1 and 2.1.

The algorithm was able to read all the boundary elements of each room, and to compare the values of the materials groups with the minimum values required by the performance level, thus automating the entire check. This approach, extended to all the other prescriptions contained in the Italian Fire Prevention Code could be the base of an automatic digital permitting, aligned with the BuildingSMART Regulatory Room's principles.



IfcGUID	Element Name
0K1l1D0AXDQxbkzzu7PNr2	Base wall:WALLPAPER:144639

Figure 8: Excel results



Room Number	Room Name	Compliance
1.1	Room	False
1.2	Room	True
1.3	Room	True
1	Corridor	True
0	Stairwell	True
2.1	Room	False
2.2	Room	True
2.3	Room	True
2	Corridor	True

Figure 9: Rooms' compliance report

Conclusions

In conclusion, the results give evidence of the high performance of the IFC format from the point of view of storing information, although they underline the difficulty of translating standards according to the logic of the IFC data format.

The study shows that, as a result of the normative translation, activities and analysis of the attributes and properties available in IFC, it is possible to use the properties already present within the IFC standard without creating new properties and to proceed to automatic code checking through the development of open algorithms. However, to achieve these results, it is necessary to have a detailed knowledge of the IFC architecture. Understanding which property to use was the biggest challenge because the definitions are not clear. Therefore, there is a need for BuildingSMART to comply with the standardization adopted by the different countries and to define a clear parameter within which the reaction-to-fire class can be entered, according to the Euroclasses. This research paves the way for many future developments, including refining what has been done about reaction to fire, digitization of other fire measures, and the mapping of regulatory requirements not currently covered by the IFC standard. The goal is the creation of an interoperable BIM database that serves both National Fire Service and fire prevention designers. The creation of an interoperable BIM database can be useful during design and verification for the issuance of the Fire Prevention Certificate.

References

Athanasopoulou, A., Sciarretta, A. ;, & Sousa, F. (2023). The status and needs for implementation of Fire Safety Engineering approach in Europe Support to policies and standards for sustainable construction. https://doi.org/10.2760/031591

buildingSMART. (2020). *Technical Roadmap buildingSMART - Getting ready for the future*.

buildingSMART International. (2022). *Fire Safety Engineering*. https://www.buildingsmart.org/standards/calls-for-participation/fire-safety-engineering/

buildingSMART International. (2024, January 4). *IFC 4.3 APPROVED as a Final Standard*. https://www.buildingsmart.org/ifc-4-3-approved-as-a-final-standard/

Dattilo, F., & Cavriani, M. (2019). *Codice di prevenzione incendi commentato. D.M. 3 agosto 2015. Norme tecniche di prevenzione incendi. Aggiornato con D.M. 12 aprile 2019 e D.M. 18 ottobre 2019. Esempi applicativi*. EPC Editore.

Dimyadi, J., Spearpoint, M., & Amor, R. (2008). Sharing building information using the IFC data model for FDS fire simulation. *Fire Safety Science*, 1329–1340. https://doi.org/10.3801/IAFSS.FSS.9-1329

Eastman, C., Lee, J. min, Jeong, Y. suk, & Lee, J. kook. (2009). Automatic rule-based checking of building designs. *Automation in Construction*, *18*(8), 1011–1033. https://doi.org/10.1016/J.AUTCON.2009.07.002

EN 13501-1. (2018).

Greenwood, D., Lockley, S., Malsane, S., & Matthews, J. (2010). *Automated compliance checking using building information models*.

Haselberger, M. (2023). *Realisierung von Fluchtwegsdaten im IFC-Format mithilfe von IfcOpenShell Realization of escape route data in the IFC format using IfcOpenShell*.

Ismail, A. S., Ali, K. N., Iahad, N. A., Kassem, M. A., & Al-Ashwal, N. T. (2023). BIM-Based Automated Code Compliance Checking System in Malaysian Fire Safety Regulations: A User-Friendly Approach. *Buildings*, *13*(6). https://doi.org/10.3390/buildings13061404

Laakso, M., & Student, D. (2012). The IFC Standard - A review of history, development, and standardization. In *Journal of Information Technology in Construction (ITcon)* (Vol. 17). http://www.itcon.org/2012/9

Lala, R., Parrini, B., Bizzarri, E., Amaro, G., Passalacqua, M., & Bertini, P. (2022). *Fire Digital Check_Tavolo 2D*.

Moreno, A. (2020). *OpenBIM per l'approvazione digitale dei progetti. A che punto siamo?*

Narinder Singh, E., Singh, H., & Singh Rai, H. (2020). Extracting Code Compliance Data from IFC With Python Language-Palarch's. In *Journal Of Archaeology Of Egypt/Egyptology* (Vol. 17, Issue 9).

Negri, R. (2022, November 25). *Fire Digital Check: il Codice di prevenzione incendi diventa digitale con il BIM - BIM Portale*. BIM Portale. https://www.bimportale.com/fire-digital-check-codice-prevenzione-incendi-diventa-digitale-bim/

Nidhra, S. (2012). Black Box and White Box Testing Techniques - A Literature Review. *International Journal of Embedded Systems and Applications*, *2*(2), 29–50. https://doi.org/10.5121/ijesa.2012.2204

Peng, J., & Liu, X. (2023). Automated code compliance checking research based on BIM and knowledge graph. *Scientific Reports*, *13*(1). https://doi.org/10.1038/s41598-023-34342-1

Ponziani, A. (2019). *Technical fire prevention standards*.

Puchovsky, M. (1996). *NFPA's Perspectives on Performance-Based Codes and Standards*.

Siddiqui, A. A., Ewer, J. A., Lawrence, P. J., Galea, E. R., & Frost, I. R. (2021). Building Information Modelling for performance-based Fire Safety Engineering analysis – A strategy for data sharing. *Journal of Building Engineering*, *42*. https://doi.org/10.1016/j.jobe.2021.102794

Simone, A., Taciuc, A., Karlshøj, J., & Dederichs, A. (2017). General rights Development of IFC based fire safety assesment tools DEVELOPMENT OF IFC BASED FIRE SAFETY ASSESSMENT TOOLS. In *Downloaded from orbit.dtu.dk on*.

Singh, N., Singh, H., & Singh, H. (2022). *Abstraction of code compliance information usin Python from the IFC model*.

Spearpoint, M. J., & Dimyadi. (2007). *Sharing Fire Engineering Simulation Data Using the IFC Building Information Model*.

Zaccarelli, G. (2020). *La prevenzione incendi in Europa: confronto fra normative*.

Zanchetta, C., Donatiello, M. G., Gabbanoto, A., & Paparella, R. (2023). *Digitization of building systems using IFC to support Performance Analysis and Code Checking: Standard limits and technological barrieres. A casa study on fire safety*. https://doi.org/10.30682/tema0901

LEVERAGING DIGITAL TWINS TOWARDS AN OCCUPANT-CENTRIC BUILT ENVIRONMENT: A REVIEW

Eduardo Navarro-Bringas [1]
[1]Heriot-Watt University, Edinburgh, UK

Abstract

One of the main purposes of building design and development is to enhance the experiences and activities of occupants. Traditionally, the longitudinal collection of occupant data has presented challenges. In this context, the emerging paradigm of Digital Twins (DTs) offers a promising solution. This systematic review investigates recent advancements in DT applications within the Architecture, Engineering, Construction and Operations (AECO) sector, focused on the integration and utilisation of real-time occupant data. Our findings categorize the types of data collected and the varying degrees of occupant involvement in these processes. The review also identifies the primary challenges encountered in data collection and integration, emphasising the need for balancing ethical and privacy concerns. Additionally, the paper explores the potential of DTs in enriching occupant-centric applications, highlighting their role in optimising building performance, enhancing comfort, and improving maintenance and safety. This review not only sheds light on the current state of occupant data in DT applications but also explores opportunities for leveraging occupant data to create more responsive and efficient built environments.

Introduction

The fundamental purpose of the built environment is to support and structure the activities and tasks of its occupants. Historically, understanding the impact of buildings on occupants has predominantly been the domain of Post-Occupancy Evaluations (POE). Initiated in the 1960s and 1970s, POEs have aimed to optimise and fine-tune building performance to better support occupants and inform the design of future structures (Preiser and Nasar, 2008). This field later expanded to include energy performance evaluations, significantly influenced by research in the 1990s and early 2000s (Leaman and Bordass, 2001). Despite significant contributions of POE to our current understanding of occupant comfort and general building performance, data collected from users has usually been limited to qualitative studies (i.e. focus groups) and one-off satisfaction surveys. As such, legacy challenges like the performance gap, the discrepancy between designed performance and actual performance persist, largely driven by occupant behaviour and usage patterns (Far *et al.,* 2022).

In recent years, the concept of Digital Twins (DTs) has emerged as a novel paradigm in built environment research. Evolving from advancements in Building Information Modelling (BIM), smart building technologies, and developments on Internet of Things (IoT) sensors, DTs offer the potential to enhance building performance through real-time data collection, granular control of building systems, and advanced operational intelligence via automation and improved simulation methods (Building Smart International, 2020). If the objective remains to design buildings that meet user needs, the role of DTs in understanding and refining performance for occupants could be substantial.

Although several systematic reviews have investigated DT applications in the built environment within recent years (Long *et al*., 2024), there has been limited focus on the specific challenges and opportunities related to integrating occupant data into DTs. This paper aims to examine how recent published research has addressed the incorporation of occupant data. Thus, the paper aims to analyse the selected articles to determine the purposes for which the identified DT applications are being developed, as well as to identify the various methods and types of occupant data that are integrated in the identified DTs.

Defining DTs

While the concept of Digital Twins (DTs) originated decades ago, primarily in the manufacturing context (Tao *et al*., 2019), its evolution in the built environment has been significantly influenced by advancements in smart buildings, the Internet of Things (IoT), enhanced sensing technologies, and the development of semantically rich digital models within Building Information Modelling (BIM). Additionally, recent reviews (Long *et al*., 2024) have shown an exponential growth in interest regarding DT applications in built environment research. Studies such as Shahzad *et al*. (2022) have highlighted a lack of common understanding among various stakeholders in the AECO sector, leading to challenges in establishing a universally accepted definition of a DT. For the purposes of this paper, a DT is defined as "(…) *a digital representation of a physical asset. Linked to each other, the physical and digital twin regularly exchange data* *throughout the Plan Build Operate Decommission lifecycle and use phase. Technologies like AI, machine learning (ML), sensors, and IoT allow for dynamic data gathering and right-time data exchange to take place*" (Building Smart International, 2020 p.2).

This definition encompasses the widely accepted view of DTs as digital replicas of physical assets, further enriched with data collected in real-time. However, to ensure consistent criteria for the article inclusion criteria used in this review, the constituents of a DT will be further defined. Drawing from the manufacturing sector, particularly the work of Tao *et al*. (2019) five layers have been identified as integral to a DT. Massafra *et al*. (2023) adapted these elements to the context of building-related DTs (Figure 1). These five constituents or layers include:

- Physical Asset: The building, equipped with IoT sensors.
- Virtual Asset: A digital representation of the building, typically a semantically rich Asset Information Model (AIM).
- Connections: The information and flows and connections that link the virtual and physical layers, as well as the other layers.
- Data: Integration of all data collected from both digital and physical asset layers to create a comprehensive and accurate information set.
- Services/Actuator: These facilitate the application of data to meet the needs of DT users. For instance, visualisation, system control, "*what-if*" scenario simulations etc.

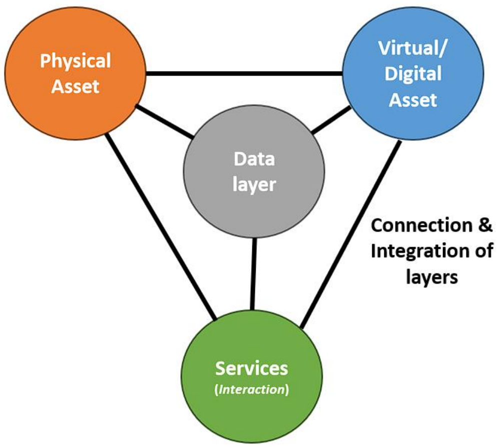


Figure 1. DT layers. Adapted from Massafra et al. (2023)

Methodology

Systematic Literature Review

This paper deploys a Systematic Literature Review (SLR) approach to establish the current state of occupant-centric DT applications within buildings. SLR was considered as a suitable approach, as the method allows for a systematic approach to collate exiting literature. Moreover, the method enables further understanding, synthesis of evidence, and identification of research gaps (Kitchenham, 2004). Additionally, the systematic approach to selection of papers enables for an unbiased approach to collating and reviewing the literature. This method is particularly valuable in uncovering themes and research directions within evolving fields, such as the integration of occupant data within DT applications. The following section will present protocol used for the SLR, as well as establish the criteria for selection of resources included in the SLR.

SLR procedure

The SLR commenced with the selection of Scopus as the primary database for literature search. Scopus was chosen over alternatives like Google Scholar due to its rigorous and well-curated nature, ensuring access to scholarly and peer-reviewed content. This decision aligns with the practices of previous SLRs in related fields, thereby upholding academic rigor and relevance. The process used to select the articles for inclusion in this review is presented in Figure 2.

The initial search, conducted on the 5th of January, utilised parameters such as "*Digital Twins*" AND "*Built Environment*" OR "*Building*" AND "*Occupant*" OR "*End user*" OR "*Post Occupancy Evaluation*". To filter these initial results, criteria were set to include sources from peer-reviewed journal articles in English language, which narrowed the initial query in Scopus to 40 articles.

Further screening involved a manual analysis of these papers by the research team. In this step, articles were excluded from the review if they did not meet the following criteria:

- Articles not focused on buildings as the unit of analysis. Through this criteria papers excluded for instance articles focused on a system of systems level, such as cities.
- Papers that did not met the previously established characteristics of a DT established in the introductory section of this paper were also omitted. These excluded papers that lacked a data layer combining real and virtual asset information.

The final sample (Table 1) of articles included in the revision included 20 articles.

SLR Results

Initial results showcase that the most common journal where occupant-centric DT research is being featured is *Building & Environment* (n=5), followed by *Energy & Buildings* (n=3), and *Sustainable Cities & Society*, *Applied Sciences* and *Buildings* with (n=2). These journals feature articles with a focus on Indoor Environmental Quality (IEQ) and energy issues within the

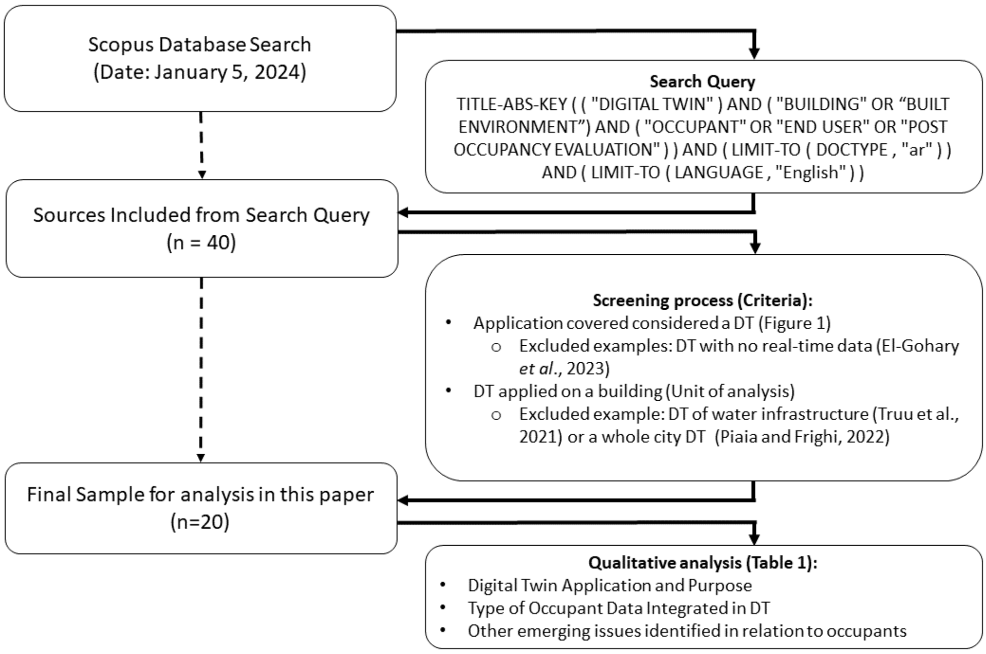


Figure 2. SLR procedure followed for the selection of the final 20 articles reviewed.

built environment. In terms of the country of affiliation of the first author, the country with the most extensive DT applications exploring occupancy issues is China with links to 4 of the reviewed studies. This is followed by 3 applications from Singapore and Italy. The remaining of the research appears quite widespread across a range of different countries, with Norway and Canada having multiple applications in the reviewed sample (n=2). Regarding publication year, the earliest article addressing

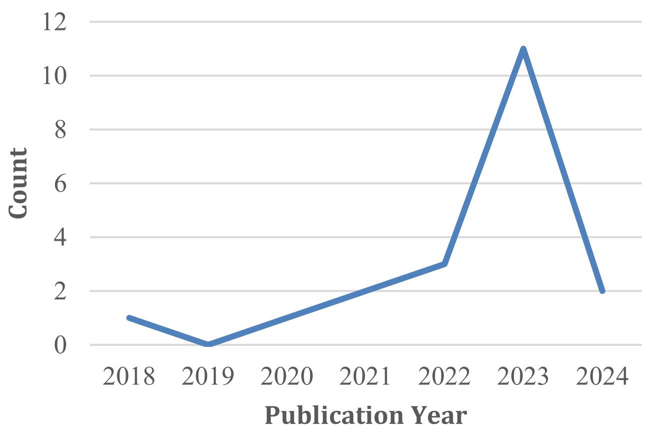


Figure 3 Publications per year (Up to 05/01/2024)

occupant data in building DTs included in the final review was published in 2018. The publication curve (figure 3) showcases an exponential growth in publications where the last 3 years (i.e. 2021-2023) have shown an increase in published articles in this area, with 55% (n=11) of the papers having been published in 2023. Table 1 presents a summary of the articles included in the final review.

Findings and themes from the SLR

Applications of occupant-centric DTs

Purpose 1: Systems & energy efficiency optimisation

One of the main applications identified refers to energy optimisation during the operation of buildings. In most cases, energy optimisation focuses on heating and cooling-related savings—i.e. mostly on the optimised use of Heating, Ventilation and Air Conditioning (HVAC) systems (Clausen *et al.*, 2021; Hosamo *et al.*, 2023b; Massafra *et al.*, 2023; Desogus *et al.*, 2023; Li *et al.*, 2023). Noticeably, the DT application presented in Seo and Yun's research (2022) places a strong focus on optimising energy consumed by lighting systems using schedules and simplified probability based behavioural models of occupants.

Regarding energy optimisation, the integration of occupant data into the DTs is expected to provide clearer insights into building operations for Facilities Management (FM). With some of the reviewed studies (Clausen *et al.*, 2021) aiming to establish an automated zone-based control that optimises the use and energy consumption from buildings. or allowing for the automation of building systems to optimise performance.

Purpose 2: Support of occupant comfort

The second purpose, while closely related to the previous, aims to develop DT applications that can better optimise the IEQ conditions to the occupants of the building. This application, often goes hand in hand with the previously discussed energy focus, presenting the DT as the key enabler of system performance improvements without the need to sacrifice occupant comfort.

Table 1. Articles included in SLR for review (n=20)

Reference	1st Author Affiliation	Journal	Purpose of DT application	Approach to occupant data
Massafra *et al.* (2023)	Italy	Sustainability	Energy optimisation in building operations	Passive: Simplified occupancy schedule data
Seghezzi *et al.* (2021)	Italy	Applied Sciences	Optimising operation & maintenance costs; supporting FM decisions	Passive: Occupancy tracked through camera- based visual sensors
Abdelrahman *et al.* (2022)	Singapore	Building & Environment	Optimisation of building systems for thermal comfort	Passive and Active: Spatial location tracking & physiological data; subjective experience feedback
Gnecco *et al.* (2023)	Italy	Energy & Buildings	Optimising buildings and systems for occupant comfort	Passive and Active: Spatial location tracking and physiological data; subjective experience feedback
Desogus *et al.* (2023)	Canada	Buildings	Early assessment/detection of occupant comfort issues	Passive: Application of Predicted Mean Vote model for occupant comfort
Wang *et al.* (2024)	China	Building & Environment	Understanding occupancy patterns for energy optimisation in historic buildings	Passive: Indoor location tracking data via Bluetooth Low Energy beacon sensors
Park *et al.* (2018)	South Korea	Applied Sciences	DT for early detection of emergencies and support of evacuation	Interactive: Occupant interaction with DT via Augmented Reality for evacuation guidance and feedback
Hosamo *et al.* (2023a)	Norway	Energy & Buildings	Early detection of failures/faults in building services and systems	No occupant data integrated
Qian *et al.* (2023)	China	Building & Environment	Optimising heating & cooling for occupant comfort	Passive and Active: Spatial location tracking; Longitudinal thermal comfort surveys
Abdelrahman & Miller (2022)	Singapore	Building & Environment	Understanding building components and elements' effects on occupant comfort	Passive and Active: Physiological data and spatial tracking; subjective thermal comfort data
Clausen *et al.* (2021)	Denmark	Energy Informatics	Optimised Zone Control DTs for energy consumption reduction	Passive: Sensing usage demands through CO^2 sensors
Huynh & Nguyen (2020)	Finland	Open Engineering	Support to FM decision-making & cost optimisation via DT interfaces	Occupant data collection not detailed; focus on DT interfaces for facilities managers
Li *et al.* (2023)	China	Energies	Energy optimisation in building operations	Passive: Occupancy data collected via sensors such as CO^2 sensors
Lee *et al.* (2023)	Singapore	Building & Environment	Optimising energy and comfort; building interactive DT interfaces	Interactive: Subjective experience data; passive physiological data; occupant DT interaction/control through interface.
Chamari *et al.* (2023)	Netherlands	IEEE Access	Establishing DT development process in systems architecture	Passive: Data integration includes CO^2 sensors and other passive occupancy sensors
Hadjidemetriou *et al.* (2023)	Greece	Sustainable Cities and Society	Visualising performance & integrating simulation tools for building operation improvement	Passive: Limited integration of occupant data via sensors (e.g., CO2)
Hosamo *et al.* (2023b)	Norway	Energy & Buildings	Early failure detection of Building Systems	Passive and Active: Passive occupant sensor data; Subjective experience surveys
Seo & Yun (2022)	South Korea	Buildings	Energy savings through optimisation of lighting systems	Passive: Behaviour modelled on schedule data; no real-time data
Tripathi *et al.* (2023)	Canada	Frontiers in Built Environment	Understanding occupant behaviour through DT-enabled POE	Active: Presented DT does not integrate user feedback, however, use case presented with subjective experience data collection.
Qian *et al.* (2024)	China	Sustainable Cities and Society	Optimising building system performance to reduce carbon emissions during building operation	Passive and Active: End user location tracking and physical sensors in buildings, additionally, survey data on subjective preferences integrated.

In this context, articles reviewed focus mostly on the thermal comfort dimension, while acknowledging the potential for DTs to expand into other (IEQ) dimensions such acoustic or visual comfort. Applications focusing on improved occupant comfort tend to integrate a wider range of occupant data into their DTs, including feedback prompts from users on their subjective experiences. Examples of these studies include the studies undertaken by Qian *et al.* (2023;2024) in residential building, as well as those in office environments conducted by Abdelrahman *et al*. (2022) and Lee *et al*. (2023).

*Purpose 3***:** *Early failure detection*

Another major purpose identified is the application of DTs for early failure detection of building systems. For instance, the DT application proposed by Hosamo *et al*. (2023a) integrates the end-user feedback with sensor data to determine faulty zones and building system. The DTs data is analysed through Bayesian networks to integrate a logic that can allow for more concrete diagnostics. The suggested model also integrates different ML algorithms to enhance and automate the failure detection process. The study also discusses the emerging topic of the use of ML in processing IoT data in DTs, while resulting analysis showcase very promising accuracy results, ranging from 94% for Fine Tree to 97% for Artificial Neural Networks, these still require substantial computing time. As such, trade-offs between accuracy of results and processing time will remain important considerations in occupant-centric DT data analytics.

Niche Purposes:

A few other purposes for the implemented DTs have also been identified. For instance, Seghezzi *et al*. (2021) model focuses on supporting the FM function through the provision of opportunities for maintenance cost optimisation using camera-based sensors to identify maintenance needs based on occupancy and usage patterns. Another niche application is proposed by Park *et al*. (2018). Their DT provides an early disaster detection system for fire safety, as well as including an interactive AR interface linked to the DT to support occupants with guidelines and feedback during evacuations.

Occupant Data Types

A focus of the SLR is to determine the methods and types of data collected in the identified applications. As such, categorising the different types of data will provide the structure for the following section. In this regard, the conceptual framework from Abdelrahman and Miller (2022) provides a good conceptualisation of data collected as objective data and passive data. The framework in this paper enhances the initial one by adding a dimension of occupant interaction, which can be considered passive (i.e. occupant data collected passively), or active (i.e. direct input from occupants). The following section will present findings regarding the main categories identified: 1) objective/passive occupant data collection, and 2) active/subjective occupant data collection (see Figure 4).

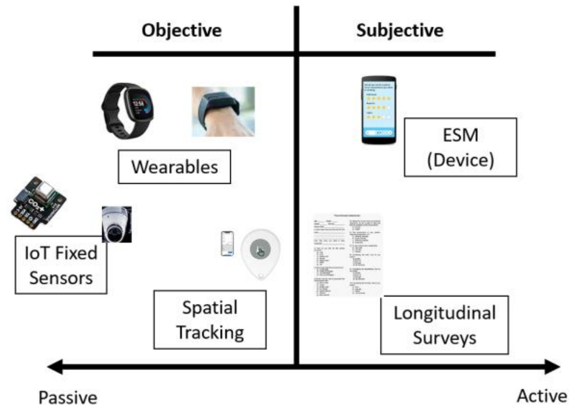


Figure 4. Clustered occupant data means

Objective-Passive: Occupant data integration

The most common tracking method across the studies is the use of passive IoT sensors. Available sensors to monitor occupancy and usage, include CO^2 concentration sensors, light and infrared sensors, camera-based sensors, wireless networks, or RFID sensors (Seghezzi *et al*., 2021). With these initial types of sensors, the aim is the monitoring of zone/area occupancy across the building in real time. Within the sampled papers, the most common occupancy sensors are CO^2 concentration change sensors (Clausen *et al*., 2021; Hadjidemetriou *et al*. 2023; Li *et al*., 2023; Chamari *et al*., 2023), which have a relatively high accuracy rate (over 90%) in detecting general fluctuations in rooms with limited issues regarding data privacy. The data collected through sensors such CO^2 or infrared sensors, while providing data on general occupancy patterns fail to achieve the accuracy of other sensing devices. Shegezzi *et al*. (2023) explore the use of camera-based sensors in their study. These types of sensors while presenting a series of calibration issues, that can affect accuracy when tracking occupants indoors, present additional opportunities for monitoring. First, data from camera-based sensors can potentially yield more granular insights on individual occupancy and usage patterns across the building, as well as provide additional insights on maintenance needs and condition of spaces. This requires further processing and analysis of video data through methods like computer vision.

Other studies present a simplified approach to occupancy monitoring (Seo & Yun, 2022; Massafra *et al*., 2023). These do not include real-time sensing of occupancy, and instead integrate occupancy considerations and room usage data through occupancy schedules. These involve the integration of data from the organisational asset management system, including scheduled room bookings within the building. While this data allows to partially integrate occupancy into DTs in a simplified manner, it fails to achieve the insights that data collected in real-time can yield. On a similar approach, using data captured through climatic stations (e.g. humidity, temperature…), Desogus *et al*., (2023) integrate issues of occupant thermal comfort in their proposed DT solution. These methods lack direct occupant input, and instead they rely on commonly used methods like the Predicted Mean Vote (PMV) to determine problematic areas in buildings.

The final set of passive sensors identified refer to personal wearables. The most comprehensive use of these is within the research studies undertaken by Abdelrahaman *et al*. (2022). In these, participants utilised smartwatches, which enabled the collection of real-time physiological data (i.e. heart rate and corporeal temperature) while inside the physical building. Additionally, perceptual, and spatial data from users was collected. At a smaller scale, and using an experimental setting, Gnecco *et al*. (2023) integrated the use of wearables to track a range of similar physiological data within their DT. While these are promising sources of data, results from co-design workshops by Lee *et al*. (2023) reveal a myriad of challenges. Workshop participants highlighted concerns regarding privacy and security of their personal data collected through wearable sensors. This raises questions about the ethics of such sensors, as well as the practical scalability at commercial scale.

Objective -Passive: Spatial Data

Spatial tracking and the integration of spatial data with semantically rich asset information models, can be considered its own category within passive and objective occupant data. These represent unique challenges for sensing and indoor positioning. As such, it has been the focus of several papers, warranting separate consideration from general IoT sensing (Wang *et al*., 2024; Abdelrahman *et al*., 2022; Gnecco *et al*., 2023).

Tracking occupants indoors has historically been challenging indoors. Unlike urban environments, where GPS can be used with high accuracy, indoor settings present a series of accuracy challenges. Often, a combination of sensors is required for precise spatial tracking. Wang *et al*. (2024) analyse various indoor tracking methods. Ultra-wideband sensors, while precise, are difficult to implement at larger scale. In contrast, more affordable options like wireless or Bluetooth Low-Energy (BLE) beacons, though less accurate (error-prone by 2-5m), are more feasible. BLE beacons, which refer to signal emitters detectable by smartphone applications, enable trilateration for indoor positioning.

Linking user positioning with building components and features can offer deeper insights into end-user behaviours and preferences, however, this integration also presents challenges. In this regard, Graph Neural Networks emerge as a promising solution, enabling mapping the proximity of occupants to building elements and zones. Thus, providing the means to link occupant data to information from these assets (Abdelrahman *et al*., 2022; Gnecco *et al.,* 2023; Wang *et al*., 2024). Recent studies, such as Tripathi *et al*. (2023), are examining GIS in the context of indoor environments. This approach, while challenging, also holds potential for enhancing understanding of spatial dynamics within built assets.

Subjective – Active occupant data integration

Another approach involves the collection of data from occupants regarding subjective and experiential dimensions. In this context, influenced by longstanding research on the measurement of perceptions of comfort from occupants, a few applications have integrated this data within their digital counterparts. Following best practices in comfort studies, Qian *et al.* (2023) conducted a 1-year longitudinal study where they collected 60-80 survey responses per month from residents in the study. These surveys were modelled on the ASHRAE's comfort scale and questions. While the exact points of entry are not clearly established, the approach while yielding additional insights raises questions regarding the labour-intensive nature of integrating data collected from surveys into DTs.

Alternatively, several studies have aimed to utilise the interfaces afforded by wearable sensors and smartphones to collect periodic experiential data from occupants. This approach builds on research on Experience Sampling Methods (ESM), a common approach used in psychology studies (Beal, 2015). Abdelrahman *et al.* (2022) utilised simple ESM to capture the subjective experience of users through a smartwatch application. Lee *et al.* (2023) utilises a similar approach in 15-day study to collect data from office users at regular times of the day, through push notifications prompting for feedback. The collection through these devices allows for automation in the integration of the data, compared with traditional surveys.

Subjective occupant data has the potential to address limitations of traditional occupant surveys (i.e. one-off surveys), by enabling collection of longitudinal datasets on user preferences, potentially leading to opportunities to provide a personalised approach to comfort in buildings (Preiser and Nasar, 2008). However, limitations are not to be overlooked. On one hand, the ethical considerations of personal data collected. Additionally, survey fatigue, a well-known phenomenon is also a consideration (Beal, 2015). Systems to collect data actively from users could be designed to be triggered only when specific events occur, limiting survey fatigue.

Emerging themes: Occupant interaction with DTs

A niche approach regarding occupants that emerged from the reviewed articles is the opportunities for end-user interaction with the DT platforms. Most of the reviewed papers, when discussing interaction with the DT, often refer to the FM or building operator (Chamari *et al*., 2023; Huynh & Nguyen, 2020; Hadjidemetriou *et al*., 2023; Li *et al*., 2023). Within these interaction with data for visualisation and control occurs through dashboards (E.g. Grafana) only accessible for FMs. Tripathi *et al*. (2023) suggest the expansion of these interaction and visualisation to the occupants. Park *et al.* (2018) expands on the end-user occupant interaction aspect. The proposed DT application, centred on early fire detection and evacuation, integrates an interactive AR application to provide occupants with real-time data regarding guidelines and evacuation procedures.

Going a step forward, Lee *et al*. (2023) in their aims to '*democratise*' future DT applications, explore the issue of limited interaction built-in in DTs faced by occupants.

The issue can be mapped out to previous research on smart buildings. They not only define interaction regarding visualisation of data, but also explore the expansion to the control dimension (i.e. ability of occupants to control building systems). Moreover, based on occupant comfort theories, particularly the adaptive comfort theory (de Dear and Brager, 2001), such interaction and control of systems might not only enhance satisfaction but might also yield significant energy savings by increasing user tolerance to discomfort.

Occupant-Centric DTs: Directions for future research

The review highlights a prevalent focus on occupant thermal comfort within occupant-centric DTs. This is evident in the dominance of DTs for energy optimisation of building systems and early failure detection. While some studies touch upon visual comfort for energy savings, other dimensions of IEQ such as acoustics remain largely unexplored. Future research should expand the focus to encompass these, particularly considering the relevance of acoustics in public environments.

The integration of occupant data beyond passive occupancy sensors remains limited. Notably, the comprehensive occupant-centric DT by Abdelrahman *et al*. (2022) stands out for its integration of physiological and spatial data through wearables, offering insights beyond thermal comfort. Leveraging such longitudinal occupant data can provide evidence-based information for design teams, opening avenues for evidence-based design in sectors like healthcare or education.

Finally, Lee *et al*.'s (2023) study underscores challenges associated with occupant-centric DTs. Participants expressed concerns regarding data privacy and ownership, highlighting the need for robust data security, especially when collecting subjective and physiological data. Addressing these challenges will remain key for the successful implementation and adoption of occupant-centric DTs at scale within real-world contexts.

Conclusions

Digital Twins (DTs) represent an emerging paradigm offering multiple opportunities in the AECO sector. This review focused on exploring how recent building DT applications are integrating occupant data. The findings highlight a variety of methods and challenges faced by developers of occupant-centric DTs. The review categorises these approaches based on the level of user involvement required. Additionally, it uncovers the challenges in integrating these data into DT platforms. The primary focus of these so far has been on building performance optimisation, especially in terms of energy efficiency. The review also sheds light on emerging applications and approaches to actively involve occupants, and provides directions for future research.

Reflecting on the early POE studies from the 1960s and 1970s, the central aim was to understand how buildings support end-users, such as understanding the influence of hospital design on patient healing rates or the impact of prison design on inmate behaviour. While common POE methods have provided valuable insights over the years, we argue that new paradigms, especially DTs, have the potential to establish these relationships with more rigorous and longitudinal evidence. Ultimately, this can lead to the design of built environments that more closely align with the needs and requirements of end-users.

The chosen method, an SLR, has several limitations. Firstly, it was restricted to peer-reviewed journals, resulting in a selection of a final 20 articles. The limitation of relying solely on peer-reviewed journal sources may exclude relevant literature from other publication types. With DTs being such a recent topic, preliminary research and applications presented in conferences might provide examples of other alternative applications. Using solely Scopus as the data source may have introduced bias, potentially overlooking relevant studies indexed in other databases (e.g. WoS). The review focused on two primary aspects: the purposes of the DT applications and the types of occupant data integrated. While this approach provided depth in analysing these, it may have overlooked broader aspects of occupant-centric DTs.

References

Abdelrahman, M.M., Chong, A., Miller, C. (2022) Personal thermal comfort models using digital twins: Preference prediction with BIM-extracted spatial–temporal proximity data from Build2Vec. *Building and Environment*, 207.

Abdelrahman, M.M. & Miller, C. (2022) Targeting occupant feedback using digital twins: Adaptive spatial–temporal thermal preference sampling to optimize personal comfort models. Building and Environment, 218.

Beal, D.J. (2015) ESM 2.0: State of the Art and Future Potential of Experience Sampling Methods in Organizational Research. *Annual Review of Organizational Psychology and Organizational Behavior*, 2, pp.383–407.

Building Smart International (2020) *Enabling an Ecosystem of Digital Twins*, Available at: https://f3h3w7a5.rocketcdn.me/wp-content/uploads/2020/06/Enabling-Digital-Twins-Positioning-Paper-Final.pdf (Accessed: 30/01/24)

Chamari, L., Petrova, E. & Pauwels, A.P. (2023) An End-to-End Implementation of a Service-Oriented Architecture for Data-Driven Smart Buildings. *IEEE Access*, 11.

Clausen, A., Arendt, K., Johansen, A., Sangogboye, F.C., Kjaergard, M.B., Veje, C.T. & Jorgensen, B.N. (2021) A digital twin framework for improving energy efficiency and occupant comfort in public and commercial buildings, *Energy Informatics*, 4(2).

de Dear, R. & Brager, G.S. (2001) The adaptive model of thermal comfort and energy conservation in the built environment. International Journal of Biometereology, 45, pp.100–108.

Desogus, G., Frau, C., Quaquero, E. & Rubiu, G. (2023) From Building Information Model to Digital Twin: A Framework for Building Thermal Comfort Monitoring, Visualizing, and Assessment. *Buildings*, 13(8).

Far, C., Ahmed, I. & Mackee, J. (2022) Significance of Occupant Behaviour on the Energy Performance Gap in Residential Buildings. *Architecture*, 2(2), pp. 424-433.

Gnecco, V.M., Vittori, F. & Pisello, A.L. (2023) Digital twins for decoding human-building interaction in multi-domain test-rooms for environmental comfort and energy saving via graph representation, *Energy and Buildings*, 279.

Hadjidemetriou, L., Stylianidis, N., Englezos, D., Papadopoulos, P., Eliades, D., Timotheou, S., Polycarpou, M.M. & Panayiotou, C. (2023) A digital twin architecture for real-time and offline high granularity analysis in smart buildings. *Sustainable Cities and Society*, 98.

Hosamo, H.H., Nielsen, H.K., Kraniotis, D., Svennevig, P.R. & Svidt, K. (2023a) Improving building occupant comfort through a digital twin approach: A Bayesian network model and predictive maintenance method. *Energy and Buildings*. 288.

Hosamo, H.H., Nielsen, H.K., Kraniotis, D., Svennevig, P.R. & Svidt, K. (2023b) Digital Twin framework for automated fault source detection and prediction for comfort performance evaluation of existing non-residential Norwegian buildings, *Energy and Buildings*, 281.

Huynh, D. & Nguyen-Ky, S. (2020) Engaging Building Automation Data Visualisation Using Building Information Modelling and Progressive Web Application. Open Engineering, 10(1), pp.434–442.

Kitchenham, B. (2004) *Procedures for Performing Systematic Reviews*, Keele University: UK.

Leaman, A. & Bordass, B. (2001) Assessing building performance in use 4: the Probe occupant surveys and their implications. *Building Research & Information*, 29(2), pp.129–143.

Lee, K.S., Lee, J.J., Aucremanne, C., Shah, I. & Ghahramani, A. (2023) Towards democratization of digital twins: Design principles for transformation into a human-building interface. *Building and Environment*. 244.

Li, C., Lu, P., Zhu, W., Zhu, H. & Zhang, X. (2023) Intelligent Monitoring Platform and Application for Building Energy Using Information Based on Digital Twin. *Energies*, 16(19).

Long, W., Bao, Z., Chen, K., Thomas Ng, S. & Yaha-Wuni, I. (2024) Developing an integrative framework for digital twin applications in the building construction industry: A systematic literature review. *Advanced Engineering Informatics*, 59.

Massafra, A., Costantino, C., Predari, G. & Gulli, R. (2023) Building Information Modeling and Building Performance Simulation-Based Decision Support Systems for Improved Built Heritage Operation, Sustainability, 15(14).

Park, S., Hoan Park, S., Won Park, L., Park, S., Lee, S., Lee, T., Hyeon Lee, S., Jang, H., Min Kim, S., Chang, H. & Park, S. (2018) Design and Implementation of a Smart IoT Based Building and Town Disaster Management System in Smart City Infrastructure. *Applied Sciences*, 8(11).

Preiser, W.F.E., & Nasar, J.L. (2008) Assessing Building Performance: Its Evolution from Post-Occupancy Evaluation. *International journal of Architectural Research*, 2(1), pp.84-99.

Qian, Y., Leng, J., Chun, Q., Wang, H. & Zhou, K. (2023) A year-long field investigation on the spatio-temporal variations of occupant's thermal comfort in Chinese traditional courtyard dwellings, Building and Environment, 228.

Qian,Y., Leng, J., Wang, H. & Liu, K. (2024) Evaluating carbon emissions from the operation of historic dwellings in cities based on an intelligent management platform. *Sustainable Cities and Society*, 100.

Seo, H. & Yun, W.S. (2022) Digital Twin-Based Assessment Framework for Energy Savings in University Classroom Lighting. *Buildings*, 12(5).

Shahzad, M., Shafiq, M.T., Douglas, D. & Kassem, M. (2022) Digital Twins in Built Environments: An Investigation of the Characteristics, Applications, and Challenges. *Buildings*, 12(2).

Seghezzi, E., Locatelli, M., Pellegrini, L., Pattini, G., Di Giuda, G.M., Tagliabue, L.C., &Boella, G. (2021) Towards an Occupancy-Oriented Digital Twin for Facility Management: Test Campaign and Sensors Assessment, *Applied Sciences*, 11(7).

Tao, F., Zhang, H., Liu, A. & Nee, Y.C. (2019) Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics*, 15(4), pp.2405–2415.

The Institution of Engineering and Technology (IET) (2019) *Digital Twins for the built environment*. London.

Tripathi, I., Froese, T.M. & Mallory-Hill, S. (2023) Applicability of BIM-IoT-GIS integrated digital twins for post occupancy evaluations. *Frontiers of Built Environment*, 9.

Wang, H., Qian, Y., Kuang, Y., Leng, J, Yang, Y. & Zhang, H. (2024), How occupant positioning systems can be applied to help historic residences manage energy consumption: A case study in China, *Building and Environment*, 249.

A SYSTEMATIC LITERATURE REVIEW OF BUILDING INFORMATION MODELLING (BIM) APPLICATION ON RAILWAY ASSET MANAGEMENT

Hyunjung Yoo[1], Weiwei Chen[1,2*], Diran Yu[1], Tim Broyd[1]
[1] The Bartlett School of Sustainable Construction, University College London (UCL), London, WC1E 6BT, UK
[2] Civil Engineering Division, Engineering Department, University of Cambridge, CB3 0FA, UK

Abstract

Building Information Modelling (BIM) has become a prominent trend in the built environment. Despite numerous countries mandating the adoption of BIM in this context, its integration into transportation infrastructure, notably railways, has progressed gradually. Asset Management (AM) continues to be emphasised in transport infrastructure. This study offers an understanding of BIM-based AM in railways via a systematic literature review (SLR). Although some studies delve into railway BIM-based applications across various life cycle phases, previous research has predominantly centred on the operation and maintenance phases. This paper presents a comprehensive view and proposes further research direction.

Introduction

Building Information Modelling (BIM) has a unique characteristic and can be found in the definitions from the standard, academic and industry fields. Nonetheless, a common thread across these definitions underscores the significance of features such as process, information, and life cycle considerations (British Standards Institution, 2013; International Organization for Standardization, 2018; National Building Specification, 2021). BIM is acknowledged in the academic field for its numerous benefits. Wang *et al.* (2013) arranged the BIM's various advantages: decreased costs, reduced errors, improved estimation, improved coordination, identifying conflicts and enhanced clients' and end-users understanding. Shin, Jung and Kim (2022) and Shin, Kim and Liao (2024) found meaningful benefits through their research on BIM implementation in the design and construction phase of the railway sector, such as cost and management aspects. Based on these advantages, governments across the globe, regardless of their geographical location, are opting to mandate the use of BIM in the built environment. Among them, the United Kingdom is positioned as a leading country globally in the application of BIM (Jaskula and Papadonikolaki, 2021).

Nevertheless, BIM is still slowly being adopted in transportation infrastructure, especially linear transportation infrastructure (e.g. railways, roads, gas pipelines, power lines, rivers and canals). Unlike buildings or other entities, transportation infrastructure's considerable size and intricacy contribute to the relatively recent initiation of BIM adoption. Notably, the application of BIM in railways is often referred to as being in its infancy (Han *et al.*, 2018).

As previously mentioned in the definition of BIM and its ultimate purpose, it is noted that it considers the lifecycle. The BIM lifecycle simply refers to the product lifecycle process, which consists of the concept, design, construction, operation, demolition and disposal (Gajera, 2017). Additionally, the term "sustainability" is a crucial global focus. Through the announcement of SDGs Goal 9, the United Nations emphasises the sustainability of infrastructure. In this paper, the perspective of sustainability is examined through an Asset Management (AM) approach. First, PAS 55- 1 defined the AM definition as 'systematic and coordinated activities and practices through which an organisation optimally and sustainably manages its assets and asset systems, their associated performance, risks and expenditures over their life cycles for the purpose of achieving its organisational strategic plan' (British Standards Institution, 2008, p.2). AM is not a novel concept, but AM is considered a relatively new discipline in the transportation infrastructure field because of the recent need to perform the lifecycle management is necessary for this domain (Garramone, Tonelli and Scaioni, 2022). Alnaggar and Pitt (2019) emphasised that BIM has a big potential in the AM sector.

Among linear transportation infrastructures, railways are environmentally friendly with low carbon emissions. They possess numerous advantages, including the ability to transport a large number of people at once and minimal traffic congestion. Consequently, they continue to garner significant global attention as transportation. Hence, the purpose of this paper is to explore a comprehensive perspective on railway infrastructure asset management using BIM.

This paper mainly comprises four sections. First, the introduction part consists of background information and general understanding. Second, the research methods part demonstrates this paper's systematic literature review (SLR) process. Third, the results and discussion section includes the SLR results and explains them. Finally, the conclusion summarises the overall contents of this paper.

Research Methods

This paper aims to investigate the BIM-applied railway infrastructure AM from a comprehensive point of view. To achieve this, this study conducted a systematic literature review (SLR), which adopted the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) methods, as illustrated in Figure 1.

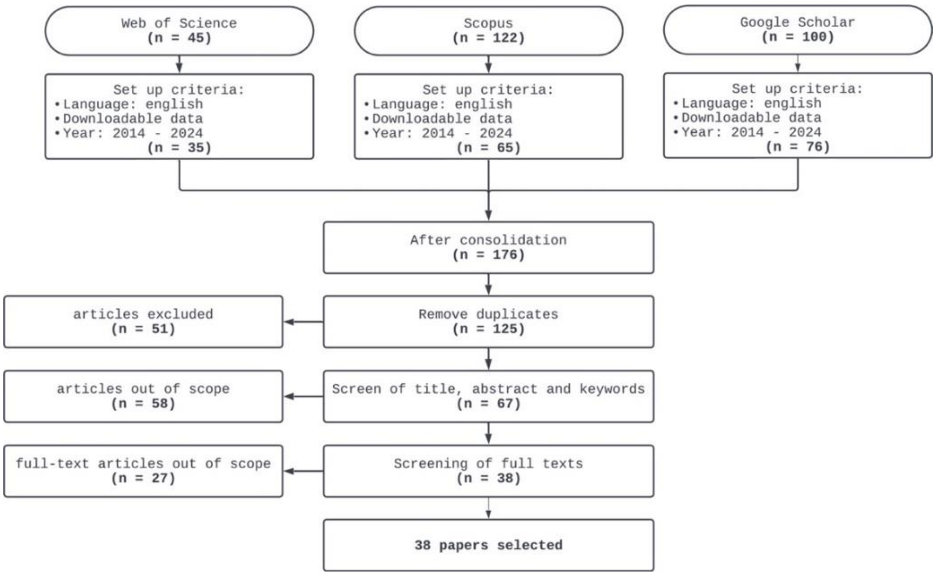


Figure 1: SLR flow chart (adopted the PRISMA method).

First, the authors decided on the keywords for BIM-based rail infrastructure AM. This study's main keywords are BIM, railway and AM. The keywords for searching in the database are set in Figure 2. The reason why the keyword "life cycle" is used not only the keyword AM, because the extended area of AM contains the consideration of the life cycle. Therefore, this review includes the term life cycle to cover all aspects of AM rather than just one specific part of it.

(BIM OR "building information modelling" OR "building information modeling" OR "building information model") AND (AM OR "asset management" OR "life cycle") AND (rail OR railway OR railroad)

Figure 2: Keywords.

This review selected three databases: Web of Science, Google Scholar and Scopus. At first, a total of 267 papers were found from three selected databases. Several criteria were used in SLR in this study. First, the language used is limited to English. Second, the article only includes the downloadable data. Third, the range of years is between 2014 and 2024. BIM is deeply related to cutting-edge technology. Hence, the literature's publication dates are only considered for the past 11 years. These years include the years when BIM exploded in popularity. Fourth, this SLR includes the grey literature (such as conference papers, books and proceeding papers). Including grey literature in this study is important because it allows us to include more current trends and ongoing research. By applying the criteria, approximately 176 papers were selected from the first 267 papers.

The combined search outcomes of three databases, as per four specific criteria, yield a total of 176 scholarly articles. After identifying 51 duplications among the 176 aggregated papers, the resultant set comprises 125 non-duplicated papers. From this point onward, authors examined titles, abstracts, and keywords to filter out relevant papers. In this procedure, 67 papers were identified after excluding the out-of-scope of 58 papers. This study examined all 67 identified documents in the last phase, verifying their relevance. Consequently, the final 38 papers align with the included scope.

Results and Discussion

Results

The 38 identified documents focus on the scope of asset management in railway infrastructure with the implementation of BIM. This SLR investigated papers from 2014 to 2024. Figure 3 shows a gradual increase in the number of papers published. The papers in this review were not published in 2015 or 2024. The years 2021 and 2023, in particular, saw a sudden drop in the published literature than the previous year. The most recently published papers were in 2023, totalling six papers.

In this SLR study, beyond the exclusive analysis of journal articles, this review encompasses grey literature, including conference papers and conference proceedings. The conference category includes both conference papers and proceedings. Specifically, there are 21 journal articles, comprising 55% of the total, and 17 conference papers, representing 45%. An observation is that more proportion of the papers analysed in this review originated from journals.

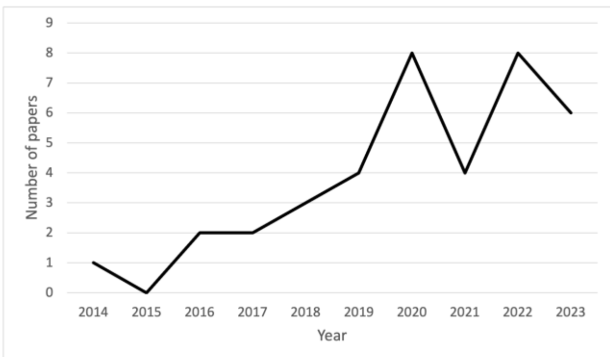


Figure 3: Trend of the published year.

Figure 4 illustrates the percentage of paper types by year. This figure consolidates the observations from published papers years, which are illustrated in Figure 3, and the types of papers published. It visually represents the proportion of papers published each year by type.

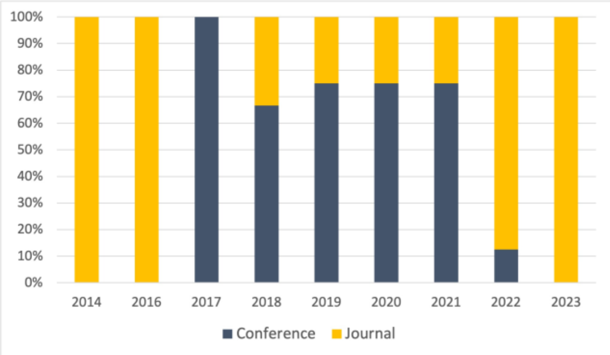


Figure 4: Percentage of paper types by year.

Conducting an SLR on Railway AM with the application of BIM, diverse applications of BIM were witnessed in Figure 5. This included not only the exclusive use of BIM but also the integration of other intelligent technology with BIM, OpenBIM, the application BIM with Industry Foundation Classes (IFC) and System Information Modelling (SIM), the combination with Geographic Information Systems (GIS), the extended application of BIM like the Digital Twins (DT) and the application of BIM in other infrastructure.

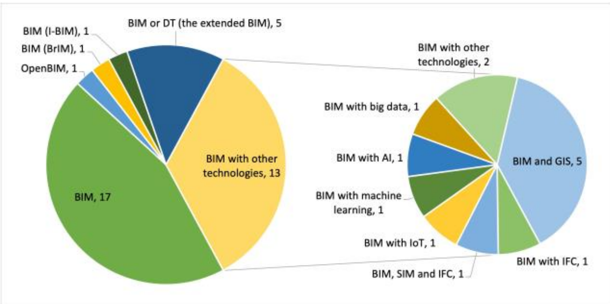


Figure 5: Used BIM classification.

Figure 6 presents a diagram designed to investigate the areas within railway infrastructure where BIM has been applied. The most extensively covered segment, accounting for 38% of instances, involves 12 papers providing a railway. The category of railway-related buildings is closely followed, representing 7 papers. Five papers delve into the application of BIM to railway bridges. Additionally, railway lines, tracks, tunnels and transportation infrastructure are investigated by applying BIM in each of the 2 papers. Lastly, the railway signal, turnout system, and light rail system were studied in one paper with BIM for AM, respectively.

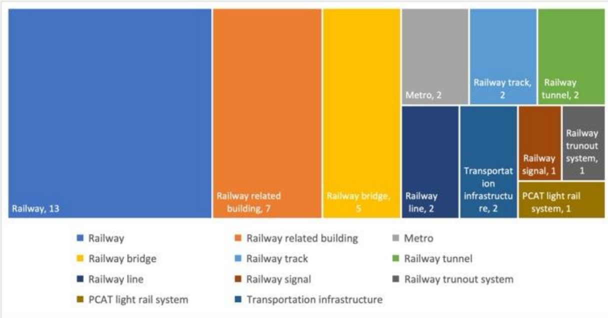


Figure 6: BIM-applicated objects.

Next, an exploration was undertaken to identify the stages or points where BIM has been applied, as depicted in Figure 7. A notable observation is that around 13 papers focused on utilising BIM for AM during all stages of the life cycle. A total of 22 papers were studied related to the design phase, and a total of 17 papers were researched in the construction phase. The operation & maintenance (O&M) phase conducted the AM based on the BIM around 27 studies. Two papers considered the rehabilitation stage.

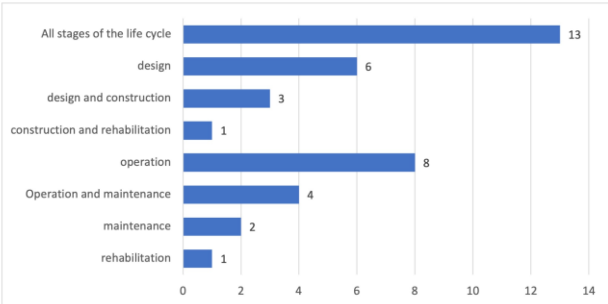


Figure 7: BIM used phases.

The papers analysed in this SLR aimed to implement BIM in railway infrastructure to facilitate AM. The different methods performed to manage AM are summarised in the literature according to the phase in which BIM is applied, as shown in Table 1. The literature shows that there are many different ways to perform AM by each phase.

Table 1: Studies employed to attain the objective at each stage.

	Asset Management	
All stages of the life cycle	Propose the BIM-based approach	(Dan, 2021; Zheleznov, 2021; Alqatawna *et al.*, 2023)
	Investigate the current situation of BIM and provide the countermeasures	(Kurwi, Demian and Hassan, 2017; Han *et al.*, 2018; Belcher and Abraham, 2023)

	Propose the new applications of DT (extended BIM)	(Kaewunruen and Lian, 2019)
	Suggest the application of BIM	(Cheng *et al.*, 2019; Kaewunruen, Peng and Phil-Ebosie, 2020; Wang *et al.*, 2022)
	Propose the data exchange model	(Gu *et al.*, 2022)
	Combine the application of BIM with other technologies	(Gebken, Drews and Schirmer, 2019; Wan *et al.*, 2020)
Design stage	Propose the BIM-based approach	(Lv, 2018; Efanov, Shilenko and Khoroshev, 2020; Pasetto *et al.*, 2020; Liu, 2022)
	Investigate the current situation of BIM and provide the countermeasures	(Sanchez *et al.*, 2014; Akponeware *et al.*, 2020)
	Introduce safety design, risk evaluation system and related theories of BIM	(Xiahou *et al.*, 2022)
	Suggest the application of BIM	(Chen, Hu and Tang, 2016)
	Provide the insight of BIM/GIS interoperability	(Floros, Ruff and Ellul, 2020)
Construction stage	Propose the BIM-based approach	(Großauer *et al.*, 2022; Liu, 2022)
	Provide the insight of BIM/GIS interoperability	(Floros, Ruff and Ellul, 2020)
	Suggest the application of BIM	(Chen, Hu and Tang, 2016)
Operation and Maintenance stages	Develop the BIM-based methodologies	(Wang *et al.*, 2020; Ciccone *et al.*, 2022)
	Investigate the current situation of BIM and provide the countermeasures	(Liu *et al.*, 2023)
	Develop a suitable platform based on BIM	(Hartung, Senger and Klemt-Albert, 2019; Hartung *et al.*, 2020)
	Introduce the new approach to digital asset management	(Love *et al.*, 2018)
	Combine the application of BIM with other technologies	(Boyes, Ellul and Irwin, 2017; Sresakoolchai and Kaewunruen, 2021, 2023; Garramone, Tonelli and Scaioni, 2022)
	Propose the new applications of DT (extended BIM)	(Kaewunruen, Sresakoolchai and Lin, 2021; Kaewunruen *et al.*, 2023)
	Develop a real-time lifecycle assessment-capable paradigm digital twin framework	(Borjigin *et al.*, 2022)
	Propose a solution within the existing CAD/BIM/GIS application	(Bartonek *et al.*, 2023)
Rehabilitation stage	Suggest the application of BIM	(McKenna *et al.*, 2017)
	Propose the BIM-based approach	(Großauer *et al.*, 2022)

Discussion

1) Current

Figure 3 captures related research trends that seem to be increasing overall. The lack of published papers in 2024 is likely due to the fact that the papers have not yet been published at the time of this review. Although the change is not dramatic and fluctuates between 2020 and 2023, it still showed a rising interest in the relevant keywords in recent years. In other words, the escalating quantity of published papers can be interpreted as a meaning of an intensifying interest in the associated keywords, signifying rising scholarly attention to these subjects. Figure 4 demonstrates the proportion of published paper types by year. The research was found to be ongoing, with a balance of journal articles and conferences being published. Figure 5 shows the current research on BIM type. While the use of only BIM remains prevalent, a noticeable adoption of BIM in more complex integrations has also been observed. For example, BIM for infrastructure, open BIM, BIM with SIM and IFC, BIM/GIS, extended BIM (DT) and BIM with other technologies. There were noteworthy attempts to expand its applications through various methods.

In the AEC sector, architecture is widely known as the domain where BIM has been most effectively applied and advanced. It was assumed that the reviewed studies would primarily focus on building research. However, Figure 6 revealed that BIM is utilised in a much more comprehensive range of areas. Likewise, the expectation was that most research would be confined to design and construction; however, it was observed to be applied across various domains, especially O&M. Figure 7 illustrates the diversity of BIM application phases. This indicates its diverse application from design to redecoration and its multifaceted implementation across multiple stages, not only the O&M phase. This illustrates the potential of BIM's diverse field of applications, further extending to all stages of the lifecycle.

Table 1 highlights the BIM-based AM. Each study was conducted using various methods at each stage to achieve its objectives. The design and construction phases have been studied in common research: suggest the application of BIM, investigate the current situation of BIM and provide countermeasures, propose the BIM-based approach, and provide insight into BIM/GIS interoperability. In the O&M phase, there were also studies such as the BIM-based approach, investigating the current situation of BIM and providing the countermeasures, but also the development of BIM with other technologies, platform construction, and DT (extended BIM) related studies were characterised differently from other phases. The methods used in all stages of the life cycle use most of the methods used in other stages. BIM application and a BIM-based approach were proposed for the rehabilitation phase. A common method used throughout the entire phase to apply BIM for AM was the proposed BIM-based approach. It can be seen that a lot of research has been done to study BIM-based approaches.

On the other hand, the challenges have been noticed in papers. In general, while the authors explained that they needed to improve their study, scholars emphasised the limitations of their studies, such as the specific performance of the application of BIM (Chen, Hu and Tang, 2016), information exchange (Floros, Ruff and Ellul, 2020; Kaewunruen, Peng and Phil-Ebosie, 2020), limitation of standards (Floros, Ruff and Ellul, 2020)and limitation of guidelines (Love *et al.*, 2018). In order to perform AM using BIM on a railway, there are obvious difficulties in terms of performance, but it is especially difficult to exchange information.

2) Future

The findings of Table 1 describe the different approaches that BIM has been used in railway infrastructure to achieve the AM. This can be connected with future research directions that are frequently mentioned in the literature's future research section. First, many of the authors described the need to develop their studies more deeply. Therefore, future research seems to study the methods that were used more deeply, as mentioned in Table 1. Second, future studies could be proposed on the combination of Table 1 methods and the literature’s mentioned limitations and future research suggestions. The keywords that are mentioned as limitations of the studies are the specific performance of the application of BIM, information exchange, and limitations of standards and guidelines. In several papers, keywords frequently mentioned in concluding remarks regarding future research directions include developing the strategy, effectiveness, efficiency, intelligence, and interoperability. In order to develop these keywords, future research might be conducted using the methods mentioned in Table 1. For example, investigating the current status of related keywords and finding the problem and countermeasure, developing a BIM-based approach, combining the application of BIM with other technologies, proposing the data exchange model, developing a new approach to digital asset management or developing the insight of interoperability.

Figure 5 demonstrates that many studies are currently conducted on BIM alone and BIM with other technologies. This indicates the potential for future applications of BIM in infrastructure AM. This shows the BIM potential of BIM integration with other complex technologies. Especially in BIM or DT and intelligent BIM, there is an emerging theme advocating for the advanced application of BIM, extending into dimensions such as 4D, 5D, 6D, and 7D, alongside the integration of various intelligent technologies. The occurrence of DT in asset management indicates a significant change towards higher integration, especially with BIM technology. This combination emphasises the need for further investigations on BIM within the context of DT. Research efforts like this are crucial for discovering innovative technologies and gaining useful knowledge about asset management. This, in turn, promotes the creation of more advanced and efficient asset management systems.

Conclusions

This paper aims to investigate the current trend and provide a comprehensive view of BIM-applied rail infrastructure AM. This paper adopts the SLR review, which includes the grey literature and uses three databases. As a result, 38 papers were selected for review. While conventional BIM applications are prevalent, research combining BIM with other technologies is also diverse. Moreover, BIM applications are conducted beyond the general railway context, encompassing a variety of targets such as buildings, bridges, systems, signals, and more. Although the majority of BIM applications for AM are concentrated in the O&M phase, they are also performed in various areas. With the current prominence of these keywords, the research trend is expected to continue, with the anticipation of further studies. Based on the studies that have been conducted, limitations of the studies and suggestions for future research are discussed. In future research, not only research focusing on the single use of BIM but also BIM integration research incorporating various intelligent technologies is expected. Furthermore, this is expected to employ real-time data integration to implement DT, enabling holistic Asset Management.

Acknowledgements

The first author (Hyunjung Yoo) is funded by the UCL Bartlett School of Sustainable Construction (BSSC) for BSSC Studentship.

References

Akponeware, A., Dawood, N., Rodriguez, S. and Dawood, H. (2020). 'EXPLORING THE DEVELOPMENT OF A BIM-ENABLED PROCESS FRAMEWORK FOR THE LCA OF RAIL TRACKS'. in *Proc. 37th CIB W78 Information Technology for Construction Conference (CIB W78), 18-20 August 2020 São Paulo, Brazil*, pp. 168–184.

Alnaggar, A. and Pitt, M. (2019). 'Lifecycle Exchange for Asset Data (LEAD): A proposed process model for managing asset dataflow between building stakeholders using BIM open standards'. *Journal of Facilities Management*. Emerald Publishing Limited, 17 (5), pp. 385–411. doi: 10.1108/JFM-06-2019-0030.

Alqatawna, A., Sánchez-Cambronero, S., Gallego, I. and Rivas, A. (2023). 'BIM-centered high-speed railway line design for full infrastructure lifecycle'. *Automation in Construction*, 156. doi: 10.1016/j.autcon.2023.105114.

Bartonek, D., Bures, J., Vystavel, O. and Havlicek, R. (2023). 'Case Study of Remodelling the As-Built Documentation of a Railway Construction into the BIM and GIS Environment'. *APPLIED SCIENCES-BASEL*, 13 (9), p. 5591. doi: 10.3390/app13095591.

Belcher, E. J. and Abraham, Y. S. (2023). 'Lifecycle Applications of Building Information Modeling for Transportation Infrastructure Projects'. *Buildings*. MDPI, 13 (9), p. 2300.

Borjigin, A. O., Sresakoolchai, J., Kaewunruen, S. and Hammond, J. (2022). 'Digital Twin Aided Sustainability Assessment of Modern Light Rail Infrastructures'. *FRONTIERS IN BUILT ENVIRONMENT*, 8, p. 796388. doi: 10.3389/fbuil.2022.796388.

Boyes, G. A., Ellul, C. and Irwin, D. (2017). 'EXPLORING BIM for OPERATIONAL INTEGRATED ASSET MANAGEMENT – A PRELIMINARY STUDY UTILISING REAL-WORLD INFRASTRUCTURE DATA'. in. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 49–56. doi: 10.5194/isprs-annals-IV-4-W5-49-2017.

British Standards Institution. (2008). 'PAS 55-1:2008: Asset management - Optimized maintenance of physical assets - Part 1: Terminology and definitions'. London: British Standards Institution.

British Standards Institution. (2013). 'PAS 1192-2:2013: Specification for information management for the capital/delivery phase of construction projects using building information modelling'. London: British Standards Institution.

Chen, Y., Hu, S. and Tang, J. (2016). 'Application research of subway station design based on BIM technology'. *RISTI - Revista Iberica de Sistemas e Tecnologias de Informacao*, 2016 (E8), pp. 411–420.

Cheng, D., Shen, P., Li, W., Wu, X., Chang, J., Zhang, M. and Luo, Q. (2019). 'Application Research of BIM in the Whole Life Cycle of Metro Station'. in. *ICIIBMS 2019 - 4th International Conference on Intelligent Informatics and Biomedical Sciences*, pp. 79–85. doi: 10.1109/ICIIBMS46890.2019.8991544.

Ciccone, A., Di Stasio, S., Asprone, D., Salzano, A. and Nicolella, M. (2022). 'Application of openBIM for the Management of Existing Railway Infrastructure: Case Study of the Cancello-Benevento Railway Line'. *SUSTAINABILITY*, 14 (4), p. 2283. doi: 10.3390/su14042283.

Dan, W. (2021). 'Organization and Management System Innovation of BIM Life Cycle Management'. in. *IOP Conference Series: Earth and Environmental Science*. doi: 10.1088/1755-1315/787/1/012193.

Efanov, D., Shilenko, A. S. and Khoroshev, V. V. (2020). 'Digital Modeling in Railway Infrastructure and Rolling Stock Objects at all Stages Life Cycle: Features'. in. *Proceedings - 2020 International Russian Automation Conference, RusAutoCon 2020*, pp. 29–34. doi: 10.1109/RusAutoCon49822.2020.9208088.

Floros, G. S., Ruff, P. and Ellul, C. (2020). 'IMPACT of INFORMATION MANAGEMENT during DESIGN &CONSTRUCTION on DOWNSTREAM BIM-GIS INTEROPERABILITY for RAIL INFRASTRUCTURE'. in. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 61–68. doi: 10.5194/isprs-annals-VI-4-W1-2020-61-2020.

Gajera, M. (2017). *Building Information Modeling-BIM Life-cycle and Facilitates Management*. *Quora*. Available at: https://cadoutsourcingservice.quora.com/Building-Information-Modeling-BIM-Life-cycle-and-Facilitates-Management (Accessed: 9 April 2024).

Garramone, M., Tonelli, E. and Scaioni, M. (2022). 'A MULTI-SCALE BIM/GIS FRAMEWORK FOR RAILWAYS ASSET MANAGEMENT'. *MEASUREMENT, VISUALISATION AND PROCESSING IN BIM FOR DESIGN AND CONSTRUCTION MANAGEMENT II*. Edited by R. Shults and A. Dlesk. *ISPRS WG V/7 Conference on Measurement, Visualisation and Processing in BIM for Design and Construction Management II*, Prague, CZECH REPUBLIC (International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences), 46–5 (W1), pp. 95–102. doi: 10.5194/isprs-archives-XLVI-5-W1-2022-95-2022.

Gebken, L., Drews, P. and Schirmer, I. (2019). 'Enhancing the building information modeling lifecycle of complex structures with IOT: Phases, capabilities and use cases'. in. *Proceedings of the Annual Hawaii International Conference on System Sciences*, pp. 5929–5938. Available at: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85108252696&partnerID=40&md5=df4dba21b0362622f990c2a5dca7882e.

Großauer, K., Huis, M., Jedlitschka, G., Matt, R., Mulitzer, G. and Zwittnig, G. (2022). 'New built and refurbished railway tunnels: Data delivery from BIM Models to OEBB's asset management'. *Geomechanics and Tunnelling*. Wiley Online Library, 15 (2), pp. 190–200.

Gu, S., Liu, B., Lv, X., Li, H. and Wang, R. (2022). 'Research on Data Exchange Schema for Railway Infrastructure'. *2022 IEEE 6TH ADVANCED INFORMATION TECHNOLOGY, ELECTRONIC AND AUTOMATION CONTROL CONFERENCE (IAEAC)*. *6th IEEE Advanced Information Technology, Electronic and Automation Control Conference (IEEE IAEAC)*, Beijing, PEOPLES R CHINA (IEEE Advanced Information Technology Electronic and Automation Control Conference-IAEAC), pp. 1793–1798. doi: 10.1109/IAEAC54830.2022.9930074.

Han, X., Yuan, F., Hou, L. and Liu, W. (2018). 'Analysis on the BIM application in the whole life cycle of railway engineering'. in. *ICRT 2017: Railway Development, Operations, and Maintenance - Proceedings of the 1st International Conference on Rail Transportation 2017*, pp. 331–338. doi: 10.1061/9780784481257.033.

Hartung, R., Senger, L., Arpe, J. and Klemt-Albert, K. (2020). 'Evaluation of structural health monitoring systems in bridge engineering for increase of safety in operations'. in. *Proceedings of the 30th European Safety and Reliability Conference and the 15th Probabilistic Safety Assessment and Management Conference*, pp. 4687–4694. doi: 10.3850/978-981-14-8593-0_4702-cd.

Hartung, R., Senger, L. and Klemt-Albert, K. (2019). 'Linking building information modeling and structural health monitoring for reliable railway infrastructure'. in *Proceedings of the 29th European Safety and Reliability Conference (ESREL)*, pp. 596–603.

International Organization for Standardization. (2018). 'ISO 19650-1:2018: Organization and digitization of information about buildings and civil engineering works, including building information modelling (BIM) — Information management using building information modelling — Part 1: Concepts and principles'. International Organization for Standardization (ISO) and British Standards Institution (BSI).

Jaskula, K. and Papadonikolaki, E. (2021). 'Blockchain use cases across entire lifecycle of a built asset: a review'. in *EC3 Conference 2021*. ETH, pp. 19–26.

Kaewunruen, S., AbdelHadi, M., Kongpuang, M., Pansuk, W. and Remennikov, A. M. (2023). 'Digital Twins for Managing Railway Bridge Maintenance, Resilience, and Climate Change Adaptation'. *SENSORS*, 23 (1), p. 252. doi: 10.3390/s23010252.

Kaewunruen, S. and Lian, Q. (2019). 'Digital twin aided sustainability-based lifecycle management for railway turnout systems'. *JOURNAL OF CLEANER PRODUCTION*, 228, pp. 1537–1551. doi: 10.1016/j.jclepro.2019.04.156.

Kaewunruen, S., Peng, S. and Phil-Ebosie, O. (2020). 'Digital Twin Aided Sustainability and Vulnerability Audit for Subway Stations'. *SUSTAINABILITY*, 12 (19), p. 7873. doi: 10.3390/su12197873.

Kaewunruen, S., Sresakoolchai, J. and Lin, Y.-H. (2021). 'Digital twins for managing railway maintenance and resilience [version 2; peer review: 2 approved]'. *Open Research Europe*, 1. doi: 10.12688/openreseurope.13806.2.

Kurwi, S., Demian, P. and Hassan, T. (2017). 'Integrating BIM and GIS in railway projects: A critical review'. Loughborough University.

Liu, W. (2022). 'Intelligent Identification and Construction System of Prefabricated Tunnel Structure Based on BIM Technology'. *ADVANCES IN MULTIMEDIA*, 2022, p. 9173929. doi: 10.1155/2022/9173929.

Liu, Y., Lin, H., Zhao, Z., Bai, W. and Hu, N. (2023). 'Research on the Visualization of Railway Signal Operation and Maintenance Based on BIM + GIS'. *Sensors*, 23 (13). doi: 10.3390/s23135984.

Love, P. E., Zhou, J., Matthews, J., Lavender, M. and Morse, T. (2018). 'Managing rail infrastructure for a digital future: Future-proofing of asset information'. *Transportation Research Part A: Policy and Practice*. Elsevier, 110, pp. 161–176.

Lv, S. (2018). 'Study on BIM modeling method of bridge railway integration based on Revit and Civil3D'. in *2018 7th International Conference on Energy and Environmental Protection (ICEEP 2018)*. Atlantis Press, pp. 214–220.

McKenna, T., Minehane, M., O'Keeffe, B., O'Sullivan, G. and Ruane, K. (2017). 'Bridge information modelling (BrIM) for a listed viaduct'. *Proceedings of the Institution of Civil Engineers: Bridge Engineering*, 170 (3), pp. 192–203. doi: 10.1680/jbren.16.00007.

National Building Specification. (2021). *What is BIM?* Available at: https://www.thenbs.com/knowledge/what-is-building-information-modelling-bim (Accessed: 17 January 2024).

Pasetto, M., Giordano, A., Borin, P. and Giacomello, G. (2020). 'Integrated railway design using Infrastructure-Building Information Modeling. the case study of the port of Venice'. in. *Transportation Research Procedia*, pp. 850–857. doi: 10.1016/j.trpro.2020.02.084.

Sanchez, A., Kraatz, J., Hampson, K. and Loganathan, S. (2014). 'BIM for sustainable whole-of-life transport infrastructure asset management'. in *Proceedings of the 2014 IPWEA Sustainability in Public Works Conference*. The Institute of Public Works Engineering Australasia (IPWEA), pp. 1–7.

Shin, M.-H., Jung, J.-H. and Kim, H.-Y. (2022). 'Quantitative and Qualitative Analysis of Applying Building Information Modeling (BIM) for Infrastructure Design Process'. *Buildings*. Multidisciplinary Digital Publishing Institute, 12 (9), p. 1476. doi: 10.3390/buildings12091476.

Shin, M.-H., Kim, H.-Y. and Liao, J.-F. (2024). 'Performance Measurement and Analysis of Building Information Modeling (BIM) Applications in the Railway Infrastructure Construction Phase'. *Applied Sciences*. Multidisciplinary Digital Publishing Institute, 14 (2), p. 502. doi: 10.3390/app14020502.

Sresakoolchai, J. and Kaewunruen, S. (2021). 'Integration of building information modeling (BIM) and artificial intelligence (AI) to detect combined defects of infrastructure in the railway system'. in *Resilient Infrastructure: Select Proceedings of VCDRR 2021*. Springer, pp. 377–386.

Sresakoolchai, J. and Kaewunruen, S. (2023). 'Track Geometry Prediction Using Three-Dimensional Recurrent Neural Network-Based Models Cross-Functionally Co-Simulated with BIM'. *SENSORS*, 23 (1), p. 391. doi: 10.3390/s23010391.

Wan, W., He, Y., Liu, J., Lu, H. and Zhang, H. (2020). 'Application of "BIM+" Architecture Based on Cloud Technology in Intelligent Management of Rail Transit'. in. *Resilience and Sustainable Transportation Systems - Selected Papers from the 13th Asia Pacific Transportation*

Development Conference, pp. 474–484. doi: 10.1061/9780784482902.055.

Wang, N., Satola, D., Wiberg, A. H., Liu, C. and Gustaysen, A. (2020). 'Reduction Strategies for Greenhouse Gas Emissions from High-Speed Railway Station Buildings in a Cold Climate Zone of China'. *SUSTAINABILITY*, 12 (5), p. 1704. doi: 10.3390/su12051704.

Wang, X., Love, P. E. D., Kim, M. J., Park, C.-S., Sing, C.-P. and Hou, L. (2013). 'A conceptual framework for integrating building information modeling with augmented reality'. *Information Technologies in Safety Management*, 34, pp. 37–44. doi: 10.1016/j.autcon.2012.10.012.

Wang, Y., Wang, Z., Ma, T., Li, G. and Tie, H. (2022). 'Research on the Realization Path of Railway Intelligent Construction Based on System Engineering'. *SUSTAINABILITY*, 14 (11), p. 6945. doi: 10.3390/su14116945.

Xiahou, X., Li, K., Li, F., Zhang, Z., Li, Q. and Gao, Y. (2022). 'AUTOMATIC IDENTIFICATION AND QUANTIFICATION OF SAFETY RISKS EMBEDDED IN DESIGN STAGE: A BIM-ENHANCED APPROACH'. *Journal of Civil Engineering and Management*, 28 (4), pp. 278–291. doi: 10.3846/jcem.2022.16560.

Zheleznov, M. (2021). 'Development of the information modeling (BIM) conceptusing Big Data technologies for the implementation of the life cycle management system for capital construction of transport infrastructure'. in. *E3S Web of Conferences*. doi: 10.1051/e3sconf/202128104010.

INCENTIVE MECHANISM FOR IPD WITH MULTIPARTY CONTRACT BY USING A PROBABILISTIC COST AND RISK SOFTWARE

Simon Christian Becker[1], Carl Philipp Friedinger[1], and Philip Sander[1]

[1]University of the Bundeswehr Munich, Neubiberg, Germany

Abstract

Participants of major projects have different interests in each tunnel project. A solution is to align the interests using the Integrated Project Delivery (IPD) with an incentive and multiparty contract (MPC). There is no paper in the literature on how to design an incentive for MPC. Starting with a deductive approach, an incentive where all participants benefit the same is designed. An inductive approach is used for validation. Furthermore, for the inductive approach, a risk software tool is used, and a separately developed Excel sheet. Finally, the mechanism's mode of action is presented and validated.

Introduction

Significant cost and schedule overruns in major projects especially in tunnel construction show that risk management is often not given the importance and necessary integration in project management that it should have (Flyvbjerg 2014). To be able to measure and control costs as well as deadlines against the defined targets, it is necessary to assess costs and risks transparently and take them into account appropriately. Nevertheless, especially in tunnel construction, there are always unforeseen and unknown risks that make this difficult.

To align the interests of all stakeholders seems like an unrealizable task. The reason is that the interests are so different. The owner aims to finish the project as soon as possible whereas the contractor is willing to do almost anything to maximize his profit. Given these almost conflicting interests, it takes all the more effort to complete tunneling projects on time and within budget. The traditional contract format fixed price transfers risks from the owner to the construction company (Becker 2022). However, the result is a high potential for claims. As described in the beginning the solution is to align the interests of all the participants and a good way to do that is an incentive contract.

Integrated Project Delivery (IPD) is used for major projects to align these objectives. Therefore, a multiparty contract is used. This means that the risks are shared and there is also an incentive mechanism to align the objectives of all parties involved.

Since there's no sufficient research on incentive mechanisms for IPD, this article will show how an incentive mechanism can be developed, and in which all participants participate in the project in proportion to their direct project costs.

To this end, the basics of IPD, multiparty contracts, and incentive contracts are first outlined. The research gap is highlighted and the software tool used for cost and risk calculation is introduced.

The description of the development of an incentive mechanism for MPC follows. First, it is described using a deductive approach.

The deductive description is succeeded by an example tunnel project (inductive). This example project is used to validate the theoretical concept.

Integrated Project Delivery (IPD)

The basic idea of IPD is to enable better handling of major projects. The aim is to achieve a faster and cheaper construction process while increasing quality. With this form of execution, an integrated execution team consists of at least the client, planners, and construction companies. They work cooperatively and project-related. All those involved in the project should be aware of the client's objectives from the outset and jointly develop project goals so that everyone involved is aware of the project requirements, and the best possible solutions to achieve the set objectives (Cheng et al. 2019).
To be able to consider identifiable risks and implement solutions cost-effectively in the planning phase, all available knowledge has to be taken into account (Warda 2019).

The core principles for IPD can be stated as follows (Ahmad et al. 2019):

- Early Contractor Involvement of core stakeholders: client, contractor, and planner prior to the start of planning (Friedinger and Becker 2023),
- joint risk management,
- incentive mechanism,
- collaborative working methods,
- joint decisions and
- conflict management.

In the following, the function of a multiparty contract is described in order to better understand the function of the delivery model.

Multiparty contract (MPC)

A multiparty contract is a contract between at least three parties: owner, planner, and construction company. This means that all parties involved in construction and planning are bound together by a uniform contract with each other. In addition, other project participants such as specialist planners, specialized finishing trades, subcontractors, or independent consultants can be included in the contract.
The multiparty contract also promotes cooperation in terms of liability and combines innovations. The incentive

contract realizes a corresponding alignment of goals between the parties. To achieve the alignment the incentive of the MPC must be designed in a way that everybody participates according to the contribution regarding the total costs of the project.
The incentive mechanism for the use case is shown below. The success of an IPD depends largely on the correct choice of the target cost. This must be done individually based on the results of the probabilistic risk analysis and the integral consideration of cost and schedules (Sander et al. 2022, Becker and Roman-Müller 2022).

Focus on Cost Plus Incentive Fee

A Cost Plus Incentive Fee (CPIF) is used when an objective relationship can be established between the fee and performance measures such as actual costs, delivery dates, and performance benchmarks. In the case of highly uncertain and speculative construction projects, it's necessary to use this kind of contract. The owner assumes the risks inherent in the contract-benefiting if the actual cost is less than the expected cost-losing if the work cannot be completed within the expected cost of performance (Becker and Sander 2023; Kerzner 2022).

Explorative Literature Analysis for finding the research gap

By using Google, Google Scholar, Scopus, and Opac+ UniBw an explorative literature analysis was made.

The search strings for this were:

- Incentive mechanism for multiparty contract
- Multiparty contract and incentive mechanism
- Incentive Mechanism for IPD
- IPD and Incentive Mechanism

Different approaches to incentive design were found during the research. These ranged from blockchain applications to theories regarding incentive design.
However, it was not possible to find a concrete implementation of an incentive mechanism for the IPD using an MPC. Therefore, there is a research gap in the area of incentive design for MPCs.
After finding the research gap the basics for the cost and risk calculation will be described. This will be later on needed for the description of the deductive and inductive approach.

Integral Modelling of Cost, Deadlines, and Risks

Cost Components

The use of cost components that build on each other aims to create cost transparency by specifying a clear cost structure that can be applied from an early planning stage to construction completion. The main cost components are (Sander and Becker 2023):

- Base Cost (B): Cost if „everything goes according to plan", without reserves for risks or approaches for escalation (price increase).
- Risk (R): Cost resulting from threats and opportunities that can occur but are not certain to occur (probability of occurrence).
- Escalation (E): Cost resulting from the forecast price increase.
- General Business Expenses (GBE): Include all costs that cannot be directly allocated to this specific construction contract but are incurred by the company as a whole.
- Profit (P): Amount that remains from sales - i.e. the total income of a company - after deducting all costs.

The sum of Base Cost, Risk, and Escalation are the direct project costs (DPC) (1).

$$DPC = B + R + E \qquad (1)$$

The sum of the DPC with the general business expenses and the profit is the Target Cost (TaC) (2).

$$TaC = DPC + GBE + P \qquad (2)$$

RIAAT (Risk Administration and Analysis Tool)

Description of the Software Application

Risk Administration and Analysis Tool (RIAAT) is a desktop application and therefore a stand-alone application.

RIAAT was developed to manage and integrate cost, risk, and schedule analysis for large-scale construction projects.
RIAAT considers the interdependence of cost and schedule. Time-related costs, risk impact, incentive fees, etc. are factored in to optimize your project in terms of cost and schedule (Sander et al. 2021; RIAAT 2024). The benefits of RIAAT are:

- Build a schedule including risks and uncertainties.
- Link schedule and cost using drag-and-drop.
- Consider cost caused by schedule risks.
- There's never only one critical path. Takes every option into account with multiple critical paths.

Figure 1 shows the connection between Cost, Risk, and Schedule and the result as an example of the delay cost.

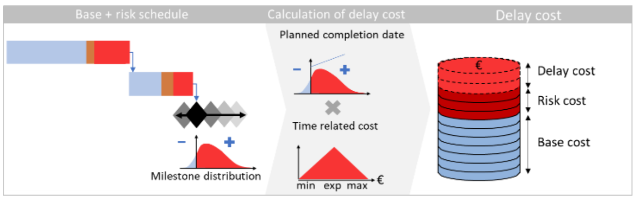


Figure 1: Example integrated calculation of delay cost (RIAAT 2024)

RIAAT bases the decisions on the best risk-benefit ratio shown in Monte Carlo simulations creating probability distributions that allow you to assess the level of probable cost or time overruns with regard to your defined budget or milestone date.

The benefits are:

- Consider uncertainties at all levels.
- Use ranges (bandwidths) instead of single deterministic numbers, and
- Bottom-up aggregation for transparent results.

Both in the business and financial sector and in the construction industry, a statistically determined fractile value of the value at risk (VaR) is used to quantify this monetary sum. To determine this fractile value, a certain probability value must be set depending on the assessment of the complexity of the project and the risk appetite or risk acceptance (Bergmeister 2021; Bergmeister 2022).

With X as a variable with the distribution function (3).

$$F_x(x) = P(X \leq x) f\text{ür } x \in \mathbb{R} \qquad (3)$$

Random losses are represented by the positive values of the random variable X inverse of the distribution F_X^{-1} and the confidence level with $\alpha \in (0,1)$.
The VaR will be defined in (4) (e.g. Figure 2).

$$VaR_\alpha(X) = F_X^{-1}(\alpha) \qquad (4)$$

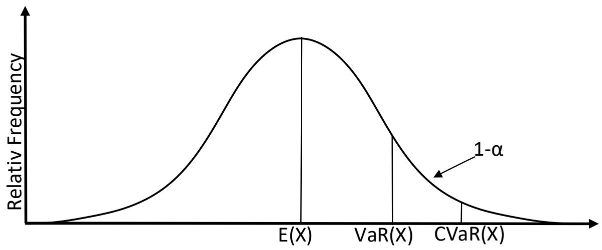


Figure 2: Value at Risk

Deductive Approach for Designing the MPC Incentive

Accordingly, it is important to design the incentive for the MPV in such a way that everyone participates equally in this contract. This means that everyone participates in the potential bonus according to their share of the production costs of the project. The following is an example of how the DPC are jointly pooled and how the bonus is subsequently shared.
Figure 3 shows all contractual partners (CP). The representation of CP 1 to CP 3 refers to the Owner. The client is not considered further in this example, as it is assumed that the client does not contribute to the costs with the MPV. First, each contractor determines its DPC. As mentioned above, this is done using probabilistic methods and in the form of a work calculation.

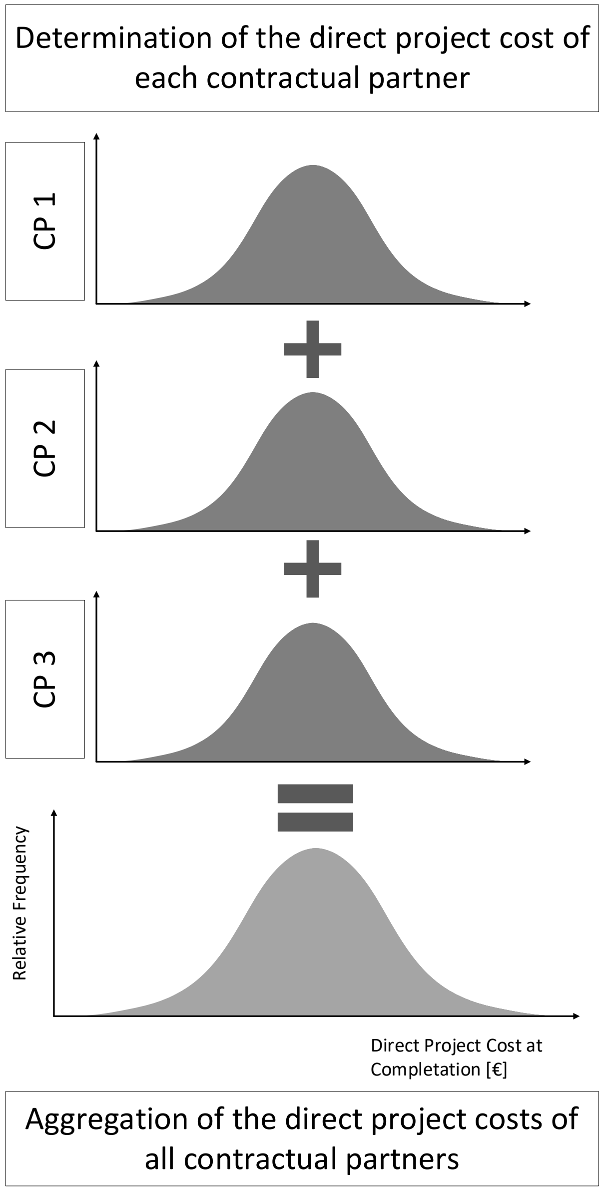


Figure 3: Determining the manufacturing costs of all contractual partners

Once the DPC for the individual contract partners have been determined, the functions (in this case three, e.g. Figure 3) are aggregated. This creates a function for the manufacturing costs. All contract partners must then decide on a P-value for the DPC. Figure 4 shows a distribution and a P-value of 50. This stands for a probability of 50% (P50) that the budget will be underrun. Conversely, there is a 50% probability that the budget will be overrun.

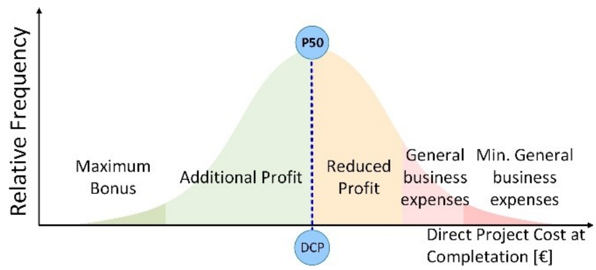


Figure 4: Determination of the P-value of the individual contractual partners

Once the P value has been set, the TaC can be assigned again for all CPs. Involved in the TaC are the GBE and P. This develops the incentive mechanism for the CPs.

Figure 5 depicts an incentive mechanism. The horizontal axis shows the potential final costs for the client, and the vertical axis shows the compensation of the contractor. The light blue dashed line shows the owner/contractor share-ratio. In this example, the ratio was set at 50/50 across all areas. This means that if the TaC is undercut or exceeded, the deviation is split equally between the two partners. Additionally, Figure 5 shows the target cost and the target profit. Both are set in the contract.

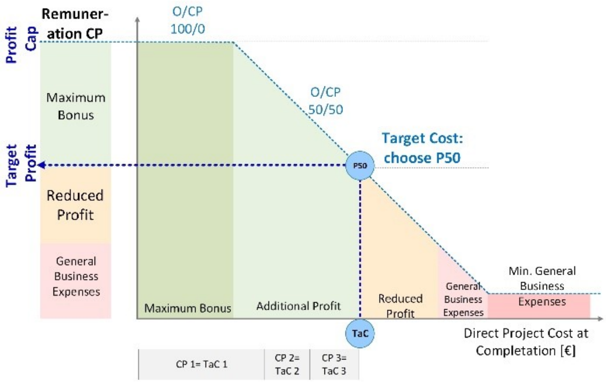


Figure 5: Aggregation of distributions and determination of target costs

Once the construction work has been carried out and the construction project has been completed, billing can take place (theoretical). Figure 6 compares the actual Construction Costs (CC) with the previously calculated DPC and the real TC. Below the horizon line is the area of the DPC.

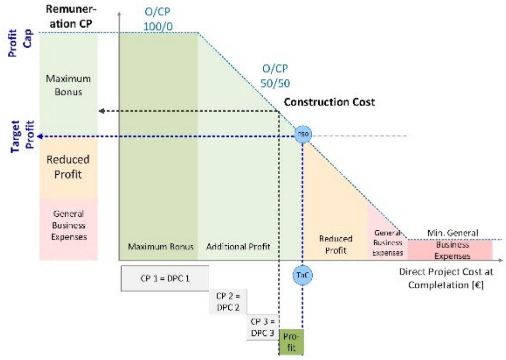


Figure 6: Settlement and distribution of the bonus

Figure 6 shows a scenario at the end of the project in which the Construction Cost (CC) is lower than the agreed TaC. Due to the lower CC, the grey dashed line shifts to the left. This can be achieved by, for example, increased efficiency, the use of new, innovative construction methods, etc. In this case, the contractor generates a bonus of 50% of the savings in addition to the target profit (increased profit). The remaining 50% of the savings goes to the client (Becker and Sander 2023).

Use Case – Tunnel Project

Project Description

A fictitious sample project in this paper is used to illustrate the process. It is based on experience from major European railway base tunnels. This 14-km twin-bore tunnel consists of several Tunnel Boring Machine (TBM) drives as well as Drill & Blast (D&B) drives in different geological formations, an access shaft, emergency stops, various cross cuttings, and inner linings. The project is separated into five lots. Lot 0 is for the access road construction, lot 1 is for the crosscut and the New Austrian Tunneling Method (NATM), lot 3 is for the access shaft, and lot 4 is for the underground refuge. So for the example project are four contractors needed. Lot 0 is not considered any further in the project description. It is no longer relevant in the consideration.

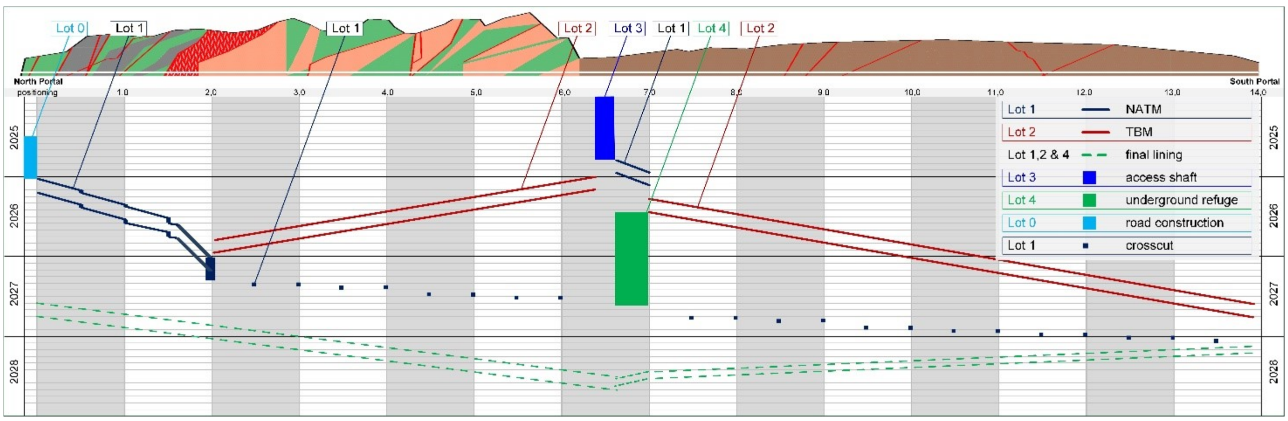


Figure 7: Path-time diagram fictional tunnel project

Inductive Approach

After the description of the deductive approach in the section above, this section shows the inductive approach by using a fictitious tunnel project. The IPD is applied with an MPC. The remuneration model is a cost-plus incentive fee. The authors created a fictitious tunnel and determined the base costs, risks, and price increases. The information was determined in a workshop and validated and adjusted accordingly by the third author. The data was entered into the RIAAT software and analyzed probabilistically. In the process, 35 risks have been included in the model as examples. The individual lots were evaluated using another specially developed Excel sheet.

The DPC (production costs) were determined for lots 1-4 and a VaR of 50% was set. After determining the DPC the GBE and the P, together a percentage premium of 8%, were multiplied by the aggregated DPC. The calculation of the DPC, the TaC, Savings, and bonus allocation are shown below.

Description of the procedure for a MPC

First, the DPC of all parties of the multiparty contract are calculated. In our case, there are four contractors. The direct project costs are composed of the base costs, risks, and price increase (Sander and Becker 2023).

Together with the determination of the direct project costs distribution functions are created, one for each member of the multiparty contract (e.g. Figure 8). Then, the P-values of the contractual partners are determined. The P-value indicates the probability of the target costs occurring. For public construction projects, as is almost always the case in tunnel construction, this value is between 50% (P50) and 90% (P90) (e.g Bergmeister 2021).

In Figure 8, the P-value was set to P=50 for this example. So, the direct project costs are for Contractor Lot 1: 90.3 Mio. €, Contractor Lot 2: 375.9 Mio. €, Contractor Lot 3: 11.1 Mio €, and Contractor Lot 4: 25.7 Mio. € (see also Table 1). The calculation was done with the software RIAAT and included all costs, risks, price increases, extra costs for schedule delays, and so on.

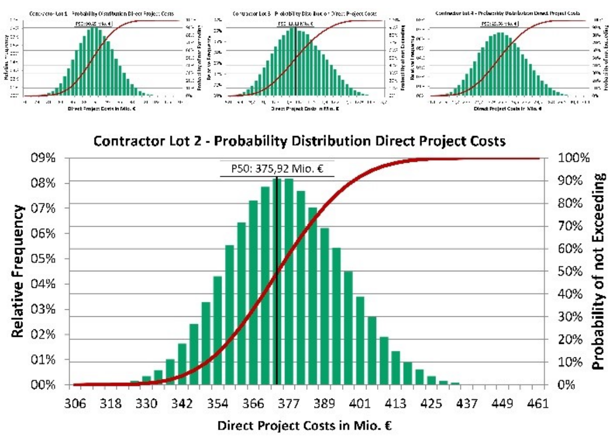


Figure 8: Distribution functions of direct project costs for each contractor

The direct project costs are determined by all members of the multiparty contract or all contractors and the client. In this example, the contractual partners would set the direct project costs at 502.9 Mio. € by aggregating all contributions of all participants.

In the next step, the TaC for the project is defined jointly and the incentive mechanism is set. In this use case, the owner and the contractors decided to add 8% to the Direct Project Cost, 5% for General Business Expenses, and 3% for Profit. The jointly agreed 8% is added to the aggregated distributions and the Target Cost is set on P50 at 543.43 Mio. € (see Figure 9). This later serves as the basis for billing.

Figure 9: Probability Distribution Target Cost

Table 1 shows all the results of the example project. As described above, the VaR was set at 50%. The distribution between the client and contractor is 50/50. For example: if €5 million is saved, the client would receive €2.5 million and the contractors would also receive €2.5 million. The Target Cost was described previously. The table also shows jointly agreed DPC in Mio. €, the share of each contractor of the DPC and the agreed general business expenses and profit. This is the foundation for the Target Cost. The CC and the cost of GBE, and Profit based on the TaC are shown as well. Together, they form the Total Cost (ToC) in Mio. €. Also, the savings regarding the DPC are depicted in € and as a percentage. Finally, the table shows the share ratio, the bonus in millions for every contractor, and the bonus share of each contractor regarding their specific Target Cost.

Table 1: Costs and Savings for the Example Tunnel Project

	Contract Lot 1	Contract Lot 2	Contract Lot 3	Contract Lot 4	∑ and Share Owner
Direct Project Costs P50 in Mio. €	90,25	375,92	11,11	25,66	**502,94**
Share of Direct Project Costs in %	*17,9*	*74,7*	*2,2*	*5,1*	*100,0*
General Business Expenses and Profit in %	*8,00*				
General Business Expenses and Profit in Mio. €	7,26	30,26	0,89	2,07	**40,48**
Probability Target Cost	*P50*				
Target Cost in Mio. €	97,52	406,18	12,01	27,72	**543,43**
Construction Costs in Mio. €	89,31	373,55	10,91	24,09	**497,86**
General Business Expenses and Profit based on Construction Costs in Mio. €	7,14	29,88	0,87	1,93	**39,83**
Total Cost in Mio. €	96,45	403,43	11,78	26,02	**537,69**
Savings in Mio. € regarding Direct Project Costs	0,95	2,37	0,20	1,56	**5,08**
Savings in % regarding Contractors Direct Project Costs	*1,05*	*0,63*	*1,80*	*6,10*	***1,01***
Share Ration Savings Owne/Contractor	*50/50*				
Bonus in Mio. €	0,46	1,90	0,06	0,13	**2,54**
Bonus in % regarding Target Cost	*0,47*	*0,47*	*0,47*	*0,47*	*0,47*

(5) shows how to calculate the Direct Project Cost (DPC) for all four contractors in the use case.

$$\sum DPC_{c1-c4} = DPC_{c1} + DPC_{c2} + DPC_{c3} + DPC_{c4} \quad (5)$$

To get the percentage share for any contract (P_{DPCci}) of the DPC see (6).

$$P_{DPCci} = \frac{DPC_{Ci}}{\sum DPC \times 100} \quad (6)$$

With (7) the total of percentage of GBE (P_{GBE}) and contractor Profit (P_{CP}) is calculated.

$$P_{GP} = P_{GBE} + P_{CP} \quad (7)$$

The premium of General Business Expenses (GBP) and Profit can calculate with (8).

$$GBP_{\sum} = DPC_{ci} \times P_{GP} \quad (8)$$

Now, to calculate the Target Cost (TaC) of the project the sum of the DPC is multiplied by sum of the percentage of the GBE and Profit.

$$TaC = \sum DPC_{c1-4} \times [1 + (GBE + P)] \quad (9)$$

The calculation of the Target Cost of each Contractor (S_{TCci}) is shown in (10).

$$S_{TCci} = TaC \times \frac{P_{DPCci}}{100} \quad (10)$$

After the calculation for the project is done, the construction phase of the project follows, and at the end the billing. The contractors will receive the Construction Cost (CC) which are occur in reality. The calculation for the total CC can be seen in (11).

$$\sum CC_{c1-4} = CC_{c1} + CC_{c2} + CC_{c3} + CC_{c4} \quad (11)$$

Now, to get the Total Cost (ToC) of the project the CC are multiplied by the GBE and Profitas shown in (12).

$$ToC = \sum CC_{1-4} \times \left(1 + \frac{P_{GP}}{100}\right) \quad (12)$$

The savings of the different contracts can be calculated by subtracting the CC from the DPC for each contractor (e.g. 13).

$$Sa_{ci} = DPC_{ci} - CC_{ci} \quad (13)$$

To get the savings in percentage for each lot, divide the Savings (Sa_{ci}) and the DPC_{ci}.

$$P_{Saci} = \frac{Sa_{ci}}{DPC_{ci}} \times 100 \quad (14)$$

At the end of the project, the bonus needs to be shared between all contractors. (15) shows the formula. With Sa_{ci}, the percentage of each contractor's P_{DPCci}, and the Share Ratio (SR) the bonus can be calculated.

$$Bonus_{ci} = \sum Sa_{ci} \times P_{DPCci} \times \frac{SR}{100} \quad (15)$$

Finally, the percentage of the Bonus P_{Bonus} results from each $Bonus_{ci}$ in relation to the Target Cost of each contractor (TaC_{ci}) (e.g. 16).

$$P_{Bonus} = \frac{Bonus_{ci}}{TaC_{Ci}} \times 100 \quad (16)$$

As this is an example project, it was assumed that only around 50% of the risks occurred in the example project. As a result, the contractors and the client were able to generate cost savings of €5.08 million. Due to the 50/50 split, €2.54 million is distributed among the contractors involved. The distribution key for the bonus payment depends on the participation of the Targets Cost. Contractor Lot 1, for example, has a share of the target costs of 17.9%. Therefore, his share of the bonus is also 17.9 %. Consequently, Contractor Lot 1 receives € 0,46 million. The distribution of Contractor Lots 2-4 is equivalent.

The distribution of the bonus in % about the target cost is thus equally distributed among all participants. Each one of them has a bonus share of approx. 0.47 % of his TaC. This clearly shows that all participants in the project participate equally in the bonus of the project and thus an actual target alignment between all participants takes place.

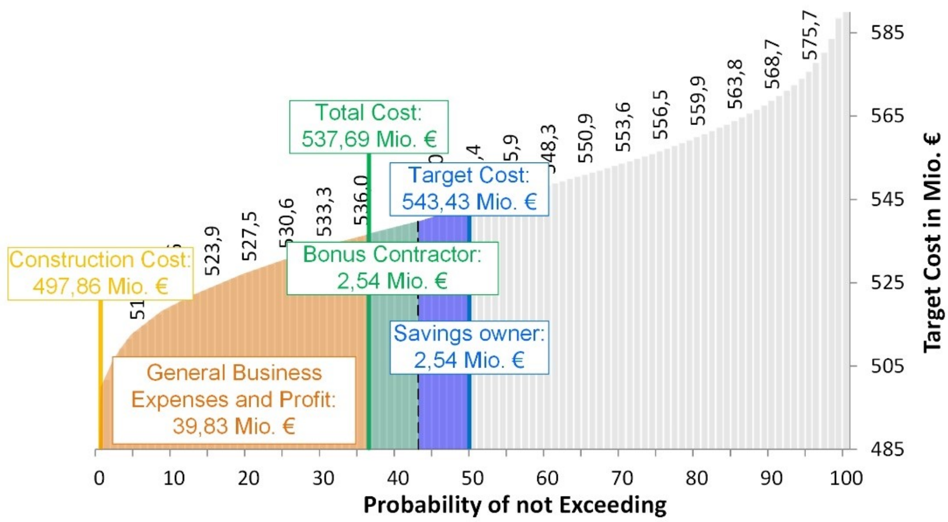


Figure 10: Results of the example construction project with a Lorenz Curve

Figure 10 shows the results with a Lorenz curve. The TaC are 543.43 Mio. € (P50) and were calculated together. The Construction Cost shows the cost that accrued for the tunnel project. The Total Cost are the cost, that the owner has to pay for the construction project (here 537.69 Mio. €). It includes the Construction Cost (here 497.86 Mio. €). and general business expenses as well as profit (here 39.83 Mio. €). In this example, the contractors' bonuses are not included in the Total Cost.
The shares of the savings and bonus are presented in more detail in Figure 11.

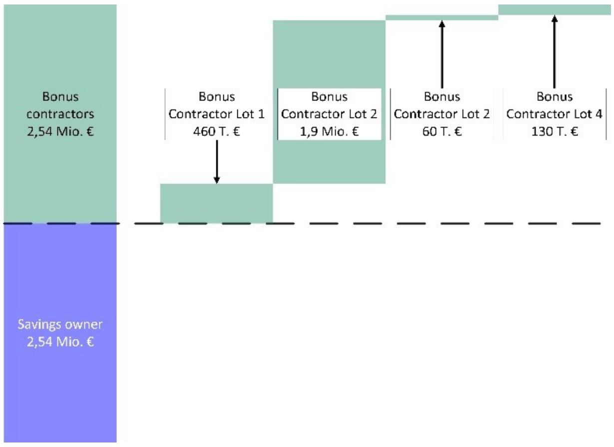


Figure 11: Shares of Savings and bonuses for the example project

According to the share ratio, the savings are split 50/50. The owner gets € 2.54 million and the contractors get the share the same amount among themselves. Every contractor gets his share according to equation (15). The calculation for Lot 1 is as follows (17).

$$Bonus_{c1} = 5.08\, Mio. € \; x \; 17.9\% \; x \; \frac{50}{100} = 460\, T. € \quad (17)$$

The calculations for the other lots are equivalent.

Conclusions

This paper shows the application of risk software for creating an incentive mechanism for the IPD with MPC. The deductive process was validated with an inductive process to check whether the incentive mechanism treats all parties equally.
The example tunnel with several lots and the use of IPD, a cost-plus incentive fee contract, and a multi-party contract were used to demonstrate the use of the software. An incentive system was developed for a multi-party contract whereby all parties involved can participate in the success of the project. This aligns the interests of all parties involved and they work together for the success of the project. By linking the incentive to the respective parties, everyone is always obliged to do their best for the project. A multi-party contract with a cohesive incentive mechanism between all parties therefore leads to joint project success.

Outlook

Incentive mechanisms can be designed in a variety of ways. Currently, the literature generally lacks simple considerations on how all parties participate equally in an MPC. This is particularly important, as otherwise, not all parties will work in the best interest of the project goals. This simple example shows an incentive mechanism for multi-party contracts in which all parties participate equally. In order to further validate the results, further use cases need to be created and the incentive effect needs further investigation. In addition, an application should be used in a large-scale project and the results should be shared and further optimizations incorporated into the research. This would further improve the mechanism of action on the one hand and on the other hand the design of the incentive mechanism.

Acknowledgments

This research paper on the project DigiPeC is funded by dtec.bw – Digitalization and Technology Research Center of the Bundeswehr. dtec.bw is funded by the European Union – NextGenerationEU.

References

Ahmad, Irtishad/Azhar, Nida/Chowdhury, Arindam (2019). Enhancement of IPD Characteristics as Impelled by Information and Communication Technology. Journal of Management in Engineering 35 (1). https://doi.org/10.1061/(ASCE)ME.1943-5479.0000670.

Becker, Simon Christian (2022). Identifizierung von Anreizen in Verträgen nach der VOB/A-EU unter Einbezug von der VOB/B. In: Arbeitsbereich Für Baumanagement, Baubetrieb Und Tunnelbau (Ed.). Tagungsband zum 31. BBB-Assistent:innentreffen Innsbruck 2022. Studia Verlag, 16–29.

Becker, Simon Christian/Roman-Müller, Horst (2022). Integrierte Projektabwicklung (IPA). Schnelleinstieg für Bauherren, Architekten und Ingenieure. Wiesbaden/Heidelberg, Springer Vieweg.

Becker, Simon Christian/Sander, Philip (2023). DEVELOPMENT OF A "PROJECT OBJECTIVE SYSTEM" (POS) TO ALIGN THE INTERESTS OF ALL THE STAKEHOLDERS AND FIND THE RIGHT DELIVERY MODEL. In: Proceedings of the Creative Construction Conference 2023, Creative Construction e-Conference 2023, Keszthely, Hungary, 20.06.2023 - 23.06.2023. Online, Budapest University of Technology and Economics, 252–257.

Bergmeister, Konrad (2021). Holistisches Chancen-Risiken-Management von Großprojekten. Unbekanntes erkennen und handeln. Berlin, Ernst & Sohn.

Bergmeister, Konrad (2022). Models for financing, cost and risk assessment. Major railway tunnel projects in Europe. Berlin, Germany, Ernst & Sohn.

Cheng, L./Osburn/Lee, L. (2019). An Action Guide for Leaders. Paris, Independently.

Flyvbjerg, Bent (2014). What you Should Know about Megaprojects and Why: An Overview. Project Management Journal 45 (2), 6–19. https://doi.org/10.1002/pmj.21409.

Friedinger, Carl Philipp/Becker, Simon Christian (2023). Early Contractor Involvement für öffentliche Auftraggeber – Chancen für eine effizientere Projektabwicklung. https://doi.org/10.17185/DUEPUBLICO/79115.

Philip Sander, Markus Spiegl, John Reilly. Incentive-Based Project Delivery with Fixed Price Incentive Fee Contracts.

RIAAT (2024). RIAAT. Available online at https://www.riaatsoftware.com/ (accessed 1/31/2024).

Sander, Philip/Becker, Simon Christian (2023). Gemeinsames Risikomanagement bei Großprojekten mit der Integrierten Projektabwicklung (IPA). Berlin. DVP-Tagung.

Sander, Philip/Becker, Simon Christian/Friedinger, Carl Philipp/Riemann, Stefan/Ditandy, Michael/Spiegl, Markus (2022). Creating Incentive Mechanisms for Integrated Project Delivery. tunnel (4), 12-23.

Sander, Philip/Becker, Simon Christian/Lammers, Martin/Uphoff, Raimund/Brodehl, Raimund/van Drogenbroeck, Arno (2021). Digital Project Risk Twin – Application for the Construction of U5 East, Hamburg. tunnel (06/2021), 20–29.

Warda, Julius (2019). Die Realisierbarkeit von Allianzverträgen im deutschen Vertragsrecht. Dissertation.

ASSESSING THE BENEFITS OF DIGITAL TWINS TOWARDS THE FORMULATION OF WHOLE LIFE CYCLE VALUE PROPOSITIONS: A SYSTEMATIC REVIEW

Omar Abbas, Nicola Moretti, and Tim Broyd
University College London, London, United Kingdom

Abstract

Digital twins have been advocated for adoption in the construction industry to tackle many of its inherent challenges. There has been a lack of clarity, however, in demonstrating their realized value with some enterprises facing difficulties justifying their investment. This study employs a systematic review method to extract empirical evidence demonstrating the value of digital twins. Some realized Digital Twins' benefits and costs were demonstrated in the articles, but there were no investment appraisal studies that provide a full account of value propositions. Future studies adopting case studies method are recommended to assess the value propositions of digital twins.

Introduction and background

Context

In the era of the fourth industrial revolution (Industry 4.0), businesses have been adopting digital technologies to drive further efficiencies in their processes. The proliferation of digital technologies has been transforming traditional workflows while promoting further productivity gains. The Architecture, Engineering, Construction, and Operation (AECO) industry, however, remains to be one of the least digitized industries (Agarwal, Chandrasekaran and Sridhar, 2016) with a very slow uptake of new technologies (Davies and Harty, 2013).

The AECO industry faces several inefficiencies such as low levels of productivity and a fragmented supply chain (Opoku et al., 2021). As a project-based industry, the delivery process is transient where the output is unique (Gann and Salter, 2000; Morris, 2004), and this inhibits the industry from adopting the benefits of repeatability like that in the manufacturing industry (Fernández-Solís, 2008). Projects are usually initiated for the purpose of creating new value (Winch, 2010).

Capitalizing on different digital technologies, a Digital Twin (DT) has been argued to be a driver of value in the AECO industry. A DT is a real-time virtual replica of a physical asset (Grieves and Vickers, 2017) with its use extending throughout the whole life cycle of the physical asset (Madni, Madni and Lucero, 2019). It employs a set of different technologies with the aim of achieving bidirectional flow of data between the physical asset and digital counterpart (Alizadehsalehi and Yitmen, 2023). This automated flow of data along with the use of data analytics can create several benefits including reduced rework, better energy management, and lower construction costs (Opoku et al., 2021). Despite their attested value and advancements in other industries such as Aerospace, and Manufacturing (Xie et al., 2023), the AECO industry has been slowly adopting this trend (Opoku et al., 2022).

In the context of their potential value, DTs can tackle many of the AECO industry's underlying inefficiencies and issues. However, due to the lack of clarity in their value propositions (Opoku et al., 2023), they have not been widely adopted in practice. Furthermore, construction enterprises find difficulties justifying the additional DT investment. It was even asserted that the value of DTs is yet to be defined (Pregnolato et al. 2022), and Shahzad et al. (2022) suggested the need for demonstrating their benefits as new technologies are adopted after their cost-effectiveness is demonstrated in real-world projects.

Research aim, question and objectives

The aim of this study is to investigate the reasons behind the lack of DT adoption in the AECO industry given the suggested lack of value propositions of DTs. Consequently, this research is motivated by the lack of a comprehensive identification of the actual value of DTs during the life cycle of built assets. Towards tackling the above problem, the proposed research question is *what is the value realized through the use of DTs in AECO projects?*

To tackle this question, the objectives of this study are threefold: (1) Investigating the value of DTs by identifying the associated benefits and costs with their use; (2) assessing the availability of empirical evidence supporting the claim of those benefits and costs; and (3) the identification of the life cycle stages during which value can be materialized.

Research background

Firstly, a clear boundary should be drawn between the difference of value and benefits. The two terms have been used interchangeably in literature (Laursen and Svejvig, 2016) while benefits are only a component of the total value. In basic terms, the realized value equates to the benefits less the cost depleted (Morris, 2013; Project Management Institute, 2019). Thus, the true value of an asset or a system involves an interplay between the accrued benefits and incurred costs. Benefits are those advantages provided to stakeholders (Ward and Daniel, 2006) that should be perceived by some stakeholders as a positive outcome (Bradley, 2010). The benefits derived from Information System (IS) investments can be either tangible, semi-tangible or intangible (Becerik-Gerber and Rice, 2009). Tangible and semi-tangible benefits are quantifiable while intangible benefits are non-quantifiable. Costs can be classified as direct or indirect.

One of the main desired benefits of the DT concept application in construction projects is the creation of value to deliver better project outcomes. Value creation has been denoted to be one of the nine fundamental properties of a DT as per the general concepts described in the Gemini Principles (Bolton et al., 2018). Within different architectures proposed for DTs (Ferré-Bigorra, Casals and Gangolells, 2022; Lu et al., 2020), the service layer is where DTs can deliver value to different stakeholders. At this layer, the user can interact with the system, and different functions can be deployed such as energy management and space utilization (Lu et al., 2020). During the design phase, DTs have the potential to reduce the possibility of rework, and deliver sustainable outcomes (Opoku et al., 2021). Construction costs can be minimized as well while improving the quality of deliverables (Opoku et al., 2021). Better decision making and predictive maintenance were some potential benefits suggested during the operation and maintenance phase (Khajavi et al., 2019). The literature was not very clear on demonstrating the cost behind the utilization of DTs. Costs were generally represented as barriers to adoption where difficulties in initial investments were stated (Madni, Madni and Lucero, 2019; Opoku et al., 2021).

For those reasons, the next section of this study will focus on collecting and analyzing literature sources related to the value of DTs realized by different stakeholders in the AECO industry. The method used for systematic collection of literature will be outlined, followed by the frameworks used for benefits and costs identification.

Methods and tools

To address the outlined research question, this paper will utilize a systematic review to organize and analyze the literature production in the realm of DT value delivery. This collation of evidence will follow the guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) (Page et al., 2021).

Scopus and Web of Science (WoS), well regarded as robust academic databases, were used for article identification. The search was carried out using the title/abstract/keywords field with three blocks of keywords:

(1) Digital twins: (digital twin*);
(2) Value: ("cost-benefit" OR "benefit-cost" OR "cost-effectiveness" OR "return on investment" OR "financial*" OR "feasibility" OR "value for money" OR cost* OR benefit* OR value* OR "business case");
(3) AECO industry: (aec* OR "construction industry" OR "built environment").

Given the nascence of DT research, and to conduct a comprehensive review of the literature, no filters were added for the date range. The results were filtered for journal articles and the search returned 99 articles in Scopus, and 87 in WoS.

The identification step was followed by the screening stage to assess the relevance of obtained articles. Initially, the obtained records were imported into a reference management software to remove duplicate records. Only articles written in English were retained for further screening. An initial screening of the abstracts and titles of the retrieved articles was followed to ensure that the focus of the publications was on DTs within the Built Environment. At this step, 45 articles were retained for a comprehensive review based on the eligibility criteria defined in Table 1 following which 12 articles were deemed eligible for the systematic review as shown in Figure 1.

Table 1: Inclusion and exclusion criteria

	Inclusion criteria	Exclusion criteria
Methodology	Case study, Survey, Experiment	Literature-based only
Focus	Studies providing some empirical or theoretical evidence regarding the potential benefits or costs of DTs	No identification or evaluation of DTs' benefits or costs
Depth	A comprehensive application of a DT	Study focusing on a particular technology rather than a comprehensive DT implementation

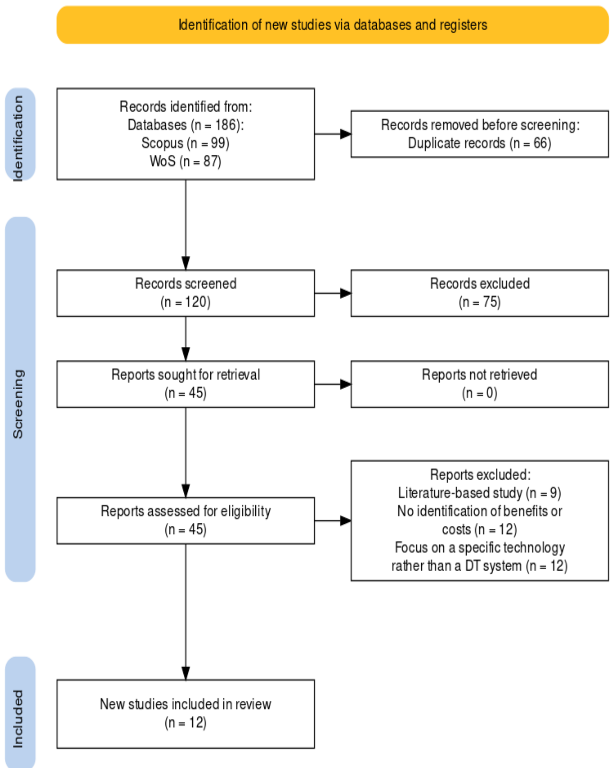


Figure 1: PRISMA screening process. Adapted from Haddaway et al. (2022).

Benefits and costs identification

The Centre for Digital Built Britain (CDBB) has adopted the Five Capitals Model developed by Porrit (2007) to profile the value of DTs (CDBB, 2021). This model has been developed to assess sustainable developments, and it is composed of five value categories namely: Natural, Social, Human, Manufactured and Financial capital. Natural capital includes the natural resources that are fused to provide goods and services. The Social capital includes the systems and institutions that help the Human capital work more productively when working collectively rather than in isolation. Manufactured capital is comprised of the assets that contribute to the production process and not the output. Financial capital is what allows the earlier categories to be traded and owned, and it encompasses their economic benefits and different costs.

To assess the value of DTs, their different benefits will be first identified from the collected literature and classified according to their tangibility. Tangible benefits can be easily quantified, while intangible benefits require a qualitative assessment that would subsequently impact one or more of the tangible benefits by means of a factor or percentage (Irani and Love, 2002). This classification of benefits was adopted by Irani and Love (2002) to develop a frame of reference for Information technology/ Information Systems (IT/IS) investment evaluation. It should be noted that this framework did not provide a taxonomy of IT/IS benefits, and hence it will be used as a reference to identify and cluster benefits. The focus of this process will be on benefits that are tangible and hence quantifiable.

For the costs, the identification process will follow the cost taxonomy developed by Irani, Ghoneim and Love (2006). The costs are classified as direct or indirect. Direct costs are associated with the use of technology components, while indirect costs capture that related to the organization and people dimension. This classification embodies both the technical and social components of a socio-technical system (Bostrom and Heinen, 1977) which could be deemed suitable for capturing a holistic DT utilization. Furthermore, this scope can capture all value categories highlighted in the Five Capitals Model (Porrit, 2007).

Investigation method of studies

Bakis, Kagioglou and Aouad (2006) outlined three main empirical investigation methods for evaluating the benefits of new technologies implementation as case studies, experiments, and surveys. Case studies were identified as the most robust of the three as they place the most emphasis on the context of the realized benefits. Experiments and surveys lack an emphasis on this context where the former imposes control factors for replication, while the latter does not depict how the benefits were created. Additionally, Hakimi et al. (2023) emphasized the need for conducting case studies to prove the business value of DT implementation.

The investigation method used in the selected articles will be highlighted. Articles using more than one investigation method will be highlighted as well. Finally, a further metric will be added to classify whether the study presented some quantifiable findings or not.

Results

Bibliometrics

To understand the broader research trends of DT publications, some quantitative bibliometric analysis was conducted on the 45 articles that were assessed for eligibility. Figure 2 illustrates an overall increasing trend in DT publications over the years.

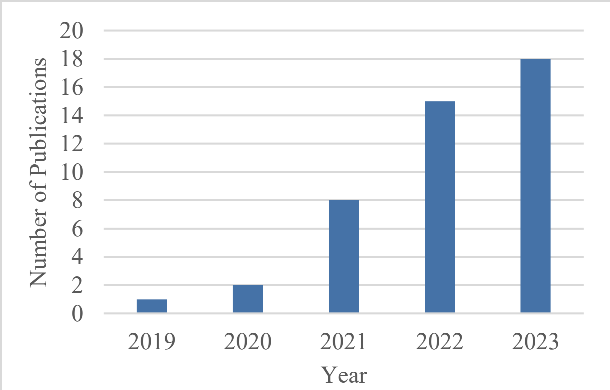


Figure 2: Bar chart showing the number of publications per year.

Based on a minimum number of two occurrences, a co-occurrence for author keywords was created as shown in Figure 3. In total, there were 24 occurrences where digital twins were the most common. There were no clusters for any of the *value* block keywords while the *AECO* industry's keywords appeared near the digital twin's cluster.

Content analysis

The classification of the selected twelve articles is illustrated in Table 2. Most of the studies followed a case study approach with only two of them employing more than one case study. Surveys were only used in three studies, and only one experiment was carried out. Four articles, adopting a case study approach, demonstrated some quantifiable benefits or costs from the use of a DT (Greif, Stein and Flath, 2020; Lin and Cheung, 2020; Love and Matthews, 2019; Zhang et al., 2023)

After reviewing the collected articles, a total of nineteen clusters of benefits were identified as illustrated in Table 3. Different benefits were matched to one cluster despite having different designations in different studies. For instance, increased transparency of information (Ammar et al., 2022), and transparency and data reliability (Esmaeili and Simeone, 2023) were matched to improved information management. The most common benefit was better environmental management followed by improved maintenance. Five other benefits were cited three times while the remaining were cited only once. Some of the presumed benefits were not matched to either of the

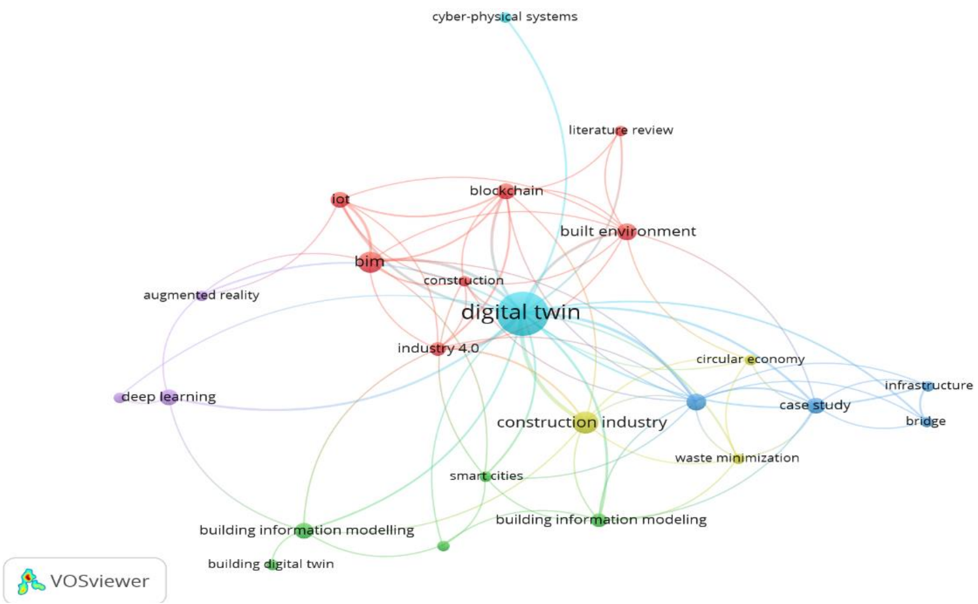


Figure 3: Author keyword co-occurrence map.

suggested benefits in Table 3 as they were essentially applications with different potential benefits but not benefits per se. For instance, Ammar et al. (2022) and Love and Matthews (2019) identified real-time reporting as a benefit while it could lead to other benefits that could be matched to more than one of the benefit clusters.

As for the benefits demonstrating some quantifiable findings, the study conducted by Greif, Stein and Flath (2020) presented a potential cost reduction of 25% in truck costs during silo movements between sites and plant. Another case study focusing on building heritage preservation demonstrated Indoor Air Quality (IAQ) improvements by an average of 9% and 1.2% for a restaurant and an exhibition room, respectively (Zhang et al., 2023). Love and Matthews (2019) illustrated different benefits across several case examples. In an Iron Ore Mine, the studied benefits were reported to result in a reduction cost of 94.25% in documentation cost. The time to produce a drawing was reduced to 2 hours from 40 hours with the use of a DT instead of a traditional Computer-aided Design (CAD) approach. Furthermore, the time required to address Request for Information (RFI) documents can be reduced by 91.67% person-hours in addition to improving the overall information management process of documents. The next case was a Magnetite Iron Ore Processing Plant, and the findings suggested an overall improvement in the information management process. This improvement has led to an improved site management since that design information was directly available for personnel on site. An improvement in installation efficiency as well as a reduction in cabling wastage were also reported, which suggests the delivery of an improved output.

In their next case of an Oil Refinery, improved information management, and improved site management were also reported. Monitoring the progress of person-hours worked on site has improved resource management. The DT was also reported to ensure asset integrity, which suggests an improved output. In their final study of a Rail project, there was a better handover of data from the construction phase to the operation and maintenance. DTs have also demonstrated their effectiveness in digital asset management.

Regarding the costs identified, only three publications outlined some relevant findings. Ammar et al. (2022) identified (1) Data understanding, preparation and usage, (2) Costs of implementation, and (3) Social costs, and it was the only source that identified indirect costs. Esmaeili and Simeone (2023) outlined data collection, platform and interoperability costs. Lin and Cheung (2020) highlighted the hardware cost, and it was the only reference that provided some minor quantification for the cost of wireless sensor networks used.

With respect to the life cycle stage of the studies, the operation and maintenance phase was the most common followed by the construction phase as illustrated in Figure 4. The benefits during the decommissioning phase were only demonstrated once.

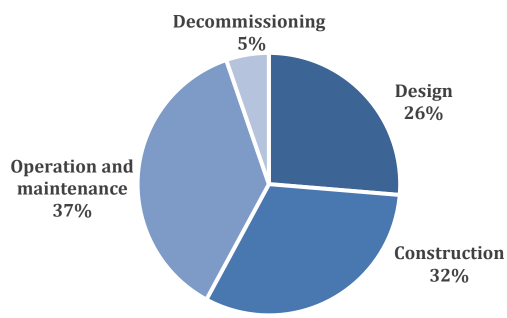


Figure 4: Distribution of Life cycle stages with demonstrated benefits from selected articles.

Table 2: Investigation method presented in the selected articles.

No. of Articles	Investigation method
6*	Case Study
2†	Case studies
1‡	Case study and survey
1§	Experiment
2**	Survey

*: (Greif, Stein and Flath, 2020; Lin and Cheung, 2020; Ogunseiju et al., 2021; Tagliabue et al., 2021; Tita et al. 2023; Zhang et al., 2023)
†: (Esmaeili and Simeone, 2023; Love and Matthews, 2019)
‡: (Ammar et al., 2022)
§: (Adibfar and Costin, 2022)
**: (Meng, Das and Meng, 2023; Shahzad et al., 2022)

Discussion and conclusions

This study has provided an overview on the realized value of DTs, based on empirical evidence, by identifying the associated benefits realized through their use. Different academic literature sources were retrieved from several databases to identify such. The findings demonstrated some benefits such as better environmental management and improved maintenance which were the most suggested benefits of this technology. Initial hardware costs, data infrastructure and handling were identified as the main cost elements. The findings suggest that the benefits of this concept can span across the whole life cycle of an asset. Most of the tested benefits were demonstrated during the operation and maintenance yet this finding does not suggest that this is the phase where stakeholders can realize most of the DTs' value. None of the studies presenting quantifiable findings investigated the benefits of DTs during the decommissioning phase.

From the selected articles, the first was published in 2019 which could suggest the impact of the Gemini principles (Bolton et al., 2018) in publicizing the value creation capabilities of DTs. The results, however, demonstrated a lack of research inquiries investigating the value realized with the use of DTs which is supported by the lack of author keywords related to value. There were also no studies providing a comprehensive record of the value realized by DTs. It was observed that all studies have either focused on benefits or costs, but not both. Other investigations focused on application areas only without providing an integral assessment of value.

Love and Matthews (2019) demonstrated some quantifiable findings in their study with respect to the benefits realized. The case examples used in this study included both vertical (building) and horizontal (infrastructure) projects which could suggest the varied applications of DTs in the AECO industry. The publication date of their paper, however, implied an early development stage for DTs in both the industry and academia, and hence their focus was on a System Information Model that can enable a DT. They also focused on providing an account of some benefits without extending appreciation to the costs involved.

Table 3: Benefits and their frequency in articles.

Benefit	References	No. of references
Better environmental management	Lin and Cheung (2020) Meng, Das and Meng (2023) Shahzad et al. (2022) Tagliabue et al. (2021) Zhang et al. (2023)	5
Improved maintenance	Ammar et al. (2022) Lin and Cheung (2020) Tagliabue et al. (2021) Tita et al. (2023)	4
Better scenario analysis	Adibfar and Costin (2022) Ammar et al. (2022) Zhang et al. (2023)	3
Improved information management	Ammar et al. (2022) Esmaeili and Simeone (2023) Love and Matthews (2019)	3
Improved output	Ammar et al. (2022) Love and Matthews (2019) Shahzad et al. (2022)	3
Improved site management	Esmaeili and Simeone (2023) Greif, Stein and Flath (2020) Shahzad et al. (2022)	3
Improved stakeholder collaboration	Ammar et al. (2022) Esmaeili and Simeone (2023) Love and Matthews (2019)	3
Better emergency and crisis management	Lin and Cheung (2020)	1
Better health and safety	Ogunseiju et al. (2021)	1
Better resource management	Esmaeili and Simeone (2023)	1
Better risk management	Shahzad et al. (2022)	1
Higher customer satisfaction	Shahzad et al. (2022)	1
Improved documentation	Love and Matthews (2019)	1
Improved efficiency	Love and Matthews (2019)	1
Improved procurement	Love and Matthews (2019)	1
Improved productivity	Love and Matthews (2019)	1
Improved progress monitoring	Esmaeili and Simeone (2023)	1
Improved supply chain management	Esmaeili and Simeone (2023)	1
Reduced execution time	Shahzad et al. (2022)	1

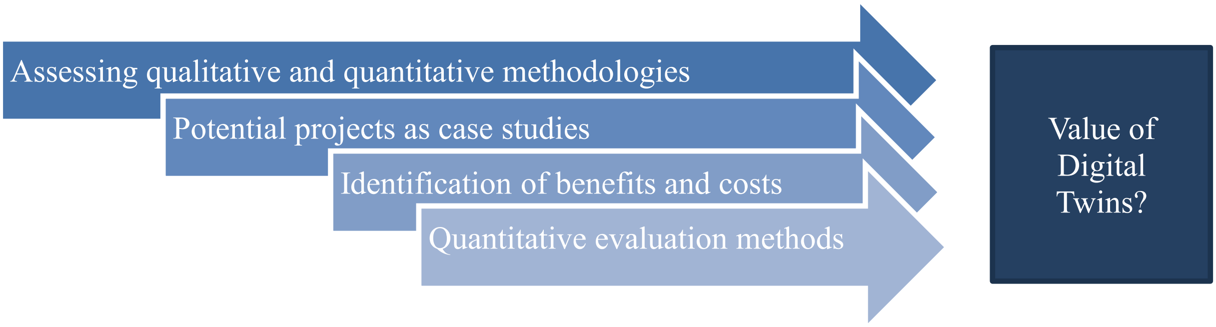


Figure 5: Research agenda.

In general, most of the studies did not present quantifiable findings which are the most sought after by practitioners when building a business case for additional investments. There was also some misclassification of benefits as an enabler of other benefits, with different designations used by different authors. The demonstration of quantifiable costs was even more limited than those of benefits. Empirical evidence coupled with quantifiable findings was only presented in four articles. This implies the need for a comprehensive assessment of the value delivered by DTs based on empirical evidence. A comprehensive cost-benefit analysis, or other quantitative evaluation methods, are recommended to tackle this issue, or to potentially identify the limitations of DTs. Further assessment of the different quantitative appraisal methods for DT investment should be conducted to recommend the most suitable one. It is also recommended to conduct research to develop a benefit and cost taxonomies for DTs to assess the viability of this innovative system. The suggested lack of DT adoption could be related to its high cost of implementation or upfront investment. This requires further appreciation to the barriers to entry experienced by enterprises at both an organizational and people level.

One limitation of this study was the adoption of IS/IT-investment evaluation frameworks due to the lack of available DT-cost or benefit taxonomies. The used framework was not prescriptive and the development of a framework with defined benefits and taxonomy can lead to a more robust identification of value constituents. Another limitation was the lack of mapping the relationship and interdependencies between different clusters of benefits. Future studies could develop a stock and flow diagram to capture any interdependencies which would ultimately lead to more refined benefit clusters.

Moving forward with this study, the research agenda will extend this theoretical framework by investigating the suitable methodologies for assessing the value of DTs in the AECO industry as illustrated in Figure 5. The findings of this study have already suggested that case studies are robust for assessing the value of new technologies, and this method would follow a qualitative inquiry. Other potential methods to assess this problem include quantitative evaluation techniques such as Cost-Benefit Analysis or Multi-Criteria Decision Analysis. In either case, the theoretical insights obtained from case studies will inform the design of the quantitative assessment.

By drawing on both qualitative and quantitative evidence, this ongoing research could result in a comprehensive analysis of the value problem of DTs. The results obtained from this study will be used to validate the outcome of the primary research. By conducting this research, valuable insights could be recommended to various stakeholders in the AECO industry including practitioners and policymakers. Furthermore, the primary research could hold potential academic implications, enriching the existing knowledge base.

Acknowledgements

This research was funded partly by Bartlett School of Sustainable Construction, UCL.

References

Adibfar, A. and Costin, A. M. (2022). 'Creation of a Mock-up Bridge Digital Twin by Fusing Intelligent Transportation Systems (ITS) Data into Bridge Information Model (BrIM)'. Journal of Construction Engineering and Management, 148 (9), pp. 1–11.

Agarwal, R., Chandrasekaran, S. and Sridhar, M. (2016). Imagining construction's digital future. Available at: https://www.mckinsey.com/industries/capital-projects-and-infrastructure/ourinsights/imagining-constructions-digital-future (Accessed: 11 June 2020).

Alizadehsalehi, S. and Yitmen, I. (2023). 'Digital twin-based progress monitoring management model through reality capture to extended reality technologies (DRX)'. Smart and Sustainable Built Environment, 12 (1), pp. 200–236.

Ammar, A., Nassereddine, H., AbdulBaky, N., AbouKansour, A., Tannoury, J., Urban, H. and Schranz, C. (2022). 'Digital Twins in the Construction Industry: A Perspective of Practitioners and Building Authority'. Frontiers in Built Environment, 8 (June), pp. 1–23.

Bakis, N., Kgioglou, M. and Aouad, G. (2006). 'Evaluating the Business Benefits of Information Systems'. in 3rd International SCRI Symposium, Salford Centre for Research and Innovation, University of Salford, Salford.

Becerik-Gerber, B. and Rice, S. (2009). 'The Perceived Value of Building Information Modeling in the U.S. Building Industry'. Journal of Information Technology in Construction, 15, pp. 185–201.

Bolton, A., Blackwell, B., Dabson, I., Enzer, M., Evans, M., Harradence, F., Keaney, E., Kemp, A., Luck, A., Pawsey, N., Saville, S., Schooling, J., Sharp, M., Smith, T., Tennison, J., Whyte, J. and Wilson, A. (2018). The Gemini Principles: Guiding Values for the National Digital Twin and Information Managment Framework. Centre for Digital Built Britain and Digital Framework Task Group. Cambridge, UK.

Bostrom, R. and Heinen, J. S. (1977). 'MIS Problems and Failures: A Socio-Technical Perspective. Part I: The Causes'. MIS Quarterly, 1 (3), pp. 17–32.

Bradley, G. (2010). Benefit Realisation Management: A Practical Guide to Achieving Benefits Through Change. Farnham, UK: MPG Books Group.

Centre for Digital Built Britain (CDBB). (2021). Digital Twin Toolkit: Developing the Business Case for Your DigitalTwin. https://digitaltwinhub.co.uk/files/file/62-digital-twin-toolkit/

Davies, R. and Harty, C. (2013). 'Measurement and exploration of individual beliefs about the consequences of building information modelling use'. Construction Management and Economics, 31 (11), pp. 1110–1127.

Esmaeili, I. and Simeone, D. (2023). 'A General Contractor's Perspective on Construction Digital Twin: Implementation, Impacts and Challenges'. Buildings, 13 (4).

Fernández-Solís, J. L. (2008). 'The systemic nature of the construction industry'. Architectural Engineering and Design Management, 4 (1), pp. 31–46.

Ferré-Bigorra, J., Casals, M., & Gangolells, M. (2022). The adoption of urban digital twins. Cities, 131, 103905.

Gann, D. M. and Salter, A. J. (2000). 'Innovation in project-based, service-enhanced firms: the construction of complex products and systems'. Research Policy, 29 (7–8), pp. 955–972.

Greif, T., Stein, N. and Flath, C. M. (2020). 'Peeking into the void: Digital twins for construction site logistics'. Computers in Industry. Elsevier B.V., 121, p. 103264.

Grieves, M. and Vickers, J. (2017). 'Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems'. in Transdisciplinary Perspectives on Complex Systems. Cham: Springer International Publishing, pp. 85–113.

Haddaway, N. R., Page, M. J., Pritchard, C. C. and McGuinness, L. A. (2022). 'PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis'. Campbell Systematic Reviews, 18 (2).

Hakimi, O., Liu, H., Abudayyeh, O., Houshyar, A., Almatared, M. and Alhawiti, A. (2023). 'Data Fusion for Smart Civil Infrastructure Management: A Conceptual Digital Twin Framework'. Buildings, 13 (11), p. 2725.

Irani, Z., Ghoneim, A. and Love, P. E. D. (2006). 'Evaluating cost taxonomies for information systems management'. European Journal of Operational Research, 173 (3), pp. 1103–1122.

Irani, Z. and Love, P. E. D. (2002). 'Developing a frame of reference for ex-ante IT/IS investment evaluation'. European Journal of Information Systems, 11 (1), pp. 74–82.

Khajavi, S. H., Motlagh, N. H., Jaribion, A., Werner, L. C. and Holmstrom, J. (2019). 'Digital Twin: Vision, Benefits, Boundaries, and Creation for Buildings'. IEEE Access, 7, pp. 147406–147419.

Laursen, M. and Svejvig, P. (2016). 'Taking stock of project value creation: A structured literature review with future directions for research and practice'. International Journal of Project Management, 34 (4), pp. 736–747.

Lin, Y.-C. and Cheung, W.-F. (2020). 'Developing WSN/BIM-Based Environmental Monitoring Management System for Parking Garages in Smart Cities'. Journal of Management in Engineering, 36 (3), p. 04020012.

Love, P. E. D. and Matthews, J. (2019). 'The "how" of benefits management for digital technology: From engineering to asset management'. Automation in Construction, 107 (July).

Lu, Q., Parlikad, A. K., Woodall, P., Don Ranasinghe, G., Xie, X., Liang, Z., Konstantinou, E., Heaton, J. and Schooling, J. (2020). 'Developing a Digital Twin at Building and City Levels: Case Study of West Cambridge Campus'. Journal of Management in Engineering, 36 (3).

Madni, A., Madni, C. and Lucero, D. (2019). 'Leveraging Digital Twin Technology in Model-Based Systems Engineering'. Systems, 7 (7), pp. 1–13.

Meng, X., Das, S. and Meng, J. (2023). 'Integration of Digital Twin and Circular Economy in the Construction Industry'. Sustainability (Switzerland), 15 (17), pp. 1–14.

Morris, P. (2004). 'Project management in the construction industry'. in Morris, P. and Pinto, J. (eds) The Wiley Guide to Managing Projects. Hoboken, N.J.: John Wiley & Sons.

Morris, P. (2013). Reconstructing Project Management. Chichester: Wiley Blackwell.

Ogunseiju, O. R., Olayiwola, J., Akanmu, A. A. and Nnaji, C. (2021). 'Digital twin-driven framework for improving self-management of ergonomic risks'. Smart and Sustainable Built Environment, 10 (3), pp. 403–419.

Opoku, D.-G. J., Perera, S., Osei-Kyei, R. and Rashidi, M. (2021). 'Digital twin application in the construction industry: A literature review'. Journal of Building Engineering, 40, p. 102726.

Opoku, D.-G. J., Perera, S., Osei-Kyei, R., Rashidi, M., Bamdad, K. and Famakinwa, T. (2023). ‘Barriers to the Adoption of Digital Twin in the Construction Industry: A Literature Review’. Informatics, 10 (1), p. 14.

Opoku, D.-G. J., Perera, S., Osei-Kyei, R., Rashidi, M., Famakinwa, T. and Bamdad, K. (2022). ‘Drivers for Digital Twin Adoption in the Construction Industry: A Systematic Literature Review’. Buildings, 12 (2), p. 113.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ, n71.

Pregnolato, M., Gunner, S., Voyagaki, E., de Risi, R., Carhart, N., Gavriel, G., Tully, P., Tryfonas, T., Macdonald, J., & Taylor, C. (2022). Towards Civil Engineering 4.0: Concept, workflow and application of Digital Twins for existing infrastructure. Automation in Construction, 141(July), 104421.

Porrit, J. (2007). Capitalism as if the World Matters. Earthscan.

Project Management Institute. (2019). Benefits Realization Management: A Practice Guide. Newton Square, PA.

Shahzad, M., Shafiq, M. T., Douglas, D. and Kassem, M. (2022). ‘Digital Twins in Built Environments: An Investigation of the Characteristics, Applications, and Challenges’. Buildings, 12 (2), p. 120. doi: 10.3390/buildings12020120.

Tagliabue, L. C., Cecconi, F. R., Maltese, S., Rinaldi, S., Ciribini, A. L. C. and Flammini, A. (2021). ‘Leveraging digital twin for sustainability assessment of an educational building’. *Sustainability (Switzerland)*, 13 (2), pp. 1–16.

Tita, E. E., Watanabe, G., Shao, P. and Arii, K. (2023). ‘Development and Application of Digital Twin–BIM Technology for Bridge Management’. Applied Sciences (Switzerland), 13 (13).

Ward, J. and Daniel, E. (2006). Benefits Management. Chichester: John Wiley & Sons.

Xie, H., Xin, M., Lu, C. and Xu, J. (2023). ‘Knowledge map and forecast of digital twin in the construction industry: State-of-the-art review using scientometric analysis’. Journal of Cleaner Production. Elsevier Ltd, 383 (November 2022), p. 135231.

Zhang, J., Chan, C. C. C., Kwok, H. H. L. and Cheng, J. C. P. (2023). ‘Multi-indicator adaptive HVAC control system for low-energy indoor air quality management of heritage building preservation’. Building and Environment. Elsevier Ltd, 246 (October), p. 110910.

Winch, Graham. (2010). Managing construction projects : an information processing approach (2nd ed.). Wiley.

LEVERAGING DIGITAL TWINS FOR ENHANCED CONSTRUCTION PROJECT DELIVERY

Obinna C. Madubuike[1], Chimay J. Anumba[1], and Eleni Papadonikolaki[2]
[1]College of Design, Construction, and Planning, University of Florida, USA.
[2]Department of Materials, Mechanics, Management, and Design, Delft University of Technology, Delft, The Netherlands.

Abstract

The construction industry increasingly recognizes the potential benefits of cyber-physical systems, particularly Digital Twin (DT) technology, within the framework of Industry 4.0 and Construction 4.0. However, there remains a lack of clarity regarding the specific contributions of DT to the construction sector, often conflated with Building Information Modeling (BIM). This paper aims to address this ambiguity by conducting a historical review of DTs, delineating their key features, and elucidating the opportunities they offer throughout the stages of constructed facilities' life. Additionally, this paper outlines potential research trajectories and practical applications for leveraging DTs to enhance construction project delivery.

Introduction

We have experienced technological developments and applications in the construction industry in recent years. The construction industry is taking advantage of the wide range of technologies to facilitate digitalization and automation, enabling the integration of the construction processes at all value chain phases, resulting in the term 'Construction 4.0' (Oesterreich and Teuteberg, 2016). Construction 4.0 is attributed to Industry 4.0 which focuses on the seamless generation, analysis, and communication of data to enable smart manufacturing (Perrier et al, 2020). However, the construction industry's adoption of digitalization is still low, particularly in the area of Digital Twin which played a huge role in realizing Industry 4.0 (Brilakis et al, 2019). However, the intent is to use a wide range of smart technologies to digitalize and automate construction processes and activities. This is done to ensure a reduction in construction costs, more accurately controlled construction activities, and time savings.

Several systems and technologies have been determined to contribute to the digitalization and automation of the construction industry. For Example, cyber-physical systems to plan and monitor mobile cranes on site (Kan et al, 2018), construction robots (Bock & Linner, 2016), and digital twin (DT) for healthcare facilities management (Madubuike & Anumba, 2022). Technological development is all attributed to enabling and ensuring that construction activities are carried out efficiently and easily. The construction industry has experienced significant project time and cost savings due to important technological developments that have improved construction operations (Son et al., 2010). Digitalizing the construction industry entails using smart technologies, systems, or data to convert physical models and tasks to virtual models and tasks, enabled by machines or robots.

This paper investigates the extent to which Digital Twin is applicable for enhanced construction project delivery. It focuses on how DT can be applied in all phases of construction projects including design, construction, operation and maintenance, and decommissioning or demolition for enhanced project delivery. To achieve this, the paper covers the following: background to construction 4.0, DT in the construction project delivery process, key features of DT and its roles in the different project delivery stages, discussions centered on characteristics of construction problems amenable to DT, construction use cases, and the potential benefits of DT application in construction. It identifies salient points in the findings and highlights future research areas.

Construction 4.0

Background to Industry 4.0

The manufacturing industry is one of the oldest and has gone through several stages seeking improvement (Madubuike et al., 2022). It has moved from Industry 1.0 to Industry 4.0 in line with the phases of the industrial revolution. The breakdown of the various phases is as follows and shown in Figure 1:

- Industry 1.0 brought about the steam engine,
- Industry 2.0 saw the assembly line concept to reduce manufacturing lead time,
- Industry 3.0 embraced computer-integrated manufacturing which replaced labor, and
- Industry 4.0 employed digitalization and automation to ensure smart manufacturing (Roy et al., 2020).

Industry 4.0 has been made possible due to many great technological and scientific advances in the 21st century. The introduction and use of computers, the Internet of Things, and cyber-physical systems contributed to the emergence of Industry 4.0 (Forcael et al., 2020). The emergence of Industry 4.0 has also brought about progress in the construction industry resulting in the term 'Construction 4.0', which is primarily based on utilizing and integrating four key concepts: digital data, automation, connectivity, and digital access to enable the digitalization of the construction industry (Berger, 2016). The current trend in technological advancement supports the enablement of Construction 4.0 which, if properly considered, would meet the standards of Industry 4.0.

Towards Construction 4.0

The construction industry is characterized by complex and peculiar features such as the nature of projects, interest in technological investments, and the nature of sites which may make it difficult to implement digitalization or automation. Despite these complexities, the construction industry has experienced some increase in technology adoption, which is enabling the implementation of the concept of Construction 4.0. There is still no consensus on the definition of the concept of 'Construction 4.0'. Similar to DT, the definition of Construction 4.0 depends on researchers and its application. However, it centers around a decentralized connection between physical space and cyberspace using ubiquitous connectivity (Berger, 2016). However, Perrier et al. (2020) opined that Construction 4.0 is not limited to existing and new technologies but also implies using real-time data to transform project management processes for smart decision-making purposes. The manufacturing industry has implemented the concept of 'Industry 4.0' which has given guidance to the construction industry on the implementation of 'Construction 4.0'. Construction 4.0 is seen as an innovative construction management technique that allows for the creation of a smart construction site as driven by Industry 4.0 (Perrier et al., 2020). However, this definition focuses on a 'smart construction site' and does not integrate the project management process which is also a key part of the construction process.

Consequently, to integrate all aspects of construction, this study defines Construction 4.0 as a smart construction technique that integrates existing and new technologies including real-time data to transform and automate the project management and construction process thereby ensuring smart decision-making that leads to efficient project delivery. The emergence of technologies amenable to construction (Madubuike et al., 2022) is paving the way for the implementation of Construction 4.0. Although there is no single technology that can develop a Construction 4.0 environment, the combination of various technologies is considered vital. Perrier et al. (2020) identified different technologies categorized under 7 groups that could be combined to create Construction 4.0. The 7 technology groups include *Group 1* (BIM), *Group 2* (Monitoring – GIS, materials tracking technologies, just-in-time (RFID), and life cycle management), *Group 3* (artificial intelligence), *Group 4* (Data Science – cloud computing, data transfer, and data protection), *Group 5* (Modeling Systems), *Group 6* (Digital Fabrication – robots and autonomous machines), and *Group 7* (Prefabrication). Some of the listed technologies also combine to develop the Digital Twin technology which is still at its nascent stage and can provide a good basis for implementing construction 4.0.

Place of Digital Twins in Construction 4.0

Digital Twinning involves the use of smart enabling technologies to virtually represent a physical asset by obtaining real-time updates and effect bidirectional coordination such that the virtual model closely or accurately represents the physical asset (Madubuike & Anumba, 2023). There is no single definition for a Digital Twin, as each definition depends on the purpose for which the DT is designed. DT combines smart technologies and data sciences similar to Construction 4.0. While DT is centered on creating a digital replica that closely represents the physical assets including all its instances, Construction 4.0 focuses on automating the construction and project management processes, ensuring smart decisions and efficient construction project delivery. In much the same way as cyber-physical systems are a critical component of Industry 4.0; they are also very important in realizing the vision of Construction 4.0.

Digital Twin can be regarded as a key component of Construction 4.0, as all DT enabling technologies are required for Construction 4.0. This helps to delineate the relationship between DT and Construction 4.0 and demonstrate how DT will facilitate the implementation of Construction 4.0. For this study, DT has been defined as the virtual representation of a physical asset using DT-enabling technologies such as sensors, communication networks, and 3D models to obtain real-time updates and effect bi-directional coordination such that the virtual model represents a replica of the physical asset in all instances.

Digital Twins in the Construction Project Delivery Process

Historical Perspective

While various literature has it that DT was developed in 2003, the origins of digital twin technology are traced back to the 1960s (Boschert et al, 2018; Madubuike et al, 2020). The NASA project in 1960 started with the twin concept where two physical space vehicles were created with one called 'the twin' which was placed on Earth to mirror the other vehicle in space (Boschert et al, 2018; Zhuang et al, 2018). However, in the NASA project, the 'digital' aspect was not included at that time. Subsequently, Micheal Grieves introduced the DT concept in his product lifecycle management course in 2003 (Grieves, 2014). Consequently, DT has been mentioned and applied in different fields and industries such as aerospace (Tao et al, 2017), automotive (Lahoti, 2021), energy (Sivalingam et al, 2018), healthcare (Laaki et al,

2019), telecommunications (He et al, 2017), and manufacturing (Roy et al, 2020). Additionally, the growth of the DT concept is beginning to be integrated into the different aspects of construction.
The flexibility in the DT application could be attributed to its features of real-time updates, bi-directional coordination, and the inclusion of digital models for its functionality. Keen (2019) identified that DT can take advantage of BIM and 3D models including other essential contractual information to create digital twin models of buildings using sensors and other smart technologies. Figure 1 provides the historical milestones with DT applications in various fields.

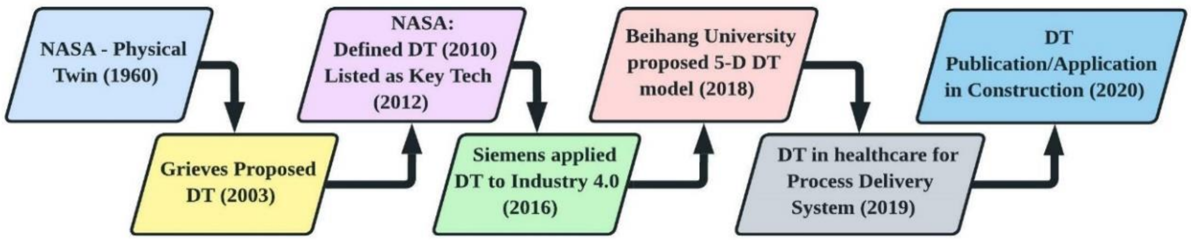


Figure 1: DT Historical Milestones (Adapted from Madubuike et al, 2020).

Key Features of DTs

DT has been identified as a promising state-of-the-art digital platform that enables the management and monitoring of physical assets (Martinez et al, 2018). It has key features that make it applicable to various fields and industrial sectors by representing the physical asset (including its operational changes) through appropriate sensors and communication platforms. Some key features of DT include a 3D digital model for visualization, real-time updates, predictive, and monitoring features, and bi-directional coordination (Madni et al, 2019).
Additionally, studies have developed DT system architectures that explain the relationship between the DT key features (He et al, 2018; Laaki et al, 2019; Madubuike & Anumba, 2022). One of the system architectures was used to develop a DT-based healthcare facilities management (Madubuike & Anumba, 2022; Madubuike et al, 2023). The system architecture was developed with six layers including the physical, sensing, application, communication, device, and virtual prototype layers. These different layers were integrated to prove the identified key features of DT.

Key Construction Applications of DT

The versatility of DT has made its application in various industries possible (Madubuike et al, 2022). Some DT applications in construction include the use of DT to monitor construction work progress (Braun et al, 2018), and using DT to monitor the state of bridges (Sacks et al, 2018). Another DT use case is Singapore's Frasers Tower constructed by Bentley Systems and Schneider Electric, providing a connected workplace for DT purposes (BIM News, 2020). DT for monitoring the state of a bridge (Sacks et al, 2018) and a DT-based system for healthcare facilities management (Madubuike & Anumba, 2021). The DT Hub to promote DT best practices for infrastructure and provide a testbed for the information management framework for the built environment that will enable future National Digital Twin (CDBB, 2020) is a huge step by Britain to provide a consensus view on the implementation of DT. Another DT implementation in construction includes DT for improved healthcare facilities management by monitoring indoor air quality (Madubuike & Anumba, 2023). Although publications on DT have focused more on the operations and maintenance phase, there are still opportunities in the construction phase.

DT Role at Construction Project Delivery Stages

The versatility of DT is seen in its applicability to different fields and industries including the construction industry. However, DT application in the construction industry is still in its nascent stage. Although publications on cyber-physical systems in construction date back to 2009, DT publications in the construction industry started in 2019 with just 2 publications which later grew to 17 in 2021 according to the Web of Science (Madubuike et al, 2022). There is an increasing volume of literature on DT applications in the construction industry. However, much of the literature on DT in construction focuses on concept proposal and prototype development with fewer cases of industry implementation. This shows that the construction industry is still exploring the best application areas for DT. The applicability of DT to the 4 high-level stages of the construction project delivery process is discussed below:

Design

The design stage is one of the earlier stages of any construction project where the planning and design of the project to be constructed are conducted. The design stage is very important as it defines the product and the associated processes needed to bring it to fruition. In Latin American countries, it was determined that 20-25% of the construction period is lost due to design deficiencies (Undurraga, 1996). Additionally, design changes during construction are responsible for the increase in project costs, delays, and in some cases, poor quality of work (Hindmarch, 2010). Several design management tools and techniques such as the Analytical Design Planning Technique (ADePT), Last Planner methodology, and Construction Design Change Management (CDCM) model have been identified (Hindmarch, 2010). However, these tools and techniques lack the essential feedback experience that would help avert any future change when construction begins. Consequently, DT can be used at the design stage to run 'what-if' simulations of real-life scenarios to ensure that most or all possible changes are fully considered at the design stage.

DT implementation at the design stage is currently being considered by construction sector companies embracing the growing trend in technology. This would entail ensuring that all requirements and conditions essential for DT implementation are considered at the design stage. It will ensure an effective and seamless integration of the digital model with the physical facility being constructed. Additionally, the use of VR and AR can be considered good tools while implementing DT in the design phase where project stakeholders can experience the 3D model and make all necessary observations and changes before construction. Negendahl (2015) proposed using 3D geometrical models to run building performance simulations (BPS) at the design stage to ensure buildability and effective building performance. Sacks et al (2018) and Brilakis et al (2019) proposed a geometrical digital twin in developing a design model. However, this was to develop digital models for already existing structures. A geometric digital model can be augmented with semantic information and is the first step to creating a digital twin (Brilakis et al, 2019). This model is then continuously updated with construction and operational data/information that enable it to adequately represent the physical facility and facilitate bi-directional coordination between the two.

Construction

The construction stage is an important phase of construction where all design ideas and plans are implemented. The construction phase is laden with complexities that may affect the adoption of modern technology. Some of the complexities include difficulty understanding the construction phase requirements and adopting modern technologies due to insufficient information (Shahrabi & Mohammadi, 2013; Anumba, 2000). However, DT has proven to be versatile and can be implemented at certain phases of the construction stage to ensure that projects are well constructed to meet the client's needs.

DT can be used to monitor construction work progress (Braun et al, 2018), and workers' productivity, and compare as-builts to the initial designs. DT can provide the opportunity for stakeholders to get involved in a project for site inspection even without being physically present. This can be possible using third-party applications that enable 3D/360-degree remote viewing. The creation and implementation of DT in the construction phase also depend on the available digital model or the Building Information Model (BIM), which contains the 3D geometric information necessary for DT creation.

One of the major purposes of every DT implementation is to create a digital model that truly represents the physical facility and enables real-time performance monitoring. Consequently, the digital model created at the design stage is developed further during construction to capture the actual as-built information, which is vital for the operations and maintenance stage.

Operation and Maintenance

The effective implementation of the design and construction stages would determine the outcome of the operation and maintenance phase. The operation and maintenance phase defines the building's performance and its lifecycle management. It is imperative that buildings or any other infrastructure should be designed and constructed with lifecycle management in mind. DT offers effective opportunities to monitor and manage facilities and ensure improved facilities' performance. The implementation of DT for the operations and maintenance stage should be considered at the inception stage (Grieves & Vickers, 2017).

DT implementation in the operation and maintenance phase has been considered by various studies. Some examples include the use of DT for bridge inspection (Sacks et al, 2018), DT for a connected workplace in one of the office towers in Singapore, and a DT-based system for healthcare facilities management (Madubuike & Anumba, 2021). These examples show that the implementation of DT in construction, particularly in the operation and maintenance phase, is beginning to grow. There are more prospects for DT implementation in the operation and maintenance phase, especially in monitoring the performance of key equipment and building components (Madubuike & Anumba, 2021; Madubuike, 2022; Asare, 2023; Kang & Mo, 2024).

Decommissioning/Demolition

Decommissioning/demolition is the last stage in the lifecycle of any constructed facility. It comes up when a facility has reached the end of its useful and economic life. Some important considerations before decommissioning and demolition include relevant components of the facility that could be salvaged given their current use/reuse value, ensuring that the demolition would not be harmful to the environment or surroundings, and the selection of the most appropriate techniques for the demolition (Anumba et al, 2008; Stevens, 2019). Currently, no study has addressed DT applicability with decommissioning or demolition. However, DT can be used to run demolition simulations of real-life scenarios to understand what better techniques to adopt and the possible effects of the techniques adopted. DT offers a range of benefits depending on what purpose it was created for.

Discussion

Bi-directional Coordination

Bi-directional coordination is the final of the DT features which enable information flow between the physical to digital model and vice versa. Figure 2 provides a proposed simple schematic representation for DT bi-directional coordination. The sensors obtain information from the physical asset which updates the DT platform and the controller. The controller detecting any anomaly in the data received, sends a control effort to the actuator. The actuator inputs a control signal that updates the physical asset. An example where this bi-directional coordination model can be applied would be in controlling systems within buildings such as the HVAC system and others.

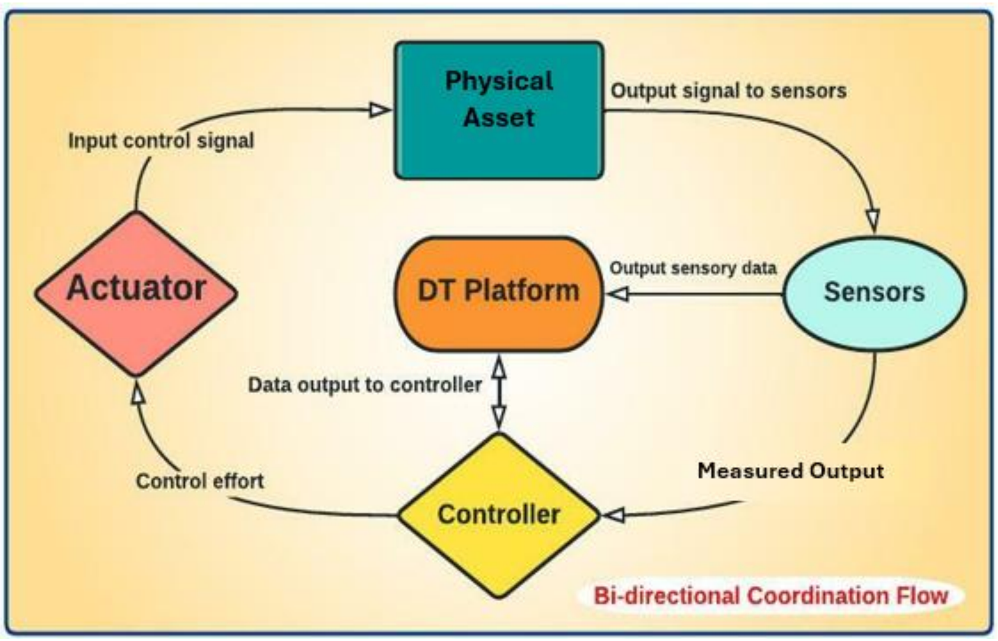


Figure 2: Schematic representation of DT Bi-directional Coordination.

Characteristics of Construction Problems Amenable to DT

The application of DT to the realization of construction 4.0 is still at the nascent stage due to the construction industry's slow adoption of emerging technologies. The construction industry (CI) is one of the major industries with unique and complex processes, interrelated activities, and high uncertainty levels. Consequently, no two similar construction projects are directly similar. This can be attributed to the volatile nature of the construction industry (Brilakis et al, 2019) which makes it difficult for the adoption of smart technology. Some of the characteristics of the CI include methods of procurement, a large number of small and informal contractors, fragmentation with too many stakeholders, no single regulatory authority, reliance on temporary skills, and cultural influence (Boadu et al, 2020). However, given the growing technological trend in the construction industry, some of the problems associated with these CI characteristics can be addressed through DT adoption t. The real-time monitoring feature of DT can be used to monitor construction work progress and construction workers, generating DT models of physical structures and enabling real-time updates and bi-directional coordination for the lifecycle performance of a facility (Braun et al., 2018; Brilakis et al., 2019).

DT Best Practices

DT has proven to be a fast-growing and versatile technology applicable in various fields. Given the versatility of DT, calls have been made to create DT best practices. Brilakis et al (2019) explored DT and its market while requesting its standardization for market demands. Other researchers have argued that its different protocols and standards affect the potential for DT standardization (Qi et al, 2019; Madubuike & Anumba, 2021). However, there has been some progress made in providing interoperability in DTs. Some of the DT best practices include the 3rd Generation Partnership Project (3GPP) for 5G communications for DT, oneM2M to standardize a service layer IoT platform, ISO/TC 184 for industrial data standards, ISO/IEEE for digital health data, and IEC TC65 for interoperability in the smart factory (Song & Le Gall, 2023). The creation of these best practices and standards may help improve DT adoption in construction which is characterized by complex and unique projects.

Potential Benefits of DT in Construction Project Delivery

DT offers several potential benefits that may be integrated into the various phases of a construction project. The potential benefits that DT has to offer in construction product delivery if properly implemented can be drawn from other fields and areas where DT has been successfully implemented. Understandably, construction project delivery (CPD) is guided by three major baselines – quality, cost, and schedule. The potential benefits of DT for construction project delivery can be weighed against these baselines. The following are DT potential benefits that can be attributed to CPD including automated progress monitoring, updated as-built drawings, safety monitoring, resource planning and logistics, optimization of equipment usage, quality assessment, monitoring and tracking of workers, monitoring of a facility, facilities management and operations, decision making and sustainable development (Madubuike & Anumba, 2022).

Conclusions

Digital Twin technology has shown flexibility and applicability in various industries particularly in the manufacturing industry leading to the emergence of Industry 4.0. Consequently, the construction industry has taken a cue from the emergence of Industry 4.0 to initiate the emergence of Construction 4.0. Similarly, Construction 4.0 adopts smart technology, digitalization, and automation to make smart decisions during the construction and operations of buildings and facilities. Although Construction 4.0 is yet to be actualized like Industry 4.0, studies show that DT application in construction is growing. This paper made efforts to discuss DT application in construction with a focus on the various stages of construction. DT applications in the design, construction, operations and maintenance, and decommissioning/demolition stages were discussed.

DT provides potential benefits to the various phases of construction which would ensure enhanced construction project delivery. In the design phase, DT can help in capturing as-built information that will fully represent the finished project. For the construction phase, it can be used to monitor project progress. During the operations and maintenance phase, DT can provide real-time updates to monitor facility performance. The use of DT to run simulations to salvage valuable building components and determine effects on the environment before demolition is another considerable benefit. In summary, DT helps in smart decision-making for improved construction project delivery.

The growing trend in smart technologies and DT studies in construction will help enhance construction project delivery through the DT approach. Understanding the concept, characteristics, and applicability of DT together with the characteristics of the construction industry amenable to DT are essential in improving construction project delivery through the DT approach. DT offers potential benefits in enhancing construction project delivery by ensuring cost reduction, avoiding design errors, minimizing delays during construction, monitoring construction progress, and real-time reporting on a facility's performance. These potential benefits of DT cut across the various stages of construction delivery. Future DT research can investigate in detail the implementation of DT at the various stages of the construction process highlighted in this paper.

References

Anumba, C. J. (2000). Integrated systems for construction: challenges for the millennium. In *Proceedings of the International Conference on Construction Information Technology, Hong Kong, 17*: 78-92. Edward Elgar Publishing Ltd.

Anumba C. J., Abdullah A. & Ruikar K. (2008): An Integrated System for Demolition Techniques Selection. *Architectural Engineering and Design Management*, Vol. 4, pp 130-148.

Asare, K. (2023): Digital Twin-Based Predictive Maintenance in Constructed Facilities, PhD Thesis, University of Florida, December 2023.

Berger, R. (2016). Roland Berger digitization in the construction industry: building Europe's road to "Construction 4.0". *Roland Berger GMBH,* Munich, Germany.

BIM News (2020, August 13). Digital twin creates the blueprint for future smart buildings. *BIM Today.* https://www.pbctoday.co.uk/news/bim-news/digital-twin-smart-building/73253/

Boschert, S., Heinich, C. and Rosen, R. (2017). Next-generation digital twin. *Proceedings of TMCE,* Las Palmas de Grad Canaria, Spain.

Braun, A., Tuttas, S., Stilla, U., and Borrmann, A. (2018): BIM-Based Progress Monitoring. In *A Borrmann, M. König, C. Koch, J. Beetz (Eds.): Building Information Modeling*: Springer.

Brilakis, I., Pan, Y., Borrman, A., Mayer, H. G., Rhein, F., Von, C., Pettinato, E., and Wagner, S. (2019). Built environment digital twinning. *Report of the International Workshop on Built Environment Digital Twinning presented by TUM Institute for Advanced Study and Siemens AG.*

Center for Digital Built Britain (January 30, 2024). Digital Twin Hub. University of Cambridge, UK. https://www.cdbb.cam.ac.uk/news/cdbb-officially-launches-digital-twin-hub

Forcael, E., Ferrari, I., Opazo-Vega, A., and Pulido-Arcas, J. A. (2020). Construction 4.0: A literature review. *Sustainability, 12*: 1-28, DOI:10.3390/su12229755

Grieves M., Vickers J. (2017). Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. *In: Kahlen F. J., Flumerfelt S., Alves A. (eds) Transdisciplinary Perspectives on Complex Systems.* Springer, Cham. https://doi.org/10.1007/978-3-319-38756-7_4

He, Y., Guo, J., and Zheng, X. (2018). From surveillance to digital twin: Challenges and recent advances of signal processing for the industrial Internet of Things. *IEEE Signal Processing Magazine*. DOI: 10.1109/MSP.2018.2842228

Hindmarch, H., Gale, A. W., and Harrison, R. (2010). A proposed construction design change management tool to aid in assessing the impact of design changes. *Paper presented at PMI® Research Conference: Defining the Future of Project Management*, Washington, DC. Newtown Square, PA: Project Management Institute.

Kan, C., Fang, Y., Anumba, C. J., and Messner, J. I. (2018). A cyber-physical system (CPS) for planning and monitoring mobile cranes on construction sites. *Proceedings of the Institution of Civil Engineers: Management, Procurement, and Law*, *171*(6), 240-250. [1700042]. https://doi.org/10.1680/jmapl.17.00042

Kang, T. W., and Mo, Y. (2024). A comprehensive digital twin framework for building environment monitoring with emphasis on real-time data connectivity and predictability. *Journal of Developments in the Built Environment*, *17*: 1-12. https://doi.org/10.1016/j.dibe.2023.100309

Keen, M. (2019). Construction technology defined. What is Digital Twin? *Autodesk Construction Cloud.* https://constructionblog.autodesk.com/digital-twin/

Laaki, H., Miche, Y. and Tammi, K. (2019). Prototyping a digital twin for real-time remote control over mobile networks: Application of remote surgery. *IEEE, 7*, 20325-20336. *Digital Object Identifier: 10.1109/ACCESS.2019.2897018*

Lahoti, N. (2021, May 7). How is digital twin technology impacting the automotive industry? *Available online:* https://mobisoftinfotech.com/resources/blog/digital-twin-technology-impacting-automotive-industry/

Madni, A. M., Madni, C. C. and Lucero, S. (2019). Leveraging digital twin technology in model-based systems engineering. *Systems, 7*, 7.

Madubuike, O. C. (2022). Investigation of Digital Twin-based Healthcare Facilities Management, PhD Thesis, University of Florida, August 2022.

Madubuike, O. C. and Anumba, C. J. (2021). Digital Twin application in healthcare facilities management. *ASCE International Conference on Computing in Civil Engineering 2021.* https://doi.org/10.1061/9780784483893.046

Madubuike, O. C. and Anumba, C. J. (2022). Digital Twin–Based Health Care Facilities Management. *Journal of Computing in Civil Engineering*, *37*(2): 04022057

Madubuike, O. C. and Anumba, C. J. and Khallaf, R. (2022). A review of digital twin applications in construction. *Journal of Information Technology in Construction (ITcon), 27*: 145-172. DOI: 10.36680/j.itcon.2022.008

Martinez, V., Ouyang, A., N., Neely, A., Burstall, C., and Bisessar, D. (2018). Service business model innovations: The digital twin technology. *Working*

Paper; Cambridge Service Alliance, University of Cambridge.

Negendahl, K. (2015). Building performance simulation in the early design stage: An introduction to integrated dynamic models. *Automation in Construction 54*: 39–53.

Oesterreich, T. D. and Teuteberg, F. (2016). Understanding the implications of digitization and automation in the context of Industry 4.0: A triangulation approach and elements of a research agenda for the construction industry. *Computers in Industry*, *83*, 121-139.

Osunsanmi, T. O., Aigbavboa, C., & Oke, A. (2018). Construction 4.0: The future of the construction industry in South Africa. *World Academy of Science, Engineering and Technology International Journal of Civil and Environmental Engineering, 12(3)*: 206-212

Perrier, N., Bled, A., Bourgault, M., Cousin, N., Danjou, C., Pellerin, R., Roland, T. (2020). Construction 4.0: a survey of research trends. *Journal of Information Technology in Construction (ITcon), 25*: 416-437, DOI: 10.36680/j.itcon.2020.024

Qi, Q., Tao, F., Hub, T., Answer, N., Liud, A., Wei, Y., Wang, L., and Nee, A.Y. C. (2019). Enabling technologies and tools for digital twin. *Journal of Manufacturing Systems,* *https://doi.org/10.1016/j.jmsy.2019.10.00*

Roy, R. B., Mishra, D., Pal, S. K., Chakravarty, T., Panda, S., Chandra, M. G., Pal, A., Mishra, P., Chakravarty, D., and Mishra, S. (2020). Digital Twin: current scenario and a case study on a manufacturing process. *The International Journal of Advanced Manufacturing Technology, 107*, 3691–3714. https://doi.org/10.1007/s00170-020-05306-w

Sacks, R., Kedar, A., Borrmann, A., Ma, L., Brilakis, I., Hüthwohl, P. and Barutcu, B. E. (2018). See Bridge as next-generation bridge inspection: overview, information delivery manual, and model view definition. *Automation in Construction, 90*, 134 - 145.

Shahrabi, M. A. and Mohammadi, H. (2013). Investigating different challenges in construction projects. *Management Science Letters 3*: 1869–1872

Sivalingam, K., Spring, M., Sepulveda, M. and Davies, P. (2018). A Review and methodology development for remaining useful life prediction of offshore fixed and floating wind turbine power converter with digital twin technology perspective. *2nd International Conference on Green Energy and Application,* pp. 197-204.

Son, H., Kim, C., Kim, H., Han, S. H. and Kim, M. K. (2010). Trend analysis of research and development on automation and robotics technology in the construction industry. *KSCE Journal of Civil Engineering*, *14*(2), 131-139.

Song, J. & Le Gall, F. (2023). Digital Twin Standards, Open Source, and Best Practices. In: Crespi, N., Drobot, A.T., Minerva, R. (eds) The Digital Twin. Springer, Cham. https://doi.org/10.1007/978-3-031-21343-4_18

Stevens, J. (2019). Demolition and decommissioning: A legal perspective. *EnviroTech – Decommissioning and Demolition Workshop Calgary*, Alberta.

Tao, F., Cheng, J., Cheng, Y., Gu, S., Zheng, T. and Yang, H. (2017). SDM-Sim: A manufacturing service supply-demand matching simulator under a cloud environment. *Robotic Computer Integrated Manufacturing, 45*, 34-46.

Undurraga, M. (1996). Construction Productivity and Housing Financing. In Spanish: La Productividad en la Construcción y Financiamiento de Vivienda. *Seminar and Workshop Interamerican Housing Union*, Ciudad de México, D.F., México.

BIM-GIS INTEGRATION THROUGH OPEN TOOLS

Eleonora Congiu[1], Emanuela Quaquero[1] and Giulia Rubiu[1]
[1]Department of Civil Environmental Engineering and Architecture (DICAAR), University of Cagliari, Cagliari, Italy

Abstract

The implementation of effective workflows based on a fully open BIM-GIS (Building Information Modelling-Geographic Information Systems) integration is still an unsolved topic. This is mainly due to the lack of freely available BIM authoring tools, as well as to still unfixed interoperability issues between software. In this context, this paper proposes a new methodology to effectively integrate BIM and GIS without using commercial tools and proprietary formats. The proposed workflow is enabled by the use of the add-on "BlenderBIM" for the free software "Blender", which allows to manage IFC (Industry Foundation Classes) models by updating non-geometric information through ".csv" files. Moreover, ".csv" files are also assumed as the main information exchange vehicle between BIM and GIS. The method is finally validated through an illustrative case study.

Introduction

Currently, the management of large building stocks is one of the most challenging activities which copes with many heterogeneous aspects and issues. In such a framework, the paper proposes a methodology capable of integrating different but strongly complementary tools, data, and information for a multiscale approach. It shows the first results of an ongoing research focused on the integration of BIM and GIS environments by taking advantage of open-source tools and non-proprietary data exchange formats. It is worth specifying that the herein proposed experimentation is part of a wider research concerning the implementation of an effective BIM-GIS integration in support of facility management of large building stocks (Congiu et al., 2024). However, the BIM-GIS workflow proposed in (Congiu et al., 2024) took advantage of open-source tools except for the BIM authoring. In this context, this contribution preliminarily attempts to replace the commercial BIM authoring in the original workflow (Congiu et al., 2024) by introducing BlenderBIM (*BlenderBIM Add-On*, n.d.) (an open-source BIM tool which is still in an unripe maturity level but really promising) and testing it through a simple abstract case study. In contrast to what has been mainly proposed in the literature concerning BIM-GIS integration so far, the methodology presented in this paper does not aim at transferring BIM data to the GIS platform. This unusual choice relies on the belief that it would be more advantageous to not wholly transfer integral BIM models into a GIS system firstly to prevent information redundancies and wastes, by keeping well distinct the two informative systems, as they have distinct purposes. Moreover, in the perspective of implementing an effective BIM-GIS integration to manage large building stocks, transferring into GIS a considerable amount of integral BIM models it would make GIS projects too large and extremely difficult to be managed. For this reason, only some necessary information (selected based on the specific application field of the BIM-GIS integration) to be handled at an urban scale should be considered to be shared between the two information systems. The work thus focuses on a bidirectional integration between the two informative systems, by linking them and allowing for an easy switch from one system database to the other. The experimental phase, reported in the final part of the paper, focuses on the validation process of the methodology developed by applying the proposed methodology to an illustrative case study.

Literature review

Open BIM workflows: use of the IFC

In accordance with the buildingSMART International definition, openBIM is intended as a collaborative vendor-neutral process, which has to ensure good interoperability, the use of open standards and workflows, reliability of data exchange flows, flexibility, collaboration and sustainability (buildingSMART International, n.d.). It is also worth remarking that the non-profit organization buildingSMART International is responsible for creating and adopting open international standards such as IFC (Industry Foundation Classes), bSDD (buildingSMART Data Dictionary), and BCF (BIM Collaboration Format) to promote a common language in data exchange. In particular, the IFC schema is a standardized, ISO-recognized (ISO 16739-1:2018), data model that codifies the identity, semantics, objects, related attribute and processes (buildingSMART International, n.d.). The IFC schema is currently used as an exchange data format. Some interoperability issues are therefore due to the fact that IFC is not treated as a native way of recording data created by specific BIM applications (Malewczyk, 2022). Unlike other BIM authoring tools, Malewczyk showed how the open-source add-on BlenderBIM for Blender (*Blender 4.0*, n.d.; *BlenderBIM Add-On*, n.d.) works as a true IFC authoring tool, thus minimizing the risk of data loss related to forced data conversions.

In this context, it is worth mentioning a comprehensive literature review on the openBIM domain, workflows, standards, software platforms, and tools, provided by Jiang et al. as an outcome of an accurate investigation of the most accredited databases (*e.g.* Scopus, Web of Science, Science Direct etc) (Jiang et al., 2019). Their study supplies indeed a detailed understanding of the progress status of OpenBIM research by also discussing the actual research gaps, including still unfixed critical issues in integrated BIM-GIS applications.

BIM-GIS integrated systems

Geospatial data in building planning and management play a very important role both because of their ability to integrate objects in a spatial context and to simplify all those operations involving a huge amount of information and data related to objects and to the territory in which these objects are located. Among geospatial data tools, Geographic Information Systems (GIS) play an important role, mainly due to their ability to perform spatial analysis and complex queries (Satria & Castro, 2016). In all cases where GIS involved a large building stock, the most crucial issue to be faced was to manage the huge amount of data and information that a building stock brings with itself. In the management of such data, a significant contribution has been made by BIM-GIS integration. GIS and BIM share similar features as they both manage spatial information but at different scales. Moretti et al. explored the combination of BIM and GIS and proposed a GeoBIM approach intended to im-prove digital Asset Management (AM), by especially focusing on outdoor and indoor condition assessment of the built environment. More specifically, Moretti et al. presented a GeoBIM model of the university Leonardo Campus of the Politecnico di Milano (Milan, Italy), by importing IFC models into a GIS environment in a 3D spatial data format by taking advantage of a specific commercial tool (FME workbench) (Moretti et al., 2021). Similarly, Sammartano et al. worked on a 3D City Model of Turin (Italy) to develop a digital twin of the city by taking advantage of a BIM-GIS integration with ArcGIS Pro (Sammartano et al., 2021a). Recently, Facility Management (FM) research is involving integrated BIM-GIS systems. FM is increasingly focused on finding solutions and tools to work at different scales more efficiently, from component, room, or building scale to spatial, regional, or national (Vankova et al., 2021). By way of example, Kang et Hong advanced a proposal of a software architecture based on an effective BIM-GIS integration using IFC and CityGML schemas, to specifically support FM (Kang & Hong, 2015). Vacca et al. showed the results of a research concerning the integration between BIM and 3D GIS, applied to a housing project in the city of Cagliari (Italy) (Vacca et al., 2018). Mangia et al. provided a comprehensive study, based on the analysis of the literature, concerning BIM-GIS integrated systems for managing large building stocks (Mangia et al., 2022). Meschini et al. focused their research on the development of an Asset Management System (ASM) for university building stocks by combining BIM and GIS with cognitive digital twins (Meschini, Accardo, et al., 2022; Meschini, Pellegrini, et al., 2022). Other interesting applications of GIS-BIM integration systems refer to the management of infrastructures such as highways (Zhao et al., 2019), airports (D'Amico et al., 2020), or railways. Wang et al. introduced a comprehensive study on the progress status of the research on BIM-GIS integration and its applications, by specifically underlining some still unsolved critical issues related to one-way IFC-to-CityGML conversion (Wang et al., 2019). The potential of BIM-GIS integration in several application fields was investigated by Liu et al. through an extensive state-of-the-art review (Liu et al., 2017). The research of Liu et al. revealed still unfixed issues related to information loss or corruption in data exchange, but also the potential of semantic web technologies as promising integration solution. Concluding, it can be stated that BIM-GIS integration still needs to be enhanced in terms of software interoperability, data exchange accuracy and use of open-source tools and datasets.

Methods and tools

In light of what has emerged from the literature review, it can be stated that although some available commercial tools allow to transfer BIM models into GIS maps by taking advantage of IFC-to-CityGML conversion (Sammartano et al., 2021b), a lack of BIM-GIS integrated workflows uniquely based on open-source tools is still noticeable. This evidence then motivated the present research which is focused on developing a new BIM-GIS integrated workflow by only taking advantage of open-source tools, add-ons and non-proprietary data formats (see Figure 1).

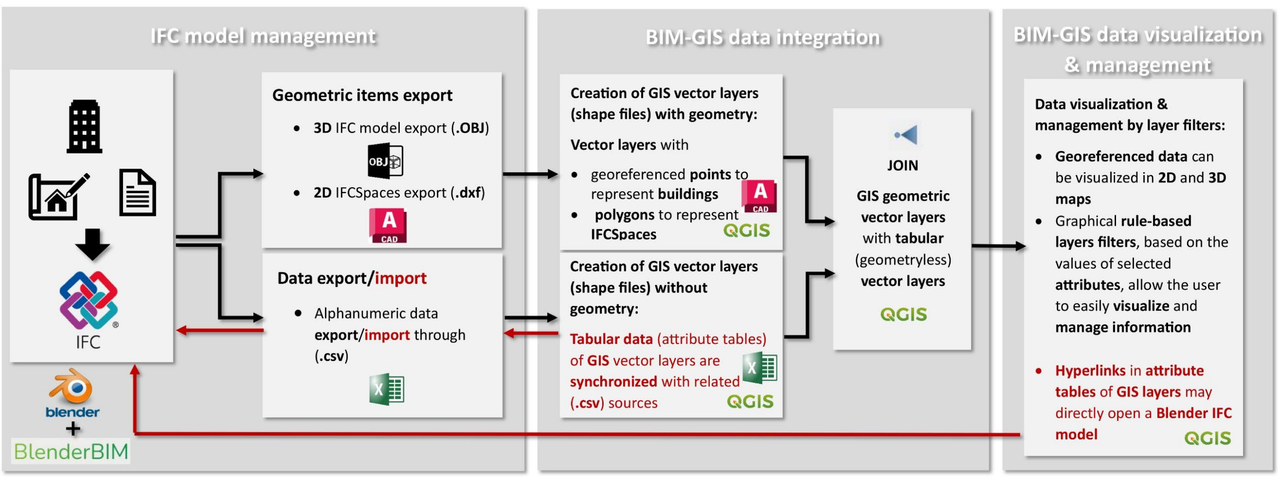


Figure 1.BIM-GIS integration workflow through open-source tools

It is also worth remarking that the proposed workflow is aimed at implementing an effective two-way integration between the two informative systems, by properly relating them and allowing for an easy switch from one system to the other, without transferring whole BIM databases into GIS maps. In this configuration, users can easily find the information which they require in the BIM model, at the scale of the specific building, and in the GIS model, at the scale of the large stock to which the specific building belongs. Moreover, the link between BIM and GIS models ensures that all the information, stored without copies or redundancies, is always updated and consequently reliable.

IFC model management through BlenderBIM

BlenderBIM (*BlenderBIM Add-On*, n.d.) is an add-on for the free but advanced modelling and rendering software Blender (*Blender 4.0*, n.d.). BlenderBIM provides an open-source IFC authoring platform which combines common abilities of modelling with additional scheduling, costing, simulating, and coordinating features. Blender and BlenderBIM add-on, unlike other BIM authoring programs, do not take advantage of object-oriented modelling tools. Blender tools allow the designer to create geometries whereas the BlenderBIM add-on allows the user to assign IfcClasses and related property sets to the modelled objects. BlenderBIM takes use of the IFC schema to classify geometries (Malewczyk, 2022). As shown in Figure 2, BlenderBIM easily allows a qualified user to model, edit and manage any sort of IFC object, as well as to manage IFC models deriving from a third BIM authoring tool. Therefore, BlenderBIM modelling is not object-oriented but rather IFC based. Elements classification and attributes are compliant with IFC standard, provided that a proper IFC class is assigned to each object. It can be therefore stated that, differing from other BIM authoring tools which use IFC standard to simply exchange data, BlenderBIM takes advantage of IFC as a native way to record information about the model (Malewczyk, 2022). As it is not among the goals of this research to transfer whole BIM models into GIS systems, the proposed workflow simply considers exporting IfcSpaces polygons as geometric items through the open-source add-on "ezdxf_exporter". The latter enables to export some selected IFC geometries to ".dxf" (Drawing Exchange Format) by properly setting custom properties of the exported layers. The choice of importing into GIS only BIM objects with simple geometry (*i.e.* IfcSpaces) is consistent with the aim of preliminarily testing the two-way effectiveness of the proposed BIM-GIS workflow on a simple dataset. In addition, IfcSpaces are suitable to be imported into a GIS environment as they can provide a proper .dxf geometric data source for a GIS shape file. The ".dxf" format is therefore assumed as the main exchange data vehicle between BIM and GIS for geometric entities. Moreover, the "Quality and Coordination" dialogue box of BlenderBIM provides a powerful "IFC-CSV Import/ Export panel" which enables two-way alphanumeric data exchange between the IFC model and ".csv" spreadsheets. This is the core of the proposed methodology as this BlenderBIM functionality allows to export IFC alphanumeric data to a ".csv" tabular form, as well as to directly update information stored in an IFC model by reimporting updated ".csv" tabular data back, provided that specific IFC selection queries are properly set up through specific panels ("Add Search Group" and "Add CSV Attribute"). It can be stated that the IFC model is the main repository of information in the proposed workflow. CSV files are used instead as main information exchange vehicle between BIM and GIS for non-geometric data.

BIM-GIS data integration

The aforementioned datasets (*i.e.* geometric and semantic IFC data) are then imported and managed through the open-source GIS software QGIS (Quantum Geographic Information System) (*QGIS*, n.d.). QGIS enables to develop a suitable GIS environment to handle spatial data. Both 2D and 3D georeferenced maps can be easily built by taking advantage of freely available raster and vector layers (provided that also elevation values are included in attributes), such as the following:

1. DTM (Digital Terrain Model) raster layer adopted as DEM (Digital Elevation Model) source for the terrain elevation (freely available for Sardinia Italian region (*SardegnaGeoportale – Aree tematiche*, n.d.)
2. orthophoto of the area
3. vector shape file of volumetric units of buildings from a geo-topographic database (GTDB) 1:2000, freely available for Sardinia Italian region (*SardegnaGeoportale – Aree tematiche*, n.d.).

2D and 3D queryable maps with basic topographic contents provide a suitable GIS-based environment to be efficiently integrated and linked with additional geometric and non-geometric information extracted from BIM models. As previously anticipated, the proposed BIM-GIS integration requires a clear distinction between geometric items and respective semantic attributes to be exported from BIM models to be transferred to the GIS environment. This clear separation is made necessary by the need to export selected geometric items to a ".dxf" format, which is suitable for generating GIS shape files but which, however, also involves the loss of most of information associated with the exported objects. Fortunately, non-geometric data associated with the selected IFC objects can be exported, in a tabular form, to a specific structured ".csv" file, which can supply a suitable data source of a GIS geometryless vector layer in the QGIS scene (by keeping information constantly and mutually synchronized). As shown in the flowchart in Figure 1, semantic alphanumeric information needs to be properly linked to the respective geometric objects in the GIS environment, without losing synchronization with the data source files. This operation is made possible by the QGIS "Joins Properties" layer tab (*QGIS*, n.d.), allowing the user to associate geometric features of the current layer (called "Target layer") with alphanumeric attributes from a geometryless layer (*i.e.* the "Join layer"). The layer "Joins Properties" functionality requires a "Join field" common to both layers to join, to accurately associate attributes of the "Join layer" with the "Target layer". In this regard, the proposed BIM-GIS integration approach provides for the assignment of IFC object names to ".dxf" layers during the related export, in order to create a suitable "join field" to allow subsequent "join" of attributes with the ".csv"-based geometryless layer in the QGIS environment. According to one of the main purposes of this research, in order to improve the bi-directionality on the BIM-GIS integration, the attribute tables of GIS layers may also include URL fields to hold specific hyperlinks to directly access useful additional files, like BIM models (*e.g.* BlenderBIM projects including IFC models).

BIM-GIS data visualization and management

The BIM-GIS workflow herein advanced, takes advantage of QGIS layer display features to enhance data visualization and management of the GeoBIM model. As previously anticipated, only IfcSpace objects have been assumed to be exported from BIM and imported into the GIS environment as 2D vector polygons. The latter may be also displayed on 3D maps as long as the related "height" values are set up as "extrusion" values in the layer 3D view tab. Also building "volume units" of the GTDB layer may be "extruded" on QGIS 3D maps to provide an effective graphical representation of urban contexts. In this regard, in order to make BIM-derived buildings more easily identifiable from the GIS urban context, a detailed 3D representation of that buildings, based on ".OBJ" geometries exported from the IFC authoring (*i.e.* BlenderBIM), may be associated with the related centroids (displayed as vector points in 2D maps) through the layer 3D view tab. In addition, QGIS rule-based filters for vector layers provide a powerful tool to ease data visualization and management in the herein advanced GeoBIM approach. More specifically, QGIS rule-based render options allow to discriminate features of a layer according to their at-tribute values by assigning them specific rendering settings. Filtering rules can be set up by taking advantage of a powerful SQL (Structured Query Language)-based "Query builder" dialog. In this regard, it is worth remarking the importance of setting up equally-named filters, based on the same rules, in both 2D and 3D view properties of layers, to ensure a perfect correspondence between 2D and 3D visualization of GIS maps. Unfortunately, QGIS does not enable to concurrently activate corresponding 2D and 3D render filters, so that users must therefore pay attention to activate 2D and 3D render options properly.

Workflow validation: a case study

The current section of this paper is aimed at validating the effectiveness of the BIM-GIS workflow, uniquely relying on the use of open-source tools and non-proprietary data formats, by simulating an illustrative example of data exchange, integration, visualization, and management, without employing a real case study. Therefore, as an example, a basic six-room house, developed on a single floor, has been modelled and adopted as an abstract illustrative case study to test the proposed BIM-GIS workflow. At this preliminary stage, the proposed methodology has not been implemented for a specific application field. For this reason, this work does not focus on specific information requirements nor on precise information sets to be transferred from BIM to GIS, as only some abstract test parameters have been used to validate the two-way effectiveness of the method.

IFC model management through BlenderBIM

A basic BIM model of the adopted application case study, managed through the IFC authoring add-on BlenderBIM, is shown in Figure 2. As highlighted through a specific object selection, the model includes six IfcSpaces, whose attributes are manageable through the "Attributes" sub-panel included in the related "Object Information" tab, available among the BIM setting tabs on the right section of the BlenderBIM Graphical User Interface (GUI). It can be noted that the "Name" attribute provides a univocal numeric code assigned to each IfcSpace, whereas the "LongName" is a textual parameter, conceived to store the space full name commonly including information on space end uses. The "GlobalId" attribute provides instead the 22-character encoded internal ID, which is automatically generated by the software for each IFC element. According to IFC standard, IfcSpaces are characterized by several property sets, among which it is worth mentioning the "Attributes" set (including the fundamental attributes cited above), the "Object Property Sets" (including various properties sub-groups), the "Object Quantity Sets" (containing all relevant parameters for quantity take-off), and so on. In addition to standard IFC attributes, IfcSpaces are herein also equipped with a custom IfcBoolean parameter, named "TestParameter", specifically created to be edited while performing experimental simulations of data exchange and management to assess strengths and weaknesses of the proposed BIM-GIS integration. In accordance with the proposed methodological workflow (Figure 1), IfcSpaces are exported to ".dxf" subject to properly set layer properties by assigning IfcSpace names to each respective exported layer through "ezdxf_exporter" add-on for Blender (see Figure 3). As far as IfcSpace alphanumeric data are concerned, some selected attributes, quantities and properties, including the aforementioned "TestParameter", are exported to a ".csv" tabular format (through the specific functionality of BlenderBIM), provided that specific IFC selection queries are defined properly (see Figure 4). It is worth remarking that the BlenderBIM sub-panel "Quality and Coordination" also allow to update the involved IFC data by simply uploading the updated version of the ".csv" file, as long as the same IFC selection queries used to export data are set up (Figure 4). This is fundamental to ensure a real bidirectional integration between BIM and GIS, as it will be shown more clearly at the next validation sections.

BIM-GIS data integration

In order to show an application of the proposed workflow, the IFC data extracted from the basic BIM-based model adopted as abstract case study, are properly imported into a GIS map. By way of example, a GIS map composed of a Digital Terrain Model, an orthophoto and the building "volume units" vector layer included in the Geo-topographic database (GTDB), all freely available as open geospatial data for the city of Cagliari, is generated. Thanks to the "elevation" attributes, it has been also possible to implement a 3D view of the map, thus providing both 2D and 3D representations of the urban context. Then, a georeferenced vector layer with geometric polygons is also added to the map as a shape file, by using the IfcSpaces ".dxf" file as data source. Consistently with the specific ".dxf" export setup herein adopted (described in the previous section), the "Layer" field in the attribute table holds the IfcSpace names.

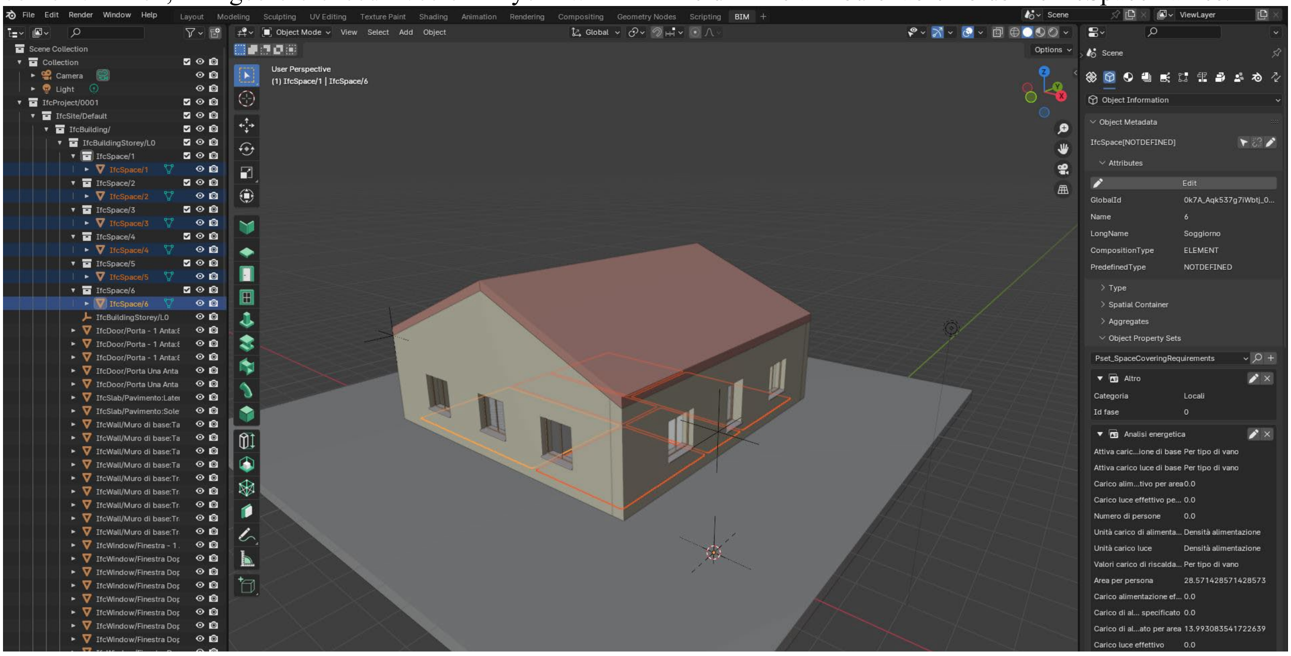


Figure 2: IFC model management through BlenderBIM (the illustrative case study)

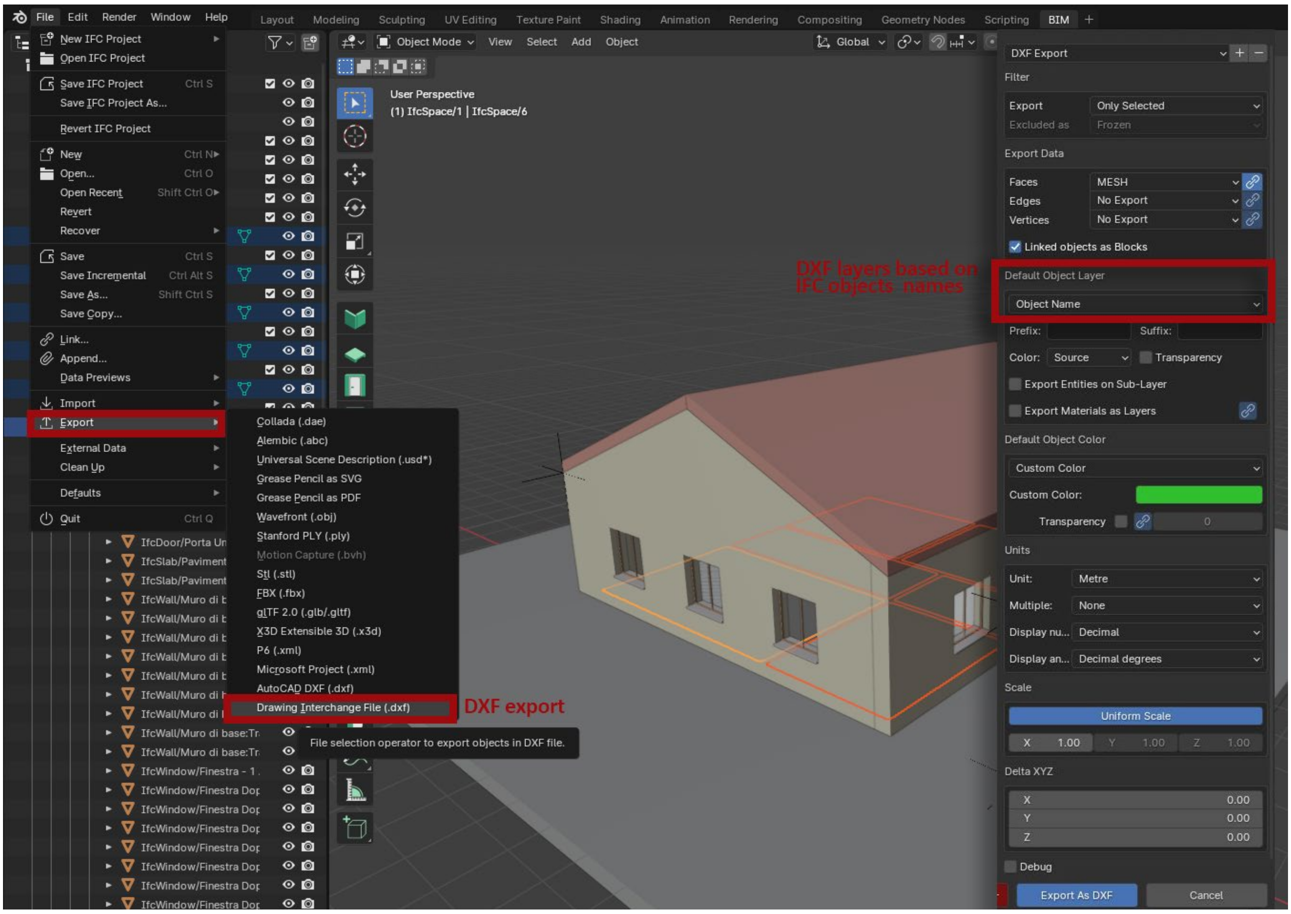


Figure 3: IfcSpaces export to .dxf (settings)

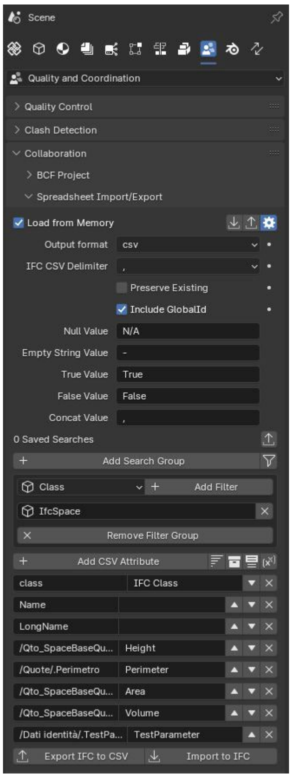


Figure 4: BlenderBIM – Export IFC to CSV and Import CSV to IFC settings for managing IfcSpaces data

The ".csv" table holding the additional alphanumeric IfcSpace data provides instead a suitable data source for a geometryless vector layer. In accordance with the advanced methodological approach, all fields of the csv-layer attribute table have been added to the attribute table of the IfcSpace layer including vector polygons by properly setting "Join Properties" of the "Target layer".

It is extremely important to underline that this operation ensures a join of attribute tables (see Figure 5) of the two involved layers without merging them, so as not to lose synchronization with the ".csv" data source of the geometry-less layer. Moreover, the IfcSpace "Name", shared by both attribute tables, provided a suitable "Join field" to correctly associate "IfcCSV" data with the respective IfcSpaces. It can be also noted that all attributes derived from the "Join layer" (*i.e.* the geometry-less csv-layer) are specifically marked with the prefix "IfcCSV_" in order to distinguish them from the layer original attributes instantly (see Figure 5). It should also be reported a critical issue related to the impossibility to directly edit attributes deriving from the "Join layer" through the attribute table of the "Target layer". Those attributes can be only managed by editing their original geometry-less layer synchronized with its ".csv" data source. Concluding, a vector layer including a georeferenced point, corresponding to the building footprint centroid, is added to the map, to hold general information about the building. To enhance the bi-directionality of the advanced BIM-GIS integration, the attribute table of the building centroid has been equipped with a special attribute (named "IFC model") to store an hyperlink (see Figure 6) to directly access the BlenderBIM project (holding the building IFC model).

BIM-GIS data visualization and management

As already declared, the GIS-based open-source tool QGIS is herein also adopted as GeoBIM data visualization and management platform. As far as data visualization is concerned, the IfcSpaces are graphically represented as vector polygons on the 2D map, whereas they are displayed as box objects on the 3D map, as long as the "IfcCSV_Height" attribute is assigned as extrusion value in the layer 3D view tab. Moreover, the building centroid is displayed as a georeferenced point on the map 2D view, whereas a detailed semi-transparent 3D hologram of the IFC model, based on the building ".OBJ" geometry exported from BlenderBIM, is associated as 3D model shape with the point 3D view properties (as shown in Figure 7). The detailed 3D hologram of the building is aimed at making BIM-derived buildings more easily identifiable from the box objects representing the built urban context. In Figure 7 some possible rule-based renderer filters are also shown. In this regard, the first exemplifying layer filters are simply based on the IfcSpace "names", which are also assigned as labels. The last two display filters, which are activated in Figure 7, are based on the Boolean value of the "IfcCSV_TestParameter", by assigning the red colour to "False" and the green colour to "True". It can be also noticed that the respective 2D and 3D layer filters, based on the same display rules, are simultaneously activated. Notwithstanding that the 3D view of the shown GeoBIM model could be beneficial in terms of objects visualization quality, the map 2D view turned out to be more easily accessible, queryable and manageable as informative system.

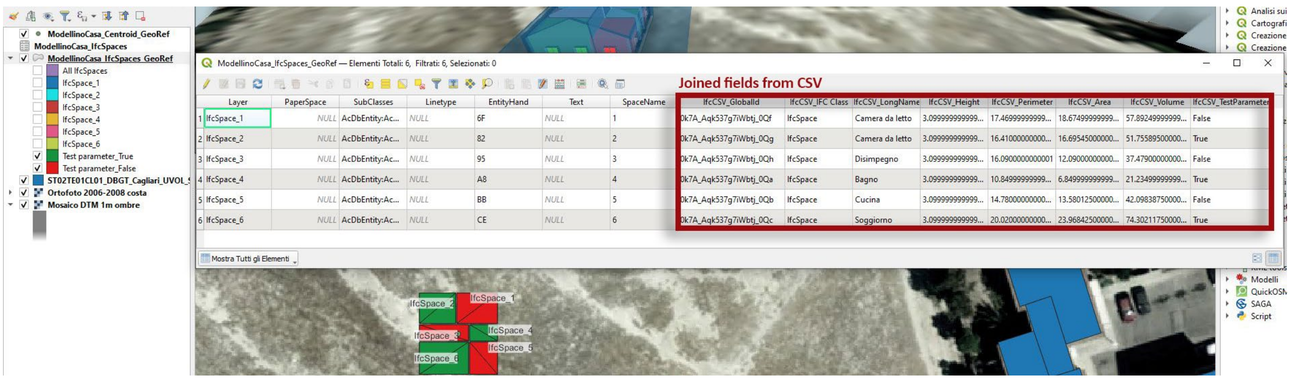


Figure 5: QGIS - Attributes table of the IfcSpaces (.dxf) polygons layer with additional "joined" attributes from .csv tabular data

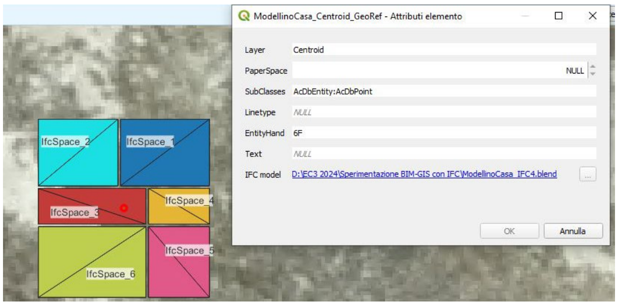


Figure 6: QGIS – IFC model hyperlink attribute

The conducted experimentation was finally concluded by validating the bidirectionality of the proposed BIM-GIS integration by first editing the "TestParameter" value of an IfcSpace on the GIS system and then successfully updating the IFC model by simply reimporting the related ".csv" file automatically updated thanks to data synchronization.

Conclusions

The integration of BIM and GIS methodologies and tools achieves high levels of effectiveness and efficiency in large building stock management. Starting from these premises, this paper shows the results of an ongoing research focused on the development of a methodological framework integrating BIM and GIS environments. The experimental phase conducted on an abstract illustrative case study allows the authors to highlight the strengths and weaknesses of the methodological framework developed. One of its strengths focuses on the exclusive use of open-source tools and open, international standards. This aspect is strictly bound with the important issues of interoperability, customization, verifiability, optimization, and low costs. The proposed methodology involves Blender-BIM, an open-source software that supports BIM approach, QGIS, an open-source software that allows users to manage geospatial information, and open standards like DXF and CSV. The work focuses on a bidirectional integration between the informative tools, by linking them and allowing for an easy switch from one system database to the other. Despite the benefits that this methodology offers, some potential critical issues have emerged. It is worth underlining the difficulties in data management and visualization through QGIS filters. Users must pay attention to manually activate both 2D and 3D display filters that are not synchronized although they are based on the same rules. Moreover, the 3D view is affected by some annoying selection bugs, which make it difficult to correctly query the 3D map. Another issue is the critical management of the attributes deriving from the "Join layer". These attributes can be only managed by editing their original geometry-less layer synchronized with its ".csv" data source. Finally, although the proposed workflow still needs to be improved and validated, it could be a solution to overcome some unfixed critical issues still affecting interoperability between commercial tools and enhance collaboration. Future work development will focus on the issues shown above.

Acknowledgments

We acknowledge financial support under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - Call for tender No.3277 published on December 30, 2021 by the Italian Ministry of University and Research (MUR) funded by the European Union – NextGenerationEU. Project Code ECS0000038 – Project Title eINS Ecosystem of Innovation for Next Generation Sardinia – CUP F53C22000430001- Grant Assignment Decree No. 1056 adopted on June 23, 2022 by the Italian Ministry of Ministry of University and Research (MUR). This manuscript reflects only the authors' views and opinions, neither the European Union nor the European Commission can be considered responsible for them. The research was also funded by "Piano Triennale della ricerca di Sistema del settore elettrico nazionale per il triennio 2022-2024», Progetto 1.5 "Edifici ad alta efficienza per la transizione energetica".

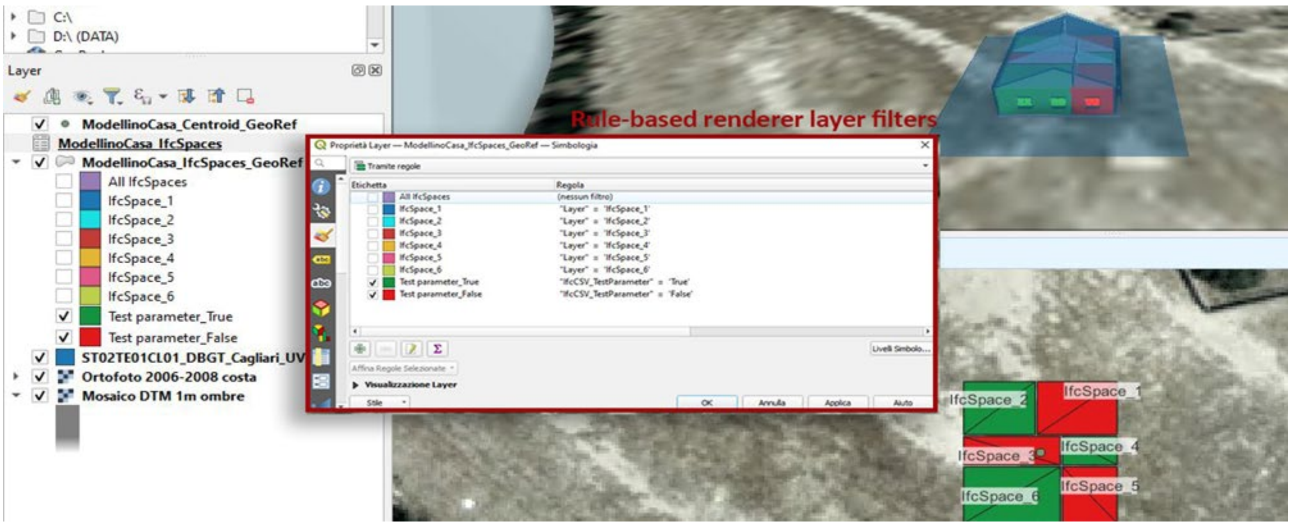


Figure 7: QGIS 2D and 3D maps with activated rule-based filters

References

Blender 4.0. (n.d.). Retrieved January 25, 2024, from https://www.blender.org/

BlenderBIM Add-on. (n.d.). Retrieved January 25, 2024, from https://blenderbim.org/

buildingSMART International. (n.d.). Industry Foundation Classes (IFC). Retrieved February 1, 2024, from https://technical.buildingsmart.org/standards/ifc/

Congiu, E., Quaquero, E., Rubiu, G., & Vacca, G. (2024) Building Information Modeling and Geographic Information System: Integrated Framework in Support of Facility Management (FM). Buildings, 14(3), 610. https://doi.org/10.3390/buildings14030610

D'Amico, F., Calvi, A., Schiattarella, E., Prete, M. Di, & Veraldi, V. (2020). BIM and GIS Data Integration: A Novel Approach of Technical/Environmental Decision-Making Process in Transport Infrastructure Design. Transportation Research Procedia, 45, p.pp. 803–810. https://doi.org/10.1016/j.trpro.2020.02.090

Jiang, S., Jiang, L., Han, Y., Wu, Z., & Wang, N. (2019). OpenBIM: An enabling solution for information interoperability. Applied Sciences (Switzerland), 9(24), 5358. https://doi.org/10.3390/app9245358

Kang, T. W., & Hong, C. H. (2015). A study on software architecture for effective BIM/GIS-based facility management data integration. Automation in Construction, 54, 25–38. https://doi.org/10.1016/j.autcon.2015.03.019

Liu, X., Wang, X., Wright, G., Cheng, J. C. P., Li, X., & Liu, R. (2017). A state-of-the-art review on the integration of Building Information Modeling (BIM) and Geographic Information System (GIS). ISPRS International Journal of Geo-Information 6(2), 53. https://doi.org/10.3390/ijgi6020053

Malewczyk, M. (2022). The usage of the openBIM idea in architectural design on the example of Blender and BlenderBIM add-on. Architectus, 2(66). https://doi.org/10.37190/arc210210

Mangia, M., Lazoi, M., & Mangialardi, G. (2022). Digital Management of Large Building Stocks: BIM and GIS Integration-Based Systems. In O. Canciglieri Junior, F. Noël, L. Rivest, & A. Bouras (Eds.), IFIP Advances in Information and Communication Technology: Vol. 639 IFIP (p.pp. 133–150). Springer, Cham. https://doi.org/10.1007/978-3-030-94335-6_10

Meschini, S., Accardo, D., Avena, M., Seghezzi, E., Tagliabue, L. C., & Di Giuda, G. M. (2022). Data integration through a bim-gis web platform for the management of diffused university assets. In L. C. Tagliabue, A. Chassiakos, D. M. Hall, & D. Nikolic (Eds.), Proceedings of the 2022 European Conference on Computing in Construction (Vol. 3). University of Turin. https://doi.org/10.35490/ec3.2022.217

Meschini, S., Pellegrini, L., Locatelli, M., Accardo, D., Tagliabue, L. C., Di Giuda, G. M., & Avena, M. (2022). Toward cognitive digital twins using a BIM-GIS asset management system for a diffused university. Frontiers in Built Environment, 8. https://doi.org/10.3389/fbuil.2022.959475

Moretti, N., Ellul, C., Re Cecconi, F., Papapesios, N., & Dejaco, M. C. (2021). GeoBIM for built environment condition assessment supporting asset management decision making. Automation in Construction, 130 (October 2021), 103859. https://doi.org/10.1016/j.autcon.2021.103859

QGIS. (n.d.). Retrieved July 7, 2023, from https://www.qgis.org/it/site/

Sammartano, G., Avena, M., Cappellazzo, M., & Spanò, A. (2021). Hybrid GIS-BIM approach for the torino digital-twin: The implementation of a floor-level 3D city geodatabase. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 43(B4-2021), p.pp. 423–430. https://doi.org/10.5194/isprs-archives-XLIII-B4-2021-423-2021

SardegnaGeoportale – Aree tematiche (n.d.). Retrieved July 7, 2023, from https://www.sardegnageoportale.it/areetematiche/

Satria, R., & Castro, M. (2016). GIS Tools for Analyzing Accidents and Road Design: A Review. Transportation Research Procedia, 18, p.pp. 242–247. https://doi.org/10.1016/j.trpro.2016.12.033

Vacca, G., Quaquero, E., Pili, D., & Brandolini, M. (2018). Integrating bim and gis data to support the management of large building stocks. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 42(4), p.pp. 717–724. https://doi.org/10.5194/isprs-archives-XLII-4-647-2018

Vankova, L., Krejza, Z., Kocourkova, G., & Laciga, J. (2021). Geographic Information System Usage Options in Facility Management. Procedia Computer Science, 196, p.pp. 708–716. https://doi.org/10.1016/j.procs.2021.12.067

Wang, H., Pan, Y., & Luo, X. (2019). Integration of BIM and GIS in sustainable built environment: A review and bibliometric analysis. Automation in Construction 103 (July 2019), p.pp. 41–52. https://doi.org/10.1016/j.autcon.2019.03.005

Zhao, L., Liu, Z., & Mbachu, J. (2019). Highway alignment optimization: An integrated BIM and GIS approach. ISPRS International Journal of Geo-Information, 8(4), 172. https://doi.org/10.3390/ijgi8040172.

DIGITAL TWINS IN BUILDING MANAGEMENT: BOOSTING EFFICIENCY IN RENEWABLE ENERGY COMMUNITIES

Giulia Rubiu [1], Emanuela Quaquero [1]
[1]University of Cagliari, Department of Civil Engineering, Environment and Architecture, Italy

Abstract

The study proposes a methodological approach for organizing, storing, and handling diverse data across disciplines using digital systems structured on the internet with a multi-level strategy. This workflow follows traditional knowledge process steps, storing and organizing outcomes in a unified web-based model using pre-existing ontologies. The goal is to distribute data among stakeholders and establish a decision support system for the energy community. The digital twin implementation integrates IoT technologies, Big Data analysis, and territorial modeling, enabling energy management solutions at both individual (residential) and larger (community) scales.

Introduction

The energy consumption of buildings is a pivotal factor in achieving sustainable development, encompassing economic, environmental, and social dimensions. This theme of sustainability is currently a high priority in global and European policies. In the European Union, buildings account for 40% of energy consumption and 36% of climate-altering emissions (Energy Efficiency in Buildings - European Commission.

In 2019, the European Community launched the European Green Deal, a comprehensive suite of strategic initiatives aimed at combating climate change and global warming. The primary objective of this plan is to achieve climate neutrality by 2050 (Fit for 55 - The EU's plan for a green transition - Consilium.

The European Green Deal highlights the need for a holistic and cross-sectoral approach to address climate challenges, where key sectors such as climate, environment, energy, transportation, industry, and agriculture work in synergy (EU 2019).

These initiatives are pivotal in reducing energy poverty, decreasing carbon emissions, and promoting energy independence(In focus: How can the EU help those touched by energy poverty?).

The emergence of new paradigms and innovative systems, such as Collective Self-Consumption models and Energy Communities, represents a significant shift in energy strategies. These community-focused energy models are transforming the landscape of renewable energy production, distribution, and consumption. Central to these models are the principles of localization and decentralization of energy generation.

Involving citizens, commercial entities, and local businesses, these models facilitate the creation, consumption, and exchange of energy with an emphasis on self-consumption and collaborative effort. Community energy systems and collective self-consumption facilities offer several advantages over traditional energy production models. These include the optimal combination of diverse consumption profiles, immediate utilization of generated energy, and a reduced load on existing grid infrastructures.

This approach not only contributes to environmental sustainability but also fosters a sense of community and collective responsibility towards energy consumption and climate action. Such initiatives, supported by policies like the European Green Deal, are essential steps towards a sustainable, energy-efficient future (Pisello et al. 2023).

Self-consumption of energy can take place at three different levels: individual, collective and community.

The concept of collective self-consumption is defined in the EU Directive 2018/2001(Directive EU - 2018/2001). The directive stipulates that the group of self-consumers (minimum two) who act collectively must be in the same building or condominium.

The text also stipulates that they can generate renewable electrical energy for their own consumption and can store or sell self-produced electrical energy. The main condition is that such activities do not constitute the main commercial or professional activity.

The consumer in the realm of renewable electricity is no longer merely a recipient of energy but has evolved into a 'prosumer.' The term 'prosumer' denotes a user who actively participates in all stages of energy production, as opposed to solely fulfilling the passive role of a consumer. A key advantage of this paradigm shift is the ability for individuals to act collectively, thereby sharing the costs and benefits of such systems, which may include joint maintenance and other pertinent practices.

Legally, participants in these systems share the energy they generate through the existing distribution network.

Each member is equipped with their own meter to measure individual consumption. However, to accurately assess the amount of energy shared for self-consumption purposes, the installation of a secondary metering system is necessary. This system is crucial for monitoring the distribution and sharing of generated energy among the participants.

Despite sharing the overarching goal of enhancing the management of renewable energy generation and usage, there are notable differences between Self-Consumption Units (SCUs) and Renewable Energy Communities (RECs). The primary distinction lies in their composition. An energy community, as defined by the EU Directives 2018/2001(Directive - 2018/2001) and 2019/944(Directive (EU) 2019/944 of the European Parliament and of the Council on common rules for the internal market for electricity and amending Directive 2012/27/EU, is typically comprised of associations of citizens, businesses, local administrations, and small enterprises. These entities collectively decide to produce, exchange, and consume energy derived from local renewable sources (Menegon et al. 2021). This collaborative approach not only fosters a sense of community but also aligns with broader objectives of sustainability and local empowerment in energy production and consumption (Notton et al. 2018).

The Digital Twin (DT) is revolutionising the energy sector by providing an innovative approach to management and monitoring. It enables the management of on-site energy exchanges, optimises the use of renewable resources, and allows users to monitor their own energy consumption (Xiong et al. 2021; Bortolini et al. 2022; Tahmasebinia et al. 2023; Testasecca et al. 2023). The future energy system finds a valuable support in the concept of the digital twin, as developed by Michael Grieves (Grieves and Vickers 2016).

According to Grieves, the Digital Twin (DT) system is based on three main elements: real-world physical products, their virtual counterparts, and the data connections between them. This setup allows the virtual model to mimic the real object's behavior, ensuring changes in the physical product are reflected in the digital one, often through IoT devices and sensors. (Barricelli et al. 2019). Digital twins in the construction sector facilitate the integration of the building's informational and physical models, thereby ensuring iterative optimization of both models. The DT leverages the geometric and parametric properties of Building Information Modeling (BIM) models, as well as streaming environmental data (e.g., temperature, humidity, consumption, etc.) collected from sensors (Brilakis et al. 2019). BIM serves as a common knowledge resource that can be employed as a singular source for data management and collection, facilitating the sharing, preservation, and provision of reliable guidance throughout the building's lifecycle. Data access is enabled through various mechanisms, including proprietary system manual interfaces, Application Programming Interfaces (APIs), and exporting via standard formats such as Industry Foundation Classes (IFC). To date, a unified standard in the IoT and BIM domain has yet to be established, yet numerous open standards are emerging. These are intended to provide a common language across different systems and devices, enabling interoperability and seamless communication among various entities. However, research on the integration between BIM and IoT remains in its nascent stages, where most studies primarily introduce conceptual theoretical propositions (Mengistu and Mahesh 2020; Tai et al. 2021). The importance of digital twins in the construction sector is underscored by the emerging demands for intelligent systems that are capable of representation, analysis, identification, and the enhancement of consumption patterns(Jradi and Bjornskov 2023). Efficient energy management necessitates a reliable data source through the transmission of data from an IoT infrastructure. Employing advanced methods such as machine learning, artificial intelligence, and deep learning, energy community managers can monitor, collect, and analyse data to produce reports, charts, and integrate informational models. Predictive analytics can be utilized to foresee potential issues and undertake preventative measures before they escalate into problems. With these tools, energy community managers make more informed decisions to the benefit of the RECs. Energy data analysis is a crucial aspect of understanding and managing consumption. By interpreting the collected data, RECs members can gain valuable insights into their energy consumption patterns and trends. This knowledge can assist them in making more informed decisions on how to best utilize their resources and reduce costs (Al-Ali et al. 2017; Adu-Kankam and Camarinha-Matos 2023). The digital twin is revolutionising the energy sector by amalgamating the prowess of IoT technology, Big Data analytics, and informational and spatial modelling. This facilitates the deployment of energy management solutions both at a granular level (individual residential units) and at a broader scale (energy communities). This endeavour seeks to delineate a methodological approach that affords RECs members facile access to data and to devise a system that aids them in making well-informed decisions. The methodology employs a variety of tools and techniques, such as statistical analysis and machine learning, to scrutinise data and generate reports.

The principal objective is to proffer an efficacious framework for the cataloguing, storage, and management of multidisciplinary data and information utilising web-based digital systems. This system ought to be scalable, secure, and user-friendly, enabling users to access and meticulously analyse data with alacrity. The chief aim of the methodology lies in the dissemination of reports among the diverse REC members, as well as in the development of a decision-support system.

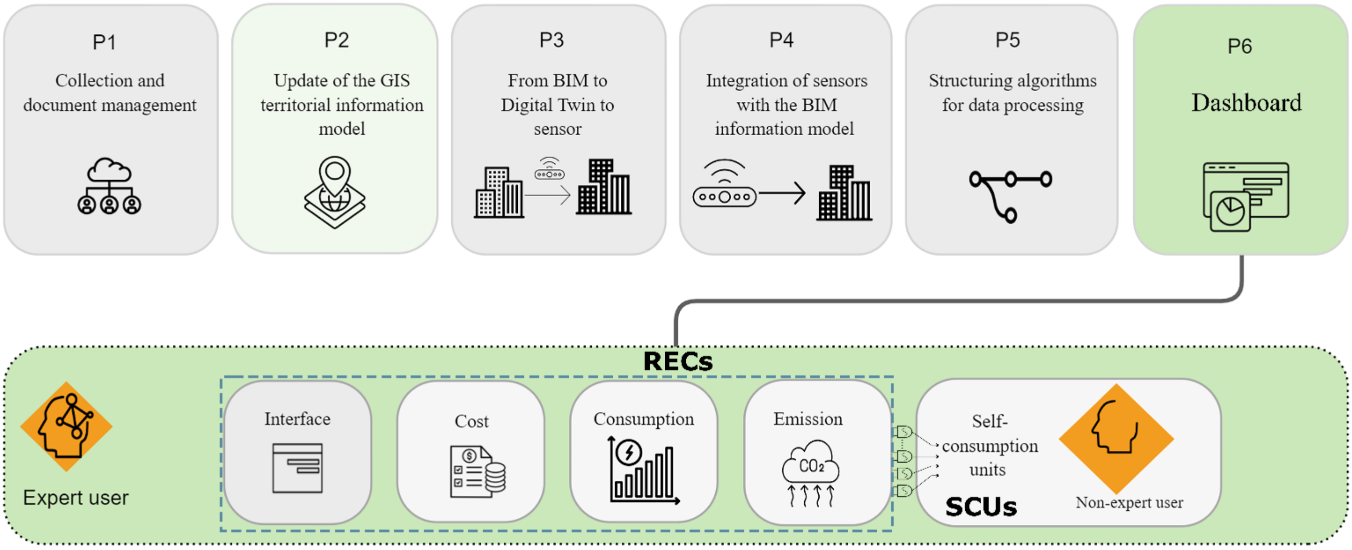


Figure 1 Methodological Framework

Methodology

A comprehensive framework designed to bolster self-consumption units and complex systems like Energy Communities. This research work provides an analysis of current tools, their features, and the best practices for their application to develop an effective management and monitoring infrastructure. The system facilitates the creation of geographical interfaces for community energy operators through the integration of digital twin technologies. Digital Twins (DTs) lack sufficient autonomy to optimise consumption based on information received from Building Information Modelling (BIM) and Internet of Things (IoT) systems, as the unpredictable behaviour of occupants plays a pivotal role. Energy consumption is significantly influenced by user intentions and goals, as well as their expectations in terms of energy savings. With the support system, all stakeholders can better manage the community's economic framework and energy efficiency programmes. REC members have access to the latest updates on the energy system and benefit from reliable recommendations to make informed decisions.

The framework designed to support the management and monitoring processes of the energy community is based on a portal (Web-GIS) containing specific interfaces for sharing information about the energy community and collective self-consumption units. The digital techniques underpinning the methodological framework include Geographic Information Systems (GIS-based), useful for georeferencing different energy communities, primary and secondary substations, followed by architectural entities. The digital techniques underpinning the methodological framework include Geographic Information Systems (GIS-based), useful for georeferencing different energy communities, primary and secondary substations, followed by architectural entities.

i. BIM digital models of buildings structured according to the proprietary guidelines of the energy community. BIM models are implemented on the platform using web-based graphic libraries.

ii. Relational databases structured for the systematisation and querying of data.

iii. Machine Learning algorithms for energy profiling divided into clusters and future predictions of consumption and costs.

iv. Web-based systems for report provision.

The proposed methodology is divided into document collection and management; updating the GIS territorial information model; from BIM to Digital Twin; integration of sensors with the BIM information model; structuring algorithms for data processing and conceptualisation of the dashboard.

P1: Collection and Document Management

This part is dedicated to the cataloguing and administration of documents within the energy community, as well as the pivotal elements that must be incorporated into the process. Entities such as local authorities, citizens, associations, and businesses are required to upload data pertinent to the RECs or collective SCUs. The assemblage and management of documents entail the structuring of a pre-established repository. The subsequent step involves their cataloguing to facilitate ease of search and retrieval. This includes the employment of metadata tags and other forms of categorisation to ensure that data are readily accessible. Once organised and catalogued, the data must be stored in a secure and reliable manner. The most critical phase of the process is data management, which entails ensuring regular updates and protection of the data, and that all

Figure 2 Illustration of a GIS spatial model applied to RECs

participants in the REC have access to the necessary information. The documents must provide clear instructions and information on different aspects of the energy community. Technical reports, safety verifications, regulatory documents, and statutes must all be included in this list. The collection of this information is essential to support the community. To ensure secure and reliable storage, the system must be hosted on the cloud or another reliable storage infrastructure. The collection and management of documents are fundamental tools to facilitate the efficient functioning of the REC.

P2: Update of the GIS Territorial Information Model

The second section is dedicated on sharing a geographical representation of the current state of the energy community. A colour-coding system identifies the various communities with individual self-consumption units. Regarding buildings, there are three distinct types of input data essential for defining the territorial information model. The first type is derived from national statistical institutes' censuses across Europe which furnish essential details regarding the age of buildings. The second type of data emanates from regional topographical databases, which include, among other details, information on the buildings' intended use. The final categorisation originates from Building Information Modeling (BIM) systems, incorporating both geometric data and consumption metrics.

The confluence of these databases facilitates the determination of specific characteristics for each building, such as its age class, surface-to-volume (S/V) ratio, number of storeys, and occupant count. Utilizing the GIS system, in conjunction with previously prepared BIM models, smart meters' analyses are conducted to glean precise consumption data. Following the delineation and consumption allocation to building clusters, the data are aggregated into summary tables. Grasping the behavioural dynamics of a community of buildings throughout the year is paramount, as it harbours significant practical ramifications for the management of RECs.

P3: From BIM to Digital Twin

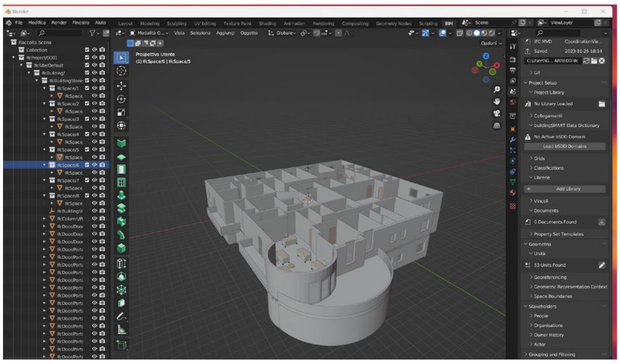


Figure 3 Building Information Model (Simplified Representation), Blender with BIM Add-on

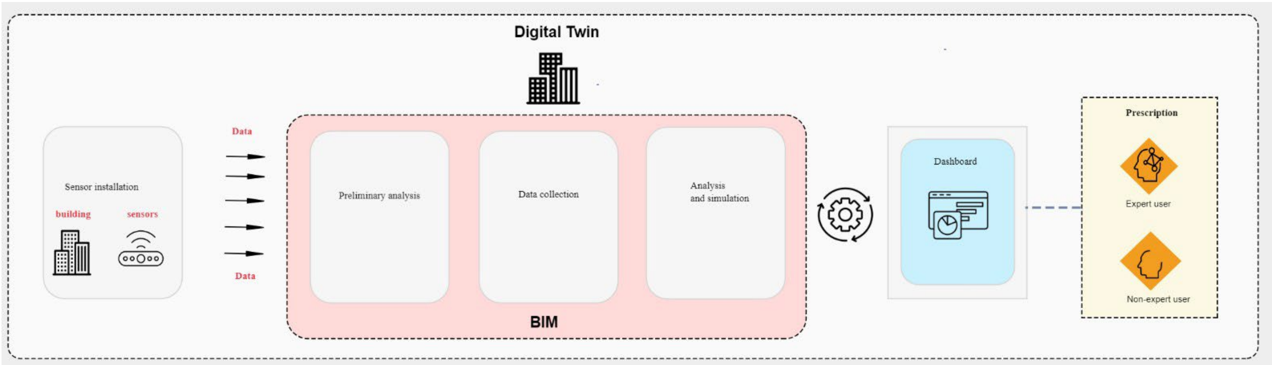


Figure 4 Flowcharts: from BIM to Digital Twin

A digital twin can be conceived in several stages, yet it is most efficacious when established in the initial design phase of self-consumption units; hence, the data amassed can serve as a foundation for managing subsequent stages.

Building Information Modelling (BIM) is the most apt approach for creating the building's digital twin, as it is integrated with parameters functional to the monitoring phase. Furthermore, it is employed as a source for the collection and management of data which is conveyed to the GIS-based territorial system. Although BIM encompasses information about the building and space, it cannot solely provide comprehensive insights into the building's behaviour upon which to base decisions for the community or individual users. To this end, this endeavour integrates the model with information gathered in the initial phase (P1), such as the building's typology, residential units, and the occupancy level of the users. To ensure the correct implementation of the digital twin, it is crucial to associate all necessary documents for the selection and installation of the sensor network with the BIM. Supported by the informative model, we are capable of precisely determining the optimal placement for each individual sensor, thereby enabling the REC to maximise the network's performance and minimise the costs associated with installation.

P4: Integration of Sensors with the BIM

The objective of this phase is to identify the most suitable type of sensors for the construction of a monitoring system, taking into account the requirements and spatial properties of the structures identified in the previous analysis (P3). As part of this step, the techniques for sampling and transmitting data from the control devices are outlined and implemented. The application of smart meters for monitoring energy consumption, along with sensors, facilitates the analysis of usage habits within individual residential units. Utilising data from both external and internal temperatures can enhance the prediction of energy consumption. It is essential to determine the type of monitoring device that best suits the specific needs of the building. This task involves a thorough analysis of performance, application limits, and costs of available technologies. Given that the expenditure required for hardware is one of the most significant concerns for households, this study leans towards a non-intrusive monitoring system with economical sensors that are easily programmable to minimise the cost of the entire system. BIM can provide significant assistance in this phase to define various sensor configurations and compare them as an additional source of information in the building.

In addition to the technical and economic considerations, the security of the data collected and transmitted by these sensors warrants careful attention. The integrity and confidentiality of data within the monitoring system are paramount, necessitating the implementation of robust encryption protocols to safeguard data in transit and at rest. User authentication and access controls are crucial to ensure that only authorized individuals can access sensitive information, thereby preserving the privacy of the residential units involved. The proposed system will also explore the adoption of blockchain technology to provide an immutable ledger for data transactions, enhancing trust and transparency within the community. By incorporating these security measures, the monitoring system not only optimizes energy consumption and operational efficiency but also ensures the protection of sensitive data, aligning with global standards for data privacy and security.

P5: Data Processing Algorithms

In the context of refining the BIM information model, it is essential to elaborate on the development of sophisticated algorithms tailored for the efficient management of data captured by monitoring sensors. The integration of sensor-generated data with BIM-centric applications is achieved through Python scripting. These scripts are not merely facilitative but are at the core of the data processing mechanism, enabling a dynamic and seamless exchange of data between the sensor array and the BIM ecosystem. Such integration is pivotal for real-time data analysis and automated responses within the BIM framework.

Further enhancing the functionality of these scripts, machine learning algorithms are deployed to scrutinize patterns in energy consumption meticulously. These algorithms are adept at sifting through vast datasets to identify anomalies and inefficiencies that may not be immediately apparent. By learning from historical data, they can predict future trends, enabling preemptive actions to optimize energy use and reduce waste. This is particularly beneficial in managing complex systems such as HVAC and in the predictive maintenance of various devices and machinery, where early detection of potential issues can prevent costly downtimes and repairs.

The strategic use of machine learning in this domain not only enhances operational efficiency but also contributes significantly to energy conservation efforts. By analyzing consumption patterns and identifying areas of excessive use, these algorithms can provide actionable insights for energy savings. This leads to the development of targeted recommendations that are both pragmatic and feasible, ensuring sustainable energy utilization across different applications. Moreover, the predictive capabilities of machine learning algorithms extend to forecasting maintenance needs, thereby alerting facility managers and consumers well before potential system failures occur, ensuring reliability and continuous operation.

P6: Dashboard Conceptualization

The dashboard is conceptualized as an interactive tool for the visualization and interpretation of data originating from a centralized storage system. It employs an analytics engine to convert raw data into intelligible reports, charts, and diagrams. The dashboard is customized to cater to diverse user groups, each with varying levels of expertise and access rights.

For novice users, the dashboard is made accessible through a cross-platform mobile application, enabling them to view summary charts that provide an overview of pertinent data related to their self-consumption units, along with a straightforward depiction of REC. Conversely, community managers are provided with a more comprehensive desktop interface for conducting detailed analyses, thereby facilitating informed decision-making processes.

The dashboard architecture is predicated on user privilege levels, thereby dictating the granularity of the presented data. Expert users are afforded extensive visibility, ranging from a holistic view of the energy community to detailed information on self-consumptions units. Novice users, on the other hand, are presented with data relevant to their specific dwelling units, complemented by summary reports that encapsulate key information about their own energy community.

The dashboard is segmented into four main sections, each designed to highlight different aspects of the energy community's data:

General Data: This section employs geographical visualization to outline the energy community's layout, identifying communities and associated substations through a color-coding scheme. An integrated alert mechanism proactively notifies expert users of necessary actions within the RECs framework. Additionally, this section provides direct access to essential REC documentation, such as community statutes and incentive management guidelines.

Costs and Consumption: Acknowledging the significance of financial and consumption metrics in energy resource management, this section provides insights into energy costs, consumption patterns, and future projections. It includes a comprehensive report detailing the energy costs of community-level and individual self-consumption units, breakdowns by energy source, incentive allocations, comparative analyses with similar communities, and future cost and consumption forecasts. Graphical alerts are utilized to signal anomalies, encouraging users to modify their consumption behaviors.

Emissions: This segment enables community members to monitor and assess the REC's contribution to CO2 emission reduction, providing a tangible measure of the community's sustainability efforts.

Self-Consumption Units (Novice User): Tailored for novice users, this section offers an interface that allows users to access summary information from the previous sections specific to them own self-consumption unit, thereby facilitating an intuitive and streamlined user experience.

The dashboard is designed not merely to present data but also to enhance user engagement and awareness of specific aspects of the energy community, thereby enabling users to make informed decisions based on the presented data.

Discussion

The deployment of this dashboard heralds a substantial advancement in the domain of renewable energy management, pertinent to both SCUs and RECs. Its primary objective is to streamline the engagement with sustainable energy practices, rendering them more accessible and intuitive to the end-users.

Within the confines of SCUs, which serve as the nucleus of individual energy production and consumption dynamics, the dashboard emerges as an instrumental tool. It leverages real-time monitoring capabilities to discern patterns of peak energy production, often coinciding with troughs in energy consumption. This synchronicity allows the dashboard to advocate for the scheduling of energy-intensive tasks during periods of abundant renewable energy generation, thereby optimizing the utility of renewable resources.

Expanding its utility to the broader ambit of RECs, the dashboard plays a pivotal role in the nuanced orchestration of energy flows among a consortium of interconnected entities. It aids in the strategic allocation of surplus energy, ensuring its efficacious distribution within the community. This strategic approach to energy management engenders a harmonious balance within the REC, mitigating the challenges posed by the inherent variability of renewable energy sources.

Furthermore, the dashboard transcends its operational utility by incorporating a predictive maintenance feature, crucial for the proactive upkeep of renewable energy systems. By continuously monitoring the health and efficiency of these systems, it preemptively identifies potential issues, thereby safeguarding the continuous optimal performance of energy assets.

In its essence, the dashboard also serves as a decision-support system, facilitating strategic planning for the expansion or augmentation of renewable energy infrastructures within RECs. By amalgamating diverse data streams, including historical energy consumption patterns, current trends, and predictive analytics, it furnishes a holistic view of the energy landscape. This empowers stakeholders with the necessary insights for informed decision-making regarding future investments in renewable energy technologies, thus reinforcing the community's trajectory towards sustainable energy practices.

Conclusions

This work outlines a comprehensive methodological framework designed to bolster the efficacy and operational efficiency of Energy Communities (RECs).

The primary objective of this initiative is to devise an advanced, user-friendly dashboard that offers real-time insights into the performance metrics of RECs, alongside granular data on each individual self-consumption unit. This is particularly focused on providing detailed information regarding costs and consumption patterns. The significance of this contribution lies in its ability to facilitate industry experts in their endeavors to effectively manage and monitor RECs by leveraging sophisticated digital models, cutting-edge geographic information systems, and robust web-based relational databases.

Addressing the pervasive issue of data fragmentation and cognitive heterogeneity in the building sector is paramount. This sector is characterized by the disparate structuring of data across numerous repositories, which inevitably leads to critical information gaps. This fragmentation hinders the seamless management and interpretation of data within energy communities or collective self-consumption systems. Therefore, it is of utmost importance to equip energy community managers with a streamlined, intuitive platform that presents essential information in a clear, accessible manner. This approach is vital for bridging the information divide and enhancing the decision-making process.

In parallel, the promotion of a smart energy system is instrumental in motivating individuals to alter their energy consumption behaviours towards more sustainable and environmentally friendly practices. This necessitates a robust support framework that not only facilitates the digital transformation of services and the automation of processes but also emphasizes the empowerment of REC members. The analysis conducted by the International Energy Agency underscores the critical role of citizen involvement in the successful adoption and optimal utilization of future technologies within the energy sector.

Looking forward, the potential for engaging members and fostering a sense of credibility and trust within local communities cannot be overstated. Hence, future research endeavors will be directed towards exploring the realms of training, knowledge dissemination, and strategic communication activities. These areas are identified as pivotal for the sustained growth and impactful operation of RECs. Such initiatives will aim to enhance the collective understanding of energy management practices, promote widespread adoption of sustainable energy solutions, and ultimately contribute to the overarching goals of energy efficiency and environmental stewardship.

Acknowledgments

AUREE Project is funded by “Piano Triennale della ricerca di sistema del settore elettrico nazionale per il triennio 2022-2024», Progetto 1.5 “Edifici ad alta efficienza per la transizione energetica”.

References

Adu-Kankam, K.O. and Camarinha-Matos, L.M. 2023. Collaborative Digital Twins: The Case of the Energy Communities. SN Computer Science 4(5), pp. 1–22. Available at: https://link.springer.com/article/10.1007/s42979-023-02050-2 [Accessed: 27 January 2024].

Al-Ali, A.R., Zualkernan, I.A., Rashid, M., Gupta, R. and Alikarar, M. 2017. A smart home energy management system using IoT and big data analytics approach. IEEE Transactions on Consumer Electronics 63(4), pp. 426–434. doi: 10.1109/TCE.2017.015014.

Barricelli, B.R., Casiraghi, E. and Fogli, D. 2019. A survey on digital twin: Definitions, characteristics, applications, and design implications. IEEE Access 7. Available at: https://ieeexplore.ieee.org/document/8901113 [Accessed: 12 January 2023].

Bortolini, R., Rodrigues, R., Alavi, H., Vecchia, L.F.D. and Forcada, N. 2022. Digital Twins’ Applications for Building Energy Efficiency: A Review. Energies 15(19), p. 7002. Available at: https://www.mdpi.com/1996-1073/15/19/7002/htm [Accessed: 19 March 2023].

Brilakis, I., Pan, Y., Borrmann, A. and Mayer, H.-G. 2019. Built Environment Digital Twinning. (December), pp. 17–18. Available at: https://aspace.repository.cam.ac.uk/handle/1810/318329 [Accessed: 31 January 2023].

Directive - 2018/2001 - EN - EUR-Lex. [no date]. Available at: https://eur-lex.europa.eu/eli/dir/2018/2001/oj [Accessed: 24 January 2024].

Directive (EU) 2019/944 of the European Parliament and of the Council on common rules for the internal market for electricity and amending Directive 2012/27/EU (recast). | FAOLEX. [no date]. Available at:

https://www.fao.org/faolex/results/details/en/c/LEX-FAOC187976/ [Accessed: 24 January 2024].

EU. 2019. European Green Deal - Consilium. Available at: https://www.consilium.europa.eu/en/policies/green-deal/ [Accessed: 24 January 2024].

Fit for 55 - The EU's plan for a green transition - Consilium. [no date]. Available at: https://www.consilium.europa.eu/en/policies/green-deal/fit-for-55-the-eu-plan-for-a-green-transition/ [Accessed: 24 January 2024].

Grieves, M. and Vickers, J. 2016. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. Transdisciplinary Perspectives on Complex Systems: New Findings and Approaches, pp. 85–113. doi: 10.1007/978-3-319-38756-7_4.

In focus: Energy efficiency in buildings - European Commission. [no date]. Available at: https://commission.europa.eu/news/focus-energy-efficiency-buildings-2020-02-17_en [Accessed: 24 January 2024].

In focus: How can the EU help those touched by energy poverty? [no date]. Available at: https://commission.europa.eu/news/focus-how-can-eu-help-those-touched-energy-poverty-2022-02-16_en [Accessed: 24 January 2024].

Jradi, M. and Bjornskov, J. 2023. A Digital Twin Platform for Energy Efficient and Smart Buildings Applications. 2023 5th International Conference on Advances in Computational Tools for Engineering Applications, ACTEA 2023, pp. 1–6. doi: 10.1109/ACTEA58025.2023.10194071.

Menegon, D., Lobosco, D., Micò, L. and Fernandes, J. 2021. Labeling of Installed Heating Appliances in Residential Buildings: An Energy Labeling Methodology for Improving Consumers' Awareness. Energies 2021, Vol. 14, Page 7044 14(21), p. 7044. Available at: https://www.mdpi.com/1996-1073/14/21/7044/htm [Accessed: 14 January 2023].

Mengistu, D.G. and Mahesh, G. 2020. Dimensions for improvement of construction management practice in Ethiopian construction industry. Available at: https://idr.nitk.ac.in/jspui/handle/123456789/10596 [Accessed: 31 January 2023].

Notton, G., Nivet, M.L., Voyant, C., Paoli, C., Darras, C., Motte, F. and Fouilloy, A. 2018. Intermittent and stochastic character of renewable energy sources: Consequences, cost of intermittence and benefit of forecasting. Renewable and Sustainable Energy Reviews 87, pp. 96–105. doi: 10.1016/j.rser.2018.02.007.

Pisello, A.L., Piselli, C. and Pioppi, B. 2023. Background: the current energy community implementation state in the EU. Net-Zero and Positive Energy Communities, pp. 19–32. Available at: https://www.taylorfrancis.com/chapters/edit/10.1201/9781003267171-2/background-current-energy-community-implementation-state-eu-anna-laura-pisello-cristina-piselli-benedetta-pioppi [Accessed: 24 January 2024].

Tahmasebinia, F., Lin, L., Wu, S., Kang, Y. and Sepasgozar, S. 2023. Exploring the Benefits and Limitations of Digital Twin Technology in Building Energy. Applied Sciences 2023, Vol. 13, Page 8814 13(15), p. 8814. Available at: https://www.mdpi.com/2076-3417/13/15/8814/htm [Accessed: 24 January 2024].

Tai, S., Zhang, Y. and Li, T. 2021. Factors affecting BIM application in China: a social network model. Journal of Engineering, Design and Technology 19(2), pp. 373–384. doi: 10.1108/JEDT-12-2019-0330.

Testasecca, T., Lazzaro, M. and Sirchia, A. 2023. Towards Digital Twins of buildings and smart energy networks: Current and future trends. 2023 IEEE International Workshop on Metrology for Living Environment, MetroLivEnv 2023 - Proceedings, pp. 96–101. doi: 10.1109/METROLIVENV56897.2023.10164035.

Xiong, T., Cheng, Q., Yang, C., Yang, X. and Lin, S. 2021. Application of Digital Twin Technology in Intelligent Building Energy Efficiency Management System. Proceedings - 2021 International Conference on E-Commerce and E-Management, ICECEM 2021, pp. 393–396. doi: 10.1109/ICECEM54757.2021.00083.

OPTIMIZING ROAD INFRASTRUCTURE DESIGN USING I-BIM TECHNOLOGY

Panagiotis Tsikas, Rebeca Giavi, Athanasios Chassiakos
Department of Civil Engineering, University of Patras, Patras, Greece

Abstract

Infrastructure Building Information Modeling (I-BIM) has emerged as an important supporting technology for managing infrastructure project design, construction, and operation. This paper presents the development of an I-BIM model that can holistically capture the components and their integration characteristics of a road project. Extending previous works, the development deems to provide detailed road design analysis including drainage, signage, and lighting, traffic lanes of variable width, inclined pavement layers, vegetation and plantation, etc. In addition, the difficulties in developing “irregular” components and aligning them with each other and with the longitudinal land terrain are examined. The development applies to a real-world road design to explore the applicability potential.

Introduction

Transport infrastructure projects, being large scale engineering ones, require significant capital investments for their construction and maintenance, in return to multiple benefits for the users and the society out of their operation (Timilsina et al. 2020). In fact, they are quite important projects playing a key role in the social and economic development of regions and countries, often echoing the economic and technological development of them (Cigu et al. 2019).

The continuous improvement and application of new methods to improve the design and operation of such projects is always a goal (Costin et al. 2018). Considering the potential and positive effect of applying BIM technologies in construction (Wong and Fan 2013), the application and use of such tools in infrastructure project management could provide an effective method to increase efficiency during the project life cycle.

In large scale infrastructure projects, cost overruns, completion time overflow, wastefulness, and performance loss constitute conventionality rather than exception (Love et al. 2014). Engineering challenges - arising mainly from the specificities of such projects - exhibit a wide variety in terms of structural, geometrical, and mechanical characteristics. On the other hand, such projects are spawned by various multidisciplinary groups (architects, engineers, contractors, operators, etc.) having different (and often conflicting) priorities and objectives at every project phase. As a result, there is a large volume of information, generated from start to finish, subject to managing deficiencies and performance degradation.

Liu et al. (2017), after studying the most important risks in engineering projects through a questionnaire shared with 500 construction companies, noted that project performance - in terms of time, cost, quality, and safety - is significantly affected by the design team capability, the information accuracy, and the design associated delays. Undoubtedly, all the above factors contribute to problems and risks that infrastructure projects face nowadays and in the past. The root cause, according to Flyvbjerg (2013), is the fact that project planners tend to systematically underestimate or even ignore project complexity risks during the project development phase and decision making. Thus, many problems are associated with project design (or communication about the design), while modifications made alongside the way are related to deficiencies or omissions by the design.

As a result, both the productivity and the efficiency of the projects remain modest (Agenda, I. 2016). In this context, the traditional ways for managing infrastructure projects cannot meet the challenges of today, where construction is becoming more and more complex and demanding, and these projects require adequate planning and prudent resource management. Teizer (2015) stresses that key issues in the direction of improving infrastructure project management include schedule obedience, material procurement, supply chain-related actions, project status control, safety and quality monitoring. Furthermore, as Chi et al. (2012) state, during the evolution of a project, interim and ad hoc works arise. Because of this, the design of such temporary works does not always receive the attention and control levels that are typically required. Therefore, these elements should be integrated within the design, construction performance and safety parameters.

The introduction and use of BIM technology in the construction industry has assisted in automating traditional processes, manifesting positive effects in quality and productivity (Ullah et al. 2019). In addition, the adoption of BIM has significantly contributed to the digital transformation of the general industry, moving from the traditional way of designing and planning typical construction projects (e.g., buildings) towards a digital ecosystem that includes 3D visualization and real-time communication. Instead, in large infrastructure projects, the application and adoption level of BIM technologies is lacking behind, resulting in lower productivity and efficiency outcomes.

In recent years, there have been steps forward in the application and use of BIM in the infrastructure sector, as the benefits of 3D modelling and the use of smart objects are increasingly recognized (Tang et al. 2020; Cantisani et al. 2022). A 3D/BIM model is not limited to the geometric and design features visualized in the simulation (Gould 2010) but it is like a digital simulation model of the actual structure, whose properties and features are fully configurable. Thus, BIM technology is a reliable design tool that provides access to a comprehensive set of knowledge about the form, materials, environment, technical characteristics, costs, etc. of the project.

Despite the numerous benefits of I-BIM and the development of implementation strategies, the application of this technology remains limited, and the overall effectiveness of BIM in real projects has not been fully studied. Therefore, there is a need to explore the practical application of Infrastructure BIM (I-BIM) technology in the planning and design of projects. In this paper, the development and description of such a system for road projects is presented and discussed.

Infrastructure BIM

The development of road infrastructure projects involves two phases, the design and the construction ones. There are several sub-stages, in which plans and documents are produced, relating, among others, the technical specifications of the design, the cost and the duration of the construction.

In the traditional approach, communication among stakeholders has been largely fragmented and based on the exchange of 2D drawings and documents. As a result, coordination is particularly difficult, as much information (data) is lost or miscommunicated in the transition from stage to stage, often leading to the need for re-creation. Further, in large-scale projects, there are always large volumes of data and, likewise, it is difficult to communicate and present them correctly. Consequently, data omissions or miscommunication lead to inaccurate information sharing among participants and project phases. As such, when errors are detected through the several steps, the project has already progressed making any needed intervention difficult, costly, time consuming, and often of inadequate effectiveness.

In contrast, in infrastructure BIM (I-BIM) management, the information flow is coordinated in real time so that the work execution can be carried out by all involved parties (of different branches) effectively, as the digital simulation of the construction is up to date to the needed modifications and project progress monitoring. Thus, decision making for changes can be carried out in the early stage where the cost of changes is low and their effectiveness is high. Hence, more efficient collaboration is achieved by all groups involved in all phases throughout the project life cycle, eventually leading to an effective design and construction decision making.

Methodology

I-BIM provides a comprehensive 3D visual database and modeling of the project under development, integrating a range of data, including 3D terrain models and road design elements, such as road geometric design, cross-sections, measurements, electromechanicals, hydraulics, and more. This allows for the effective design of road systems and the optimization of project outcome. By integrating more data into the existing 3D model, i.e., work scheduling and construction cost, project managers can validate road project conformance to standards before construction commencement, leading to enhanced project efficiency and effectiveness. Therefore, the use of I-BIM in the initial design phase of large and complex horizontal geometry projects, can significantly contribute to different scenarios evaluation by providing visual, qualitative, and quantitative information that helps in the direction of selecting the best scenario that minimizes the time and cost and improves the quality of the work.

The design methodology of the road project through I-BIM includes the following steps, in accordance with the conventional way of doing things:

- Capture the terrain surface along the road alignment.
- Model the road and its individual elements (road deck, positions of sidewalks, roundabouts, islands, crossings, overpass bridge).
- Visualize the core road elements and layers in a 3D representation.
- Model and present secondary road elements that are necessary for road operation (signs, lighting fixtures, planting objects, etc).

Unlike traditional techniques and programs, all elements and characteristics of the project are present in a single 3D/I-BIM model. This enables the visualization of the road design in a realistic environment, which is helpful in identifying any potential conflicts that may arise. Most importantly, alternative design scenarios can be easily simulated and evaluated to realize the best layout solutions. Further, the I-BIM technology facilitates collaboration among the design team members by allowing them to access and modify the single digital model in real-time.

Interoperability and data standardization, using openBIM technology and Industry Foundation Classes (IFC), are vital for efficient information exchange among different stakeholders in the construction process. In this direction, the IFC 4.3 format is a significant step toward automating and optimizing facility management within the project lifecycle. In the road design study, interoperability helps transferring topographical mapping from AutoCAD Civil 3D to Revit for further processing and enhancing decision-making and stakeholder communication in project design.

Case study

As part of the development, this work presents as a case study the employment of the I-BIM technology to the design of the new ring road in Nemea (Greece). The 3D modeling of the road and its elements (road surface, sidewalk, roundabouts, electric lighting, technical works, etc.) has been realized using the AutoDesk REVIT software. Such kind of software have been developed with an initial focus on building design and construction, at least in construction industry applications. It appears, however, that they can effectively be used for other types and large-scale projects.

The terrain design of the area has been based on topographic CAD diagrams, which are first adjusted to their actual dimensions during modeling in REVIT. The adoption of BIM in this project has largely focused on design activities. The existing road passes through the urban area of the city of Nemea, as shown in Figure 1. As a result, traffic congestion is observed within the city that results in increased travel times and accident risk. The latter is more evident at an intersection where the new road is going to commence, in which case the intersection angles of the current geometry of the converging roads provide very poor visibility conditions (Figure 2). The bypass follows a new alignment, totaling 1.4 km, starting at location (1) and ending at location (2), as shown in Figures 2 and 3. The new road design includes two roundabouts at its ends to ensure high safety conditions and smooth traffic flow.

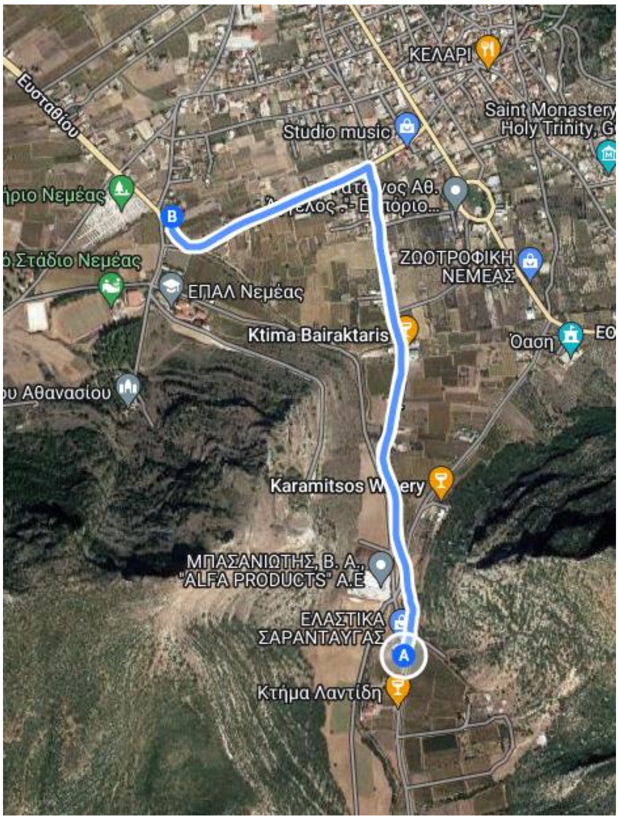


Figure 1: Existing road through the urban area

To select the optimal layout, several design alternatives have been developed and evaluated along the land zone around the potential layouts, with special emphasis to the roundabouts that link the ring road with the existing road segments. The proposed layout is shown in Figure 4.

The new road design features a nominal speed V_e = 60 km/hr, horizontal radius of R $\geq$ 140 m, and maximum longitudinal gradient 3.80%. The road consists of two lanes (one per direction) with an effective road width of 7.00 m plus 2.00 m of sidewalk. Further, at the roundabout access sections, appropriate widening of the roadway is designed. Along the entire length of the road, guard rails of 0.60 m in height have been laterally designed in both directions.

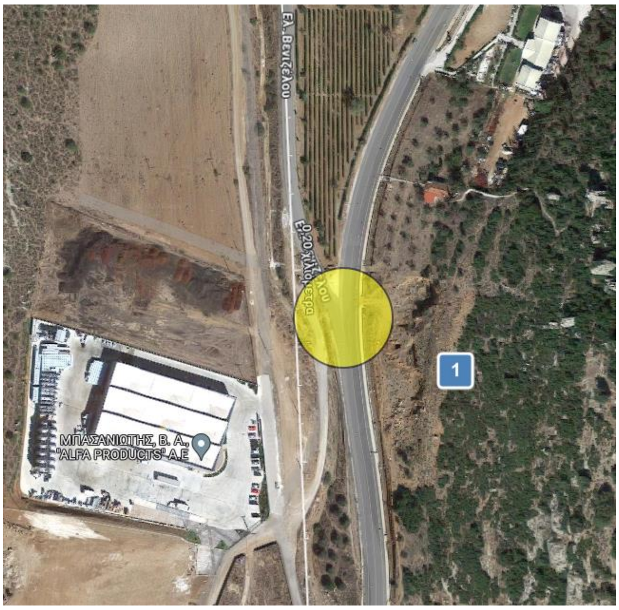


Figure 2: A high accident risk segment of the existing road design (point 1, same as point A in Figure 1)

The full project model, encompassing all the information, is depicted in Figure 5 and allows efficient project analysis and management. As stated before, the utilization of such a model facilitates the effective management of changes at any project component (e.g., drainage system) and the direct communication of such changes at all management levels and stakeholders.

The cross-section of the carriageway comprises multiple layers that include (from the lowermost to the uppermost level): two layers of sub-base of 10 cm thickness each, two layers of base of 10 cm likewise, asphalt pre-coating, asphalt base layer of 5 cm thickness, asphalt adhesive coating, asphalt traffic layer of 5 cm thickness, asphalt adhesive coating, and asphalt anti-slip layer of 4 cm thickness. All layers have been simulated in the 3D/BIM model (Figure 6).

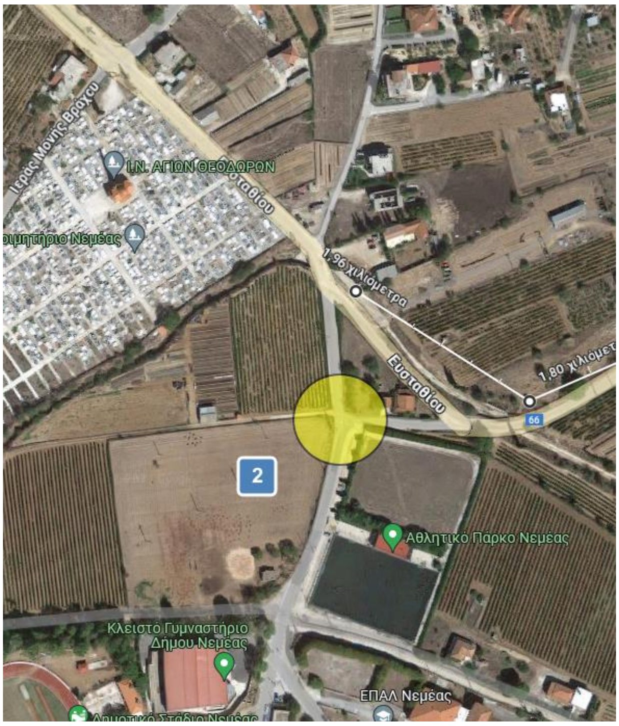


Figure 3: Endpoint of the ring road (point 2, same as point B in Figure 1)

Figure 4: Ring road alignment design

The design of the roundabouts has been carried out in accordance with the "Geometric Design of Highways and Streets" (Hancock et al. 2013) and "Roundabouts: An Informational Guide" (Robinson et al. 2000). The proposed roundabout layouts allow for convenient and safe pedestrian and vehicle access on the ring road. Figures 7 and 8 display the configuration of roundabouts 1 and 2 respectively. To reinstate the connection with local roads, two level junctions (N1 and N2, in Figure 5) have been designed. The junctions, as they have been designed and simulated in the 3D/BIM model, are illustrated in Figures 9 and 10 respectively.

Horizontal and vertical signage has been designed and installed along the entire length of the road to guide the drivers and effectively regulate the traffic. An indicative case is shown in Figure 11 for roundabout 2.

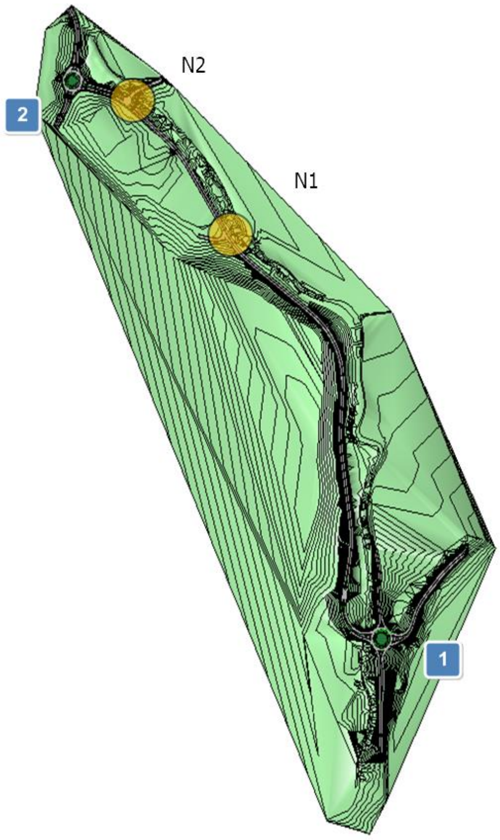


Figure 5: 3D/BIM model of the ring road

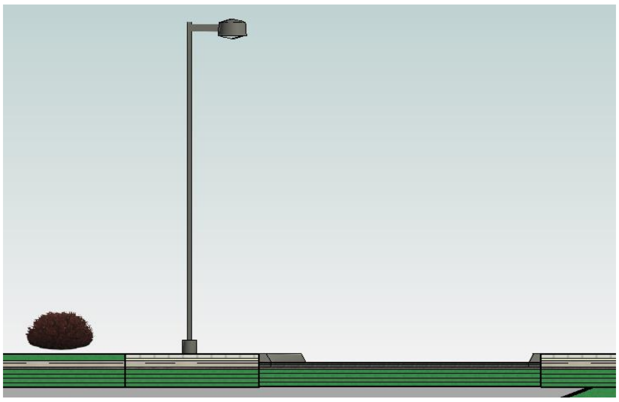


Figure 6: Cross section of pavement layers

Figure 7: Visualization of roundabout 1

Figure 8: Visualization of roundabout 2

Figure 9: Visualization of road junction N1

Figure 10: Visualization of road junction N2

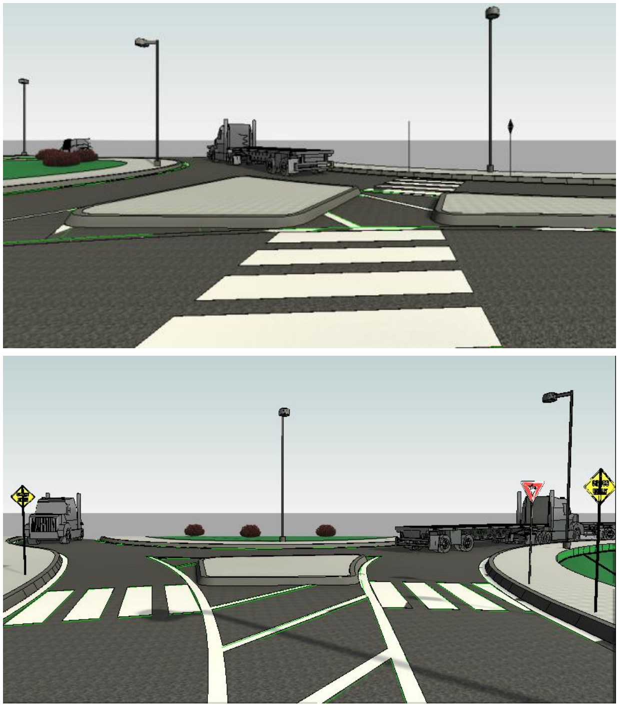


Figure 11: Horizontal marking and vertical signs at roundabout 2

Further, road lighting design has been performed through the 3D/BIM model, which pointed that a total of 51 luminaires of 150-Watt power and 16,100 lumen luminous flux are required to be installed. Figure 12 indicatively presents the electric lighting scheme at the roundabout 2 zone.

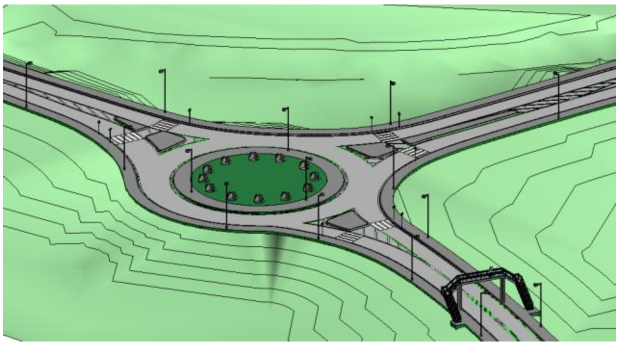


Figure 12: Electric lighting system in roundabout 2.

Analyses and cost estimation of the project have been carried out through the I-BIM model. The estimates can be either for single elements or groups of elements of the project. As an example, the costs for signs and for railings are presented in Tables 1 and 2 respectively, as reported by the software. Similar tables have been developed for all project components. Considering typical unit cost values for Greece, a total cost of 2.32 million € has been estimated for the whole project construction.

Discussion and Conclusions

This study describes the potential and advantages of BIM technology in the design of large infrastructure projects, in this case the ring road of Nemea city in Greece. Opposite to the common belief that BIM technology is mainly applicable to buildings, it can be proven to be equally applicable to infrastructure projects. In this work, I-BIM simulation is particularly effective as it combines parametric 3D simulation and high-level information codification to create a virtual digital model of the project, the 3D/BIM multi-level model, the use of which highly improves the processing, design, planning and management of the project construction and operation.

The importance of using BIM technologies is undoubtedly related to participant coordination and information sharing, the identification and resolving of conflicts at early stage or the evaluation of alternative design solutions. These aspects of BIM tools are rather well known and not widely discussed in this paper. The main goal of this work is to dig into the details of a road infrastructure project and explore whether BIM software can represent them in an integrated way for the design, construction, operation, maintenance, and management to increase interoperability. In addition, the difficulties and barriers in developing the I-BIM model of a project with "irregular" components and in aligning the different elements with each other and with the longitudinal land terrain are explored.

The work outcomes indicate that the project design through Revit (or any other BIM software) is feasible to a high detail, increasing interoperability and allowing all designers to have access to updated project plans at any time and to work collaboratively, thus significantly reducing effort and errors in design. In addition, the 4D and 5D representation of the project, through the time and cost showcase, provides aforehand information on the expected project progress and cost evolution and enhances decision-making and stakeholder communication during construction, leading to higher quality at lower cost.

Table 1: Cost estimate for signs

A	B
Family and Type	Cost
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Street_Signs_8473:	77.40€
Grand total: 16	1238.40€

Table 2: Cost estimate for railings

A	B	C	D
Family and Type	Length	Cost	Cost Railing
Railing: bares	338	45.00€	15200.06€
Railing: bares	388	45.00€	17463.72€
Railing: bares	81	45.00€	3650.95€
Railing: bares	24	45.00€	1061.31€
Railing: bares	133	45.00€	6006.63€
Railing: bares	12	45.00€	540.15€
Railing: bares	103	45.00€	4629.79€
Railing: bares	380	45.00€	17110.02€
Railing: bares	80	45.00€	3600.00€
Railing: bares	16	45.00€	720.00€
Railing: bares	138	45.00€	6206.97€
Railing: bares	8	45.00€	360.00€
Railing: bares	152	45.00€	6831.49€
Railing: bares	58	45.00€	2626.98€
Railing: bares	67	45.00€	3030.74€
Railing: bares	29	45.00€	1283.57€
Railing: bares	349	45.00€	15718.60€
Railing: bares	28	45.00€	1278.81€
Railing: bares	490	45.00€	22059.06€
Railing: bares	3	45.00€	119.19€
Railing: bares	98	45.00€	4387.92€
Railing: bares	171	45.00€	7674.50€
Railing: bares: 22			141560.44€
Railing: Glass Pane	8	25.00€	194.64€
Railing: Glass Pane	10	25.00€	245.01€
Railing: Glass Pane	8	25.00€	194.64€
Railing: Glass Pane	10	25.00€	245.01€
Railing: Glass Pane	8	25.00€	195.89€
Railing: Glass Pane	8	25.00€	195.89€
Railing: Glass Panel - Bottom Fill: 6			1271.09€
Grand total: 28			142831.53€

Unlike typical building design, in which components are of "regular" (e.g., rectangular) shapes, infrastructure project design presents extensive challenges in terms of irregular component shapes, material and manufacturer specificities, and component alignment needs within an integrated design. For example, pavement layers need to be aligned among them and along the road sections (e.g., joins at main road and entrance/exit sections or road approaches and roundabouts). Conflict resolution needs also to be addressed in reference to the road design adjustment to landscape, the incorporation of drainage system, the pavement cross-section design at curved parts of the road, based on large-truck trajectory analysis, etc. The BIM-software database often needs to be enhanced with specific road design elements and potential vendor information, e.g., horizontal and vertical signage, lighting, vegetation and plantation, etc.

In the present work, the design dives into greater detail, in comparison to previous works, indicating the feasibility potential of such a design. However, some limitations should be expected in practical applications. A main difficulty refers to the exact fitting of the road to the landscape, which is mainly due to the inaccuracies or approximations in the digital representation of this landscape. Another challenge is related to the effective placement of overlapping layers of components that are curved and inclined in the longitudinal and transverse directions (e.g., pavement layers). This can be done by considering smaller sections of uniform shape, which are then aligned at their joins and integrated to the full model.

The I-BIM technology can be successfully used in the design and management of large-scale infrastructure projects. In road projects, in particular, the design can improve traffic flow and safety as well as contribute to drainage, signage, lighting, and aesthetic upgrading. The use of BIM technology in large infrastructure projects reduces the effort required during design and construction, while offering effective tools to manage cost and completion time, to reduce waste, and to improve construction quality performance.

References

Agenda, I. (2016). Shaping the future of construction, a breakthrough in mindset and technology. In World Economic Forum.

Cantisani, G., Panesso, J. D. C., Del Serrone, G., Di Mascio, P., Gentile, G., Loprencipe, G., & Moretti, L. (2022). Re-design of a road node with 7D BIM: Geometrical, environmental and microsimulation approaches to implement a benefit-cost analysis between alternatives. Automation in Construction, 135, 104133.

Chi, S., Hampson, K., & Biggs, H. (2012). Using BIM for smarter and safer scaffolding and formwork construction: a preliminary methodology. In Proceedings of the CIB WO99 International Conference on Modelling and Building Health and Safety 2012 (pp. 64-73). Department of Building, School of Design and Environment, National University of Singapore.

Cigu, E., Agheorghiesei, D. T., Gavriluță, A. F., & Toader, E. (2018). Transport infrastructure development, public performance and long-run economic growth: a case study for the Eu-28 countries. Sustainability, 11(1), 67.

Costin, A., Adibfar, A., Hu, H., & Chen, S. S. (2018). Building Information Modeling (BIM) for transportation infrastructure–Literature review, applications, challenges, and recommendations. Automation in Construction, 94, 257-281.

Flyvbjerg, B. (2011). Over budget, over time, over and over again: Managing major projects. The Oxford Handbook of Project Management, 321-344.

Gould, L. (2010). What is BIM… and should we care?. Construction Research and Innovation, 1(2), 26-31.

Hancock, M. W., & Wright, B. (2013). A policy on geometric design of highways and streets. American Association of State Highway and Transportation Officials: Washington, DC, USA, 3.

Liu, J., Xie, Q., Xia, B., & Bridge, A. J. (2017). Impact of design risk on the performance of design-build projects. Journal of Construction Engineering and Management, 143(6), 04017010.

Love, P. E., Sing, C. P., Wang, X., Irani, Z., & Thwala, D. W. (2014). Overruns in transportation infrastructure projects. Structure and Infrastructure Engineering, 10(2), 141-159.

Robinson, B. W., Rodegerdts, L., Scarborough, W., Kittelson, W., Troutbeck, R., Brilon, W., ... & Mason, J. (2000). Roundabouts: An informational guide (No. FHWA-RD-00-067; Project 2425). United States. Federal Highway Administration.

Tang, F., Ma, T., Zhang, J., Guan, Y., & Chen, L. (2020). Integrating three-dimensional road design and pavement structure analysis based on BIM. Automation in Construction, 113, 103152.

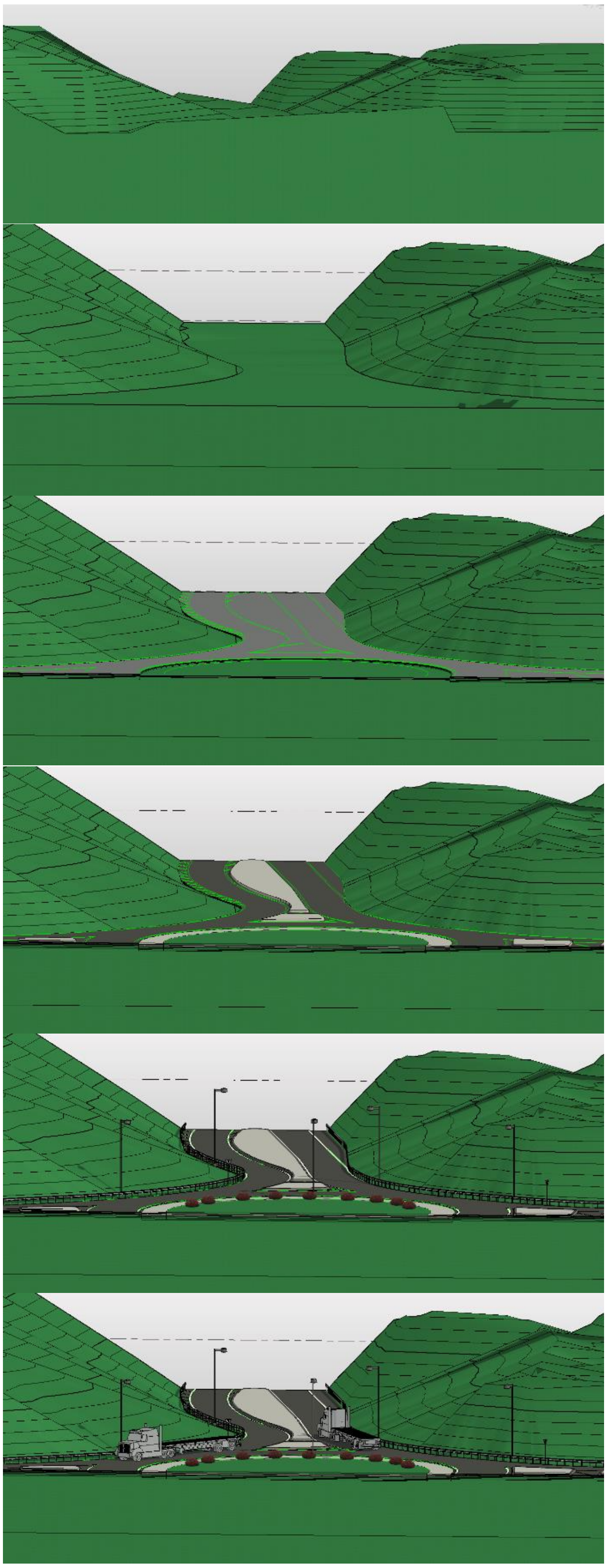
Teizer, J. (2015). Status quo and open challenges in vision-based sensing and tracking of temporary resources on infrastructure construction sites. Advanced Engineering Informatics, 29(2), 225-238.

Timilsina, G. R., Hochman, G., & Song, Z. (2020). Infrastructure, economic growth, and poverty: A review. World Bank Policy Research Working Paper 9258.

Ullah, K., Lill, I., & Witt, E. (2019). An overview of BIM adoption in the construction industry: Benefits and barriers. In 10th Nordic Conference on Construction Economics and Organization, Emerald Publishing Limited, 297-303.

Wong, K. D., & Fan, Q. (2013). Building information modelling (BIM) for sustainable building design. Facilities, 31(3/4), 138-157.

Appendix: Layers in the 3D/VIM model (design phases)



A DIGITAL FRAMEWORK FOR GENERATING AND EVALUATING DIGITAL CIRCULAR TWINS

Anastasia Wieser[1], Sophia Pibal[2], Stefan Schützenhofer[1], Matteo Bosco[3] and Iva Kovacic[1]
[1]Department of Integrated Planning and Industrial Building, TU Wien, Vienna, AUT
[2]AIT Austrian Institute of Technology GmbH, Vienna, AUT
[3]Department of Visual Computing and Human-Centered Technology, TU Wien, Vienna, AUT

Abstract

The Architecture, Engineering, and Construction industry is crucial in promoting circular economy principles to mitigate resource scarcity and negative environmental impacts. Taking the lead in the global extraction of raw materials and making significant waste contributions, the sector incorporates circular economy assessments into building design decisions. This paper introduces a concept using visual programming language for End-of-Life algorithms linked to Building Information Modelling Data. The goal is to create a decision-support tool for the early design stage using common architectural software coupled to visual programming. The results display a computational solution for implementing relevant parameters through utilised software.

Keywords: Circular economy, digital twins, generative design, virtual reality, digital ecosystem, end of life, material building passports, EU taxonomy

Introduction

The growing interest in the circular economy (CE) represents a paradigm shift from the prevailing linear economy, driven not only by the need for assessment for several certifications and EU regulations but also because of its significant advantages compared to the traditional linear model (United Nations Environment Programme, 2022). These benefits extend to environmental, social, and economic concerns, addressing crucial challenges such as resource scarcity and environmental degradation (Kovacic et al., 2020a; Ellen MacArthur Foundation, 2024; Gorgolewski, 2018; Korhonen et al., 2018; Cramer, 2022). Defining the basic concepts of circular economy and sustainability is essential to provide a framework for this research. The Brundtland Report (1987) offers a widely recognised definition of sustainable development as ”development that meets the needs of the present without compromising the ability of future generations to meet their own needs.” (Brundtland, 1987, p.37). Barbier (1987) was among the first to describe sustainability's three pillars: economic, environmental and social, with this paper concentrating on the environmental aspect (Barbier, 1987). Cramer (2022) introduces the 10 R's, which ranks circularity actions by their impact, placing reuse higher than recycling due to its more significant contribution to circularity (Cramer, 2022). Moreover, Geissdoerfer et al. (2017) define Circular Economy as ”a regenerative system in which resource input and waste, emission, and energy leakage are minimised by slowing, closing, and narrowing material and energy loops. This can be achieved through long-lasting design, maintenance, repair, reuse, remanufacturing, refurbishing, and recycling.” (Geissdoerfer et al., 2017, p.759), positioning it as essential for sustainability.

A transition to a circular system is particularly crucial in the construction industry, contributing to around 37% of global energy consumption and process-related CO2 emissions (United Nations Environment Programme, 2022). Furthermore, this industry's significant role in resource depletion, being the largest contributor to global raw material extraction (Almeida et al., 2016), highlights the need to transition to circular systems. However, this transition challenges existing design processes and policy frameworks, making revisiting and altering current approaches necessary (Korhonen et al., 2018). Therefore, there is a need for a framework to better include circular economy principles, such as reuse or reduce and especially guide cooperation between designers, fabricators and clients in the construction industry (Svilans et al., 2019; Çetin, 2023; Geissdoerfer et al., 2017; Çetin et al., 2021).

Digital tools are seen as critical enablers for efficient data sharing, monitoring, optimisation and enhanced communication between stakeholders throughout the design and construction process (Antikainen et al., 2018; Antova and Tanev, 2020; De Wolf et al., 2020; Martínez Rocamora et al., 2021; Çetin, 2023). Moreover, initiatives like Europe's Digital Decade (European Commission, 2021) and the 2020 EU CE Action Plan highlight the importance of digital tools in promoting sustainability and supporting the circular economy (European Commission, 2021; Çetin, 2023).

As a result, finding computational solutions to embrace a circular economy is among the most recent topics in the Architecture, Engineering and Construction (AEC) industry, as it can tackle today's pressing environmental problems (De Wolf et al., 2020; Kovacic et al., 2020a; Cramer, 2022). Building Information Modelling (BIM) and Material Passports, for example, promise to address sustainability barriers and facilitate the automatic inclusion of reliable datasets (Honic et al., 2019; Çetin, 2023; Martínez Rocamora et al., 2021; Santos et al., 2019). However, there are challenges in integrating these tools, particularly when considering one of the most commonly used methods, Life Cycle Assessment (LCA) (Santos et al., 2019). The reliability of LCA results is affected by uncertainties coming from assumptions, such as energy consumption and product lifespan. Additionally, challenges in integrating specific, local information from manufacturers and the lack of detailed data on material composition in the early design stages further affect the accuracy of LCA outcomes (Martínez Rocamora et al., 2021; Santos et al., 2019; Honic et al., 2019). This highlights the need for more flexible and comprehensive tools within the BIM environment to simplify the integration of environmental analysis in the design process (Martínez Rocamora et al., 2021; Çetin, 2023). Research into BIM-based LCA tools shows that most require a combination of BIM and other software to measure environmental impacts effectively. This process, while beneficial, often requires manual input and thus is time-consuming, contrary to the goal of automating these assessments (Martínez Rocamora et al., 2021; Honic et al., 2019). The need for a multidisciplinary approach comes with its hurdles, as the link between digital tools and circular economy principles requires skills spanning environmental science, architecture, engineering, and programming - raising another limitation: the interoperability issue between different software (Santos et al., 2019). Nevertheless, this collaborative effort is vital for developing tools that are both functional and beneficial (Martínez Rocamora et al., 2021; Ilhan and Yaman, 2016). Despite the identified potential of digital tools to integrate circular economy principles in the AEC industry, there remains a significant gap in creating user-friendly, integrated solutions that effectively bridge the gap between theoretical concepts and their practical application. This highlights the ongoing academic and practical need for research to develop accessible tools that enable sustainable construction practices within the circular economy framework(Ilhan and Yaman, 2016; Çetin, 2023; Martínez Rocamora et al., 2021).

The ongoing research presented in this paper addresses these challenges by aiming to optimise architectural designs in terms of circularity and sustainability in the early design stage, as it promises fundamental advantages, representing a vital phase where materials can enter (or re-enter) a new life cycle, and design concepts are still in their initial adjustable stage, thus more flexible to adapt (Gorgolewski, 2018; Honic et al., 2019; Çetin, 2023; Çetin et al., 2021). This paper emphasises the need to flexibly document, validate and benchmark environmental impacts to ensure current activities support a sustainable future and bridge the gap between abstract principles and practical implications (Çetin, 2023; Geissdoerfer et al., 2017). Therefore, this paper considers key indicators, including the building's overall mass, Global Warming Potential (GWP-total), Acidification Potential (AP), and Primary Energy Non-Renewable, Total (PENRT). These indicators were chosen because they are used in the Austrian klimaaktiv OI3 Eco Index calculation to assess the environmental impact of materials (IBO – Österreichisches Institut für Bauen und Ökologie, 2023), and are also relevant for compliance with the EU Taxonomy classification system (European Commission, 2024).

The concept of this research merges Algorithm Aided Design (AAD), Building Information Modeling (BIM) (Pibal et al., 2022), and Virtual Reality (VR). It supports sustainable architectural planning by automating variant generation, conducting circularity assessments, and using a VR platform interface to facilitate decision-making in the early design stage, aligning with Circular Economy (CE) principles and objectives. The framework of the project is based on:

1. a BIM-based architectural model (in software ArchiCAD)
2. an End-of-Life (EoL) design algorithm to conduct the variant study generating an assessable geometric twin, an identical digital illustration of the variant in the VR platform
3. an EoL- assessment algorithm for attributing and evaluating different building variants
4. VR platform with included Virtual Agent for planning support
5. An enriched spreadsheet-based object catalogue containing relevant data from material databases such as the Austrian database "baubook" (baubook, 2024), essential for conducting defined assessments

This research builds on the previous publication, Wohnen 4.0 (Pibal et al., 2023), Digital Ecosystem to enable Circular Buildings – The Circular Twin Framework Proposal (Schützenhofer et al., 2024) and Digital Platform for Affordable Housing - a Framework Proposal (Kovacic et al., 2020b). The overall context established is visualised in Figure 1, which provides an overview of the proposed framework. This paper focuses on Module 2 and Module 3, shown in Figure 1. These EoL algorithms are implemented in the visual programming environment Grasshopper 3D within Rhinoceros 3D and their connection to the architectural design model in ArchiCAD. It discusses the challenges of incorporating these EoL algorithms to enhance the effectiveness of circular design approaches to incorporate crucial data regarding building materials and their reuse potential, as well as Life Cycle Assessment (LCA) and EU taxonomy compliance (European Commission, 2024).

The paper is structured as follows: First, the methodology section reviews the different digital tools used and outlines the intended workflow between Grasshopper and ArchiCAD. Followed by illustrating in more detail the steps created in the EoL algorithms to generate the needed data flow between the programs. Finally, the paper discusses findings and suggests ways for further improvement in integrating these algorithms effectively.

Methodology - Data Integration for EoL- Algorithms

The method is an interdisciplinary research approach. A specific set of digital tools has been selected to facilitate the research objectives. ArchiCAD, employed as a BIM-model and digital planning tool (IFC data/OpenBIM), was selected not only because it is one of the most used digital tools in the AEC industry but also for its ability for efficient data storage and its inclusion of essential data required

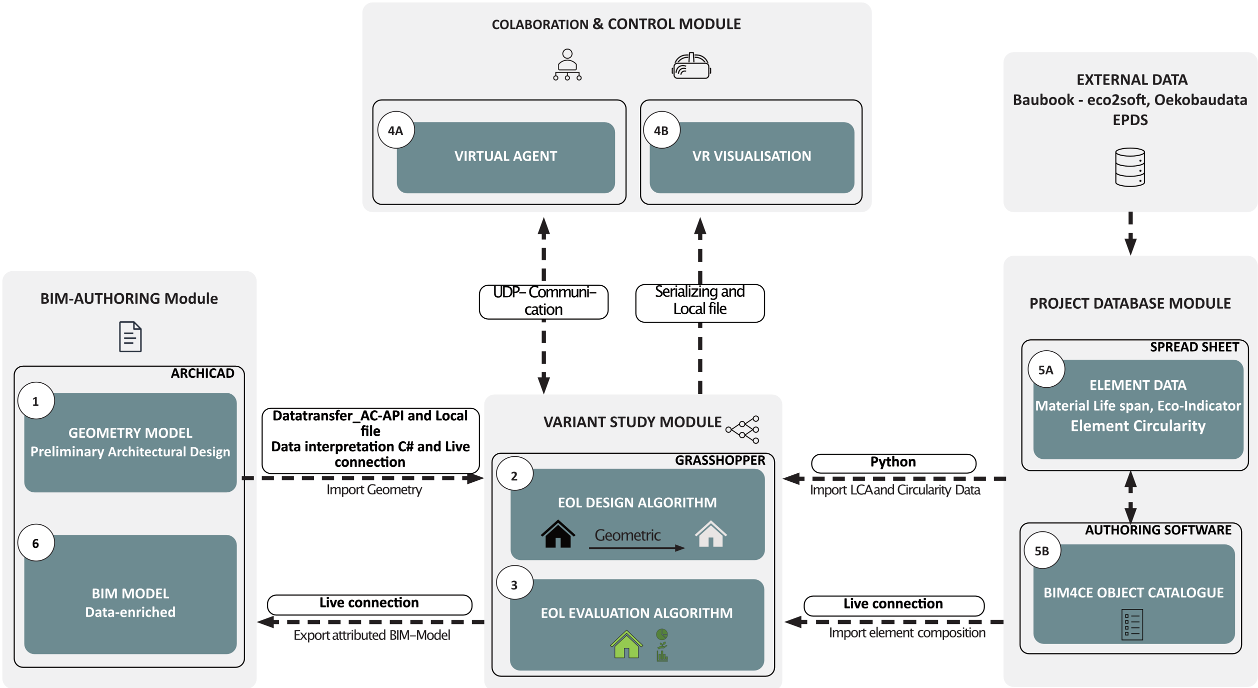


Figure 1: Overview of the proposed framework and software interfaces

for diverse Life Cycle Assessment calculations. Additionally, Grasshopper as a visual programming tool, with an established Live Connection to ArchiCAD (Graphisoft Deutschland, 2024), handling spreadsheet-based documents and including Python/C# script, has been incorporated because of its multifaceted benefits. Furthermore, Grasshopper enables the Virtual Agent in the VR platform to access and modify various components, a crucial requirement for this project. This selection of tools was analysed, leading to the development of an initial prototype of End-of-Life (EoL) algorithms by utilising these diverse interfaces. The conducted literature review led to assessment collection (LCA), which defined the necessary data and basic information and connections required to create the Grasshopper script obtained from the BIM model, are shown in Figure 2 and described below.

Data Input and Model Reconstruction

- Extracted from the architectural design model: Basic geometric data, the quantity of elements essential for subsequent calculations, and spatial reference line accurately generating the geometric twin in Grasshopper for access by the Virtual Agent in the VR platform.
- From ArchiCAD composites (products): The composite name, crucial for efficient product search within the spreadsheet-based object catalogue and the retrieval of accurate data for LCA calculations, and the dimensions of the (new) elements, necessary for precise volume calculations used in mass determination.

Using basic geometric data, the visual geometric twin of the preliminary architectural design is established within the Rhino/Grasshopper environment. This is achieved through the Grasshopper – ArchiCAD Live Connection (Graphisoft Deutschland, 2024) and direct model integration. Recreating the architectural design in Grasshopper is crucial because it cannot perform basic operations with ArchiCAD components. As a result, the geometric information must be filtered, and the model reconstructed accordingly in the Grasshopper environment.

Algorithm Development

a) EoL- design algorithm
After successfully incorporating the geometric form of the case study and the subsequent visualisation of the 'circular twin', the process continues with the extraction of design parameters derived from the data imported from the case study and the BIM-based object database created in ArchiCAD. This process uses several components from the ArchiCAD Plug-in within the Grasshopper environment while maintaining a live connection to ArchiCAD. This allows users to select different materials and compositions from the BIM object database to visualise and generate different variants of the initial case study.
b) EoL- assessment algorithm
In the following part, connecting the user-selected data from the EoL- design algorithm and assessment information, such as the Key Performance Indicator GWP(total) for an LCA calculation through an EoL- assessment algorithm, is essential.

Visualisation - Virtual Reality Integration

The visualisation aspect of the conducted research is not the main focus of this paper. However, it is essential to emphasise that during the development of the algorithm, we needed to consistently consider that the Agent within the VR platform, who interacts with the user, must be able to access and control changeable parameters in Grasshopper.

Framework

To better understand the implementation within the visual programming environment Grasshopper, Figure 3 displays a simplified version of the Grasshopper script and its structure, focusing on wall elements within a test building. Achieving a seamless intersection with the BIM model data required the utilisation of a custom Plug-in within Grasshopper, tailored to receive and process data from ArchiCAD dynamically via the bi-directional Grasshopper-ArchiCAD Live connection was necessary (Graphisoft, 2023).

The framework described will be further validated through its application to a specific use case, showcasing its practical utility and effectiveness in a real-world scenario. This use case involves an architectural design project where the integration of circular economy principles is paramount. By implementing the digital framework within this context, the seamless interaction between Grasshopper and ArchiCAD facilitates a comprehensive analysis and reconfiguration of wall elements, demonstrating the framework's capacity to adapt and optimise based on circular design criteria. This application not only highlighted the robustness of the custom Plug-in and bi-directional live connection but also underscored the potential of the developed EoL algorithms to significantly enhance the sustainability of architectural projects through informed decision-making and design optimisation. The following section describes parameter mapping, variant and assessment generation of our status quo.

Concept of Parameter Mapping

In order to integrate ArchiCAD element types, such as walls, facades, or columns, from the BIM model, it is necessary to manually select each architectural design element per type for every floor within ArchiCAD, done by right-clicking on the parameters used for including a collection of ArchiCAD element types. This method enables the user to subsequently change each element type on each floor separately. However, this initial step cannot be achieved using the Agent in the VR platform, as it cannot access components through right-clicks.

After successfully incorporating all elements, the subsequent phase involves extracting crucial information. The custom components within the ”Deconstruct” group of the Live Connection Plug-in were required to achieve this. Each element type parameter was linked to its corresponding ”Deconstruct” component to obtain further details, including Brep (Polysurfaces) and reference lines. The obtained information is used to generate new variants at the exact location and assemble the visual digital geometric twin of the architectural design by merging all ”Brep” outputs, which can be accessed by the Agent and displayed in the VR platform.

However, it turned out that retrieving the net volume, crucial for assessment methods, was not achievable using this method. Thus, a custom property set (CPset) had to be included within ArchiCAD to automatically retrieve the net volume for every element via the ”Get Property Settings” component in Grasshopper. This workaround enabled the inclusion of the missing geometrical data in the script to calculate the overall mass of the building design.

Notably, during the development of the Grasshopper script, to ensure the accurate functionality of new connections, minimise dataflow time, and maintain proper data structuring. Initially, the focus was on a single ArchiCAD element type (wall) and one floor. Later, the other ArchiCAD element types were integrated for each floor.

Concept of Variant Generation

The tailored Grasshopper component ”Composite” represents products from the current ArchiCAD project and is used to include a selection of products from the object catalogue in the algorithm, for which information is available for the evaluation calculation. In order to make it accessible to the Virtual Agent, it was essential to link components that the Agent could modify to the customisable information in the script. In this case, the Virtual Agent could use the ”Number Slider” to run through predefined products to select from a list, facilitated by the Grasshopper ”List Item” component.

Once all the required information had been imported and retrieved, it was used to generate new ArchiCAD type elements using the 'Design' group components to create new ArchiCAD elements, such as Walls or Columns along the acquired reference data, creating new design variations. Figure 3 shows a simplified representation of a variant in the Grasshopper environment.

Concept of Assessment Generation

The final segment of the Grasshopper script involved incorporating crucial data to evaluate the circularity potential and environmental impacts of the user-generated variants, such as environmental product declarations (EPDs) and material densities, which could not be retrieved through BIM (Honic and Wolf, 2023). The spreadsheet-based object catalogue based on material databases was incorporated through a Python script developed for the EoL- assessment algorithm. This script imports the density required for calculating the overall mass of the building, along with the net volume of the products obtained through the customised CPset in ArchiCAD. Additionally, it provides the necessary data for evaluating the designs in terms of the circular economy. The Python script contains the following inputs:

- The names of the new products of all the elements modified or not modified as a list.
- The location and name of the external spreadsheet-based database

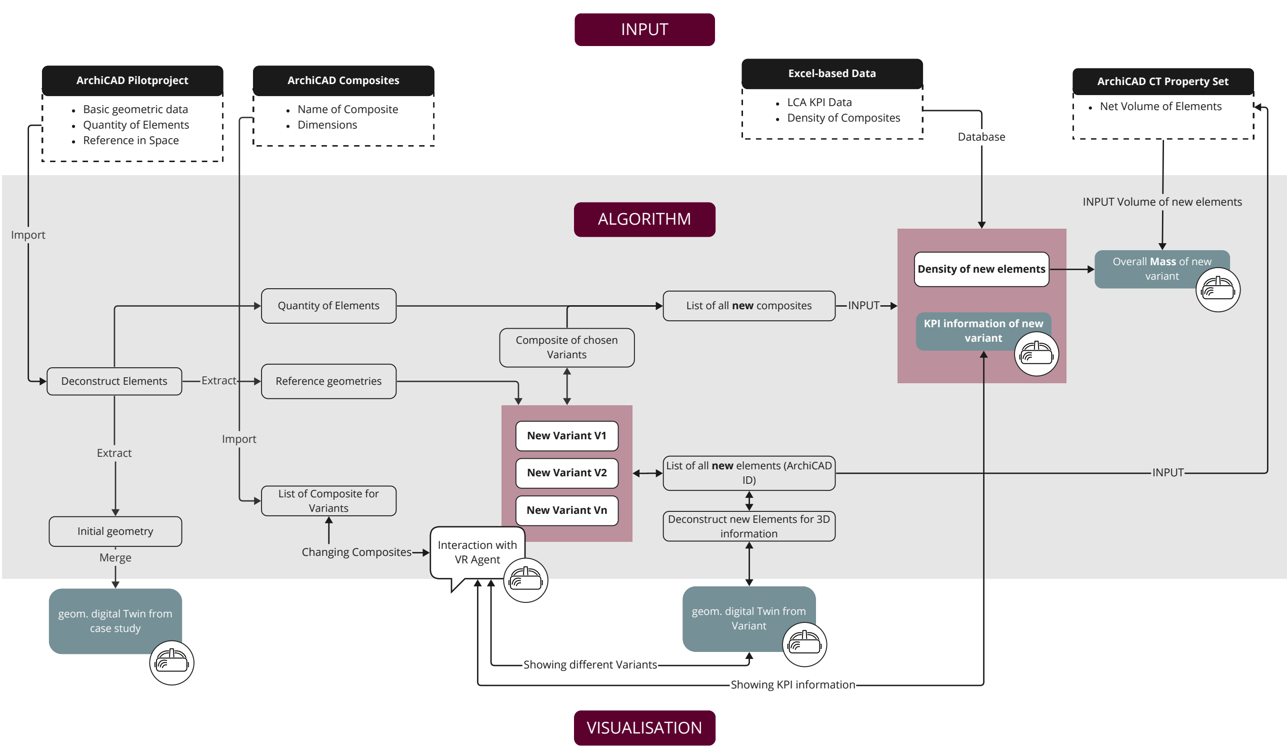


Figure 2: The process and data requirements to develop the EoL algorithms

- The Key Performance Indicator (KPI) the script should search for, e.g., Total Global Warming Potential (GWP - total) and the Material Density

Given the constant requirement for material density, the Python script has been duplicated. This allows the Virtual Agent to use a ”Number Slider” to modify the search input for the calculations in the first Python script while providing a stable estimate of the total mass in the second. However, the outputs of both scripts consist of lists outlining the requisite values for each product. Consequently, these lists must first be aggregated using the Grasshopper ”Mass addition” script before being made accessible to the Agent, enabling user visualisation in the VR platform.

Discussion

Throughout the development of the EoL- assessment algorithm, a consistent pattern emerged: the need for practical solutions to overcome various challenges concerning the interaction between the software used. These challenges primarily revolved around accessing and managing diverse data formats and flows.

One notable solution involved integrating spreadsheet-based circular economy assessment data and material density values using a Python script. This approach was necessary because of the inherent limitations of extracting such information directly from the BIM file (Honic and Wolf, 2023) using the Live connection and its ArchiCAD Grasshopper Toolset components or the customised CPset. Nevertheless, while effective, this solution revealed a vulnerability: the dependence on the spreadsheet-based file's location. However, having an external database for updates can be beneficial as more and more products on the market are expected to receive environmental information (baubook, 2024). This allows easy access to the latest data for design decisions and variant generation with the EoL algorithms.

To address the challenges associated with the extended processing time of the script, which arises from heavy data flow, including constant API interaction, visualisation of 3D elements, and execution of Python scripts for accessing and searching external data, the generation of variants becomes excessively time-consuming. To tackle this issue, exploring innovative approaches becomes imperative. One option worth considering is integrating the ” API - Grasshopper Plug-in” from (Wilk et al., 2023), which promises faster integration of the architectural design model and its data. However, as it relies primarily on textual input and output, it will require a hybrid script to ensure 3D visualisation in the VR platform.

Another challenge is the need for manual selection of ArchiCAD elements, as explained in the section ”Grasshopper Customisation - Parameter Mapping”. This requirement places a disadvantage on the user, requiring access to the Grasshopper script and understanding the basic concept and workflow of the EoL algorithms. However, one method currently under development involves a C# script in Grasshopper, connected via an Add-On in ArchiCAD, to collect and export the required data from ArchiCAD, which must be integrated into the developed algorithms structured according to the Grasshopper-script requirements. This approach would make the steps of

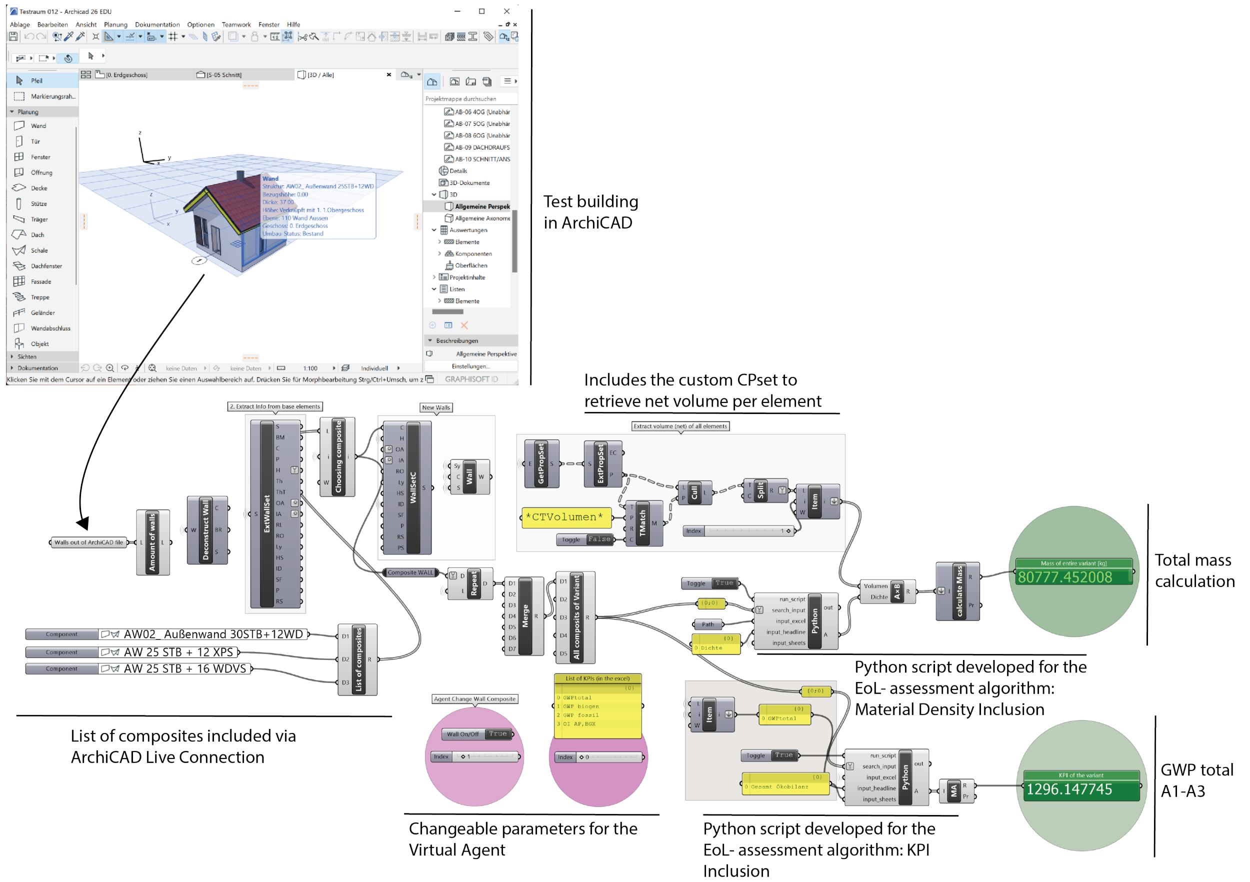


Figure 3: Simplified representation of the Grasshopper script using a test building focusing on four existing walls, showing how the mass [kg] and, in this example, the embodied carbon (GWP-total) for phases A1-A3 (product stage) [kg CO2e] are calculated

picking ArchiCAD elements manually per floor obsolete, enhance the script processing, and mitigate the risk of overlooking elements in ArchiCAD. It may improve user-friendliness, and other initial architectural designs can used.

Moreover, to increase the efficiency and versatility of the presented framework's application, the architecture model must be created following certain rules and structures. This involves the creation of specific ArchiCAD layers to more clearly distinguish between element types, such as internal and external walls. This step is essential due to the different structural functions of these elements and, therefore, different dimensions and product specifications. In addition, an ArchiCAD template should facilitate the seamless integration of the EoL algorithms into other architectural designs. However, it is important to note that doors and windows have been excluded from the EoL- design algorithm for this paper. This decision is due to the fact that the door/window ArchiCAD library part can only be accessed and modified by right-clicking on the ”Window Settings” component in Grasshopper, which is inaccessible to the Agent.

The complexity is additionally increased by successfully storing selected variants to ensure the comparability of generated design decisions. Whenever the Agent modifies the algorithm, the previously selected variant is replaced with the new one. In an initial attempt to address this issue and preserve the variants, a Python script was developed to save generated information through a repetitive loop. However, as the results of the loop were insufficient, either the Python script needs adjustment or an alternative solution must be pursued.

Table 1 outlines each limitation with its respective description and proposed future steps to address challenges that occurred during framework conceptualisation.

Conclusion

This paper outlines the possibility of automated decision support through EoL algorithms via a VR platform for design optimisation in terms of circularity and sustainability. It successfully implemented a case study and its building elements and obtained outcomes related to circular economy. While modifying the elements and generating a variant were successful, numerous technical and design challenges remain. However, assuming the intended changes and workarounds are well implemented, the potential for success is promising. It can provide a basis for further research to combine digital tools to better determine the feasibility of the respective interdisciplinary methods. This way, a computational approach and strategy for continuous environmental impact assessment in the construction industry can be developed. By addressing challenges head-

Table 1: Outline of limitations with descriptions and proposed future steps

Limitation	Description of Limitation	Future Steps
Integration Challenges with Diverse Data Formats and Flows	Managing and accessing various data formats between software tools presents significant challenges	Develop and implement middleware or APIs that can automatically translate and integrate data across platforms
Dependence on External spreadsheet-based File Locations	Utilising spreadsheet-based data for assessment data introduces a dependency on the file's location	Transition to cloud-based databases for dynamic access and updates, reducing reliance on static spreadsheet-based files
Extended Processing Time Due to Heavy Data Flow	The extensive data flow leads to time-consuming variant generation processes	Optimise algorithms for efficiency and explore parallel processing to manage and process data more quickly
Manual Selection of ArchiCAD Elements	Manual selection of elements in ArchiCAD for every floor adds complexity and time to the design process	Automate the selection process with AI or develop intuitive UIs that simplify and expedite the selection process
Structural Requirements for Architectural Models	Specific rules and structures must be followed for effective integration of EoL algorithms	Create templates and guidelines that standardise model structures, facilitating smoother integration
Exclusion of Doors and Windows from the EoL-Design Algorithm	Doors and windows are excluded due to limitations in accessing and modifying ArchiCAD library parts	Develop enhancements or Plug-ins for the VR platform to allow direct manipulation of doors and windows
Difficulty in Preserving Selected Variants for Comparison	Storing selected variants becomes complex when the Agent modifies the algorithm, leading to loss of previously selected variants	Implement a versioning system within the software that automatically archives and tracks changes to design variants

on and continuously iterating on the process, the EoL algorithms move closer to facilitating decision-making in the early design process, enabling circular building structures and integrating End-of-Life concepts.

Acknowledgment

The research described in this paper was carried out as part of the project ”Circular Twin - Ein digitales Ökosystem zur Generierung und Bewertung kreislauffähiger Digitaler Zwillinge”, funded by the Austrian Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology through the Austrian funding agency (FFG), grant number 899167, which facilitated the conduct of this research.

References

Almeida, P., Solas, M., Renz, A., Bühler, M., Gerbert, P., Castagnino, S., and Rothballer, C. (2016). Shaping the future of construction: A breakthrough in mindset and technology.

Antikainen, M., Uusitalo, T., and Kivikytö-Reponen, P. (2018). Digitalisation as an enabler of circular economy. Procedia CIRP, 73:45 – 49. 10th CIRP Conference on Industrial Product-Service Systems, IPS2 2018, 29-31 May 2018, Linköping, Sweden.

Antova, G. and Tanev, V. (2020). Creation of 3d geometry in scan-to-cad/bim environment. IOP Conference Series: Earth and Environmental Science, 609.

Barbier, E. B. (1987). The concept of sustainable economic development. Environmental Conservation, 14(2):101–110.

baubook (2024). Web-portal für bauprodukte, bauteile und tools, das ökologisches und gesundes bauen vereinfacht. https://www.baubook.info/de [Accessed: 2024-02-11].

Brundtland, G. H. (1987). Report of the world commission on environment and development: Our common future. Technical report, the World Commission on Environment and Development (WCED).

Cramer, J. (2022). Building a Circular Future - Ten Takeaways for Global Changemakers. Amsterdam Economic Board.

De Wolf, C., Hoxha, E., and Fivet, C. (2020). Comparison of environmental assessment methods when reusing building components: A case study. Sustainable Cities and Society, 61:102322.

Ellen MacArthur Foundation (2024). What is the circular economy? https://www.ellenmacarthurfoundation.org/circular-economy/what-is-the-circular-economy[Accessed: 2024-02-11].

European Commission (2021). Europe's digital decade: Commission sets the course towards a digitally empowered europe by 2030. https://ec.europa.eu/commission/presscorner/detail/en/IP_21_983 [Accessed: 2024-04-02].

European Commission (2024). Eu taxonomy for sustainable activities. https://finance.ec.europa.eu/sustainable-finance/tools-and-standards/eu-taxonomy-sustainable-activities_en#related-links [Accessed: 2024-02-11].

Geissdoerfer, M., Savaget, P., Bocken, N., and Hultink, E. (2017). The circular economy – a new sustainability paradigm? Journal of Cleaner Production, 143:757–768.

Gorgolewski, M. (2018). Resource Salvation: The Architecture of Reuse. John Wiley & Sons, Ltd.

Graphisoft (2023). Grasshopper-archicad live connection user guide for archicad 27. https://help.graphisoft.com/AC/27/INT/GC.pdf[Accessed: 2024-02-11].

Graphisoft Deutschland (2024). Rhino – grasshopper – archicad toolset. https://graphisoft.com/at/downloads/addons/interoperability/rhino[Accessed: 2024-02-11].

Honic, M., Kovacic, I., and Rechberger, H. (2019). Improving the recycling potential of buildings through material passports (mp): An austrian case study. Journal of Cleaner Production, 217:787 – 797.

Honic, M. and Wolf, C. D. (2023). Applying eu level(s) framework indicators to improve circularity: A case study. IOP Conference Series: Earth and Environmental Science, 1196(1):012041.

IBO – Österreichisches Institut für Bauen und Ökologie (2023). Oekoindex oi3. https://www.ibo.at/materialoekologie/lebenszyklusanalysen/oekoindex-oi3[Accessed: 2024-04-02].

Ilhan, B. and Yaman, H. (2016). Green building assessment tool (gbat) for integrated bim-based design decisions. Automation in Construction, 70:26–37.

Korhonen, J., Honkasalo, A., and Seppälä, J. (2018). Circular economy: The concept and its limitations. Ecological Economics, 143:37 – 46.

Kovacic, I., Honic, M., and Sreckovic, M. (2020a). Digital platform for circular economy in aec industry. Engineering Project Organization Journal, 9.

Kovacic, I., Pibal, S., Reisinger, J., and Lorbek, M. (2020b). Digital platform for affordable housing - a framework proposal. In Working Paper Proceedings (EPOC 2020), page 1–16).

Martínez Rocamora, A., Rivera-Gómez, C., Galán-Marín, C., and Marrero, M. (2021). Environmental benchmarking of building typologies through bim-based combinatorial case studies. Automation in Construction, 132:103980.

Pibal, S., Kovacic, I., Lorbek, M., Jakoubek, R., Reisinger, J., Temel, R., Ilcik, M., Wimmer, M., Kerbl, B., Travas, D., Bajric, A., Hagmann, E., and Hödl, C. (2023). Wohnen 4.0 - digitale plattform für leistbares wohnen.

Pibal, S. S., Khoss, K., and Kovacic, I. (2022). Framework of an algorithm-aided bim approach for modular residential building information models. International journal of architectural computing, 20(4):777–800.

Santos, R., Aguiar Costa, A., Silvestre, J., and Pyl, L. (2019). Integration of lca and lcc analysis within a bim-based environment. Automation in Construction, 103:127–149.

Schützenhofer, S., Pibal, S., Wieser, A., Bosco, M., Fellner, M., Petrinas, V., and Kovacic, I. (2024). Digital ecosystem to enable circular buildings – the circular twin framework proposal. J.sustain. dev. energy water environ. syst., 12(2),1120500.

Svilans, T., Tamke, M., Thomsen, M. R., Runberger, J., Strehlke, K., and Antemann, M. (2019). New Workflows for Digital Timber, pages 93–134. Springer International Publishing, Cham.

United Nations Environment Programme (2022). 2022 global status report for buildings and construction: Towards a zero-emission, efficient and resilient buildings and construction sector.

Wilk, G., Beneitez, J., and Enzyme APD (2023). Archicad api - grasshopper plugin. https://www.archicad-api.com [Accessed: 2024-02-11].

Çetin, S. (2023). Towards a circular building industry through digitalisation: Exploring how digital technologies can help narrow, slow, close, and regenerate the loops in social housing practice. A+BE | Architecture and the Built Environment, 13(22):236.

Çetin, S., De Wolf, C., and Bocken, N. (2021). Circular digital built environment: An emerging framework. Sustainability, 13.

ENTERPRISE DIGITAL TWINS FOR STRATEGIC DATA UTILISATION FROM CONSTRUCTION SITES

Abdulmalik Badamasi[1], Xiang Xie[2], and Mohamad Kassem[3]
[1]Newcastle University, Newcastle, UK

Abstract

The progressive digitalisation of construction sites, driven by the deployment of digital technologies, presents organisations with the significant challenge of effectively extracting value from the surge in data production. This paper introduces the concept of '*Enterprise Digital Twin*' (EDT) as a new form of digital enterprise in the construction sector, which extends beyond the scope of individual construction sites and projects to enable a variety of data-driven operational and strategic insights for construction organisation. Despite its importance, there exists a notable scarcity of research in this area. This paper introduces the EDT concept, presents a review of the available literature, proposes an early conceptualisation, and discusses its challenges and opportunities. These are early contributions to this nascent area and lay the foundation for future research.

Introduction

The construction industry is increasingly incorporating different digital tools and methods, such as Building Information Modelling (BIM), Wireless Sensor Networks (WSN), and Internet of Things (IoT) related technologies across the project lifecycle, including the construction phase (Statsenko et al., 2023). As a result, a large amount of operational data is produced in the construction industry and managed across geographically dispersed construction sites, supporting various functions such as procurement, quality assurance, supply chain, and logistics (Chen et al., 2024). Valuable insights can be untapped at both the project and enterprise levels through the exploitation and analysis of such extensive and diverse repositories of data, encompassing both structured and unstructured (You & Wu, 2019).

One of the pivotal technologies conceived for capturing and analysing data to deliver insights is Digital Twin. A Digital Twin (DT) is a digital replica of a physical entity, process, or system that enables a bidirectional information exchange between the digital and physical counterparts (Botín-Sanabria et al., 2022). Parvin (2021) stated that the information standard from enterprise to project is crucially important for digital twins to work in the construction industry. Research indicates a positive relationship between the degree of digitalisation of an enterprise through the use of DT and enterprise performance (Li & Liu, 2023). When augmented with other technologies, such as Internet of Things (IoT) and Artificial Intelligence (AI) , DT has the potential to provide real-time accurate status information and optimise ongoing design, plan, and production (Yitmen et al., 2023). Data integration is crucial because it allows businesses to utilise important operational data, which can yield new strategic and operational insights across diverse domains and functional areas. Yet, investigating how to effectively utilise this data from construction projects at the enterprise level is an emerging paradigm for both research and industry practice.

This paper explores the concept of EDTs within the construction sector. It highlights the necessity for EDT deployment and introduces a low-granularity conceptual architecture for EDTs. The methodology used to develop the EDT conceptual architecture is based on an extensive review of related concepts in literature and the guidelines established in existing reviews, such as this one. Furthermore, the paper examines the challenges associated with implementing EDTs, alongside the potential opportunities arising from their adoption.

Literature Review

Construction Digital Twin

Large contracting companies involved in major infrastructure projects usually manage construction projects with multiple geographically dispersed construction sites. Managing these construction sites at once within the same project can be challenging in terms of data generation, storing, processing, and visualisation (Hasan & Sacks, 2023). In the contemporary context, it is problematic that contractor organisations persist in relying on traditional methods to manage substantial volumes of data, owing to the absence of platforms capable of integrating data from diverse sources and types, making it very challenging to collect, manage, and exploit the generated data (Shamshiri et al., 2024). This is further challenged by the growing increase in size and complexity of data generated (Yan et al., 2020).

The collection of vast quantities of data also bears sustainability implications for the organisations involved. Storing large amounts of data in the cloud is not only expensive but also impacts the planet`s sustainability due to the emissions produced from data storage facilities (Vlăduțescu & Stănescu, 2023). The inadequate quality of visualisations and continued dependence on paper-based documents prevents managers and decision-makers from gaining clear insights and applying knowledge (Pan & Zhang, 2023).

Incorporating data into strategic management and decision-making processes can provide companies with numerous benefits (Ragazou et al. 2023). Although a large amount of data is available, many organisations fail to utilise it to their benefit (Ranjan & Foropon, 2021).

Digital Twin has emerged as a new technological concept that can address the management of smart construction site data for many reasons. DT has been adopted thus far only for a limited number of construction applications, and its concept has not been extended to the enterprise level. For instance, Sacks et al. (2020) proposed a comprehensive digital twin construction (DTC) system that integrates data from different activities to achieve closed-loop control systems. The DTC system extends beyond traditional activity-based monitoring and incorporates information from various sources. However, it is only activity-based and not operation-based. DT enables the integration of data, as highlighted by Salem and Dragomir (2022), indicating its role in integrating information technology and IoT for automating processes and managing activities related to building and urban structure operations. However, at the enterprise level, the exploration of this concept remains insufficiently addressed, with merely a handful of studies exploring its incorporation into Business Intelligence (BI) systems. For instance, Lopes and Boscarioli (2021) studied the implementation of BI and Analytics tools in the construction industry to improve management and decision-making. Rodrigues et al. (2022) studied the integration of BI with BIM in the construction industry, providing real-time data analysis for sustainable construction management and decision making. Thus, to address this important gap, in this paper we propose the introduction and early conceptualisation of the concept of 'Enterprise Digital Twin'.

The Need for Enterprise Digital Twin

The importance of DT at the enterprise within the construction industry presents unique opportunities and poses critical challenges. While digital transformation presents the potential to attain excellence, it is crucial to establish connections between digital technologies, information, and strategic decision-making in order to tackle construction challenges (Balzano & Marzi, 2023). Previous research has put forward the importance of establishing a sustainable innovation ecosystem for industry and enterprise knowledge management and practices by interconnecting technological innovation, the business model, and the market (Yan, 2015; Yun et al., 2017; Yang & Yan, 2019). However, the success of this proposition heavily relies on the ability of collecting the necessary data to generate insights. An enterprise level digital twin is conceived as a socio-technical paradigm with the ability of capturing data from diverse resources. Mêda et al. (2021) focused on the concept of an incremental digital twin construction (DTC) system, which aims to integrate data from different activities to achieve closed-loop control systems. Merino et al. (2023) proposed a method for integrating data from construction activities, Building Automation Systems (BAS), and IoT using federated data models and ontologies. However, these studies concentrated solely on activity-based aspects, excluding considerations of operational foresight and business intelligence. This study will discuss the Enterprise Digital Twin concept and its role in leveraging data from different construction sites and organisation operations in the generation of operational foresight and business intelligence.

Enterprise Digital Twin (EDT)

In the context of manufacturing, an Enterprise Digital Twin is defined as the utilisation of business-wide information to enable strategic decisions utilising asset and process digital twins across an organisation (Yan et al., 2022). Enterprise Digital Twins enable manufacturing managers to replicate all activities and interactions in the production chain, regardless of the number and location of assets, sites, suppliers, contractors, and sub-contractors involved (Kubelskiy, 2021). While some studies on EDT are available within the manufacturing sector, there exists a clear gap within the construction sector. This section introduces the concept of EDT in construction. EDT in construction is proposed as a virtual replica of an enterprise that incorporates people, processes, data, and technology to produce business outcomes. It is created by integrating technologies such as digital twins, cloud computing, analytics, artificial intelligence, and simulation.

An enterprise comprises multiple organisational units that collaborate to offer services that cannot be provided by a single organisational unit on its own. An Enterprise Architecture (EA) framework plays an essential role in defining principles and practices for creating and employing the information systems and technology infrastructure (Haki & Legner, 2021). EA frameworks form the foundation of the EA approach and serve the purpose of making the intricacies of the real world comprehensible and manageable for stakeholders (Scheer, 2023). Presently, various EA frameworks exist, including the Zachman framework, which was first introduced in 1987 as an early framework for EA (Zachman, 2003). Another prevalent EA framework is The Open Group Architecture Framework (TOGAF), which was initially developed as a methodology for the deployment of technical architectures but has since shifted its focus to EA over the years (Harrison, 2011). According to Riege and Aier (2009), Enterprise Architecture (EA) is widely recognised as an effective method for managing transformations and ensuring business/IT alignment. Therefore, an EA framework is employed in this research for initiating the early conceptualisation of EDT (Figure 1). This approach is guided by Zachman (2003) emphasis on addressing the what, how, when, who, where, and why aspects of information systems, mirroring the layers within the EDT conceptualisation that address the sourcing, integration, storage, analysis, and utilisation of data for decision-making in construction enterprises. Additionally, Harrison (2011) proposed that the Architecture Development Method (ADM) of The Open Group Architecture Framework (TOGAF) provides direction for developing architectures by following a series of stages, such as business architecture, information systems architectures, technology architecture, and others. A similar approach is mirrored in the conceptualisation of the Enterprise Digital Twin (EDT), where repetitive improvements are made at each layer to address the operational requirements of construction enterprises.

Data Source

Nowadays, data generation has become a prominent aspect across all types of construction activities, resulting in an increase in the volume of data that needs to be integrated (Alaloul et al., 2021). The data layer, which harvests information from IoT sensors and various other sources including models, programs, among others, in addition to organizational processes like procurement, estimation, and logistics, gathers extensive volumes of real-time, near-real-time, and static data. Such data is available from various sources on construction sites, such as IoT sensors, drones, BIM data, and project management software (Merino et al., 2023) and can be used to generate important foresights about the effectiveness and efficiency of construction operations (Wang et al., 2023). This data can include real-time performance metrics of machinery, worker productivity, material usage, environmental conditions, and more. For an enterprise digital twin, collecting high-quality, granular data is crucial as it forms the foundation of both the operational and business intelligence insights.

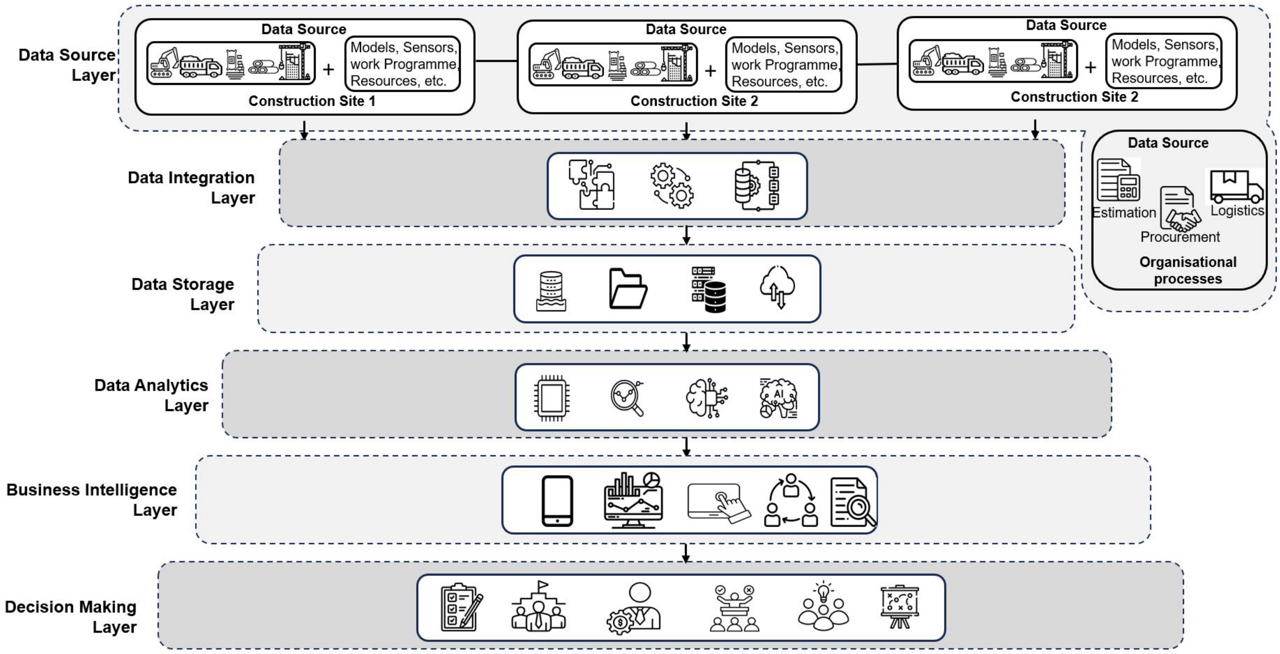


Figure 1: Conceptualisation of Enterprise Digital Twin

Data Integration

The significance of data integration is highlighted by Zhang et al., (2022) in the context of developing a digital twin that accurately mirrors the physical assets of construction sites. Within the context of EDT, the primary function of the data integration layer is to enable the seamless sharing and consolidation of data across various sources, ensuring that data and information is accessible, accurate, and consistent across the enterprise. Data integration layer has grown to encompass many different technologies and capabilities beyond Extract Transform Load (ETL) and ELT, which is just one-use case of data integration (Luengo et al., 2020), such as data virtualisation, data federation, cloud-based data integration services (e.g. Azure data factory, AWS Glue, etc.). These technologies are designed to gather, integrate, and transforms data from multiple sources into consistent, conformed, comprehensive, clean, and current information (Sherman, 2014). The data managed in the integration layer is critical to understand the status and performance of various enterprise components.

Data Storage

The data storage layer plays a crucial role in an enterprise architecture. It is responsible for persisting, managing, and utilising data within the organisation (Jones et al., 2022). Data warehousing is a data storage process that involves storing and preparing information separately from an organisation's routine transaction processing activities. It is designed to optimise the data for efficient access and analysis within the enterprise (Bharadiya, 2023). The layer in question serves as a repository for integrated data, including both structured data from data warehouses and unstructured or semi-structured data from data lakes. In this process, data from construction sites is transmitted to the data warehouse, where it undergoes data transformations such as cleansing, filtering, and aggregations that prepare the necessary to suit the enterprise required view of the data (Sherman, 2014).

Digital twin data, which are characterised by being sourced from multiple sources and having a large volume, necessitate the utilisation of big data storage technologies (Wang et al., 2023). The selection of a storage database for managing large amounts of data is contingent upon its ability to provide accessibility, scalability, high performance, and effective management. data storage solutions (e.g., cloud storage, database management system (DBMS), data warehouse and data lakes) are diverse, catering to different needs in terms of scale, performance, security, and accessibility. The selection of a data storage option is critical for supporting the varied applications and services an organisation relies on. Most studies have utilised cloud-based computing platforms for storing data which offer flexible and excellent backend access for computing applications (Betti et al., 2024). The scalability and flexibility provided by cloud computing enable the storage, processing, and accessibility of substantial amounts of data produced by the digital twin (Knebel et al., 2023). The selection of a data storage option is critical for supporting the varied applications and services an organisation relies on.

Data Analytics

The data analytics layer encompasses a suite of technologies, including Artificial Intelligence (AI) with Machine Learning (ML) as a subset, which are employed to conduct comprehensive analyses on accumulated data (Raschka et al., 2020). These technologies have become indispensable for organisations of all sizes and across all industries, as analytics and data-driven decision making have proven to be of paramount importance, enhanced decision-making is achieved through analytics and machine learning by offering practical insights derived from digital twin data (Arsiwala et al., 2023). Digital twin data can be utilised for multiple user services through advanced data analytics technologies. The utilisation of AI and ML technologies is important for the processing and analysis of data in the within the construction sector (Kazeem et al., 2023). AI and ML algorithms are indispensable for managing and analysing the substantial volumes of data produced by deep learning technologies (Kaur et al., 2020). These computational tools excel in discerning patterns and trends within datasets, thus enabling the extraction of pivotal insights (Bharadiya, 2023). DTs use AI/ML for predictive modelling, optimisation, and real-time monitoring, offering organisations valuable information for data-driven decisions and enhanced performance (Omrany et al., 2023). The analytics layer in an EDT utilises data analytics, AI, and ML algorithms to transform raw data into actionable intelligence to supports decision-making processes and identifies patterns and predictions. The processed data, which are supported by visualisation technologies (Wu et al., 2021), are ultimately accessible to end users in a straightforward and interactive manner through data visualisation.

Business Intelligence

The Business Intelligence (BI) layer aims to visualise information using tools such as dashboards and reports, transforming complex datasets into accessible formats that facilitate understanding and decision-making. Effective visualisation is a crucial component within the construction industry, as it plays a key role in facilitating communication and decision-making among team members (Lucchi, 2023). Lopes and Boscarioli (2021) highlights the role of BI in facilitating data-driven decisions, where stakeholders can quickly understand key performance indicators (KPI) and trends, enabling timely and informed decisions. One of the strongest aspects of digital twins is the ability to display sensor data in a virtual environment (Petri et al., 2023). Two commonly employed techniques for visualising digital twin data are colour coding in 2D and 3D schematics, performance dashboards, and time-series graphs, which are facilitated by visualisation platforms (Pal et al., 2023). Boje et al., (2020) posit that the implementation of data visualisation techniques in project management at smart and distributed sites alleviates the challenge of data inundation. According to Salem and Dragomir (2022), incorporating data from various devices and models can provide improved insights for enterprise decision-making through cross-checking and cross-referencing. For instance, visualisation tools in the EDT can be extended to enable an understanding of operations across all sites. This layer enables the evaluation of performance among different sites and the recognition of optimal strategies that can be implemented at the enterprise level.

Decision Making

The decision-making layer, on the other hand, is where strategic, tactical, and operational decisions are made based on the insights provided by the BI layer and other sources. Yue (2023) describes how the decision support ecosystem consumes analytical insights with business goals to guide decision-making. The decision support system (DSS) employs the results from business intelligence (BI) tools, such as dashboards and reports, to guide decisions that align with the company's strategic plan. This involves not only analysing past data but also utilising predictive analytics to forecast potential future scenarios. The system aims to provide a comprehensive view of the enterprise's operations and enable informed decision-making that supports the company's long-term goals (Yue, 2023). The abilities provided by these allow decision-makers to predict potential problems, assess alternative solutions, and take measures. Yan et al. (2022) highlight its impact on strategic planning, where predictive insights can shape the direction of construction projects and broader enterprise initiatives. They argue that predictive insights derived from the digital twin's analytics can significantly influence the planning and execution of construction projects. For instance, by predicting the impact of various factors on project timelines and costs, enterprises can make informed decisions that optimise resource allocation, mitigate risks, and enhance overall project outcomes (Yan et al., 2022). Sacks et al. (2020) state that the insights derived from the lifecycle of the digital twin significantly contribute to operational excellence and competitive advantage. Therefore, the digital twin at the enterprise level can become an indispensable tool for continuous improvement and strategic decision-making using data from across smart construction sites and business functions.

EDT: Challenges and Opportunities

The implementation of EDTs in construction projects involves different technical challenges that need to be addressed to fully leverage their capabilities. Construction projects often generate large amounts of data in different formats due to the application of various systems and technologies (Himeur et al., 2023).

Advanced data management and interoperability solutions are necessary to integrate this data into an interconnected digital twin. The challenge is made more difficult due to the requirement for timely updates of data to keep the digital twin accurate and relevant (Borrmann et al., 2015). The potential to address this challenge lies in the adoption of recognised and accepted formats (such as XML, CSV, IFC, JSON, etc.) and protocols for storing and exchanging data that ensure compatibility and ease of use across different systems and applications employed by organisations in construction projects. However, as organisations pursuing EDTs are likely to be face complexity and scale challenges, standardisation, and integration across the various areas of the business become key (Faith and Tori, 2020). Semantic Web concepts such as the Enterprise Knowledge Graphs (EKG), a formal model to represent and manage enterprise information at a semantic level, can play a key role in integration.

Privacy issues emerge when encountering confidential data regarding project sites, designs, and staff members (Zhang et al., 2022). The nature of the data gathered and utilised by EDTs may encompass sensitive information, giving rise to issues regarding individuals with authorised entry to such data and the way it is used. Defining the roles and responsibilities for data management and ensuring compliance with relevant regulations can assist in mitigating privacy concerns by instituting well-defined policies for data access, usage, and sharing (Coupaye et al., 2023).

Security is an additional issue in relation to EDTs, as these digital systems are susceptible to cyber threats, like any other digital system (Faleiro et al., 2022). Potential interconnectedness between digital twins can present a major security risk, as a breach in one system has the potential to compromise the entire network, making it susceptible to cyberattacks (Faleiro et al., 2022). Protecting EDTs from such threats requires the implementation of advanced cybersecurity measures, including encryption, secure data storage and transmission protocols, and regular security audits (Lu et al., 2017). This includes the use of firewalls, intrusion detection systems, and regular vulnerability assessments by organisations.

Overcoming technical challenges and addressing organisational challenges are both necessary to leverage EDTs in the construction industry. The workforce must possess a high level of digital literacy and technical skills owing to the intricate nature of DTs (Hou et al., 2020). This encompasses proficiency in data analysis, an understanding of digital twin software and platforms, and the ability to comprehend and act upon the insights provided by DT at the enterprise level (Himeur et al., 2023). To address this challenge, it is necessary to implement training programs and recruit individuals with the necessary digital skills. Moreover, fostering a culture that promotes continuous learning and adaptation to stay abreast of technological advancements is important (Howard et al., 2017).

EDTs require organisations to develop and implement customised training programs to enhance the digital proficiency of their existing workforce and acquire new talents to address potential gaps in competencies (Agrawal et al., 2023). Continuous training is essential to maintain relevance in the face of fast-paced technological advancements. Implementing EDTs in construction initiatives frequently requires changing normal procedures and operational patterns (Riss et al., 2020). One approach to improving project management is to incorporate real-time data analysis and decision-making, as well as utilising the digital twin's predictive capabilities to refine procurement, scheduling, and risk management processes (Ragazou et al., 2023). Finally, due to the likely need for process reengineering that have the potential to cause disruptions and face opposition, the transformation require effective leadership and transparent communication regarding the advantages to all parties involved (Müller et al., 2024).

The implementation of EDTs in the construction industry presents many opportunities for innovation, construction practices, and the development of new business models (Kulkarni et al., 2019). New business models can be developed by leveraging the insights and efficiencies obtained from EDTs. Construction firms can enhance their value proposition by providing data-driven services that diversify revenue streams and foster long-term client relationships by offering value beyond project completion. For example, EDTs facilitate the construction of sustainable buildings by improving planning and resource allocation. The optimisation of energy use, selection of sustainable materials, and integration of renewable energy sources into new projects can be facilitated by EDT, thereby aligning with global sustainability goals (Riss et al., 2020).

To take advantage of these opportunities, construction companies need to cultivate an environment that encourages innovation, collaboration, and standardisation, that are key enablers to transformative projects (Hook et al., 2022) such as EDTs. They also require exploring partnerships within their supply chains to enhance the connectedness that is key to EDT as well as allocating the resources to acquire essential technologies and develop and acquire the necessary competencies.

Conclusions

The concept of Enterprise Digital Twins (EDTs) in the construction industry is emergent, driven by the exponential increase in data availability due to the widespread adoption of digital technologies on construction sites and beyond. Recognising the emergent nature of EDT within the construction industry, this paper attempted to provide a balanced perspective about EDTs. The paper presented a low-granularity architecture model with six layers for EDTs for construction organisations. It highlighted the importance of EDT in the transition towards a more interconnected and data-driven method for managing and executing construction projects and organisations. It also outlined some key challenges such as interoperability, privacy, and security, as well as some of the organisational challenges.

Acknowledgments

This work was supported by the Petroleum Technology Development Fund (PTDF).

References

Agrawal, A., Thiel, R., Jain, P., Singh, V., & Fischer, M. (2023). Digital Twin: Where do humans fit in? *Automation in Construction*, *148*, 104749.

Alaloul, W. S., Qureshi, A. H., Musarat, M. A., & Saad, S. (2021). Evolution of close-range detection and data acquisition technologies towards automation in construction progress monitoring. *Journal of Building Engineering*, *43*. Scopus. https://doi.org/10.1016/j.jobe.2021.102877

Arsiwala, A., Elghaish, F., & Zoher, M. (2023). Digital twin with Machine learning for predictive monitoring of CO2 equivalent from existing buildings. *Energy and Buildings*, *284*, 112851. https://doi.org/10.1016/j.enbuild.2023.112851

Balzano, M., & Marzi, G. (2023). Exploring the pathways of learning from project failure and success in new product development teams. *Technovation*, *128*, 102878.

Betti, G., Tartarini, F., Nguyen, C., & Schiavon, S. (2024). CBE Clima Tool: A free and open-source web application for climate analysis tailored to sustainable building design. *Building Simulation*, *17*(3), 493–508.

Bharadiya, J. P. (2023). The role of machine learning in transforming business intelligence. *International Journal of Computing and Artificial Intelligence*, *4*(1), 16–24.

Boje, C., Guerriero, A., Kubicki, S., & Rezgui, Y. (2020). Towards a semantic Construction Digital Twin: Directions for future research. *Automation in Construction*, *114*, 103179. https://doi.org/10.1016/j.autcon.2020.103179

Borrmann, A., Kolbe, T. H., Donaubauer, A., Steuer, H., Jubierre, J. R., & Flurl, M. (2015). Multi-scale geometric-semantic modeling of shield tunnels for GIS and BIM applications. *Computer-Aided Civil and Infrastructure Engineering*, *30*(4), 263–281.

Botín-Sanabria, D. M., Mihaita, A.-S., Peimbert-García, R. E., Ramírez-Moreno, M. A., Ramírez-Mendoza, R. A., & Lozoya-Santos, J. de J. (2022). Digital twin technology challenges and applications: A comprehensive review. *Remote Sensing*, *14*(6), 1335.

Chen, Z.-S., Liang, C.-Z., Xu, Y.-Q., Pedrycz, W., & Skibniewski, M. J. (2024). Dynamic collective opinion generation framework for digital transformation barrier analysis in the construction industry. *Information Fusion*, *103*, 102096. https://doi.org/10.1016/j.inffus.2023.102096

Coupaye, T., Bolle, S., Derrien, S., Folz, P., Meye, P., Privat, G., & Raïpin-Parvedy, P. (2023). A Graph-Based Cross-Vertical Digital Twin Platform for Complex Cyber-Physical Systems. In *The Digital Twin* (pp. 337–363). Springer.

Faith Tim & Tori Denis. “Gartner”. (2020). The Future of ERP Is Composable. https://www.gartner.com/doc/reprints?id=1-25D2ZFKL&ct=210302&st=sb

Faleiro, R., Pan, L., Pokhrel, S. R., & Doss, R. (2022). Digital twin for cybersecurity: Towards enhancing cyber resilience. *Broadband Communications, Networks, and Systems: 12th EAI International Conference, BROADNETS 2021, Virtual Event, October 28–29, 2021, Proceedings 12*, 57–76.

Haki, K., & Legner, C. (2021). The mechanics of enterprise architecture principles. *Journal of the Association for Information Systems*, *22*(5), 1334–1375.

Harrison, R. (2011). *TOGAF® 9 Certified Study Guide*. Van Haren.

Hasan, S., & Sacks, R. (2023). Integrating BIM and Multiple Construction Monitoring Technologies for Acquisition of Project Status Information. *Journal of Construction Engineering and Management*, *149*(7). Scopus. https://doi.org/10.1061/JCEMD4.COENG-12826

Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2023). AI-big data analytics for building automation and management systems: A survey, actual challenges and future perspectives. *Artificial Intelligence Review*, *56*(6), 4929–5021.

Hook, J., Nielsen, L., & Nyhuis, P. (2022). Introducing a Fast Lane to Multi-Project Environments in Factories to Focus on Digital Transformation. *IFIP International Conference on Advances in Production Management Systems*, 157–164.

Hou, L., Wu, S., Zhang, G., Tan, Y., & Wang, X. (2020). Literature review of digital twins applications in construction workforce safety. *Applied Sciences*, *11*(1), 339.

Jones, L., Adams, K., Stoller, S., Lemaire, M., Vierhaus, E., & Frank, I. (2022). *Data storage architecture for an enterprise communication system*. Google Patents.

Kaur, M. J., Mishra, V. P., & Maheshwari, P. (2020). The convergence of digital twin, IoT, and machine learning: Transforming data into action. *Digital Twin Technologies and Smart Cities*, 3–17.

Kazeem, K. O., Olawumi, T. O., & Osunsanmi, T. (2023). Roles of Artificial Intelligence and Machine Learning in Enhancing Construction Processes and Sustainable Communities. *Buildings*, *13*(8), 2061.

Knebel, F. P., Trevisan, R., do Nascimento, G. S., Abel, M., & Wickboldt, J. A. (2023). A study on cloud and edge computing for the implementation of digital twins in the Oil & Gas industries. *Computers & Industrial Engineering*, *182*, 109363.

Kubelskiy, M. (2021). Application of semantic networks and enterprise architecture approaches for creation of digital twins of organizations. *Networks in the Global World V: Proceedings of NetGloW 2020 5*, 144–161.

Kulkarni, V., Barat, S., & Clark, T. (2019). Towards adaptive enterprises using digital twins. *2019 Winter Simulation Conference (WSC)*, 60–74.

Li, J., & Liu, L. (2023). The Influence of Digitalization Degree on Enterprise Performance in the Era of Big Data. *Journal of Investment and Management*, *11*(4), 63–68. https://doi.org/10.11648/j.jim.20221104.11

Lopes, A. B., & Boscarioli, C. (2021). Business intelligence and analytics to support management in construction: A systematic literature review. *Revista Brasileira de Computação Aplicada*, *13*(1), 27–41.

Lu, Y., Wu, Z., Chang, R., & Li, Y. (2017). Building Information Modeling (BIM) for green buildings: A critical review and future directions. *Automation in Construction*, *83*, 134–148.

Lucchi, E. (2023). Digital twins for the automation of the heritage construction sector. *Automation in Construction*, *156*, 105073.

Luengo, J., García-Gil, D., Ramírez-Gallego, S., García, S., & Herrera, F. (2020). Big data preprocessing. *Cham: Springer*.

Mêda, P., Calvetti, D., Hjelseth, E., & Sousa, H. (2021). Incremental digital twin conceptualisations targeting data-driven circular construction. *Buildings*, *11*(11), 554.

Merino, J., Xie, X., Moretti, N., Chang, J. Y., & Parlikad, A. (2023). Data integration for digital twins in the built environment based on federated data models. *Proceedings of the Institution of Civil Engineers-Smart Infrastructure and Construction*, *176*(4), 194–211.

Müller, S. D., Konzag, H., Nielsen, J. A., & Sandholt, H. B. (2024). Digital transformation leadership competencies: A contingency approach. *International Journal of Information Management*, *75*, 102734.

Omrany, H., Al-Obaidi, K. M., Husain, A., & Ghaffarianhoseini, A. (2023). Digital twins in the construction industry: A comprehensive review of current implementations, enabling technologies, and future directions. *Sustainability*, *15*(14), 10908.

Pal, A., Lin, J. J., Hsieh, S.-H., & Golparvar-Fard, M. (2023). Automated vision-based construction progress monitoring in built environment through digital twin. *Developments in the Built Environment*, 100247.

Pan, Y., & Zhang, L. (2023). Integrating BIM and AI for smart construction management: Current status and future directions. *Archives of Computational Methods in Engineering*, *30*(2), 1081–1110.

Parvin, S. (2021). Success Factors of an Enterprise-Wide Digital Twin Strategy. *Abu Dhabi International Petroleum Exhibition and Conference*, D041S125R003.

Petri, I., Rezgui, Y., Ghoroghi, A., & Alzahrani, A. (2023). Digital twins for performance management in the built environment. *Journal of Industrial Information Integration*, *33*, 100445. https://doi.org/10.1016/j.jii.2023.100445

Ragazou, K., Passas, I., Garefalakis, A., Galariotis, E., & Zopounidis, C. (2023). Big Data Analytics Applications in Information Management Driving Operational Efficiencies and Decision-Making: Mapping the Field of Knowledge with Bibliometric Analysis Using R. *Big Data and Cognitive Computing*, *7*(1), 13.

Ranjan, J., & Foropon, C. (2021). Big data analytics in building the competitive intelligence of organizations.

International Journal of Information Management, *56*, 102231.

Raschka, S., Patterson, J., & Nolet, C. (2020). Machine learning in python: Main developments and technology trends in data science, machine learning, and artificial intelligence. *Information*, *11*(4), 193.

Riege, C., & Aier, S. (2009). A contingency approach to enterprise architecture method engineering. *Service-Oriented Computing–ICSOC 2008 Workshops: ICSOC 2008 International Workshops, Sydney, Australia, December 1st, 2008, Revised Selected Papers 6*, 388–399.

Riss, U. V., Maus, H., Javaid, S., & Jilek, C. (2020). Digital twins of an organization for enterprise modeling. *The Practice of Enterprise Modeling: 13th IFIP Working Conference, PoEM 2020, Riga, Latvia, November 25–27, 2020, Proceedings 13*, 25–40.

Rodrigues, F., Alves, A. D., & Matos, R. (2022). Construction management supported by BIM and a business intelligence tool. *Energies*, *15*(9), 3412.

Sacks, R., Brilakis, I., Pikas, E., Xie, H. S., & Girolami, M. (2020). Construction with digital twin information systems. *Data-Centric Engineering*, *1*, e14.

Salem, T., & Dragomir, M. (2022). Options for and Challenges of Employing Digital Twins in Construction Management. *Applied Sciences*, *12*(6), 2928.

Scheer, A.-W. (2023). From Process and Enterprise Architecture to Digital Enterprise Twin in the Metaverse. In *The Composable Enterprise: Agile, Flexible, Innovative: A Gamechanger for Organisations, Digitisation and Business Software* (pp. 29–49). Springer.

Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. *Automation in Construction*, *158*, 105200.

Sherman, R. (2014). *Business intelligence guidebook: From data integration to analytics*. Newnes.

Statsenko, L., Samaraweera, A., Bakhshi, J., & Chileshe, N. (2023). Construction 4.0 technologies and applications: A systematic literature review of trends and potential areas for development. *Construction Innovation*, *23*(5), 961–993.

Vlăduțescu, Ștefan, & Stănescu, G. C. (2023). Environmental Sustainability of Metaverse: Perspectives from Romanian Developers. *Sustainability*, *15*(15), 11704.

Wang, J., Chen, R., & Lv, Z. (2023). DNA Computing-Based Multi-Source Data Storage Model in Digital Twins. *ACM Transactions on Multimedia Computing, Communications and Applications*, *19*(3s), 1–16.

Wu, A., Wang, Y., Shu, X., Moritz, D., Cui, W., Zhang, H., Zhang, D., & Qu, H. (2021). Ai4vis: Survey on artificial intelligence approaches for data visualization. *IEEE Transactions on Visualization and Computer Graphics*.

Yan, H., Yang, N., Peng, Y., & Ren, Y. (2020). Data mining in the construction industry: Present status, opportunities, and future trends. *Automation in Construction*, *119*, 103331.

Yan, M.-R. (2015). Project-based market competition and policy implications for sustainable developments in building and construction sectors. *Sustainability*, *7*(11), 15423–15448.

Yan, M.-R., Hong, L.-Y., & Warren, K. (2022). Integrated knowledge visualization and the enterprise digital twin system for supporting strategic management decision. *Management Decision*, *60*(4), 1095–1115.

Yang, T.-K., & Yan, M.-R. (2019). Exploring the enablers of strategic orientation for technology-driven business innovation ecosystems. *Sustainability*, *11*(20), 5779.

Yitmen, I., Kovacic, I., & Tagliabue, L. C. (2023). Editorial: Cognitive digital twins for facilitating construction 4.0: Challenges and opportunities for implementation. *Frontiers in Built Environment*, *9*. https://doi.org/10.3389/fbuil.2023.1130115

You, Z., & Wu, C. (2019). A framework for data-driven informatization of the construction company. *Advanced Engineering Informatics*, *39*, 269–277.

Yue, J. (2023). A deep learning method for intelligent decision-making in enterprise management based on the Internet of Things. *Journal of Computational Methods in Sciences and Engineering*, *23*(2), 617–627. Scopus. https://doi.org/10.3233/JCM-226613

Yun, J. J., Won, D., Park, K., Yang, J., & Zhao, X. (2017). Growth of a platform business model as an entrepreneurial ecosystem and its effects on regional development. *European Planning Studies*, *25*(5), 805–826.

Zachman, J. A. (2003). The zachman framework for enterprise architecture. *Primer for Enterprise Engineering and Manufacturing.[Si]: Zachman International*.

Zhang, J., Cheng, J. C., Chen, W., & Chen, K. (2022). Digital twins for construction sites: Concepts, LoD definition, and applications. *Journal of Management in Engineering*, *38*(2), 04021094.

Virtual and Augmented Reality

ON-SITE VISUALIZATION USING BIM AND EXTENDED REALITY

Kaja Pevec and Zoran Pučko
University of Maribor, Faculty of Civil Engineering, Transportation Engineering and Architecture, Slovenia

Abstract

The BIM approach is gaining significant attention in Slovenia's construction industry, soon becoming legally mandatory for publicly funded projects. Our study focuses on merging Building Information Modeling (BIM) and Extended Reality (XR), particularly Mixed Reality (MR), to improve construction processes. Focusing on the device Trimble XR10 with HoloLens 2, the visually compliance of the plans and the BIM model with the actual situation on the construction site was monitored. Additionally, a survey involving participants from local construction firms assessed user experiences with the device and contributed to analyzing advantages and disadvantages of traditional versus advanced approach (BIM and MR).

Introduction

Due to the increasing digitalization around the world, and consequently also in construction, there has been an heightened focus on advanced technologies such as BIM and XR. In Slovenia, the impending legal mandate for BIM implementation in publicly funded projects has spurred discussions on the industry's readiness and the potential benefits of integrating these technologies. The introduction of these technologies and devices faces many barriers, such as traditional mindsets, a culture of mistrust, limited knowledge of the workforce, low awareness of advanced technologies, complexity of the approach, the need to upgrade IT infrastructure, and it is therefore necessary to illustrate the benefits for construction companies. From the literature review (Alizadehsalehi et al., 2020; Safikhani et al., 2022; Sidani et al., 2021), research mainly conducted in the design and operational phase, off-site, by architects and real estate agents, and includes 3D models for visualization for owners, where VR and AR technologies predominating. AR is used during the construction phase to visualize and train workers. MR, which enables the merging of the virtual and real worlds and has the greatest potential for interactive collaboration between stakeholders on the construction sites is rather underrepresented. Our study delves into the practical implications of combining BIM and MR, specifically utilizing the Trimble XR10 helmet with the HoloLens 2 visor (*Trimble XR10 with HoloLens 2*, 2019), to enhance visualization on construction sites. Visualization is becoming a trend in the construction industry, as it enables easier understanding when monitoring construction projects and processes. With the help of advanced technologies such as AR (Augmented Reality) and MR (Mixed Reality), visualization on the construction site is more practical and efficient. An experiment was conducted on the construction site of the Retirement Home Šmarje pri Jelšah – dislocated unit Kozje, involving the creation and placement of a customized 3D BIM model with the advanced devise. The Trimble Connect AR + MR program ("Trimble Connect AR + MR," n.d.) was used for the experiment. Simultaneously, in order to compare the advanced and traditional approach in construction projects, a survey was conducted, involving 25 employees from local construction firms to capture user experiences with the advanced device and opinions on advanced approaches in construction, including the answer to the question of whether it makes sense to invest in advanced devices.

On-site visualization

Visualization

Visualization means creating a visual representation of information for easy understanding and interpretation. Visualization with the help of computer graphics and image processing enables the visual representation of forms and processes that we perceive and imagine more easily. With new technologies, immersive visualization has also appeared, which enables users to have an in-depth experience with the help of various interfaces, screens and devices. By combining enhanced sensory perception and tangible interaction, immersive visualization aims to reduce the user's perception of the real world so that their attention is focused solely on the immersive, non-real environment, thereby improving analytical and decision-making skills. This is where extended reality technologies come to the fore. By using immersive or advanced devices, it is possible to provide users with multi-sensory feedback that allows them to perceive experiences that closely resemble reality. Through sensory channels, the sense of control and realism increases for users, which usually improves the efficiency of users in obtaining data (Zhang et al., 2023).

BIM approach

Saving time, costs and improved collaboration between project participants can be achieved by collecting information about the project in one place, accessible anytime and anywhere. The above is a simplified definition of the BIM approach (Sacks et al., 2018). BIM stands for Building Information Modelling and it is a process supported by a number of tools and technologies that involve the creation and management of digital representations of the physical and functional properties of the built environment (Marc et al., 2018). In the research two dimensions of BIM, namely 3D and 4D BIM were used. 3D BIM is a geometric model of a building, which captures all the geometric data of the model and the individual building asset in an interconnected manner. The 4D BIM dimension is the 3D BIM model upgraded with a specific time dimension, with a time schedule for the construction of the building asset of the BIM model. For the purpose of the experiment, a basic 3D BIM model of the building was obtained and adapted to facilitate the treatment with the advanced device. In the practical part, a 3D BIM model over time (sequences) using MR, were presented, which already mainly belongs to the field of 4D BIM.

Extended Reality

XR is a technology that combines both the digital and the real world. It augments the physical world with digital/computational elements using advanced devices. It is divided into three types: VR - virtual reality, AR - augmented reality and MR - mixed reality. VR is a technology that creates a virtual, digital environment. AR allows digital elements such as graphics, sound to be added to the real world. Virtual elements are placed in the real environment through interfaces, enriching it, while the user still maintains contact with the physical environment. MR is a combination of VR and AR, combining the real and virtual worlds. It allows the integration of virtual elements into the real environment. This technology allows the user to interact with digital elements in a real environment while maintaining awareness and contact with the real world. MR is a great tool for good on-site visualization in combination with the BIM approach, which will be presented in the next section.

Experiment

The case

The building under consideration, the Retirement Home Šmarje pri Jelšah – dislocated unit Kozje, is designed in a semicircular form, with three floors and comprises 2555.47 m^2 of net usable floor area. The foundation slab, the intermediate floor slabs and external walls are of reinforced concrete. The internal walls are of brick construction. All partition walls in the building are dry-walled - a system of two-layer plasterboard panels in a metal substructure with mineral wool insulation. The roof structure is a reinforced concrete panel. The building is designed to accommodate 49 residents. It is designed as a zero-energy building according to the Design&Build principle. The 3D BIM model of the building is presented in Fig. 1.

Figure 1: 3D BIM model of the building

3D BIM model preparation

The 3D BIM model of the building in question contains only the structural elements and interior fittings, without the necessary installations, with the focus on the individual structural elements when monitoring the progress of the work on site with the advanced device. For further consideration of the building model and placement in the space, the model was scaled down as the full model file was too large to export and use on an advanced device. For the purpose of this paper, the limitation is given to the first floor and therefore all unnecessary elements have been deleted in the 3D BIM model. The ArchiCAD program was used to edit the model. For greater visualization or transparency of the layout of the 3D BIM model on the location in a 1:1 scale, the mezzanine panels with the terrace were also removed. The other elements on the first floor were preserved. The lift shafts are visible in their entirety, as they were created individually as one element. Fig. 2 shows the adjusted 3D BIM model of the building's first floor displayed in the Trimble Connect program.

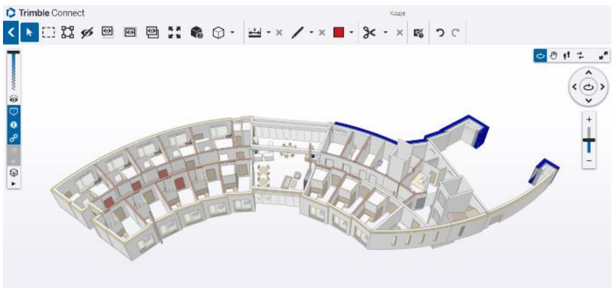


Figure 2: Customized 3D BIM model of the first floor

Equipping a 3D BIM model with a QR code

To facilitate the placement of the 3D BIM model in the real environment, a QR code or marker needed to be generated using the Trimble Connect software tool. The software generates a unique QR code, which connects to the advanced device while using the application Trimble Connect MR+AR. One or more QR codes can be generated for a single 3D BIM model of a set of elements or an object. For an easier determination of the location of the sheet with the QR code on the construction site a starting point on the 3D BIM model was chosen (presented in Fig. 3).

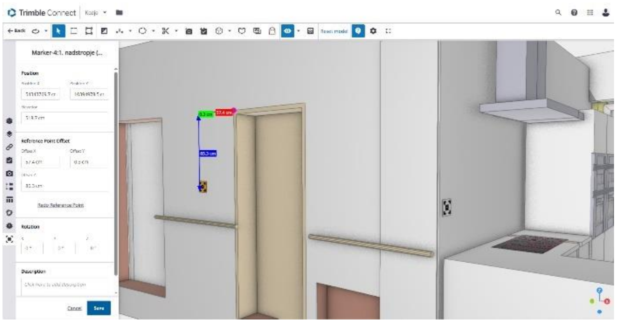


Figure 3: QR code locations and determining the starting point

The QR codes were printed in 100 % custom ratio and attached to selected locations on the object (see Figure 4). When placing the QR code on real elements, it is important to correctly determine the starting point, which was predetermined in the software, and place the lower left edge of the QR code exactly at this point. This requires a flat and clean surface to attach the QR code, taking into account orthogonality and using a quality mounting agent.

Figure 4: Attached QR codes on the building

Model placement in a real environment

An advanced device to carry out the experiment on site was used. The Trimble XR10 with HoloLens 2 is an advanced device that uses extended reality technology to bring digital elements from the screen into the physical world and clearly present the work of all participants, allowing projects to be completed on time and at no extra cost. It is the product of a successful collaboration between Microsoft Corporation and Trimble Inc. The device consists of two parts (see Figure 5) the Trimble XR10 protective work helmet and the HoloLens 2 visor.

Figure 5: Trimble XR10 with HoloLens 2

The Trimble Connect AR+MR application allows you to place a 3D BIM model into any environment in several different ways, in different sizes and views. It can simply open the model in the application and place it in any place, in any scale. Figure 7 illustrates the basic display of the entire 3D BIM model of the building on the construction site. For better accuracy and faster placement of the model on the construction site, the "Marker" subcommand was used, to scan the previously prepared and pasted QR code on the building. The model placement process is shown in Figure 6.

Figure 7: Display of the entire 3D BIM model of the building with its surroundings on the construction site

The first step was choosing the subcommand “Marker” in the menu of the advanced device. Then, with the help of the camera integrated in the advanced device, the QR code was scanned and uploaded a digital 3D BIM model using MR technology and projected it in a real environment. This allowed us to place the model on a real object at 1:1 scale. In the process of integrating the 3D BIM model at the construction site, occasional inaccuracies arose in projecting and aligning the model with the real structure, resulting in maximum deviations of 10 cm, when aligning along a non-straight wall. The minimal misalignment was observed when positioning the model using a QR code.

Preparation of the Sequence

Sequences in the Trimble Connect desktop application were created to show the construction phases of each structural element on the first floor of the building. One sequence will be presented in more detail in a selected room on the ground floor of the building in question (presented in Fig. 8). The sequence consists of seven steps (Fig. 9-15) and shows the construction of the vertical structural elements of the room in question according to the prepared 3D BIM model. The first step is the construction of the external walls, the second step is the construction of the internal brick walls, followed by the third step of the plastering. The fourth step is the installation of the internal panels of the internal drywall for the bathroom, the fifth step is the installation of the insulation on the drywall, the sixth step is the construction

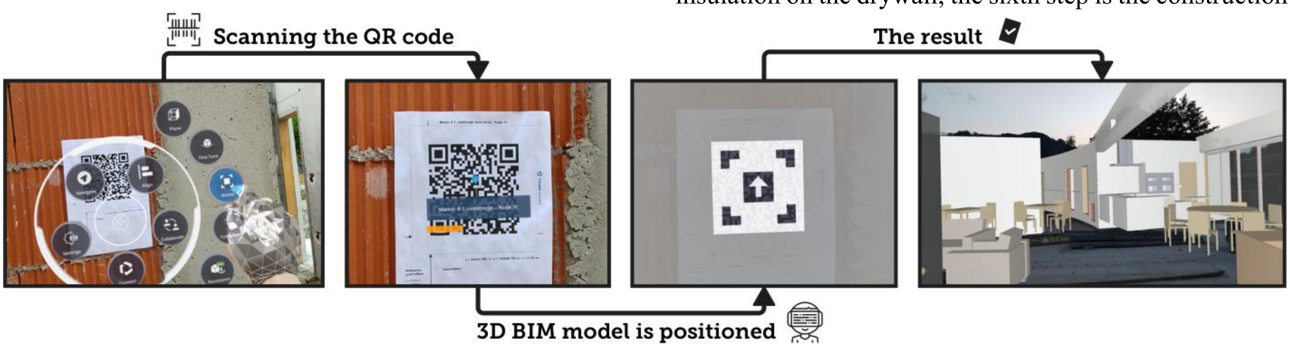


Figure 6: The process of placing a 3D BIM model with a QR code in a real environment

of the external panels of the internal drywall and the last step is the fixing of the tiles on the drywall in the bathroom. The sequence does not take into account the additional steps in the construction of the external wall (formwork, concreting, dismantling formwork), but is shown in one step, as “the construction of the external walls” (Fig. 9). The selected elements are shown and labelled individually in the figures, step by step, while the remaining elements are hidden. If a sequence were to be played, each previous element would remain displayed. For better understanding, each step and its corresponding element has been marked with a specific color, e.g. walls are marked in green.

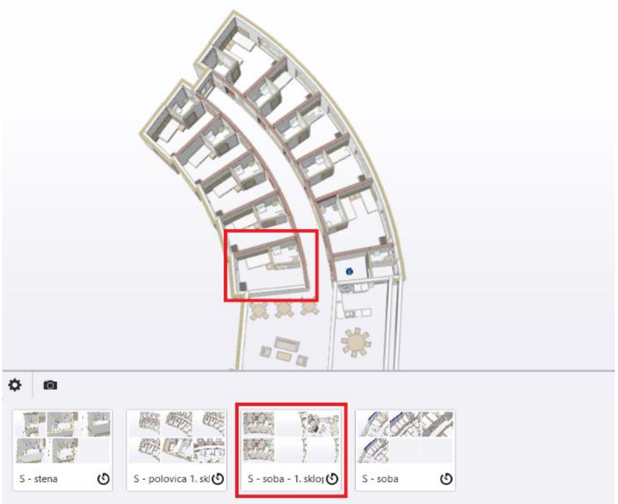


Figure 8: Selected room and sequence in the 3D BIM model

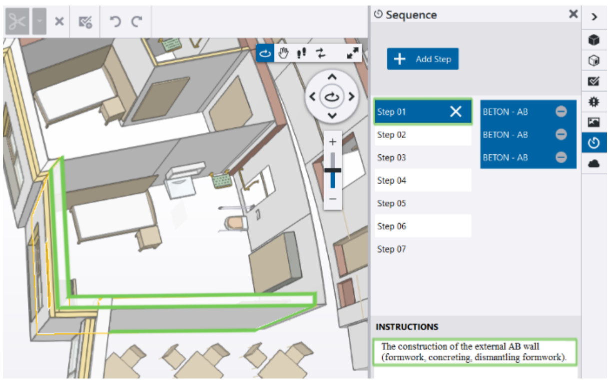


Figure 9: Step 1

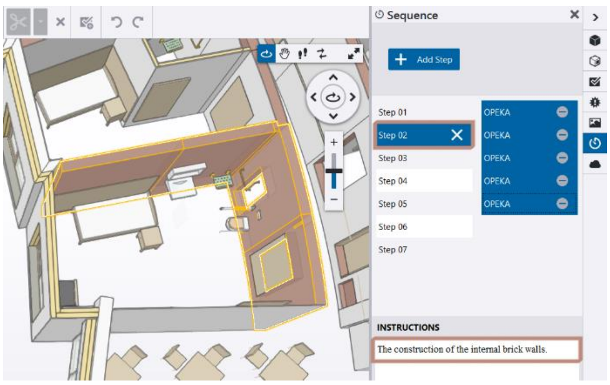


Figure 10: Step 2

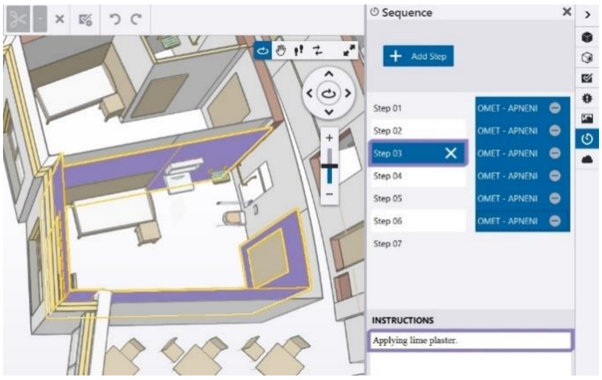


Figure 11: Step 3

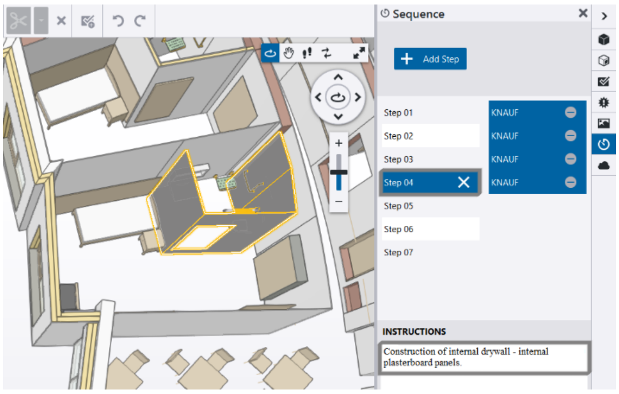


Figure 12: Step 4

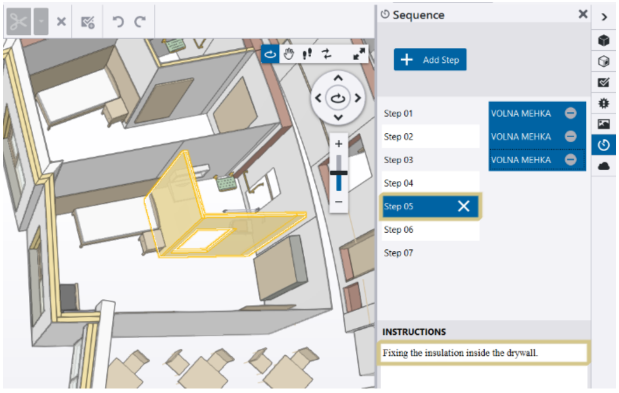


Figure 13: Step 5

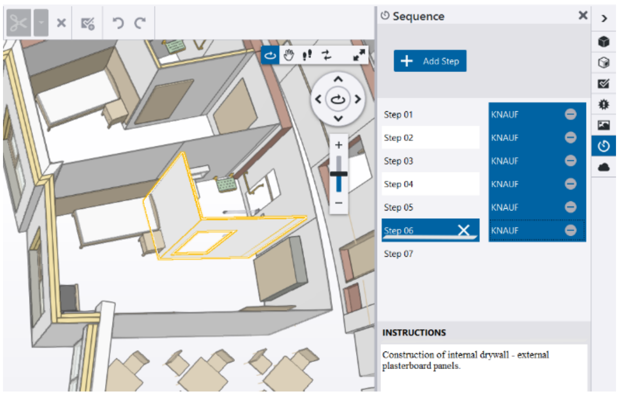


Figure 14: Step 6

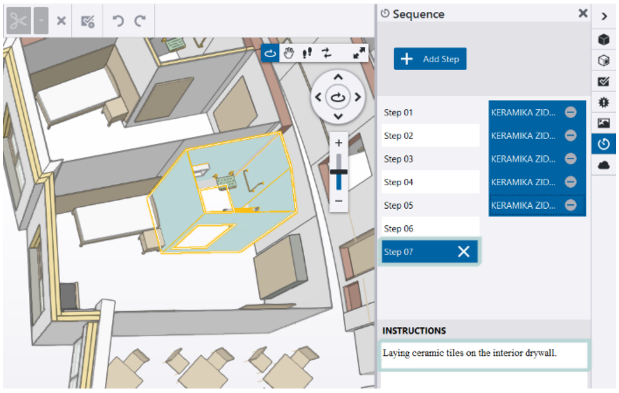


Figure 15: Step 7

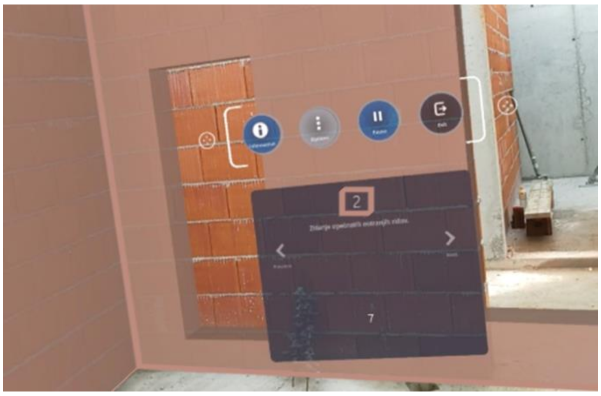


Figure 17: Step 2 of the on-site sequence

On-site sequence display

In the previous subsection Preparation of the Sequence, the construction sequence of vertical construction elements of one room on the first floor of the building was demonstrated. In the following, the sequences will be shown on a 3D BIM model with the advanced device on the construction site.

The sequence is displayed in a real environment by selecting the command "Sequence" in the basic menu of the Trimble AR+MR application and then selecting the desired project, model. After the basic placement of the model, the "Collaborate" command and the "Sequence" subcommand were selected. A window appeared showing the steps of the selected sequence. You can move the steps manually or play the sequence. Visualization of the individual steps of the sequence on the construction site is shown in Figures 16-22 below. Each displayed 3D element of the sequence step is colored with a specific color, as in the previous subsection. On the model, a hole in the bathroom is present where the purpose is to show the area where the installation will take place and be kept. This hole was subsequently made on the building, but in a smaller size. It is also visible that the balcony and entrance doors in the model are placed higher than the actual floor, as the remaining materials and elements on the floor have not yet been installed.

Figure 16: Step 1 of the on-site sequence

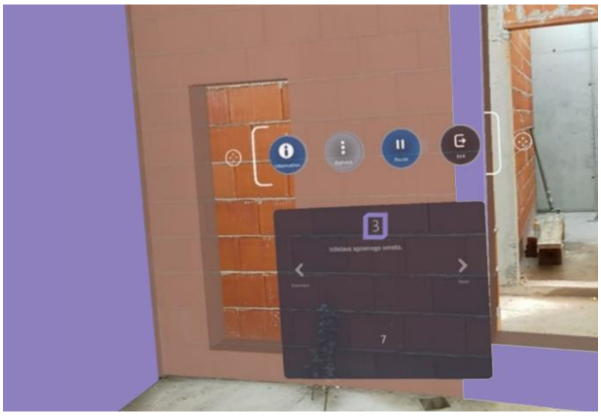


Figure 18: Step 3 of the on-site sequence

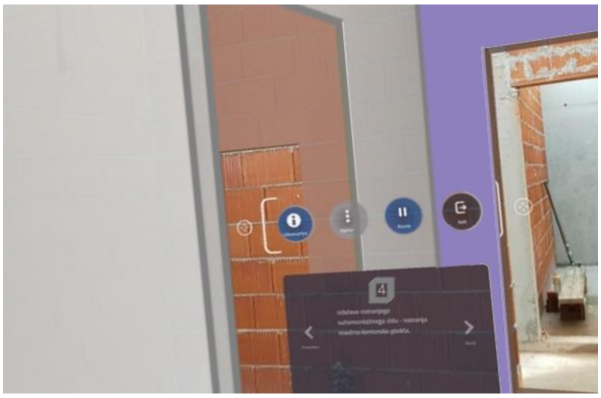


Figure 19: Step 4 of the on-site sequence

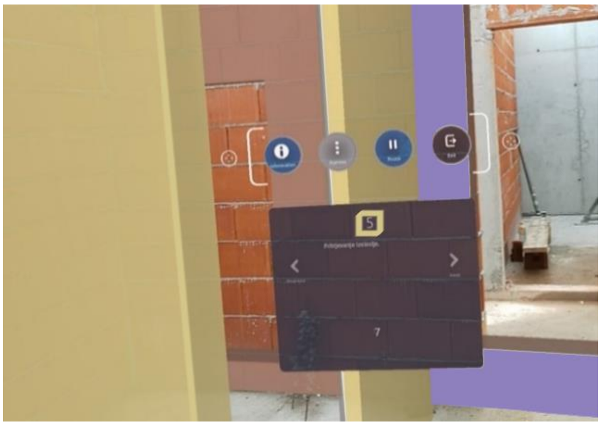


Figure 20: Step 5 of the on-site sequence

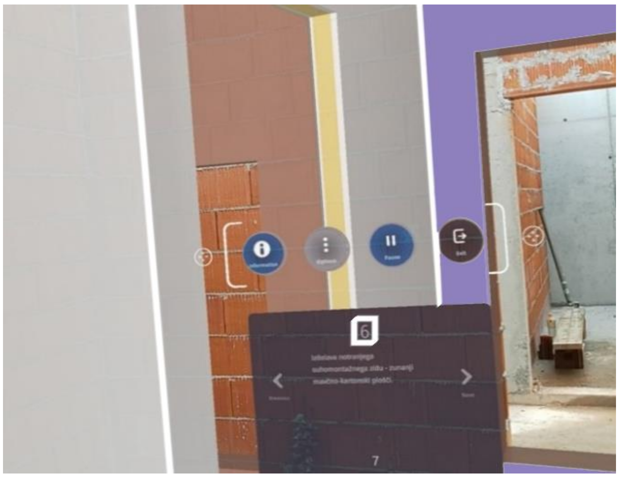


Figure 21: Step 6 of the on-site sequence

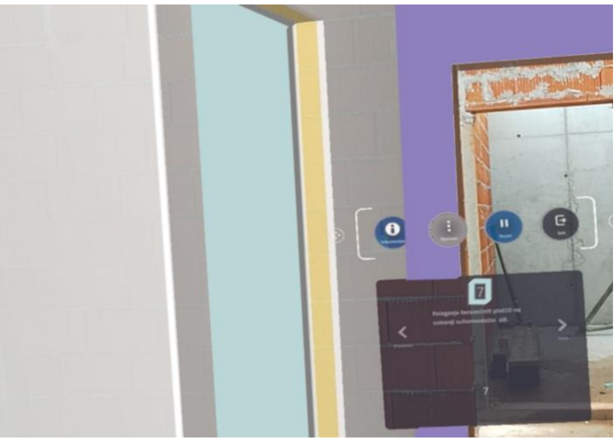


Figure 22: Step 7 of the on-site sequence

For an easier presentation and understanding of the sequence in a real environment, some steps of the sequence (steps of construction elements, without final cladding), through the construction process of the actual building, on the construction site are presented in Figure 23. These steps are: construction of the RC external walls (1), internal brick walls (2) and internal drywalls (4&6).

After the first two steps, work on the construction site continued on the second section of the first floor, where the vertical structural elements were constructed. Then the work for the installation of the slab between the floors was carried out. The next step of the sequence started after the construction of the external and internal walls on the second floor has been completed.

Figure 23: Steps 1, 2, 4 and 6 of the sequence shown with the actual object (Source: Josip Pongračić)

Survey

The traditional approach of looking at plans and monitoring progress in the field has started to be replaced by a modern approach using advanced devices such as the Trimble XR10 with HoloLens 2, which allows the projection and interaction with digital models in real time, in a physical environment. In order to gain insight into the user experience of construction employees for this device, a survey was conducted to obtain relevant information on the pros and cons of the advanced device and to determine the rationale for use and investment. For the purpose of the research, a questionnaire was prepared, entitled "The use of advanced devices for extended reality in construction projects". At the time of the experiment, there were 40 workers at the construction site who were invited to try the Trimble XR10 with HoloLens 2. 15 individuals refused to participate, while 25 took part in the survey, representing 62,5 % of those invited. The questionnaire consisted of 14 closed-ended questions and one open-ended question. For the six questions on user experience, a 5-point descriptive scale, the Likert scale, was used to better analyze the responses. The results were then processed and displayed in Excel. In the following, the findings and the analysis of the respondents' answers will be presented.

The first four questions related to the general analysis of respondents' statistics: age, education, type of employment and workplace. 25 people took part in the questionnaire, 20 male and 5 female. This can be attributed to the fact that more male workers are employed in the construction sector due to the physical labor involved. Of all respondents, 10 have completed secondary school, 5 have completed a master's degree, 3 have completed a vocational or professional higher education, 2 have completed a bachelor's degree and 2 have completed primary school. None of the respondents had completed a PhD or other course. When asked about the type of employment, the majority of respondents answered that they were in regular employment, only 4 were in student services, 2 were self-employed and 1 answered contract work.

Since the paper focuses on visualization on the construction site, the type of employment of the respondents is also interesting. In order to get as diverse answers and opinions as possible, several different professions in the construction industry are covered in the survey. Most respondents (36 %) chose the answer "Other" to the question "Job" and gave the following answers: architectural engineer, mechanical engineer, mechanical development engineer, independent purchaser, electrician, IT technician (2 persons), blacksmith and project associate. 28% of all persons surveyed were site workers, followed by civil engineers or project managers with 12%. Two persons were foremen and one person was a supervisor.

The fifth question was: Have you had any experience with VR/AR/MR devices before using an advanced device? 72% of the respondents (18) had no previous experience of using VR, AR or MR devices prior to using the Trimble XR10 with HoloLens 2. This was the first time that the majority of the respondents had seen and used an advanced device in person and were familiar with new technologies in construction.

Table 1 below shows the results of the level of agreement with the statements related to the respondents' user experience using the Trimble XR10 advanced device with HoloLens 2. The statements related to the ease of use of the advanced device, the comfort of wearing the advanced device, how satisfactory the display and sound is, whether the device responds well to commands, and the perception of the usefulness of using the device on the construction site. Each statement was rated on a 5-point descriptive scale, where I don't agree at all, 2 is I partially agree, 3 is I neither agree nor disagree, 4 is I agree and 5 is I strongly agree. The columns for each statement show the number of respondents and their proportion of each answer, the level of agreement.

Table 1: Respondents' user experience with the advanced device

5-point descriptive scale	1	2	3	4	5
	Number (%)	Number (%)	Number (%)	Number (%)	Number (%)
It is easy to use.	0 (0 %)	2 (8 %)	8 (32 %)	**10 (40 %)**	5 (20 %)
It is comfortable to wear.	0 (0 %)	1 (4 %)	5 (20 %)	**12 (48 %)**	7 (28 %)
The display is satisfactory.	0 (0 %)	2 (8 %)	5 (20 %)	8 (32 %)	**10 (40 %)**
The sound is satisfactory.	0 (0 %)	1 (4 %)	8 (32 %)	**12 (48 %)**	4 (16 %)
Good responsiveness to commands.	0 (0 %)	1 (4 %)	5 (20 %)	**13 (52 %)**	6 (24 %)
Use on the construction site is very useful, welcome.	0 (0 %)	0 (0 %)	7 (28 %)	7 (28 %)	**11 (44 %)**

The majority (40%) of respondents found the use of the advanced device to be straightforward, easy, with no respondents completely disagreeing.

Regarding comfort, over half of the participants (48%) agreed that wearing the advanced device was comfortable, and 28% fully agreed.

A significant factor in device usability is satisfactory content display. Eight percent disagreed with the display quality, while 40% expressed complete agreement, indicating a prevailing positive view.

Concerning sound quality, 48% agreed that the device provided satisfactory sound, while only one person disagreed, emphasizing the importance of clear and loud audio on construction sites.

Evaluating responsiveness to commands, the vast majority (76%) agreed that the advanced device responded well to both voice and gesture commands.

The final statement, "The use of the advanced device on the construction site is very beneficial and welcome," received overwhelming agreement, with 72% fully supporting its usefulness on construction sites. Respondents generally praised the advanced device for user-friendliness, comfort, display quality, sound clarity, responsiveness, and unanimously recognized its substantial benefits for construction site use.

It should be borne in mind that the respondents had only used the device for a short period of time.

The most important and the thirteenth question was: What do the respondents consider to be the biggest advantage of using an advanced device on site compared to the traditional approach? The vast majority of respondents (17), as much as 68%, answered that better visualization is the biggest advantage when using an advanced device on a construction site compared to a traditional approach. 4 people felt that the biggest benefit was better collaboration, and 2 people chose faster construction as the answer. 1 person answered that higher accuracy is the biggest advantage, while 1 person identified it as lower cost. No one recognized greater safety as a sufficiently important factor, nor did they highlight other examples.

When asked whether they would like to use the advanced device they tested in their work, as many as 17 people out of 25 respondents expressed a desire to use the advanced device in their work. The remaining 8 people, which represents 32%, do not. It can be concluded that the advanced device left a good impression on the respondents and that they see sufficient advantages in its use.

The final open-ended question was: What is your opinion about the usefulness of an advanced device in an advanced approach in construction compared to a traditional approach? Respondents varied in their opinions on the utility of the advanced device. While some declined to respond, the majority viewed the device as highly beneficial for on-site visualization during construction, emphasizing its role in providing a clearer understanding of the project, facilitating control, and comparing the actual state with the planned one. Two respondents highlighted its usefulness in swiftly resolving issues on-site. Improved communication between clients and contractors emerged as a significant advantage, aiding those struggling with 3D visualization. One respondent favored the advanced approach, citing error reduction, cost savings, and increased profitability. However, concerns were raised about the limited presence of such devices in construction, undervaluing their potential. One respondent emphasized the device's usefulness only when all details in the model are precise; otherwise, it is deemed a waste of time. Ensuring adherence to plans and instructions was seen as crucial for successful implementation. Traditional approaches were criticized for deviations, causing delays, unforeseen costs, errors, and tensions among participants. Participants found it challenging to comment on the downsides of advanced technology due to limited experience, but highlighted high costs, potential discomfort during use, and the risk of damage as concerns. The unfamiliarity with advanced technology among participants raised questions about collaboration efficiency, potentially leading to conflicts and misinterpretations.

Result analysis

Visualization on construction sites brings numerous advantages, fostering better understanding of projects and facilitating coordinated workflow. The Trimble XR10 with HoloLens 2 proves invaluable in envisioning the final structure, understanding project elements, and determining necessary construction tasks. Its user-friendly interface and versatility contribute to its ease of adoption, with the Trimble Connect AR+MR application being central to the study. While highlighting the device's benefits, the text explores its applications, advantages, and its potential impact on the construction industry.

The advanced device offers several advantages, such as ease of use, good visualization, seamless integration of 3D BIM models into real environments, display of work sequences, improved communication, and quick access to data, and archives. The rising accessibility of such devices due to affordability and mass production signals a shift towards incorporating advanced technologies in construction, gradually replacing traditional methods.

The paper also discusses the intricacies of model placement using QR codes and acknowledges limitations, such as imprecise projections without QR codes. External factors, like weather conditions, impact the success of QR code applications, necessitating laminated codes for durability. Practical challenges include language barriers, as commands are predominantly in English, posing a hurdle for non-English-speaking construction workers.

The importance of sequences using the advanced device becomes evident for project supervisors, managers, and other participants. Creating sequences proves crucial for specific and intricate tasks, ensuring precision in the construction process. The ability to convey the work sequence aids in organizing diverse workgroups, clarifying task logistics, and addressing critical points. Adhering to established sequences promotes timeliness, well-organized workgroups, enhanced collaboration and better task comprehension.

However, challenges emerge, particularly related to the device's weight and discomfort during prolonged use. Issues arise in strong sunlight, impacting content visibility. The device's limited battery life, approximately 2 hours, poses a constraint but can be mitigated with additional batteries. Technical glitches, encountered during the application update process, temporarily interrupted the experiment, emphasizing the need for continuous software improvements and user support.

Despite initial skepticism and language barriers, worker feedback on the Trimble XR10 has been positive. The device's ability to bridge communication gaps between leaders and workers is seen as advantageous, with potential benefits outweighing drawbacks. The implementation of advanced technologies is seen as transformative, enhancing work conditions, reducing costs and expediting construction processes.

Conclusions

This paper explores the integration of BIM and XR for construction visualization, focusing on the significance of visualization, the BIM approach, XR technologies, and the device Trimble XR10 with HoloLens 2. An experiment conducted at a retirement home construction site, utilizing a customized 3D BIM model on the existing structure, provided valuable insights into the practicality of advanced technologies in construction. Despite some drawbacks, the Trimble XR10 showcased numerous advantages, emphasizing user-friendly operation, effective workflow and improved communication. The results highlighted positive responses from workers, acknowledging the benefits of advanced technology, particularly improved visualization using advanced devices over traditional approaches. The study suggests that investing in XR technologies yields significant benefits, improving productivity, communication, and project understanding in construction projects. Further research is recommended to analyze the use of XR devices under controlled laboratory conditions and to compare user experiences in ideal scenarios with actual building conditions. Validating investments in XR devices in combination with BIM approach would promote a deeper understanding of their impact on the construction industry.

Acknowledgments

The research was conducted in collaboration with the company GIC GRADNJE, d. o. o. and we extend our gratitude for providing the documentation and the 3D BIM model. We also express our gratitude to the survey respondents for their participation and for sharing their insights.

References

Alizadehsalehi, S., Hadavi, A., Huang, J.C., 2020. From BIM to extended reality in AEC industry. Autom Constr 116. https://doi.org/10.1016/j.autcon.2020.103254

Marc, K., Medved, S.P., Štravs, B., Tibaut, A., Žibert, M., Brus, G., Lah, M., Janjić, V., 2018. Priročnikza pripravo projektne naloge za implementacijo BIM-pristopa za gradnje. Ljubljana. Sacks, R., Eastman, C., Lee, G., Teicholz, P., 2018. BIM

Handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors. BIM Handbook. https://doi.org/10.1002/9781119287568

Safikhani, S., Keller, S., Schweiger, G., Pirker, J., 2022. Immersive virtual reality for extending the potential of building information modeling in architecture, engineering, and construction sector: systematic review. Int J Digit Earth. https://doi.org/10.1080/17538947.2022.2038291

Sidani, A., Matoseiro Dinis, F., Duarte, J., Sanhudo, L., Calvetti, D., Santos Baptista, J., Poças Martins, J., Soeiro, A., 2021. Recent tools and techniques of BIM-Based Augmented Reality: A systematic review. Journal of Building Engineering. https://doi.org/10.1016/j.jobe.2021.102500 Trimble Connect AR + MR [WWW Document], n.d. URL https://connect.trimble.com/storefront (accessed 12.14.23). Trimble XR10 with HoloLens 2, 2019.

Zhang, Y., Wang, Z., Zhang, J., Shan, G., Tian, D.,

2023. A survey of immersive visualization: Focus on perception and interaction. Visual Informatics 7. https://doi.org/10.1016/j.visinf.2023.10.003

APPLICATIONS OF IMMERSIVE TECHNOLOGIES IN CONSTRUCTION COMPUTING: A SYSTEMATIC LITERATURE REVIEW

Diran Yu[1], Weiwei Chen[1*], Tao Meng[2], Haiying Yu[2], Abdul-Majeed Mahamadu[1]
[1] The Bartlett School of Sustainable Construction, University College London (UCL), London WC1E 6BT, UK
[2] College of Civil Engineering and Architecture, Zhejiang University
866 Yuhangtang Road, Xihu District, Hangzhou 310058, China

Abstract

This systematic literature review (SLR) examines the role of augmented reality (AR) and virtual reality (VR) in construction computing. It highlights how these immersive technologies enhance visualization, communication, and decision-making in construction in the whole life cycle. The review covers the evolution and applications of AR and VR in architectural modelling, building simulation, construction education, robotics, and site management. It also discussed the challenges and future directions of these technologies in construction computing, finding the bridge between technological advancements and practical application. The findings offer strategic recommendations for developing immersive technologies in building computing, aiming for a smarter, more efficient construction future.

Introduction

The historical trajectory of augmented reality (AR) and virtual reality (VR) in the realm of construction computing is a narrative of progressive integration and innovative breakthroughs. Mixed Reality (MR) is of increasing interest within technology but is not yet used widely. Hence, this study is not going to be concerned about MR or other immature immersive technology. Tracing back to the late 20th century, the inception of AR and VR technologies was primarily driven by the entertainment and military sectors (Lenoir, 2000; Lele, 2013). However, their potential for transforming the architectural and construction industries was quickly recognized (Prabhakaran, Mahamadu and Mahdjoubi, 2022; Safikhani et al., 2022). The progression from simple wireframe models to fully immersive simulations marks a period of rapid development in these technologies, reflecting a trajectory from nascent experimentation to sophisticated application.

As immersive technologies entered the building sector, they brought forth unprecedented tools for visualization and interaction. The application categories of immersive technology within construction have diversified over time. Initially, AR was utilised to overlay digital information on physical environments, enhancing the understanding of complex structures (Li, Nee and Ong, 2017). VR, on the other hand, allows for complete immersion within virtual constructs, providing a platform for thorough review and planning without the constraints of physical models (Seth, Vance and Oliver, 2011). In contemporary practice, these technologies have found various applications. For example, in architectural modelling, AR and VR facilitate the exploration of design alternatives in real time, allowing architects and clients to visualize changes instantly (Delgado et al., 2020; Shouman, Othman and Marzouk, 2022). In building simulation, they enable the assessment of building performance under different conditions, aiding in the optimization of energy efficiency and structural integrity (Niu, Pan and Zhao, 2016). The use of AR in construction education has revolutionized the learning experience, providing interactive and engaging training environments that surpass traditional classroom settings (Sepasgozar, 2022). On construction sites, VR has been instrumental in safety training, providing workers with a risk-free environment to practice and understand complex scenarios (Chan, 2012). Furthermore, AR has been utilized for on-site construction management, where it aids in precise installations and maintenance through visual guidance and information overlay (Li et al., 2018).

Despite these advancements, the application of AR and VR technologies in construction computing is not without challenges (White, Schmidt and Golparvar-Fard, 2014; Wang *et al.*, 2018). The industry faces a disconnect between the evolving capabilities of these technologies and their practical implementation. Therefore, this study will systematically investigate and analyse the various issues encountered by immersive technology applications within the field of construction computing referring to the application of computational technologies and methods in the construction industry to improve the design, planning, management, and operation of construction projects., summarizing and identifying their underlying causes.

The primary objective of this study is to critically evaluate the impact of AR and VR on building construction. The nearest related literature review study was finished in 2018 and only focuses on construction safety. This study will gather the latest technology and the literature review focus on technological innovation and new applications. It will scrutinize how these immersive technologies are currently being harnessed to enhance the efficiency, accuracy, and safety of construction projects. The research will systematically categorize the different applications of AR and VR in construction, focusing on their roles in project planning, on-site execution, and post-construction evaluation. The study will delve into the technological nuances of these applications, assessing their operational merits and identifying the factors that limit their wider adoption and effectiveness in the field.

By doing so, the paper aims to unearth the root causes of the disparity between the potential of these technologies and their actual application outcomes. It seeks to provide insights that could drive the removal of barriers to adoption and optimize the application of AR and VR in the construction industry.

Methodology

This study endeavours to execute a balanced and analytical examination of the literature pertaining to the application of AR and VR technologies in construction computing. To fulfil this aim, the research adopts a methodical approach that combines both systematic (PRISMA) and narrative literature review techniques. Originating in medical science, the systematic review method strives for a comprehensive collation of information from diverse sources, ensuring transparency and rigor (Tranfield, Denyer and Smart, 2003). Renowned for its meticulousness and replicability, this method is instrumental in guiding decision-making (Moher *et al.*, 2009; Okoli and Schabram, 2010). The narrative review will concentrate on explicating and discussing various dimensions of immersive technology in construction, encapsulating advanced knowledge in the sector, including technological innovations, operational methods, and strategic directives.

((TITLE-ABS-KEY ("Immersive Technologies") OR TITLE-ABS-KEY ("Augmented Reality") OR TITLE-ABS-KEY ("Virtual Reality")) AND TITLE-ABS-KEY ("Computing") AND TITLE-ABS-KEY ("Construction")) AND ((TITLE-ABS-KEY ("Architectural design") OR TITLE-ABS-KEY ("Simulation") OR TITLE-ABS-KEY ("Construction management")) OR (TITLE-ABS-KEY ("Devices") OR TITLE-ABS-KEY ("Facilities")) OR (TITLE-ABS-KEY ("Trend") OR TITLE-ABS-KEY ("Challenge"))) AND (LIMIT-TO (SRCTYPE , "j")) AND (LIMIT-TO (PUBSTAGE , "final")) AND (LIMIT-TO (DOCTYPE , "ar")) AND (LIMIT-TO (PUBYEAR , 2019) OR LIMIT-TO (PUBYEAR , 2020) OR LIMIT-TO (PUBYEAR , 2021) OR LIMIT-TO (PUBYEAR , 2022) OR LIMIT-TO (PUBYEAR , 2023)) AND (LIMIT-TO (LANGUAGE , "English"))

Figure 1: Variable search string

Data collection

In the research, the initial phase was to delineate the pertinent literature's boundaries, thus informing the refinement of our search parameters. The current phase meticulously maintains core search terms while encompassing a spectrum of conceptual terms related to immersive technologies within construction computing. This amalgamation of terms (immersive technologies, augmented reality, virtual reality) with pertinent domains (architectural design, simulation, construction management) and challenges yields a search query both exhaustive and nuanced. Reflecting the evolution of construction computing since the early 2010s and acknowledging publication lags, our temporal search parameters span from 2019 to 2023 to ensure the inclusion of the most recent and relevant findings. Additionally, to ensure the retrieved literature's relevance and scholarly impact, we filtered for articles published in English and excluded early access reviews, focusing on articles with empirical evidence that illuminate the application challenges of AR and VR in architectural computing.

In the initial phase of our literature review, the search strategy outlined in the mind map identified 66 articles post-consolidation from both Scopus and Web of Science databases. This process parallels the systematic methodologies observed in other architectural computing studies (e.g., Smith and Jones, 2019; Doe et al., 2020). The following multi-stage evaluation was implemented:

i. We excluded 7 duplicates, refining the selection to 59 articles.
ii. Accessibility checks further narrowed the pool, with 27 articles excluded due to restricted access, insufficient impact factor, or lack of an abstract, leaving 32.
iii. Relevance to the topic was the next criterion, with 8 articles deemed of low relevance and 24 classified as highly relevant.
iv. The final assortment comprised articles that not only aligned closely with the thematic core of immersive technologies in construction but also promised significant contributions to the field, ensuring a robust foundation for our analysis.

In the final

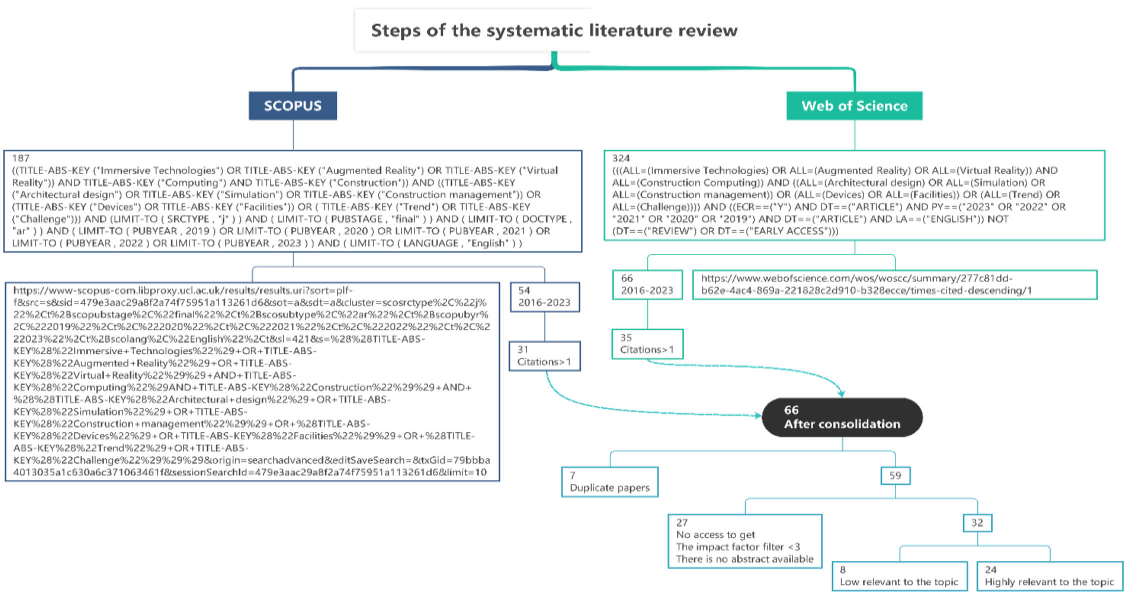


Figure 2: Steps of the systematic literature review

selection phase, a dual review approach, integrating both systematic and narrative methodologies, was employed. Through this mode, 24 papers of high relevance were meticulously identified, resonating with the central inquiry ‘What are the challenges and deep-seated reasons for the issues arising from the application of immersive technologies in construction computing?’ This thorough and discerning process entailed an in-depth examination and critical assessment of the literature, ensuring the inclusion of only the most significant studies that would contribute to the analysis and discussion of the identified challenges in the field.

Data analysis

To delineate the current landscape and progression within the field of immersive technologies in construction computing, this study conducted a scient metric analysis using VOSviewer. The initial visualization, presented in Figure 3, illustrates the clustering of keywords into three main groups. The first cluster is centred on the educational aspect, indicated by terms like 'student', 'training', and 'behaviour', suggesting a focus on the role of immersive technologies in learning and skill development. The second cluster revolves around the construction industry with 'construction site' and 'industry' as prominent nodes, pointing to discussions on the application of technologies in construction settings. The third cluster relates to the technological framework, including 'AR', 'information', and 'experiment', highlighting the emphasis on technology deployment and testing within the industry.

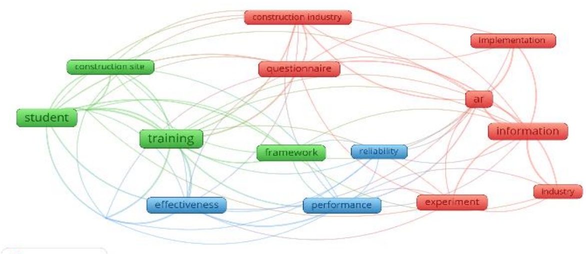


Figure 3: Co-occurrence of keywords in the literature

Figure 4 represents the temporal distribution of keyword occurrences, with yellower hues indicating more recent discussions. This shows that over time, there is an increased emphasis on the application and practical effects of certain technologies, with ‘implementation’, ‘construction site’ and ‘effectiveness’ being the latest topics to gain traction.

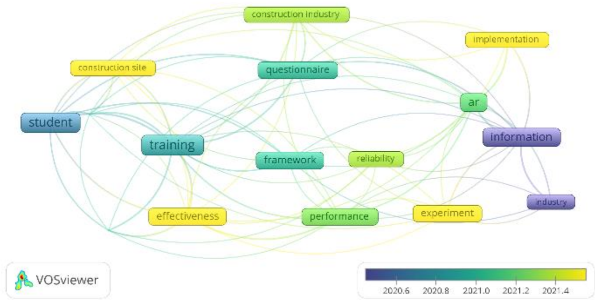


Figure 4: Timeline of keywords

Figure 5 focuses on the network of terms associated with 'immersive virtual reality'. This network shows connections to 'effectiveness', 'performance', and 'training', indicating that immersive virtual reality is primarily discussed in the context of its efficacy and performance in training scenarios within the construction sector.

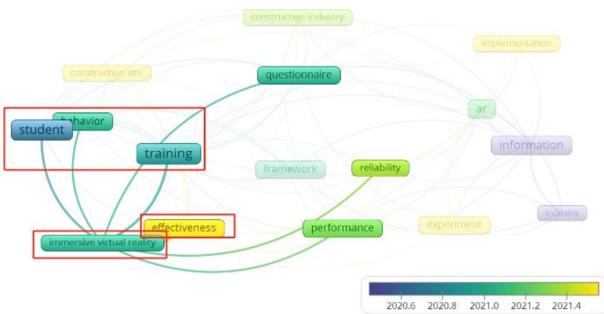


Figure 5: The main network of immersive virtual reality

Figure 6 depicts the relationships surrounding 'AR' technology. Here, 'AR' is linked to 'implementation', 'information', and 'experiment', suggesting that the current discourse is concerned with how AR is implemented in practice, the information processing capabilities of AR, and its experimental application in the field.

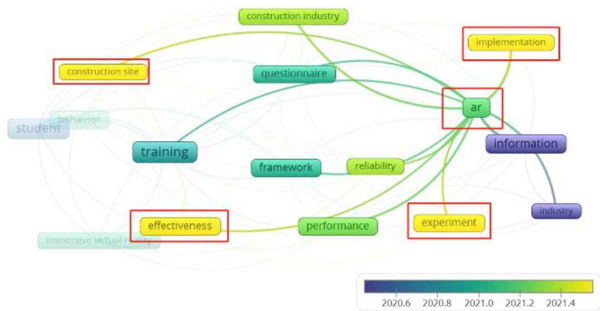


Figure 6: The main network of AR

Finally, Figure 7 reveals the density of keyword occurrences within the analysed papers, with the most frequently appearing terms encased in red. 'Student', 'training', 'AR', and 'information' appear as the most prominent terms, indicating these are central to the current research landscape within the selected body of literature.

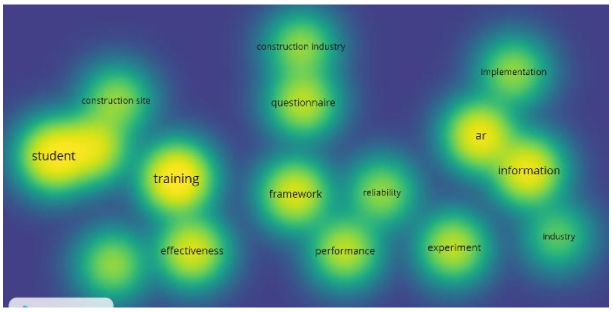


Figure 7: Heatmap of keywords

In conclusion, after analysing the key information of 24 articles, it is found that people are focused on the application of immersive virtual reality and AR technology in training and information dissemination. Density analysis highlights the centrality of educational applications and technology implementation as key areas of current research interest. Therefore, in the subsequent Narrative Literature Review, this study will focus on these contents.

Discussion and Result Analysis

The results of this systematic literature review illuminate the multifaceted roles that immersive technologies, specifically AR and VR play in the field of construction computing. The analysis, grounded in a review of 24 scholarly articles, reveals three pivotal areas: current developments, technological challenges, alongside future development shown in Table 1.

Table 1: Analysis result index table

Current Developments	Immersive technology applications	Education and Training	(Huang et al., 2020; Kim et al., 2020; Lucas, 2020; Sedlák et al., 2022; Sepasgozar, 2022)
		Auxiliary construction	(Sakib, Chaspari and Behzadan, 2021; Gath-Morad et al., 2022)
		Operation, Maintenance and Safety	(Maharjan et al., 2021; Raimbaud et al., 2021)
	Breakthrough cases	VR training to construction robots	(Huang, Zhu and Zou, 2023)
		VR remote control robot	(Wang et al., 2021)
Technological Challenges	Visual Challenges	360-degree panoramas have limited quality	(Ahn, Han and Al-Hussein, 2019; Eiris, Wen and Gheisari, 2022)
		High latency and slow rendering	(Hamzeh et al., 2019; Mutis and Ambekar, 2020; Zhou, Zhu and Du, 2020; Fu et al., 2023)
	Precision Challenge	Positioning accuracy	(Mutis and Ambekar, 2020)
		Location tracking accuracy	(Mutis and Ambekar, 2020)
		Image recognition and collection accuracy	(Wu et al., 2020; Zhou, Zhu and Du, 2020; Wang et al., 2022)
	Human experience	Lack of comfort, safety and reliability	(Maharjan et al., 2021; Raimbaud et al., 2021)
Future Developments	Technological Development	Optimization algorithms	(Zhou, Zhu and Du, 2020; Jahanshahloo and Ebrahimi, 2022; Zhang et al., 2023)
		Customer-specific mixed reality systems	(Wu et al., 2019)
		AR should be extended to other procedure	(Ahn, Han and Al-Hussein, 2019)
	Ecosystem Development	The government makes policy or compels it	(Shahzad et al., 2022; Arowoiya et al., 2023)
		Workers and operators involved in experiments	(Arowoiya et al., 2023)

Current Developments/Applications

1) Immersive technology applications

The adoption of VR and AR in the realm of education and training within the construction industry has been significant. Huang et al. (2020) and Kim et al. (2020) have demonstrated that VR can greatly enhance the training process by simulating realistic construction environments, allowing trainees to practice without the physical risks associated with construction sites. Lucas (2020) and Sedlák et al. (2022) further support this, noting that such immersive experiences lead to improved retention of information and a better grasp of complex structural designs. Moreover, Sepasgozar et al. (2022) emphasize the increased engagement and interaction that VR and AR technologies foster, which are essential for effective learning.

Meanwhile, in the auxiliary construction phase, the application of immersive technologies has been recognized for its contribution to preconstruction visualization and design verification. Sakib, Chaspari, and Behzadan (2021) discuss the utilization of AR for overlaying digital models on physical spaces, thus providing an immediate sense of scale and relation that 2D plans cannot. Gath-Morad et al. (2022) add to this by identifying how these technologies assist in identifying design conflicts before construction begins, potentially reducing costly errors and delays.

Moreover, addressing operation, maintenance, and safety, Maharjan et al. (2021) have found that AR applications improve maintenance tasks by enabling operators to visualize machinery components and operational data in real time. Raimbaud et al. (2021) highlight the impact of VR in safety training, where workers can be exposed to hazardous scenarios in a controlled virtual environment, thereby enhancing safety awareness without actual risk.

2) Breakthrough cases

The training of construction robots using VR is a significant advancement as well, as noted by Huang, Zhu, and Zou (2023). They provide insights into how robots can be programmed to perform complex tasks in a virtual setting before being deployed on-site, which could lead to increased precision and efficiency in construction activities. In addition, Wang et al. (2021) have introduced the concept of remotely controlling construction robots via VR, which allows operators to manipulate equipment from a distance. This application not only improves safety by reducing human presence on-site but also enhances the precision of tasks performed by the robots.

Technological challenges

1) Visual challenges

Visual fidelity is a significant concern in the deployment of immersive technologies. Studies by Ahn, Han and Al-Hussein (2019), Eiris, Wen and Gheisari (2022) emphasize that despite the advancements, 360-degree panoramas in VR still face limitations in quality, which can detract from the user experience and the intended realism. The visual challenges extend to the resolution and rendering speeds, as noted by Mutis and Ambékar (2020), where high latency and slow rendering can impact the effectiveness of training and operational simulations in VR environments.

2) Precision challenges

Precision in immersive environments is crucial, particularly in construction applications where spatial accuracy is paramount. Mutis and Ambékar (2020) discuss the challenges in positioning accuracy, where even minor discrepancies can lead to significant errors in the field. Similarly, location tracking accuracy is essential for AR applications where digital information must align precisely with physical elements. Wu et al. (2020), Zhou, Zhu and Du (2020), and Wang et al. (2022) highlight the need for improved image recognition and data collection accuracy to ensure that AR and VR systems can be reliably used for complex construction tasks.

3) Human experience challenges

Adoption of any new technology comes with the need to address human-centric concerns. Maharjan et al. (2021) and Raimbaud et al. (2021) have identified a lack of comfort, safety, and reliability as barriers to the widespread adoption of AR and VR in construction. The physical discomfort associated with prolonged use of headsets, the psychological unease some users feel in virtual environments, and the reliability of the systems in delivering consistent experiences are areas that require further research and development.

4) Other challenges

Eiris (2022) presents a challenging case where collaborative student behaviours did not directly lead to successful problem resolution, signifying that collaborative acts must be preceded by effective problem-solving actions. On a more positive note, Fu (2023) acknowledges data security challenges in the evolving metaverse but suggests that the advancement of blockchain and smart networking technologies holds promise for establishing a trustworthy and interactive digital environment. Wu's (2019) research offers a success story, indicating that despite a lack of expertise, students, as novices, can exhibit behaviour patterns akin to professionals and achieve comparable design review results through VR and MR models, showcasing the acceptability of these technologies in training. Furthermore, Huang (2020) documents the success of using VR to create a safe, cost-effective, and sustainable welding skills learning environment, which has been positively received by most students. These examples highlight the varying outcomes and acceptance of immersive technologies within educational settings.

Limitations of the study

This study establishes a profound understanding of the current research and practice regarding the application and challenges of immersive technologies in construction computing. However, the scope and methodology choices bring certain limitations. Firstly, although this paper systematically reviews literature from 2019 to 2023 to capture the latest trends, this time frame might have limited our consideration of earlier research outcomes, which are crucial for understanding the long-term development trajectory of immersive technologies. Secondly, the language limitation in the literature selection process (including only English literature) might have excluded significant studies in other languages that could offer different perspectives and insights.

Moreover, this study focuses primarily on AR and VR technologies and does not fully explore the potential of related technologies such as Mixed Reality (MR) and Extended Reality (XR), which are also showing increasing application prospects in the construction field. In terms of case analysis, despite efforts to select a diverse range of successful and unsuccessful cases, the quality and quantity of publicly available information might have not fully revealed the complexity and multidimensional factors behind these cases.

In discussing technological challenges, the paper focuses on visual challenges and accuracy issues, possibly without fully considering the organizational and management challenges encountered when implementing immersive technologies, such as difficulties in cross-departmental collaboration, differences in technology acceptance, and the impact on existing workflows. Furthermore, although the paper attempts to propose directions for future research and industry applications, these suggestions may be limited by the current research perspective and depth. Future work should further explore how to effectively address these challenges and achieve widespread application of immersive technologies in the construction industry.

Future developments

1) Technological Development

Future forecasts suggest an emphasis on the development of sophisticated optimization algorithms. Zhou, Zhu, and Du (2020), Jahanshahloo and Ebrahimi (2022), and Zhang et al. (2023) indicated the potential for these algorithms to enhance the performance of AR and VR applications in construction. Such advancements are anticipated to streamline data processing, improve rendering times, and provide more accurate simulations, thereby increasing the reliability and effectiveness of immersive technologies.

Furthermore, Wu et al. (2019) envision the customization of mixed reality systems to meet specific client needs. These systems are expected to integrate seamlessly with individual construction projects, providing tailored solutions that align with the unique requirements and challenges of each project. The focus will likely be on creating adaptive systems that can cater to varying scales and complexities of construction tasks.

In addition, the scope of AR is expected to broaden, as noted by Ahn, Han, and Al-Hussein (2019), to encompass other construction-related procedures beyond the current applications. This could include tasks like real-time project monitoring, on-site worker assistance, and integration with Internet of Things (IoT) devices to provide a holistic and interactive construction management ecosystem.

2) Ecosystem Development

Shahzad et al. (2022) and Arowoiya et al. (2023) suggest that government policy and regulatory frameworks will likely play a critical role in the adoption of immersive technologies in construction. The development of standards and policies can encourage the use of AR and VR, ensure data privacy, and promote safety standards in their application. Besides, the future development of the immersive technology ecosystem in construction is also predicted to involve greater participation from the workforce. Arowoiya et al. (2023) emphasize the importance of engaging construction workers and operators in the development and testing of AR and VR systems. This inclusion is crucial to ensure that the ecosystem developed are user-friendly, meet the practical needs of the users, and are effectively adopted in construction practices.

Conclusions

This study highlights the transformative potential of AR and VR in construction computing, alongside the need for targeted advancements and strategic implementations for their full realization. These immersive technologies promise to revolutionize construction through various applications, offering unparalleled opportunities to enhance training, safety, operational efficiency, and design verification by simulating realistic environments and integrating digital models into physical spaces. Addressing the challenges identified requires a multifaceted approach in future research. Key strategies include establishing a cross-disciplinary framework for collaboration across fields like computer science, architectural design, and human-computer interaction to improve user experience and technology integration. Additionally, developing quantitative evaluation models is crucial for assessing immersive technologies' impact on construction projects, paving the way for their broader adoption.

Understanding the construction industry's technology acceptance and adaptability is vital, with focus areas including training needs, interface design, and perceived utility to foster user-friendly applications. Research should also explore interoperability solutions between AR/VR and digital tools like BIM, cloud computing, and IoT to enhance data integration and project management. For industry implementation, creating industry standards and policies in collaboration with associations and government bodies is essential for the regulated use of immersive technologies. Public-private partnerships and professional training programs are also pivotal in promoting AR and VR's widespread application in construction.

In conclusion, realizing the full potential of AR and VR in transforming construction computing requires overcoming technological and human-centric hurdles. Future efforts should concentrate on tailored solutions for bridging theoretical potential with practical application, including cross-disciplinary collaboration, technology acceptance studies, and advanced interoperability solutions. Simultaneously, supportive policies, partnerships, and training programs will facilitate immersive technologies' adoption, significantly impacting the construction industry.

Acknowledgement:

Thanks to Global Engagement Fund- UCL-ZJU Strategic Partner Fund 2023-2024, which strongly supports our research in 2023-2024. This work is also supported by the Horizon Europe project, INHERIT (Project No: 101123326).

References

Ahn, S., Han, S. and Al-Hussein, M. (2019). '2D Drawing Visualization Framework for Applying Projection-Based Augmented Reality in a Panelized Construction Manufacturing Facility: Proof of Concept'. *Journal of Computing in Civil Engineering*, 33 (5). doi: 10.1061/(ASCE)CP.1943-5487.0000843.

Arowoiya, V., Oke, A., Akanni, P., Kwofie, T. and Enih, P. (2023). 'Augmented reality for construction revolution - analysis of critical success factors'. *INTERNATIONAL JOURNAL OF CONSTRUCTION MANAGEMENT*, 23 (11), pp. 1867–1874. doi: 10.1080/15623599.2021.2017542.

Chan, K. C. G. (2012). 'The use of virtual reality for visualizing construction safety management process'. Hong Kong Polytechnic University. Available at: https://theses.lib.polyu.edu.hk/handle/200/6583 (Accessed: 29 November 2023).

Delgado, J. M. D., Oyedele, L., Demian, P. and Beach, T. (2020). 'A research agenda for augmented and virtual reality in architecture, engineering and construction'. *Advanced Engineering Informatics*. Elsevier, 45, p. 101122.

Eiris, R., Wen, J. and Gheisari, M. (2022). 'iVisit-Collaborate: Collaborative problem-solving in multiuser 360-degree panoramic site visits'. *Computers and Education*, 177. doi: 10.1016/j.compedu.2021.104365.

Fu, Y., Li, C., Yu, F. R., Luan, T. H., Zhao, P. and Liu, S. (2023). 'A Survey of Blockchain and Intelligent Networking for the Metaverse'. *IEEE Internet of Things Journal*, 10 (4), pp. 3587–3610. doi: 10.1109/JIOT.2022.3222521.

Gath-Morad, M., Melgar, L., Conroy-Dalton, R. and Hölscher, C. (2022). 'Beyond the shortest-path: Towards cognitive occupancy modeling in BIM'. *AUTOMATION IN CONSTRUCTION*, 135. doi: 10.1016/j.autcon.2022.104131.

Hamzeh, F., Abou-Ibrahim, H., Daou, A., Faloughi, M. and Kawwa, N. (2019). '3D visualization techniques in the AEC industry: The possible uses of holography'. *Journal of Information Technology in Construction*, 24, pp. 239–255. doi: 10.36680/j.itcon.2019.013.

Huang, C., Lou, S., Cheng, Y. and Chung, C. (2020). 'Research on Teaching a Welding Implementation Course Assisted by Sustainable Virtual Reality Technology'. *SUSTAINABILITY*, 12 (23). doi: 10.3390/su122310044.

Huang, L., Zhu, Z. and Zou, Z. (2023). 'To imitate or not to imitate: Boosting reinforcement learning-based construction robotic control for long-horizon tasks using virtual demonstrations'. *AUTOMATION IN CONSTRUCTION*, 146. doi: 10.1016/j.autcon.2022.104691.

Jahanshahloo, A. and Ebrahimi, A. (2022). 'Reconstruction of the initial curve from a two-dimensional shape for the B-spline curve fitting'. *EUROPEAN PHYSICAL JOURNAL PLUS*, 137 (3). doi: 10.1140/epjp/s13360-022-02604-y.

Kim, K., Oertel, C., Dobricki, M., Olsen, J., Coppi, A., Cattaneo, A. and Dillenbourg, P. (2020). 'Using immersive virtual reality to support designing skills in vocational education'. *BRITISH JOURNAL OF EDUCATIONAL TECHNOLOGY*, 51 (6), pp. 2199–2213. doi: 10.1111/bjet.13026.

Lele, A. (2013). 'Virtual reality and its military utility'. *Journal of Ambient Intelligence and Humanized Computing*. Springer, 4, pp. 17–26.

Lenoir, T. (2000). 'All but war is simulation: The military-entertainment complex'. *Configurations*. Johns Hopkins University Press, 8 (3), pp. 289–335.

Li, W., Nee, A. Y. and Ong, S.-K. (2017). 'A state-of-the-art review of augmented reality in engineering analysis and simulation'. *Multimodal Technologies and Interaction*. MDPI, 1 (3), p. 17.

Li, X., Yi, W., Chi, H.-L., Wang, X. and Chan, A. P. (2018). 'A critical review of virtual and augmented reality (VR/AR) applications in construction safety'. *Automation in Construction*. Elsevier, 86, pp. 150–162.

Lucas, J. (2020). 'Rapid development of virtual reality based construction sequence simulations: A case study'. *Journal of Information Technology in Construction*, 25, pp. 72–86. doi: 10.36680/j.itcon.2020.004.

Maharjan, D., Agüero, M., Mascarenas, D., Fierro, R. and Moreu, F. (2021). 'Enabling human-infrastructure interfaces for inspection using augmented reality'. *STRUCTURAL HEALTH MONITORING-AN INTERNATIONAL JOURNAL*, 20 (4), pp. 1980–1996. doi: 10.1177/1475921720977017.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and PRISMA Group. (2009). 'Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement'. *PLoS medicine*, 6 (7), p. e1000097. doi: 10.1371/journal.pmed.1000097.

Mutis, I. and Ambekar, A. (2020). 'Challenges and enablers of augmented reality technology for in situ walkthrough applications'. *Journal of Information Technology in Construction*, 25, pp. 55–71. doi: 10.36680/j.itcon.2020.003.

Niu, S., Pan, W. and Zhao, Y. (2016). 'A virtual reality integrated design approach to improving occupancy information integrity for closing the building energy performance gap'. *Sustainable cities and society*. Elsevier, 27, pp. 275–286.

Okoli, C. and Schabram, K. (2010). 'A guide to conducting a systematic literature review of information systems research'.

Prabhakaran, A., Mahamadu, A.-M. and Mahdjoubi, L. (2022). ‘Understanding the challenges of immersive technology use in the architecture and construction industry: A systematic review’. *Automation in Construction*, 137, p. 104228.

Raimbaud, P., Lou, R., Danglade, F., Figueroa, P., Hernandez, J. and Merienne, F. (2021). ‘A Task-Centred Methodology to Evaluate the Design of Virtual Reality User Interactions: A Case Study on Hazard Identification’. *BUILDINGS*, 11 (7). doi: 10.3390/buildings11070277.

Safikhani, S., Keller, S., Schweiger, G. and Pirker, J. (2022). ‘Immersive virtual reality for extending the potential of building information modeling in architecture, engineering, and construction sector: systematic review’. *International Journal of Digital Earth*, 15 (1), pp. 503–526. doi: 10.1080/17538947.2022.2038291.

Sakib, M., Chaspari, T. and Behzadan, A. (2021). ‘Physiological Data Models to Understand the Effectiveness of Drone Operation Training in Immersive Virtual Reality’. *JOURNAL OF COMPUTING IN CIVIL ENGINEERING*, 35 (1). doi: 10.1061/(ASCE)CP.1943-5487.0000941.

Sedlák, M., Šašinka, Č., Stachoň, Z., Chmelík, J. and Doležal, M. (2022). ‘Collaborative and individual learning of geography in immersive virtual reality: An effectiveness study’. *PLOS ONE*. Edited by S. Triberti, 17 (10), p. e0276267. doi: 10.1371/journal.pone.0276267.

Sepasgozar, S. M. E. (2022). ‘Immersive on-the-job training module development and modeling users’ behavior using parametric multi-group analysis: A modified educational technology acceptance model’. *Technology in Society*, 68, p. 101921. doi: 10.1016/j.techsoc.2022.101921.

Seth, A., Vance, J. M. and Oliver, J. H. (2011). ‘Virtual reality for assembly methods prototyping: a review’. *Virtual reality*. Springer, 15, pp. 5–20.

Shahzad, M., Shafiq, M. T., Douglas, D. and Kassem, M. (2022). ‘Digital Twins in Built Environments: An Investigation of the Characteristics, Applications, and Challenges’. *Buildings*, 12 (2). doi: 10.3390/buildings12020120.

Shouman, B., Othman, A. A. E. and Marzouk, M. (2022). ‘Enhancing users involvement in architectural design using mobile augmented reality’. *Engineering, Construction and Architectural Management*. Emerald Publishing Limited, 29 (6), pp. 2514–2534.

Tranfield, D., Denyer, D. and Smart, P. (2003). ‘Towards a methodology for developing evidence-informed management knowledge by means of systematic review’. *British journal of management*. Wiley Online Library, 14 (3), pp. 207–222.

Wang, D., Wang, X., Ren, B., Wang, J., Zeng, T., Kang, D. and Wang, G. (2022). ‘Vision-Based Productivity Analysis of Cable Crane Transportation Using Augmented Reality-Based Synthetic Image’. *JOURNAL OF COMPUTING IN CIVIL ENGINEERING*, 36 (1). doi: 10.1061/(ASCE)CP.1943-5487.0000994.

Wang, P., Wu, P., Wang, J., Chi, H.-L. and Wang, X. (2018). ‘A critical review of the use of virtual reality in construction engineering education and training’. *International journal of environmental research and public health*. MDPI, 15 (6), p. 1204

Wang, X., Liang, C., Menassa, C. and Kamat, V. (2021). ‘Interactive and Immersive Process-Level Digital Twin for Collaborative Human-Robot Construction Work’. *JOURNAL OF COMPUTING IN CIVIL ENGINEERING*, 35 (6). doi: 10.1061/(ASCE)CP.1943-5487.0000988.

White, J., Schmidt, D. C. and Golparvar-Fard, M. (2014). ‘Applications of augmented reality’. *Proceedings of the IEEE*, 102 (2), pp. 120–123.

Wu, W., Hartless, J., Tesei, A., Gunji, V., Ayer, S. and London, J. (2019). ‘Design Assessment in Virtual and Mixed Reality Environments: Comparison of Novices and Experts’. *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, 145 (9). doi: 10.1061/(ASCE)CO.1943-7862.0001683.

Wu, W., Sandoval, A., Gunji, V., Ayer, S., London, J., Perry, L., Patil, K. and Smith, K. (2020). ‘Comparing Traditional and Mixed Reality-Facilitated Apprenticeship Learning in a Wood-Frame Construction Lab’. *JOURNAL OF CONSTRUCTION ENGINEERING AND MANAGEMENT*, 146 (12). doi: 10.1061/(ASCE)CO.1943-7862.0001945.

Zhang, Q., Li, C., Huang, Y. and Luo, Y. (2023). ‘Effective multi-controller management and adaptive service deployment strategy in multi-access edge computing environment’. *Ad Hoc Networks*, 138, p. 103020. doi: 10.1016/j.adhoc.2022.103020.

Zhou, T., Zhu, Q. and Du, J. (2020). ‘Intuitive robot teleoperation for civil engineering operations with virtual reality and deep learning scene reconstruction’. *Advanced Engineering Informatics*, 46. doi: 10.1016/j.aei.2020.101170.

EVALUATING MULTIUSER AUGMENTED REALITY FOR 4D CONSTRUCTION PLAN REVIEW

Saleh Alghamdi[1,2], John Messner[1], and Robert Leicht[1]
[1]The Pennsylvania State University, University Park, USA
[2]King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia

Abstract

Construction plan review meetings benefit from visualizing 4D construction models using various media such as Immersive Technologies (ImT). However, ImT, such as Augmented Reality (AR), have not been thoroughly investigated for their effects on collaborative reviews. This paper aims to evaluate the effect of using a developed 4D Multiuser AR (4DMAR) application on task performance using a comparative study against a screen-based environment. A questionnaire and observational data were used to measure the participants' experience and preferences. Sixty-two students found the interaction in both environments easy but perceived 4DMAR as more interactive and suitable for collaborative plan reviews.

Introduction

Construction projects are highly complex and dynamic and involve many interdependencies between activities (Sterman, 1992). In addition, almost every construction project is unique in design, materials, construction methods, and the surrounding environment. Therefore, many challenges exist during the construction process that take time to identify in the design process. For instance, planning the site logistics layout, especially in congested areas, can be difficult without simulating the construction process.

Incorporating visualization techniques to enhance the planning process for a construction project can be highly beneficial. As an example, Guo et al. (2017) found that visualization can enhance the identification of safety hazards before and during the construction phases of a project. It also increases our efficiency in dealing with complicated issues requiring high cognitive levels (Johnson, 1998). The most appropriate visualization method will depend on several factors, including the scope and depth of the project, the experience levels of the team members involved in the planning process, and the available resources. By carefully considering the unique needs of each project, it is possible to identify the most effective visualization methods for enhancing planning and achieving the required results.

While traditional 2D drawings and 3D models can provide a valuable visual representation of aspects of the construction process, they may not always be the most efficient way to navigate the sequence of construction activities. Adding time as another dimension in a 4D model enables the visualization of the construction sequence. Visualizing the construction sequence can provide insights into unforeseen problems and the site layout, considering the transportation cost (Bortolini et al., 2019). Hartmann & Fischer (2007) provided a case study where 3D/4D models support constructability reviews and convey general information to all stakeholders. The models were used to convey the details of the construction plan to all stakeholders, even those who are not engineers, and it is one of the examples of using 4D models in the construction industry. 4D modeling can also yield better coordination between disciplines (Staub-French & Khanzode, 2007).

Immersion affects how people interact with the environment, including increased attention (Silva et al., 2017). One of the most promising options to improve construction project planning is to use ImT, such as AR (Davila Delgado et al., 2020). This can be especially valuable in complex construction projects with many stakeholders with diverse perspectives and interests. Using AR can improve experts understanding of the design and construction of a project (Meža et al., 2015). Virtual Reality (VR) can visualize 4D construction plans and support constructability analyses (Haymaker & Fischer, 2001; Al-Adhami et al., 2018).

Immersive technologies offer new possibilities for multi-user interactive visualization to review construction plans. However, collaborative reviews in AR have not been thoroughly studied. This study aims to assess the impact of immersive, multi-user AR on task performance when reviewing 4D construction plans. To accomplish this, several objectives were pursued. First, a multi-user AR solution, termed 4DMAR, was developed to visualize 4D models through head-mounted AR headsets. Subsequently, a quasi-experiment was conducted to evaluate the effectiveness of the 4DMAR solution in facilitating multi-user construction plan reviews. This experiment involved conducting interactive multi-user sessions and employing various data collection methods. The study involved 62 participants who completed a questionnaire to collect perceptual data on their overall experience and preferences along with the collected observational data.

Related Works

Several papers that have presented an AR construction application were identified (see Table 1). In general, two different types of AR displays were used, including screen-based, such as using the camera on a tablet PC and projecting augmented information on the screen, and head-mounted, where augmented content was directly placed in the user's view of the space. Very few examples of head-mounted AR were identified in the literature.

Lin et al. (2020) developed a prototype for using screen-based AR and 4D modeling to enhance the operations of a tower crane. The paper did not provide an assessment of the usability of this prototype. However, it demonstrated how AR could translate 4D model information into showing the lifting path sequence. Lin et al. (2019) developed another prototype for construction progress monitoring using AR to show the construction sequence.

Another early study demonstrated a detailed approach to visualizing construction activities using AR and GPS (Behzadan & Kamat, 2013). This research provided a technical evaluation and demonstration of how a platform could be developed to visualize objects in their future physical location. The paper showed the possibility of a multiuser interactive AR construction model. Wang et al. (2014) used a 4D AR model to discuss construction issues through multiple stakeholders in a collaborative environment. The model helped visualize a natural gas plant's complex construction. The presented model was interactive and provided insights into safety issues. One of the comparative examples is the one provided by Kim et al. (2012). They provided an interactive multiuser and immersive AR tool to visualize equipment operation on the construction site. The tool provides multiple pieces of equipment to visualize different scenarios. Lin et al. (2015) provided an example of using screen AR for site plan review in a multiuser environment. The paper evaluates the use of AR using the time spent reviewing the plan. Also, it provides some perceptual feedback using a user-experience questionnaire.

Different evaluation approaches have been used to consider the impact of using AR/VR in design reviews. Most studies used questionnaires focusing on the user experience regarding comfort or how they enjoyed the experience. For example, Amrollahibuki (2019) used a questionnaire to evaluate the user's overall experience. One of the assessment methods that can be used is to compare the learning outcomes between the use of immersive technologies and traditional environments. A similar example was provided by Wolfartsberger (2019), where added flaws to the design were used to evaluate the VR value. However, this application was specific to electric equipment (power units), and flaws were quantified based on the design elements being reviewed, such as ergonomic or design issues. In the construction context, this can be difficult to quantify and categorize due to the lack of design standardization, which makes this method difficult to apply. Wang et al. (2022) used a weighted evaluation metric to control an assembly process's AR experience. For example, the presented metric can evaluate how the users felt about manipulating the objects.

Two gaps were identified throughout the literature. First, many previous examples provided concepts or prototypes of using AR for construction planning; however, these were often confined to proof-of-concept implementations due to technological limitations of early AR devices like poor computing performance, inaccurate synchronized alignment of objects, and lack of multi-user capabilities. With recent advancements, building fully functional multi-user AR environments is viable. These technologies include advanced alignment methods that use image recognition and the increased performance of AR/VR headsets to handle larger amounts of information. Second, most previous examples were limited to feedback about user experience, with limited applications of an objective evaluation. Evaluating the methods is essential to consider this technology construction planning. Therefore, a comprehensive example of using AR as a review environment for collaborative work was evaluated in this experiment in the context of construction planning.

Developing A 4D Multi-User Augmented Reality (4DMAR) Visualization System

The 4D Multiuser Augmented Reality (4DMAR) system was developed following the framework in Table 2. Some elements in the framework, such as implementing

Table 1: Research Trend Summary

Source	Technology	Application	Assessment	Multiuser
Lin et al. (2020)	Screen AR (Markerless)	Equipment Operation Monitoring		
Lin et al. (2019).	Screen AR (Markerless)	Construction Progress Monitoring		
Behzadan & Kamat (2013)	Head Mounted AR (Marker)	Construction Visualization		✓
Wang et al. (2014)	Screen AR (Marker)	Construction Control		
Kim et al. (2012)	Head Mounted AR (Marker)	Equipment Operation		✓
Lin et al. (2015)	Screen AR (Markerless)	Site Plan Review	✓	✓

Autodesk Navisworks® and Autodesk 3ds Max®, are adopted from Boton (2018).

The primary use case was to enable multi-user, real-time visualization and interaction with 4D models in an augmented reality context. Key functional requirements included multi-user synchronization of the 4D model state, navigation of the model sequence using a time slider, and physical exploration of the model. The system architecture centered around utilizing the Autodesk Navisworks® platform for 4D building modeling capabilities, the Unity game engine for multiuser networking and augmented reality functionalities, and the Meta Quest Pro headsets as the AR delivery devices. Autodesk Navisworks® was selected for its robust 4D modeling toolset and integration with common BIM authoring tools. Unity provided a framework for spatial computing applications along with its multi-platform deployability. The Meta headsets offered a powerful option for an AR multiuser experience using the passthrough functionalities.

The system development started with creating the 4D model using Autodesk Navisworks®. Next, the model geometry was optimized using Autodesk 3ds Max® to reduce the polygon count. The next three steps were performed to develop the interaction in the Unity Realtime Development Platform, and then the app was deployed to the Meta Quest Pro.

Table 2 AR Experience Development Process

Steps	Resources	Description
1. Build the 4D model	• Autodesk Navisworks®	*Autodesk Navisworks® was used to build the 4D model.*
2. Optimize the 3D models	• Autodesk 3ds Max®	*The geometry of the 3D models was optimized for enhanced performance in the AR view.*
3. Customize the AR environment	• Unity Realtime Development Platform Oculus Package Manager	*The AR experience was enabled using the passthrough feature. The alignment of the model for multiuser was developed using a Shared Spatial Anchor (SSA)*
4. Create interactivity features	• Unity Realtime Development Platform • Oculus Package Manager	*Interaction features and User Interface (UI) were developed using Unity and Oculus UI.*
5. Build the multiuser platform	• Unity Realtime Development Platform Photon Unity Networking	*Photon Unity Networking was used to create rooms for multi-users.*
6. Deploy the app	• Meta Quest Pro	*The app was deployed on multiple headsets.*

A 4D model of a real constructed project with construction additional flaws added was developed. The purpose of adding flaws was to enable conversation elements and determine whether those flaws were noticed and in which review environment. The building used for this activity is a service and office building. The building consists of three floors and a basement. Only the structural and architectural models were used in the 4D model. The excavation tasks were included, and some exterior scaffolding for the masonry facade was added. Fifty-seven activities were represented in the 4D model. Figure 1 illustrates four major stages of the project, from excavation to completion. The first image on the top left shows the excavation activities of the basement and the footings. The top right image illustrates the finished basement, some steel erection activities, and the start of one concrete floor activities. The bottom left image shows the façade activities with the scaffolding highlighted in yellow, and the last image shows the completed project.

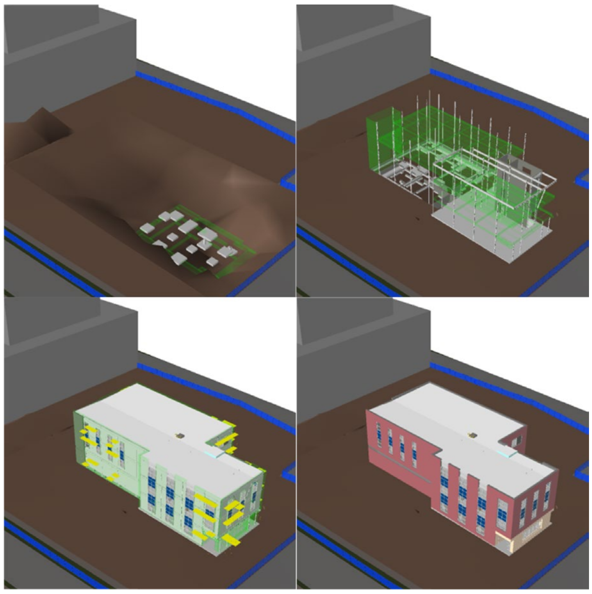


Figure 1: 4D Model Major Stages

The 3D geometry of the model was optimized to increase the system's performance by reducing the number of polygons. Reducing the number of polygons in a scene can increase the Frames Per Second (FPS) for displaying content in the AR head-mounted display. The Meta Quest Pro headsets perform better within a certain number of triangle counts or polygons, which are around 750k-1m triangles, to allow acceptable FPS as recommended by Oculus (2022). Oculus suggests that a minimum of 72 FPS should be targeted for optimal content viewing. Therefore, the model was optimized to increase the system's performance by reducing the number of polygons using tools incorporated into Autodesk 3ds Max®. However, this does cause quality reductions in the model for visualization, in which some elements lost some features like missing or reduced quality of a geometry. Because the model size was small when viewed in the physical environment, this issue was insignificant and did not affect the visualization of the building elements.

Customizing the environment was performed by adjusting the Unity Realtime Development Platform settings making it compatible with the Meta Quest Pro. The AR experience was enabled by using the passthrough video feature of the display. The alignment of the model for multiuser was developed using a Shared Spatial Anchor (SSA). Other alignment methods were considered, such as manual alignment using the controller. However, SSA provides a fast and relatively accurate geospatial positioning. The main interaction features added to the system were time navigation using a slider and moving the model using pinching by the controller side button. The slider was linked to an animation file that contains the construction schedule. However, using the animation file adds a limitation if there are future schedule changes, which should be addressed in the future development of the 4DMAR. The collaborative environment was built using the Photon Unity Networking Platform. The platform enables multiple users to enter the same rooms and share the same synchronized 4D model. The 4DMAR was tested on three different headsets simultaneously, two Meta Quest Pros and one Meta Quest 2, and it performed adequately. Figure 2 shows the built application of the AR review environment.

Figure 2: Examples of Users in the 4DMAR Application, each looking at a synchronized model instance

Assessment of the 4DMAR

To assess the performance of the 4DMAR, a quasi-experiment was performed. The assessment started with designing an experiment to reflect the assessment approach and a questionnaire to measure the participant's perception of the experience. Three pilot tests were performed to adjust and refine the experiment design.

The current practice used for reviewing 4D construction models in the industry is typically using a monitor(s) or a projector(s), referred to as a screen-based method in this study. Therefore, this research compared the 4DMAR as a review environment to the screen-based review, providing a shared resource in both options, see Figure 3. This led to the question: Can 4DMAR provide a suitable environment for reviewing 4D construction plans compared to screen-based reviews?

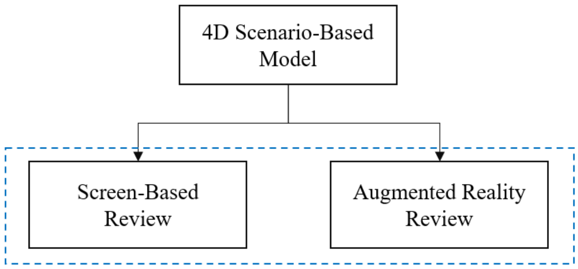


Figure 3: Evaluation Approach

The ability to review the 4D models was performed on two platforms: 1) the screen-based (SB) using Autodesk Navisworks® and 2) AR using 4DMAR. The laptop screen was rear-projected on a 6'x8' screen to ensure an appropriate view for a pair of participants. Similar interaction features, timeliner, and navigation were used in both environments. Also, to maintain the consistency of interactions between the two environments, the toolbar of Autodesk Navisworks® was hidden, and only the Timeliner was enabled in the screen-based 4D model review.

An experiment path was designed to test both environments, as seen in Figure 4. Since this is a multiuser activity, two participants performed each experiment. This was a deliberate choice to create a manageable multiuser environment while minimizing technological and data collection complexities that could arise with larger groups. To test whether the order of the experiment would affect the overall preference, another path was added, switching both environments, and it was annotated as path B. After each review session, a questionnaire was answered by both participants separately.

Each path started with a script that introduced the purpose of 4D modeling and what kind of flaws are usually found in a 4D model. It also introduced the design of the building being reviewed and how to interact with both environments. However, the script did not ask the participants to complete a specific task while reviewing the model and left the review open for participants based on the mentioned flaws.

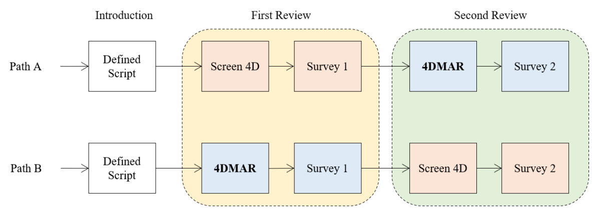


Figure 4: Experiment Paths

A questionnaire was designed to collect perceptual data. The questionnaire was answered individually after each review environment to address the second research question of whether the order of using the AR would affect the preference. The questionnaire was designed to test three factors: 1) overall experience, 2) interaction experience, and 3) overall preference between the two environments. Both usability and interactivity questions were rated on a 6-point Likert scale since participants will answer this for both environments and to make the comparison more distinctive. However, the overall preference includes both environments in the same questions, and a no-preference choice was added to reduce bias. More details are provided in Table 3.

Table 3: Questionnaire Design

Collected data	**Survey** 1	2	**Description**	**Question type**
Background information	✓		*Collected information includes sex at birth, industry experience, AR/VR experience, and Autodesk Navisworks® Experience. Collected as a subpart of survey 1 to minimize the number of surveys.*	*Open-ended questions and multiple choice*
Overall experience	✓	✓	*The rating included comfort, familiarity, enjoyment, ability to find flaws and opportunities, and need for technical support.*	*Likert scale Strongly agree – Strongly disagree.*
Interactivity	✓	✓	*Rating the interactivity experience of the interactive features, time navigation, model exploration, and ability to interact with a teammate.*	*Likert scale Easy – Hard*
Overall preference		✓	*The last part of Survey 2 asked about the overall preference*	*Likert scale AR – no preference (mid) – Screen-Based*

Figure 5 shows the two review environments' physical spaces. The blue boundary represents the guardian set for all headsets. During pilot testing, system functionality was assessed, along with clarity and adequacy of the questionnaire and the experiment flow. The pilot was performed on the apparatus of the experiment and how the flow of the other experiment was maintained.

Results and Discussion

Questionnaire Results

Sixty-two participants were involved in this experiment; 16 groups followed Path A, while 15 groups followed Path B. Consent was obtained from all experiment participants. Each group had two participants. The participants included 23 female and 39 male participants.

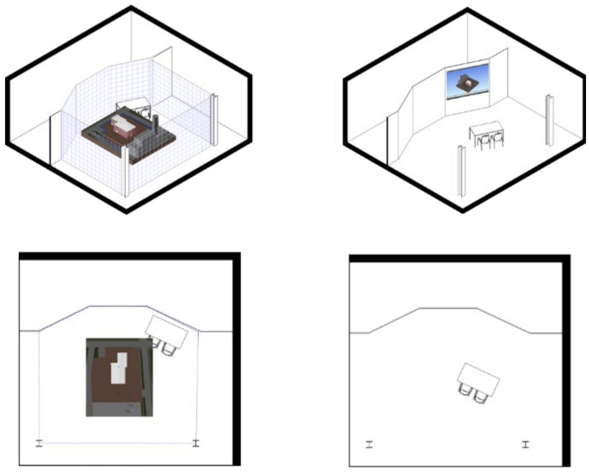


Figure 5: Review Environments Physical Space (left is 4DMAR and right is screen-based)

Forty-seven (76%) participants have had an internship experience in the construction industry. Most participants (82%) had limited experience using Autodesk Navisworks®, having used the software once or twice. In addition, most participants (71%) have experienced AR/VR headsets before, but a majority only experienced them once or twice.

Six questions were answered for both the SB and AR reviews to evaluate the user's overall experience with the review. All responses slightly favor AR as a review environment for the overall experience (see Figure 6).

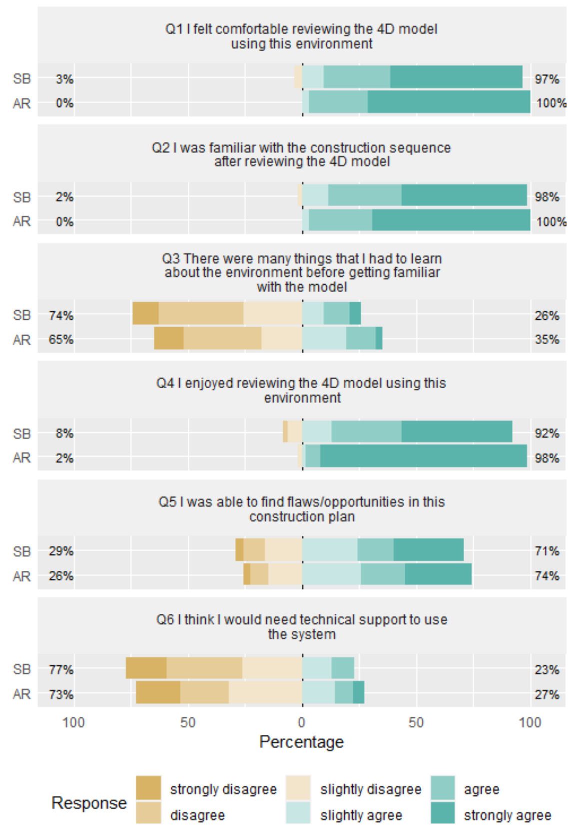


Figure 6: Overall Experience Response, Screen-Based (SB) top bar vs Augmented Reality (AR) bottom bar

Only one result showed a majority preferring AR: question 4, in which most participants strongly felt they enjoyed the AR reviews.

Three questions were used to rate the interactivity features in both environments. Similarly, the overall response slightly favors AR. Participants typically felt that it was relatively easy to interact with both environments. However, the questions were particular to the shared interaction features. See Figure 7 for a graphical representation of the Likert scale.

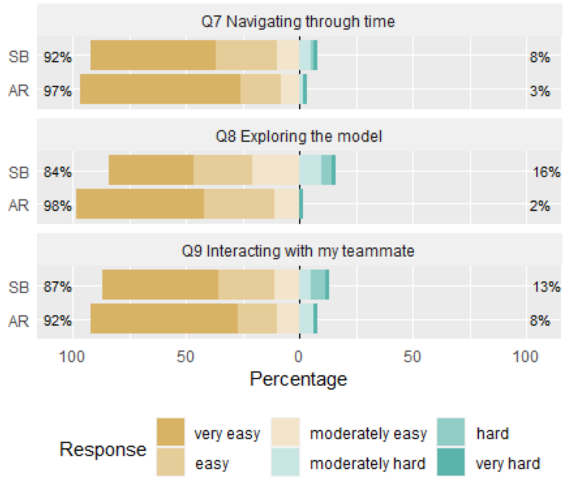
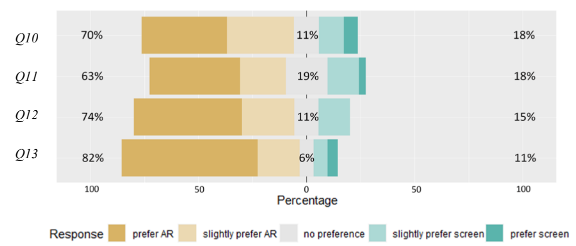


Figure 7: Interaction Response, Screen-Based (SB) top bar vs Augmented Reality (AR) bottom bar

Lastly, each participant answered four questions on their overall technology preference after using both environments (see Figure 8). This data is further analyzed in the next section, considering the experimental path. Participants indicated they would prefer AR to be used in their next review in Q10. They perceive AR as having more potential to be used in the future, as answered in Q11. Also, in Q12, they perceive AR as giving them a stronger sense of depth, spatial awareness, and realism. In the last question, which helps answer the collaboration question, 63% of participants felt AR is more suitable for collaborative reviews, 19% had no preference, and 18% preferred screen-based reviews.



Q10: Next time, I will review the 4D model using:
Q11: In your opinion, which environment has more potential for construction plan reviews in the future?
Q12: Which environment provided a stronger sense of depth, spatial awareness, and overall realism for the models being presented?
Q13: Which environment do you think is better suited for collaborative work and group activities?

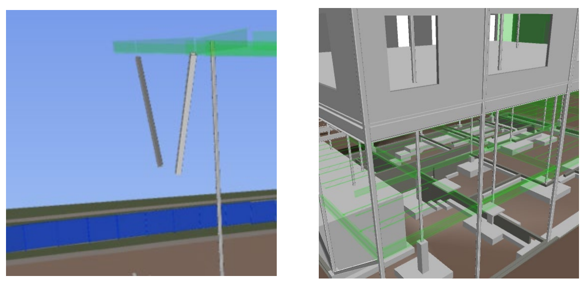
Figure 8: Preference Responses, Screen-Based (SB) vs. Augmented Reality (AR)

Observational Data

While performing the experiments, observational data were collected using video recordings. In addition, videos were analyzed to study participants' behavior using the two environments.

The experiment's script outlined the goal of the 4D modeling reviews, but there was no emphasis on participants recording any flaws or opportunities in the construction plan. This was intended to reduce the pressure on participants in a new environment regarding finding flaws. However, the performance of finding flaws was similar in both environments. Participants did not discuss any opportunities for improving the construction plan in either environment. It was also noticed that some participants assumed that the model was relatively accurate, although the experiment aimed to have participants critically review it. This might be related to their limited industry experience, lack of 4D modeling reviews, or lack of clear tasks.

Participants only identified flaws that were related to sequence activities. Only 13 of the 31 participating groups were able to identify sequencing flaws in the construction plan. These primarily focused on two sequencing issues that participants discussed, shown in Figure 9. The left (a) image shows a flaw in the sequence where two steel braces were installed ahead of the supporting members. The other flaw on the right (b) image shows the installation of the second floor before the start of the first-floor grade-beam install. The first flaw was found most frequently. The number of instances when participants found these flaws was similar for both 4DMAR and the screen-based. Seven groups identified at least one sequencing issue using the SB review, while six identified sequencing issues in the 4DMAR. When participants found a flaw in the environment, they usually did not discuss it again in the second environment. Whether there is an advantage to using the 4DMAR or SB cannot be demonstrated, but both can be used to identify flaws. Participants found no missing object flaws or flaws related to the model's construction methods in both environments. This was documented using the notes taken during the experiment when the participants discussed a sequencing issue. The participant sample consisted mainly of students with limited professional experience in construction planning and 4D modeling practices. Most

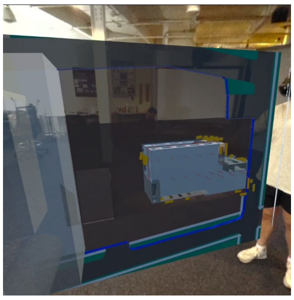


a) Steel structure sequence. b) Installation of the second-floor.

Figure 9: Sequencing Issues Found by Participants

had only used 4D authoring tools like Navisworks once or twice before. This lack of deep domain expertise may have caused many to assume the provided 4D model was largely accurate. While the experiment instructions mentioned that 4D models can reveal construction flaws, participants were not explicitly tasked to identify flaws or sequencing errors. Moving forward, we plan to enhance the experiment design by instructing participants to document the observed 4D model flaws or sequencing issues as part of the review exercise.

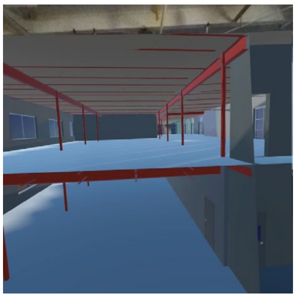
The 4DMAR has initiated different interactions compared to the screen-based. Figure 10 shows examples of the interaction methods unique to the 4DMAR, such as exploring the model's interior.



a) Rotating the model sideways.

b) Flipping the model upside down.

c) Going underneath the model.



d) Going inside the model.

Figure 10 Examples of the Interaction Methods

The primary research question investigated in this study was: Can 4DMAR provide a suitable environment for reviewing 4D construction plans compared to screen-based reviews? As indicated by the questionnaire results in Figure 8, 63% of participants felt AR was better suited for collaborative work and group activities compared to only 18% preferring the screen-based environment, while 19% had no preference. While task performance in identifying construction plan flaws was similar across both environments, the overall user preference for AR's interactive and immersive qualities supports its potential as an engaging platform. By counterbalancing the experiment paths (A and B), an order effect on preferences was not detected based on the collected data. This suggests that familiarity with the AR system did not significantly sway participants' judgments after experiencing both visualization methods.

Conclusions

Planning for construction sequencing is a complicated process that requires careful consideration. Visualizing the construction sequence is beneficial in determining possible challenges and opportunities in a construction plan, which heavily relies on the planners' experience. Collaborative environments that support model interaction and improved or novel visualization have previously shown benefits to the review process.

Participants generally preferred the 4DMAR regarding collaboration activities in the review quasi-experiments. Participants felt comfortable and enjoyed using the AR headsets and felt they could communicate with their teammates. The effect on task performance and collaboration was evaluated using an immersive, multi-user AR to review 4D construction models. Several objectives were achieved in this research. First, a multi-user, collaborative, immersive solution using AR, named 4DMAR, was developed for visualizing a 4D model. An assessment approach was designed and performed to evaluate the performance of the 4DMAR and show its potential to provide a collaborative environment. A comparative study was conducted to measure the performance of 4DMAR. 63% of participants felt AR was better suited for collaborative work and group activities than only 18% preferring the screen-based environment. By counterbalancing the experiment paths (A and B), an order effect on preferences was not detected based on the collected data.

While the results showed a user preference for AR for construction plan review, this has some limitations based on the participant pool and task definition. Future research should examine this technology with industry professionals to critically assess the potential benefits and challenges of adopting AR and address the limited industry experience of the participants to give a better judgment. Another key limitation identified was the low rate of participants detecting sequencing flaws in the 4D construction plan despite this being a primary use case. Incorporating explicit flaw identification tasks will be a critical focus for further developing the 4DMAR system. In addition, complex models that include different building systems should also be used to reflect more complicated challenges within extensive model reviews. Solutions such as performing more systems-selective task reviews to enable more value-based approaches can be explored. However, with the current limited computational power of AR headsets, it is very difficult to visualize large building information models without optimizing the meshes or using other proximity solutions.

References

Al-Adhami, M., Ma, L., Wu, S., 2018. Exploring Virtual Reality in Construction, Visualization, and Building Performance Analysis. https://doi.org/10.22260/ISARC2018/0135

Amrollahibuki, R., 2019. Modeling Construction Equipment in 4D Simulation and Application in VR Safety Training (masters). Concordia University.

Behzadan, A.H., Kamat, V.R., 2013. Enabling discovery-based learning in construction using telepresent augmented reality. Autom. Constr., Augmented Reality in Architecture, Engineering, and Construction 33, 3–10. https://doi.org/10.1016/j.autcon.2012.09.003

Bortolini, R., Formoso, C.T., Viana, D.D., 2019. Site logistics planning and control for engineer-to-order prefabricated building systems using BIM 4D modeling. Autom. Constr. 98, 248–264. https://doi.org/10.1016/j.autcon.2018.11.031

Boton, C., 2018. Supporting constructability analysis meetings with Immersive Virtual Reality-based collaborative BIM 4D simulation. Autom. Constr. 96, 1–15. https://doi.org/10.1016/j.autcon.2018.08.020

Davila Delgado, J.M., Oyedele, L., Beach, T., Demian, P., 2020. Augmented and Virtual Reality in Construction: Drivers and Limitations for Industry Adoption. J. Constr. Eng. Manag. 146, 04020079. https://doi.org/10.1061/(ASCE)CO.1943-7862.0001844

Guo, H., Yu, Y., Skitmore, M., 2017. Visualization technology-based construction safety management: A review. Autom. Constr. 73, 135–144. https://doi.org/10.1016/j.autcon.2016.10.004

Hartmann, T., Fischer, M., 2007. Supporting the constructability review with 3D/4D models. Build. Res. Inf. 35, 70–80. https://doi.org/10.1080/09613210600942218

Haymaker, J., Fischer, M., 2001. Challenges and Benefits of 4D Modeling on the Walt Disney Concert Hall Project.

Johnson, S., 1998. What's in a representation, why do we care, and what does it mean? Examining evidence from psychology. Autom. Constr. 8, 15–24. https://doi.org/10.1016/S0926-5805(98)00062-4

Kim, B., Kim, C., Kim, H., 2012. Interactive Modeler for Construction Equipment Operation Using Augmented Reality. J. Comput. Civ. Eng. 26, 331–341. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000137

Lin, T.-H., Liu, C.-H., Tsai, M.-H., Kang, S.-C., 2015. Using Augmented Reality in a Multiscreen Environment for Construction Discussion. J. Comput. Civ. Eng. 29, 04014088. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000420

Lin, Z., Petzold, F., Hsieh, S.-H., 2020. 4D-BIM Based Real Time Augmented Reality Navigation System for Tower Crane Operation, in: Construction Research Congress 2020. Presented at the Construction Research Congress 2020, American Society of Civil Engineers, Tempe, Arizona, pp. 828–836. https://doi.org/10.1061/9780784482865.088

Lin, Z.Y., Petzold, F., Ma, Z.L., 2019. A Real-Time 4D Augmented Reality System for Modular Construction Progress Monitoring, in: ISARC. Proceedings of the International Symposium on Automation and Robotics in Construction. IAARC Publications, Waterloo, Canada, pp. 743–748. https://doi.org/10.22260/ISARC2019/0100

Meža, S., Turk, Ž., Dolenc, M., 2015. Measuring the potential of augmented reality in civil engineering. Adv. Eng. Softw. 90, 1–10. https://doi.org/10.1016/j.advengsoft.2015.06.005

Silva, T. de S., Marinho, E.C.R., Cabral, G.R.E., Gama, K.S. da, 2017. Motivational Impact of Virtual Reality on Game-Based Learning: Comparative Study of Immersive and Non-Immersive Approaches, in: 2017 19th Symposium on Virtual and Augmented Reality (SVR). Presented at the 2017 19th Symposium on Virtual and Augmented Reality (SVR), pp. 155–158. https://doi.org/10.1109/SVR.2017.28

Staub-French, S., Khanzode, A., 2007. 3D and 4D modeling for design and construction coordination: issues and lessons learned. J. Inf. Technol. Constr. ITcon 12, 381–407.

Sterman, J.D., 1992. System Dynamics Modeling for Project Management 246.

Testing and Performance Analysis | Oculus Developers [WWW Document], 2022. URL https://developer.oculus.com/documentation/unity/unity-perf/ (accessed 6.19.23).

Wang, J., Hu, Y., Yang, X., 2022. Multi-person Collaborative Augmented Reality Assembly Process Evaluation System Based on HoloLens, in: Chen, J.Y.C., Fragomeni, G. (Eds.), Virtual, Augmented and Mixed Reality: Applications in Education, Aviation and Industry, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 369–380. https://doi.org/10.1007/978-3-031-06015-1_25

Wang, X., Truijens, M., Hou, L., Wang, Y., Zhou, Y., 2014. Integrating Augmented Reality with Building Information Modeling: Onsite construction process controlling for liquefied natural gas industry. Autom. Constr. 40, 96–105. https://doi.org/10.1016/j.autcon.2013.12.003

Wolfartsberger, J., 2019. Analyzing the potential of Virtual Reality for engineering design review. Autom. Constr. 104, 27–37. https://doi.org/10.1016/j.autcon.2019.03.018

A MIXED REALITY HUMAN ROBOT INTERACTION INTERFACE USING HAND TRACKING AND STEREO VISION FOR CONSTRUCTION APPLICATIONS

Damith Tennakoon[1], Mojgan Jadidi[1], and SeyedReza RazaviAlavi[2]
[1]York University, Toronto, Canada
[2]Northumbria University, New Castle, UK

Abstract

The modern day construction site is able to complete complex tasks using highly skilled on-site workers and robust machinery. On-site workers face severe physical injuries from unforeseen accidents due to construction site dynamics. This project proposes and implements a mixed reality human robot interaction system, utilizing stereo vision systems and hand tracking algorithms, to teleoperate a robot arm within a simulation. Results demonstrate enhanced depth and situational awareness in the robot workspace versus traditional teleoperation. This research allows operators to safely perform complex construction tasks remotely with a robot arm using a head-mounted display.

Introduction

The modern day construction site completes large-scale tasks by training on-site workers and utilizing heavy-duty machinery. However, success is not achieved with efficiency or safety as a major priority since construction work is prone to surpassing timelines and facing potential delays due to low productivity rates Davila Delgado et al. (2019). On-site workers are often subject to serious health consequences, that are frequently overlooked. Some of these hazards include exposure to high concentrations of particulate matter, workers colliding with heavy machinery, and falls off unsecured high altitude sites Yang et al. (2023); Castro-Lacouture (2009).

Various industries continue to adopt automation techniques, leveraging robotics solutions, to improve task efficiency, productivity, and product quality. For instance, the manufacturing industry utilizes robotics to automate assembly lines, with the advent of performing repetitive, hazardous, and time-constrained tasks with ease, to produce high volumes of products. Not only are robotics solutions very efficient, but are also highly competent with precision-seeking tasks as made evident by their applications within the medical field. The healthcare industry deploys robots to perform surgical procedures with the advent of providing minimally invasive procedures and higher precision than human surgeons. The advantages of uninterrupted work is highlighted by the application of robotics in agriculture where tasks including planting, harvesting, and monitoring crops are automated and carried out by multiple robots Campilho and Silva (2023). Huang and Sakurai (1990) concluded that the implementation of factory automation (FA) for well-established automotive companies, including Mitsubishi, Nissan, and Mazda, displayed a noticeable reduction in labor costs, an increase in vehicle quality, and increased flexibility for adapting tasks. An extensive study performed on the implications of introducing industrial robots into manufacturing firms showed increased worker safety; specifically, 1 standard deviation of an increase in robots resulted in a 4% reduction in physical job intensity and a 5% reduction in job worker disabilities Gihleb et al. (2022).

With the clear benefits that automation and robotics provide, many attempts have been made to integrate this technology into the construction industry. A study demonstrated the successful implementation of robots into performing construction tasks require the adaptation of robots to the unpredictable and dynamic nature of construction sites. This has been achieved in modern robotics in two ways: (1) pre-programming robots to handle many different scenarios and rigorously testing them prior to deployment and (2) by training robots, using sophisticated real-world datasets and machine learning techniques, for on-site applications. The fast-paced and dynamic construction environment prevents either of these requirements from being met; pre-programmed robots are not able to handle unforeseen tasks or changing tasks and quality training data for construction applications are limited Zhang et al. (2023). Mukherjee et al. (2022) suggests a more intuitive method to deploy robotics into the construction sector is to create a human-robot collaboration (HRC) system that integrates the on-site worker's experience with the advantages of autonomous robots, to increase task productivity and reduce labor shortage. Advent of such as system alleviates the need for dynamic robot programming and multiple task execution.

In order to perform any intricate task, a complex manipulator is required. Robot arms are used to interact with and manipulate objects within an environment via various teleoperation methods. Teleoperated robots are specially useful where the robot's physical target workspace is challenging or impossible to access. The most commonly used control interfaces for these instances are keyboards, joysticks, and portable touch screen devices. However, these traditional modes of teleoperation are ineffective as the control interfaces are not intuitive. One of the emerging solutions for this problem is the use of mixed reality technologies which facilitate full immersion into the robot's workspace, creating a more natural human-robot interaction (HRI) interface Meng et al. (2023); Wonsick and Padir (2020a). Mixed reality (MR) is a combination of augmented reality (AR) and virtual reality (VR) technologies. MR combines the real and virtual worlds to construct a hybrid environment where interactable 3D virtual models exist Aloqaily et al. (2023). The advent that mixed-reality HRI interfaces provides over traditional teloperational methods is the ability to view and interact with robots in 3D environments Wonsick and Padir (2020a). The ability to view in 3D is due to the emergent properties of 3D cameras and immersive head-mounted displays (HMDs). Compared to 2D video-based visualization, 3D real-time visualization coupled with HMDs has shown to improve localization, maneuverability, and control in robotics applications Wonsick and Padir (2020b).

In this paper, we introduce a novel HRI interface that leverages MR technology, state-of-art hand tracking methods, coupled with stereo vision systems to teleoperate a 4-DoF robot arm in an immersive first-person view (FPV) using a Meta Quest 2 HMD. By providing a mode to operate a robot arm using hand gestures from a safe distance to the manipulator, on-site construction workers can perform complex tasks, utilizing their experience, while taking advantage of the strength, precision, and deployabality of robots.

Related Works

Much research and development has accumulated over the past few years specifically on leveraging AR, VR, and MR technologies to effectively improve human-robot interaction and collaboration. Puljiz et al. (2019) proposes and implements an HRI interface to teleoperate an industrial grade robot arm utilizing the Microsoft HoloLens AR HMD. The approach presented utilizes the built-in spatial mapping of the HoloLens to identify the target robot arm, and its end effector. This is achieved by storing a previously known model of the robot arm and matching the known model with the detected mesh of the spatial mapping output. The known mesh and the HMD is localized with the detected robot arm, of the physical space, and hand detection techniques are used to translate the end effector. This interface enables the teleportation of a robot arm without the need for physical interaction with the manipulator.

Yim et al. (2022) emphasized the importance of full immersion into the robot workspace through means of stereo vision. They implement a method to teleoperate a 6-DoF robot arm by fix mounting an Intel RealSense stereo camera to the platform the robot arm is fixated on, with a view facing down and towards the end effector. The stereo images are streamed to a Meta Quest 2 VR HMD to replicate human vision. Hand held controllers are localized with the robot arm to manipulate its end effector. This HRI interface enables a robot to be controlled from a safe distance with human-like perception.

Chang et al. (2023) tackles on the ongoing challenge of effective robot workspace visualization and perception by their design of a 2D/3D multi-view teleoperated robot arm. Specifically, they have devised an approach to view a robot's workspace via an FPV real-time 3D point cloud (PCD) model, from a stereo vision camera mounted to the end effector of the manipulator, and a 2D view from a global RGB camera. This effectively enabled for the visualization of the direct workspace, via the PCD, as well as the global workspace, via the RGB camera. The developed MR interface enabled the user to switch between the two visual systems in their HMD, using a handheld controller, providing a comfortable teleoperation process.

Many other similar approaches have been proposed and implemented showcasing positive results towards achieving a natural HRI interface. Though, these HRI systems do not come without their own setbacks. With the implementation of the HoloLens AR approach by Puljiz et al. (2019), though the interface provides robot arm teleoperation without required physical interaction with the manipulator, it is constrained by the requirement of a physical operator needing to be present near the robot's workspace. This limits the control bandwidth as the operator needs to have a direct line-of-sight of the robot arm for teleoperation. In the case of the immersive stereo vision teleoperated robot arm developed by Yim et al. (2022), many safety issues arise when the user's view is obstructed by the robot arm itself or other objects, within the robot's workspace, due to the vision system being fixed mounted. Chang et al. (2023) is on the right track with their multi-view MR HRI interface, however, the cognition load required for the operator increases when frequently needing to switch between a 3D view and a fixed 2D view; the operator will need significant training to be comfortable localizing their coordinate system with the robot arm's coordinate frames, in the two views. Many lessons were taken away from all aforementioned approaches and has aided the development of the HRI interface discussed in this paper.

Methodology

Approach

The design of the HRI system that is being developed revolves around creating an immersive, hands-free, and safe interface to teleoperate a robot arm towards performing construction tasks. There are 4 main categories that are required to achieve this interface; (1) stereo vision system to replicate human-like perception; (2) immersive wearable technology to construct an immersive MR interface; (3) real-time hand-tracking algorithms for intuitive teleoperation; (4) tracking systems to provide real-time operator situational awareness.

The stereo vision system used for this project is the ZED Mini (ZEDM) stereo camera, developed by StereoLabs. Stereo vision is a visualization system that enables 3D point extraction of an environment, through images obtained via two cameras separated by a fixed distance. The general concept is founded on the basis of finding the disparity between the two images to extract the missing depth information, through means of geometry DANDIL and ÇEVİK (2019). The ZEDM camera replicates the human-eyes, enabling real-time depth perception, with extreme accuracy. The approach is to mount the ZEDM camera on top of the end effector of the 4-DoF (degrees of freedom) robot arm. The left and right videos will be streamed, in real-time, to the left and right displays of the Meta Quest 2 VR HMD. This enables the operator, wearing the HMD, to perceive the robot's point-of-view (POV) with human-like vision. Open source libraries utilize the 4 on-board cameras of the HMD, used for spatial mapping, to track the user's hands in real-time. The tracked hand models are superimposed to the FPV stereo view on the HMD. To provide full situational awareness, the ZEDM camera is mounted on a servo actuated two-axis gimbal that tracks the orientation of the HMD, in real-time. As the operator wearing the HMD looks around, the servo motors are actuated to map the HMDs look angles with the ZEDM camera. This will provide a natural one-to-one orientation system enabling full situational awareness of the robot's workspace. The ZEDM vision system has a noteworthy advantage of facilitating pass-through reality, allowing seamless integration of 3D virtual models into the real-time stereo images. This is a fundamental requirement for this HRI system since the robot arm will be controlled via an MR control sphere; in the FPV view of the HMD, a translucent spherical virtual object will be fixated onto the visualized end effector, using pass through reality. The operator will be able to place their hands into this tracking sphere towards the desired direction for the robot to move towards. Hand gestures, such as pinching and releasing two fingers together, will be detected to provide gripper controls.

This paper covers the development of the proposed architecture, in a simulated environment, and preliminary results from experiments conducted, comparing traditional teleoperation methods with the developed MR HRI interface, when performing simple tasks.

Simulation Environment Setup

The simulation environment chosen for this project was Unity Game Engine. Unity is a cross-platform game development engine that enables developers to build video games, simulations, and mixed reality applications, by utilizing 3D objects and scripting via the C# objected-oriented programming language Haas (2014). Unity was the ideal candidate for this simulation as it provides open-source libraries for AR, VR, and MR development. It is also built on the capable NVIDIA PhysX 3.4 physics engine, providing the ability to run real-time physics simulations.

The first step in designing a real-time simulated robot arm was selecting a 3D model that represented a real-world industrial grade robot arm, namely the manipulators developed by KUKA. The model used for this simulation was retrieved from the open-source 3D modelling website Sketchfab, by a developer named Yuki. The 3D model features a 4-DoF robot arm, in FBX format. In the Unity environment, the design underwent many modifications to organize the multi-component model into a structured hierarchical set of objects, where each rotating joint, connecting each of the links together, were aligned relative to a global coordinate system. To actuate the robot arm during run-time, a joint-space controller was scripted, enabling each joint to rotate to an input angle with a constant angular velocity. Although joint-space control for robot arm's are used in real-world scenarios, it is not representative of how robot arm's are often operated majority of the time. Rather, a target position vector, a 3D coordinate in space, is provided to the manipulator. The system then computes a set of ideal joint angles to achieve the desired end effector position. In robotics, this mathematical process is referred to as inverse kinematics, a key-part of robot trajectory planning. One of the challenges with inverse kinematics is for a given target end effector position, there may be infinite many solutions, leading to computation efficiency problems. Many open-source inverse kinematic solver libraries exist such as Ikpy library. Ikpy is a Python library that performs inverse kinematic solutions efficiently, given a kinematic chain, as a Unified Robotics Description Format (URDF) file, and a target position and orientation. A URDF file is a low level description file, written in the Extensible Markup Language (XML), which describes the geometry, mobility, and their corresponding relationships, of a robot. Utilizing the Unity 3D model, a URDF file was developed for the robot arm. To connect the Python inverse kinematic scripts to the Unity simulation environment, a Transmission Control Protocol/Internet Protocol (TCP/IP) server-client system was designed. TCP/IP is a data link protocol that enables the transmission & reception of messages, between devices, over an internet connection. A three-way handshake connection, between two devices, enables the host server to send packets of information to the client, where the client transmits a message back to the server, confirming data reception. TCPs error handling capability ensures reliable point-to-point communication, without data loss. Python TCP server scripts, integrating the inverse kinematic solvers, and the corresponding C# client scripts, were developed to establish a connection between the computing server and the simulation environment client. This simulation architecture is summarized in Figure 1 in which the 3D coordinates of a target object, in the Unity environment, is transmitted to the server which returns a set of joint angles to achieve the desired target position. The result is a 4-DoF robot arm that tracks and follows a target sphere object, in real-time. To further actualize the simulation, a prismatic gripper tool was designed, using the CAD modelling software Autodesk Fusion 360, and subsequently affixed to the end effector joint of the manipulator. To devise experiments relating to construction work, as discussed in later sections, 3D models of pallets and bricks were placed into the virtual environment. This simulated environment, ascribed in Figure 2, is an ideal representation of the physical robot arm, and its implementation, that is underway at the time of writing.

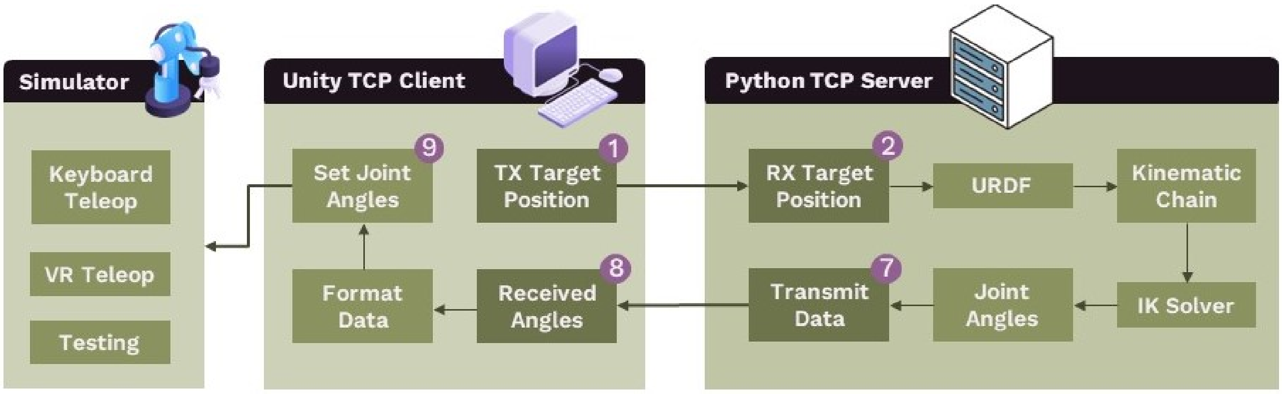


Figure 1: Real-time 4-DoF robot arm inverse kinematic simulation workflow diagram.

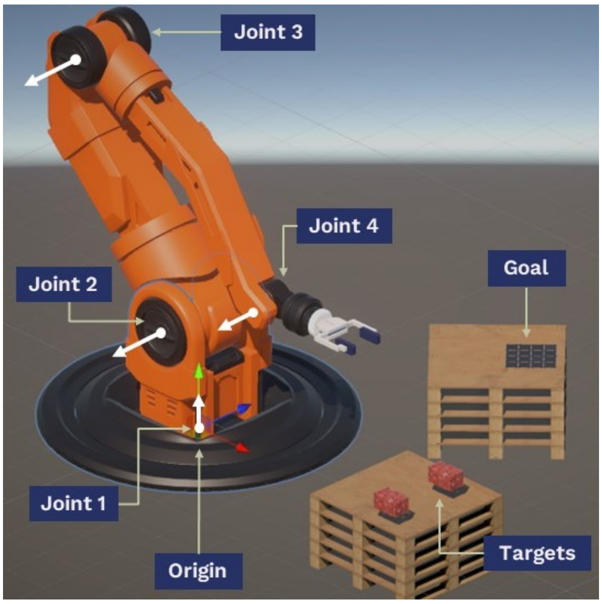


Figure 2: Simulated virtual environment with key labels.

Traditional Teleportation Interface

Prior to developing the proposed MR HRI interface, traditional modes of robot teleoperation were simulated. One of the common modes a robot arm is controlled via keyboard inputs. The developed control algorithm enables the robot's end effector to translate along the horizontal and vertical, while the gripper can be manipulated via two separate keys, as documented on Table 1. The underlying control algorithm takes the input from the keys and translates the manipulator's end effector target vector, in 3D space, by at a constant rate. This effectively provides real-time end effector control via keyboard inputs. Teleportation is commonly coupled with visual input systems to provide a real-time view of the robot's workspace, when it is operating at an inaccessible setting. In the real-world, this is often done by fix mounting an RGB camera that has a direct line-of-sight to the robot's workspace. A monocular (single-lens) camera was integrated into the virtual environment, presented in Figure 3, to provide the operator with a view of the robot arm's workspace for teleoperation via keyboard inputs. This achieves a simulation of a traditional teleoperation method which will be used for comparison with the MR HRI interface.

Table 1: End effector keyboard control inputs.

Horizontal	Vertical	Gripper
I, J, K, L	O, P	Q, E

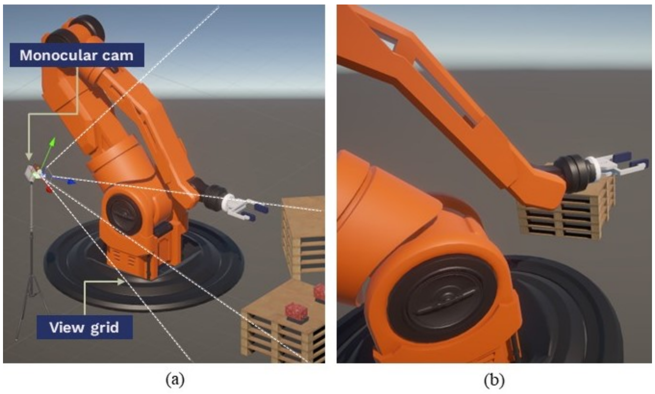


Figure 3: Monocular camera system: (a) Camera rig setup within the virtual environment; (b) Real-time view from the operator's computer screen during run-time.

Mixed Reality Teleoperation Interface

The first sub-component of developing the proposed MR HRI interface was to simulate a real-time FPV stereo view from the robot arm's end effector, utilizing the Meta Quest 2 VR HMD. The ZEDM stereo camera, by StereoLabs, is able to provide human-like perception. This stereo camera system was integrated into the simulation by hierarchically attaching a 3D model of the camera to a CAD modelled, and maneuverable, servo actuated two-axis gimbal, shown in Figure 4. The objective of this system is to receive real-time orientation data, of the HMD, and map its rotations about the simulated servos y-z axis. This enables real-time head tracking of the operator, presented in Figure 5. Natural head-tracking is of significance as MR wearable technology is prone to creating discomfort for operators. This occurs when the physical movements of the operator do not align with their sensory input systems (vision, sound, touch). As a result, constructing a one-to-one human-like vision system was a key priority.

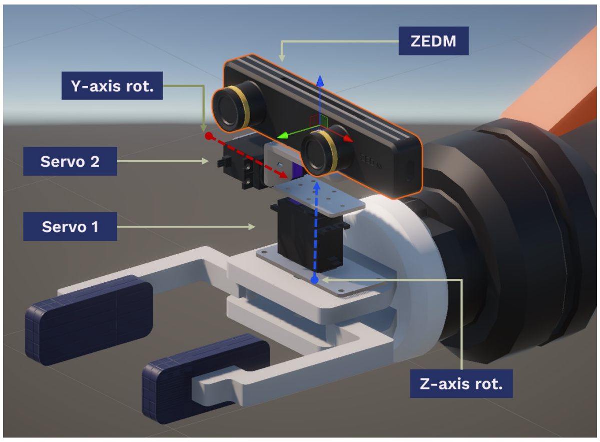


Figure 4: Simulated head tracking 2-axis gimbal system for ZEDM camera.

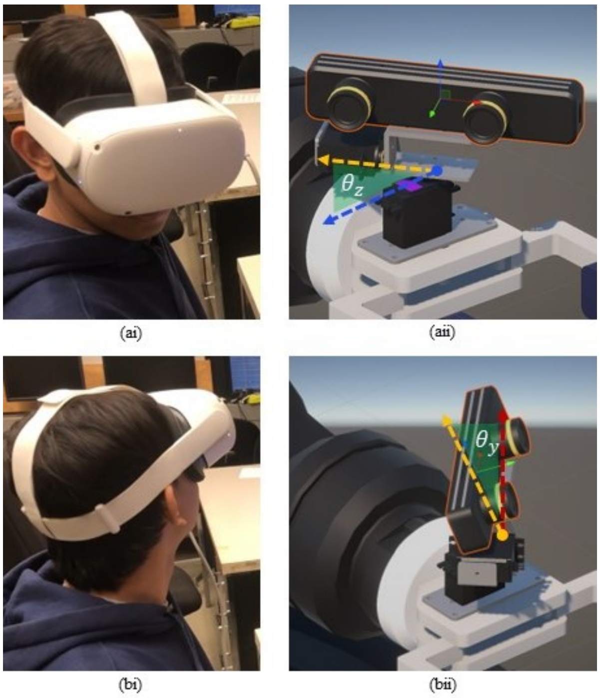


*Figure 5: HMD head tracking system during run-time: **(ai)** operator turning their head towards the right, rotating about the z-axis by angle θ_z ; **(aii)** 2-axis gimbal rotating the z-axis servo by corresponding angle; **(bi)** operator turning their head upwards, rotating about the y-axis by angle θ_y; **(bii)** 2-axis gimbal rotating the y-axis servo by corresponding angle.*

The second sub-component was the integration of real-time hand tracking and gesture detection. The objective here was to provide a mode to manipulate the robot arm using nothing more than the operator's hands, which was achieved utilizing the open-source Oculus Integration package. This is a well-developed library, by Oculus, that provides many features for the HMD including access to IMU data, hand tracking, and spatial mapping. The hand tracking functions by utilizing the 4 onboard multi-camera system of the HMD. Utilizing this library, a VR rig was setup in the virtual environment, with the integrated hand models, to render the tracked hand data in real-time. To perform gesture detection, such as pinch recognition, algorithms were written to measure the relative separation of individual joints in 3D space. Figure 6 diagrammatically shows the real-time hand tracking and how simple gestures are detected.

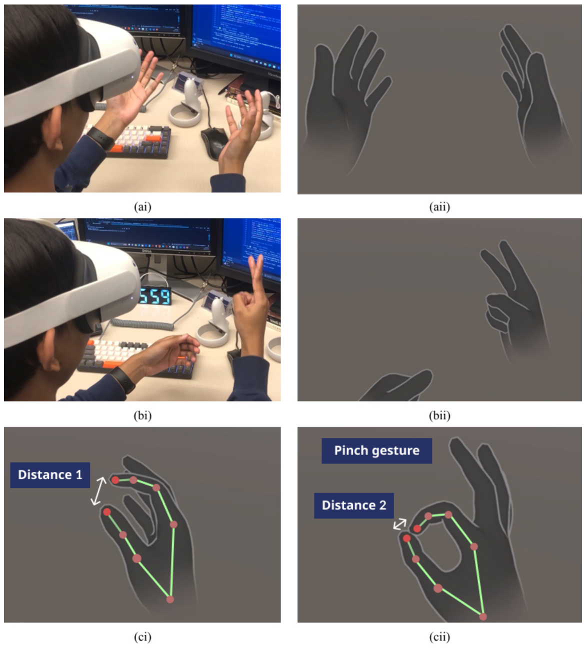


Figure 6: (ai, aii, bi, bii) demonstrates the implemented real-time hand tracking system with the Meta Quest 2 HMD. (ci, cii) demonstrates the pinch gesture being detected by measuring the separation between two joints.

To manipulate the end effector, the tracked hand models were fused with the stereo camera vision system. The result is a full situational aware FPV of the robot arm's end effector with the users tracked hand models. To move the robot arm using the hand tracking system, a Control Sphere was designed. The Control Sphere consists of two 3D virtual models of translucent spheres, one with a smaller radius than the other, anchored at the center of the larger sphere, spatially mapped onto the robot arm's end effector. The outer sphere, with radius R_2, is the Movement Sphere, and the inner sphere, with radius R_1, is the Gesture Sphere, described in Figure 7. When a tracked hand model is inside the Movement Sphere but outside the Gesture Sphere, a vector, from the 3D position of the hand model to the Control Sphere center, is computed and normalized. The resulting unit vector is indicative of the direction the operator desires to move end effector. The robot arm's target tracking algorithm is updated by taking the sum of the current target position with the product of the unit vector and a constant multiplier. The result is the end effector moving in the vector direction, relative to the center of the Control Sphere, as input by the operator's tracked hand models, providing omnidirectional control. When the tracked hand models are within the Gesture

Sphere, the tracking target's movement speed is set to zero, effectively pausing motion control. Users are then able to input hand gestures within this sphere, such as pinching the thumb and index fingers together, to close the gripper and releasing these fingers to open the gripper. Equation 1 computes the relative position vector between the tracked hand model to the center of the Control Sphere while equation 2 normalizes the position vector. In equation 3, the variable Δt corresponds to the elapsed time between the current and previous frame in the simulation while C is a constant multiple to set the movement speed of the manipulator. The updated target tracking position, $\vec{\mathbf{T_{t+1}}}$, is computed, taking into account its current position $\vec{\mathbf{T_t}}$. Figure 8 shows the described mixed reality HRI system, at simulation run-time, for moving the manipulator and grabbing objects, using the hand tracking system.

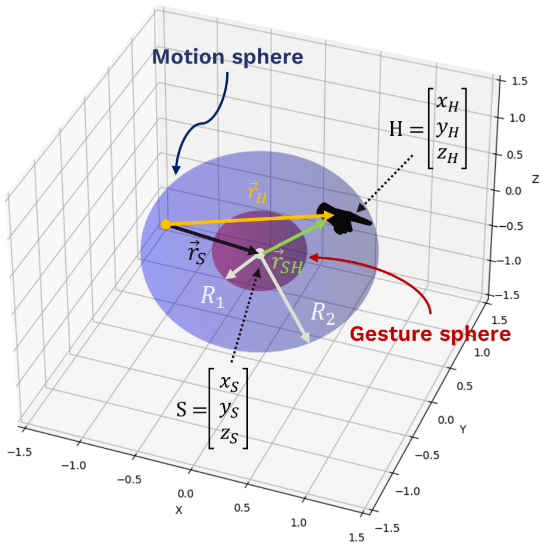


Figure 7: Diagram of the Control Sphere structure used to control the robot arm.

$$\vec{\mathbf{r_{SH}}} = \vec{\mathbf{r_H}} - \vec{\mathbf{r_S}} = \begin{bmatrix} x_H - x_S \\ y_H - y_S \\ z_H - z_S \end{bmatrix} \tag{1}$$

$$\hat{r_{SH}} = \frac{\vec{\mathbf{r_{SH}}}}{|\vec{\mathbf{r_{SH}}}|} = \frac{< x_H - x_S, y_H - y_S, z_H - z_S >}{\sqrt{(x_H - x_S)^2 + (y_H - y_S)^2 + (z_H - z_S)^2}} \tag{2}$$

$$\vec{\mathbf{T_{t+1}}} = \vec{\mathbf{T_t}} + C\Delta t \times \hat{r_{SH}} = \begin{bmatrix} x_t + C\Delta t \times \hat{r_{SH_x}} \\ y_t + C\Delta t \times \hat{r_{SH_y}} \\ z_t + C\Delta t \times \hat{r_{SH_z}} \end{bmatrix} \tag{3}$$

Experiment Setup

An experiment was devised to compare the implemented MR HRI system to the traditional monocular fixed-view keyboard teleoperation method. The goal of the experiment was to observe trends relating to depth awareness, overall situational awareness, and task completion efficiency. The experiment consisted of driving the robot arm, from its starting position, to a brick, that is placed on a pallet, with the goal of moving to the position where the brick would be in between the gripper, as shown in Figure 9. To collect data during run-time, C# scripts were written to store the coordinates of the end effector, along with the elapsed time, between iterations. Each teleoperation method was executed by the same operator 5 times. The raw flight path data will be indicative of depth perception and situational awareness of the operator while the time to minimize the distance to the target will be indicative of efficiency.

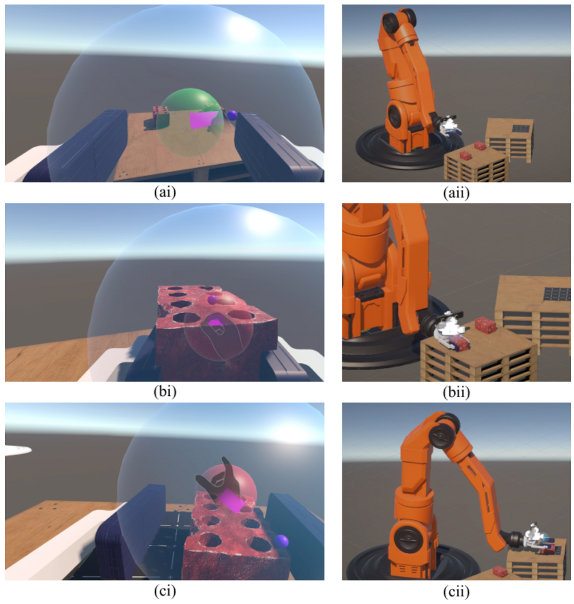


Figure 8: The robot arm being teleoperated by the MR HRI system using hand tracking for motion control and gesture recognition to control the gripper. Left: hand tracking input, right: robot's response.

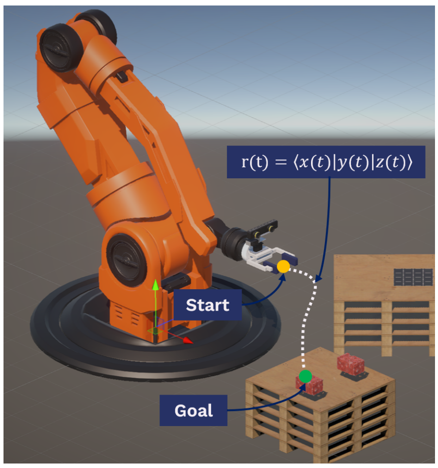


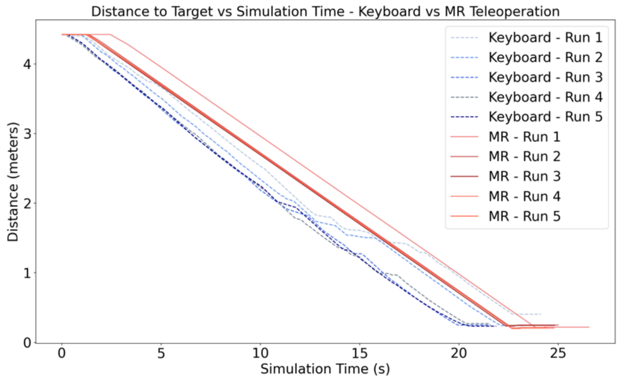
Figure 9: Experiment Setup

Results

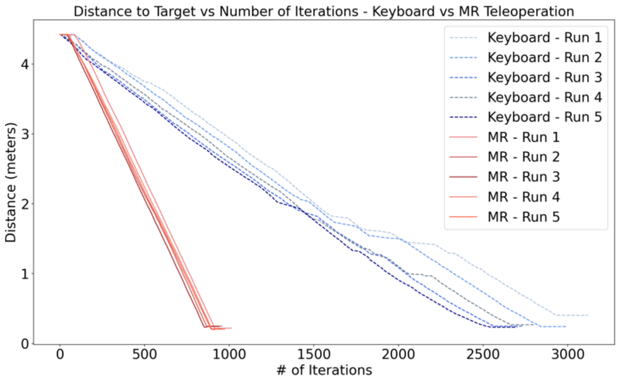
The trajectories of the robot arm's end effector for each of the 5 runs of the monocular fixed-view keyboard teleoperation is shown in Figure 11a). The trajectories of the robot arm's end effector for each of the 5 runs of the MR

HRI teleoperation is shown in Figure 11b). A side-by-side comparison of both teleoperation modes, for each run, is shown in Figure 11c). Inspection of the trajectories of the monocular fixed-view keyboard teleoperation method demonstrates irregular flight paths on each of the runs. The flight path trends show two distinct patterns: simultaneously moving forward and laterally towards the target followed by an immediate singular lateral movement. The inability to maintain a smooth trajectory indicates a lack of depth perception of the robot's workspace. Inspection of the trajectories of the MR HRI teleoperation method demonstrates consistent and smooth flight paths on each of the runs. Each of the trajectories has the properties of a direct vector, from the starting pose to the goal pose. This suggests that there is a high level depth-awareness and situational awareness of the robot's workspace. The side-by-side view of both teleoperation methods demonstrates this difference in operator awareness.

The time to complete the task for each teleoperation method and runs are shown in Figure 10a). The number of iterations in the simulation each teleoperation method took to complete the task for each of the runs is shown in Figure 10b). Upon initial inspection, both teleoperation methods are observed to complete the task, per run, approximately around the same duration of time. The results show the traditional monocular fixed-view keyboard teleoperation method completes the task quicker, in some runs, than the MR HRI method. This is contradictory to the trajectory results as the MR HRI method moves in a direct path to the goal pose while the traditional method has an indirect path. Furthermore, for both teleoperation methods and all simulation runs, the robot arm was set to move at the same constant velocity. Inspection of the number of iterations for each of the runs shows that the MR HRI system completes the task in approximately 2.5x fewer iterations than the traditional teleoperation method. Therefore, the longer task completion duration for the MR HRI method is hypothesized to be due to its higher computational demand compared to the simplistic keyboard teleoperation method. The MR HRI simulation takes longer to execute a single iteration, or a lower frame rate, resulting in a longer simulation run-time. However, it is evident that the MR HRI system is able to minimize the same distance to the target as the keyboard teleoperation method with fewer iterations. To truly understand the relationship between task completion time and the two teleoperation methods, a real-world implementation is required.



(a)



(b)

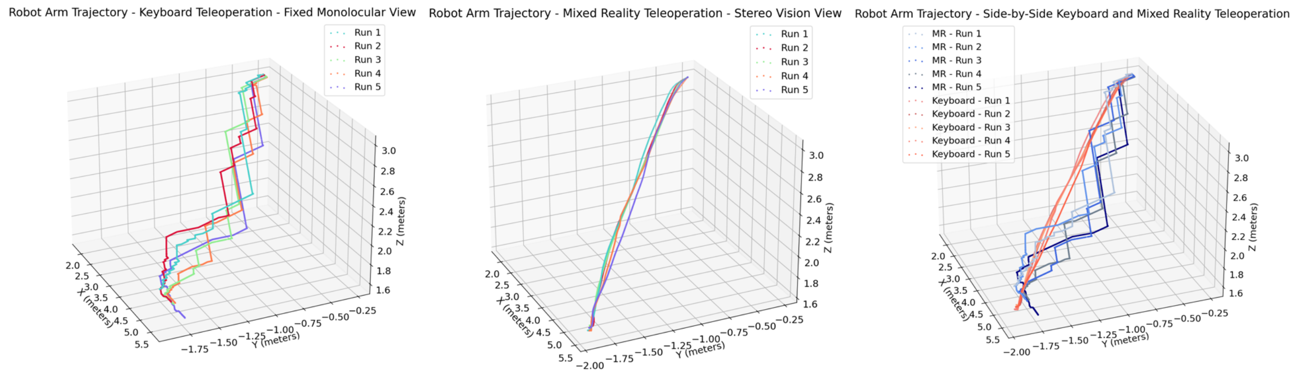
Figure 10: (a) Distance to target vs simulation run-time data and (b) Distance to target vs number of simulation iterations to complete task data.

Conclusion and Future Work

We proposed and implemented a mixed reality human robot interaction system for a 4-DoF robot arm. The results of the our architecture demonstrate a higher operator depth perception and situational awareness of the robot's workspace compared with a traditional monocular fixed-view keyboard teleoperation method. Supported by stereo vision systems, a head mounted display, and hand tracking algorithms, the operator is able to manipulate the robot arm's end effector safely and efficiently to perform simple construction tasks, such as picking up bricks in a simulation environment. The modularity and mechanics of the MR HRI system will enable the control of various real-world robot arms, rather than a select few. The system alleviates the need for on-site construction workers, greatly reducing accidents and serious injuries in construction sites. Rather, the system enables operators to utilize their knowledge and expertise from a safe distance, performing intricate tasks while simply wearing an HMD. Future work will require the extension of the MR HRI system to a physical robot arm. At the time of writing, a successful implementation of real-time hand tracking with the real-time stereo vision of the physical ZEDM camera has been completed. Next steps will include the development of a two-axis gimbal, to mount the vision system, to demonstrate HMD tracking. Additionally, control algorithms to interface between a physical 4-DoF robot arm and the MR HRI system will be developed. A test with a physical robot arm will provide important data on task completion efficiency, operator training, and robot control accuracy.

Acknowledgments

This research is funded by the Natural Sciences and Engineering Research Council of Canada (NSERC DG 485794).



(a) (b) (c)

Figure 11: Trajectory data of the robot arm's end effector for multiple runs. (a): monocular fixed-view keyboard teleoperation, (b): mixed reality teleoperation, (c): both teleoperation methods.

References

Aloqaily, M., Bouachir, O., and Karray, F. (2023). Chapter 3 - digital twin for healthcare immersive services: fundamentals, architectures, and open issues. In El Saddik, A., editor, Digital Twin for Healthcare, pages 39–71. Academic Press.

Campilho, R. D. S. G. and Silva, F. J. G. (2023). Industrial process improvement by automation and robotics. Machines, 11(11).

Castro-Lacouture, D. (2009). Construction Automation, pages 1063–1078. Springer Berlin Heidelberg, Berlin, Heidelberg.

Chang, Y.-W., Hsu, S.-T., Yeh, H.-H., and Liu, Y.-C. (2023). An intuitive human-manipulator interface with mixed reality. In 2023 23rd International Conference on Control, Automation and Systems (ICCAS), pages 629–634.

DANDIL, E. and ÇEVİK, K. K. (2019). Computer vision based distance measurement system using stereo camera view. In 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pages 1–4.

Davila Delgado, J. M., Oyedele, L., Ajayi, A., Akanbi, L., Akinade, O., Bilal, M., and Owolabi, H. (2019). Robotics and automated systems in construction: Understanding industry-specific challenges for adoption. Journal of Building Engineering, 26:100868.

Gihleb, R., Giuntella, O., Stella, L., and Wang, T. (2022). Industrial robots, workers' safety, and health. Labour Economics, 78:102205.

Haas, J. K. (2014). A history of the unity game engine.

Huang, P. and Sakurai, M. (1990). Factor automation: the japanese experience. IEEE Transactions on Engineering Management, 37(2):102–108.

Meng, L., Liu, J., Chai, W., Wang, J., and Meng, M. Q.-H. (2023). Virtual reality based robot teleoperation via human-scene interaction. Procedia Computer Science, 226:141–148. Proceedings of International Conference on Biomimetic Intelligence and Robotics.

Mukherjee, D., Gupta, K., Chang, L. H., and Najjaran, H. (2022). A survey of robot learning strategies for human-robot collaboration in industrial settings. Robotics and Computer-Integrated Manufacturing, 73:102231.

Puljiz, D., Stöhr, E., Riesterer, K. S., Hein, B., and Kröger, T. (2019). General hand guidance framework using microsoft hololens. In 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 5185–5190.

Wonsick, M. and Padir, T. (2020a). A systematic review of virtual reality interfaces for controlling and interacting with robots. Applied Sciences, 10(24).

Wonsick, M. and Padir, T. (2020b). A systematic review of virtual reality interfaces for controlling and interacting with robots. Applied Sciences, 10(24).

Yang, X., Yu, Q., Zhang, Y., and Ma, W. (2023). Occupational health risk assessment of construction workers caused by particulate matter exposure on construction sites. Heliyon, 9(10):e20433.

Yim, L. S., Vo, Q. T., Huang, C.-I., Wang, C.-R., McQueary, W., Wang, H.-C., Huang, H., and Yu, L.-F. (2022). Wfh-vr: Teleoperating a robot arm to set a dining table across the globe via virtual reality. In 2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pages 4927–4934.

Zhang, M., Xu, R., Wu, H., Pan, J., and Luo, X. (2023). Human–robot collaboration for on-site construction. Automation in Construction, 150:104812.

INPUT STRATEGIES FOR PRECISE CAD MODELLING WITH MODERN INTERACTION DEVICES

Markus Berger[1], Jörn Plönnigs[1], and Martin Wogan[2]

[1]University of Rostock, Rostock, Germany
[2]TU Dresden, Dresden, Germany

Abstract

The release of ever more capable wearable display and interaction devices, like Apple's Vision Pro or the Meta Quest line of products, and the continuous change in the mobility of engineering work, enables new means of interaction with Computer Aided Design (CAD) software. This includes new input modalities that go beyond traditional mouse and keyboard interfaces. In this paper, we want to analyse these new input devices in the context of 3D CAD modelling tasks, and compare them to more conventional interfaces to see in which scenarios they may be most productive and where challenges lie on the way ahead.

Introduction

Computer Aided Design (CAD) was one of the initial innovation drivers behind the development of novel human-computer-interfaces like the mouse from Engelbart et al. (1965) or the stylus (light pen) by Sutherland (1964). Since then, the combination of mouse and keyboard established itself as the common way to model objects in CAD software. It allows designers and engineers to select objects, draw shapes and input numbers and text data quickly and with minimal movement affecting fatigue. However, Engelbart et al. (1965) never intended it to be used in 3D. Any modelling task that requires 3D inputs relies on either special key-combinations to add 3D translations and rotations, or on dedicated 3D-input devices, like 3D-mouses or trackballs, which add additional axes of inputs. Sans such devices, the 2D mouse is still the most common input device, simply because of its ease-of-use and due to the fact that most displays are still 2D.

This situation is shifting nowadays, with modern immersive display hardware, like VR (Virtual Reality), AR (Augmented Reality) or MR (Mixed Reality) headsets or high-quality stereoscopic 3D screens being more commonly available. Here, the models are actually shown in 3D space and mouse inputs quickly fail to compensate the missing third dimension. New specialized input devices become a necessity. Similar issues are observable with portable touch devices. Tablets and phones initially just emulated the mouse and keyboard and provided mouse-like touch inputs and virtual keyboards. However, most of the efficiency of these input methods vanished in practise. Specific multi-touch input gestures and swift keyboards developed, which proved more appropriate for most modelling tasks. But none of the finger-based touch interaction concepts have become established in CAD, primarily due to their inherent precision issues (Benko and Wigdor, 2010).

AR and VR systems on the other hand are recently experiencing a move away from physical controllers and towards fully hand-based interactions. These hand tracking systems are either used to simulate a full hand within the virtual space, in which the physicalized live shape of the hand is pressing buttons and touching elements (Meta Quest), or they recognize specific hand gestures and translate these into space, for example by combination with eye gaze (Apple Vision Pro, Microsoft HoloLens). While selection tasks and other simple interactions can be effectively represented, inputs that require high precision can be an issue. Additionally, the more specific interaction techniques are, the more difficult it is to achieve broad support for them in 3D CAD software.

Nevertheless, 3D headsets and touch devices have huge potential to enable new approaches to immersive design, where people in different physical locations can collaborate on a virtual model, and make changes by interacting with virtual objects through the same direct means they would use with physical models. How intuitive would it be if we could move weight bearing pillars in CAD as simply as furniture in the real world, in the process getting a feeling for the weight, the pressure, the sound absorption, and the touch of the material, and then still have the capability to position it down to millimeter precision?

To achieve this, we need to first solve the precision problem. Only once these novel interaction techniques can reliably achieve the same, or better, results than traditional interfaces will we see broad adoption in professional CAD software. For this, we have to learn from the dominant paradigms of mouse and keyboard and touch and stylus interfaces, which crucially are *bi-manual combinations* of two input modes that seems to net appropriate levels of precision. In the following sections, we want to answer the question of why that is, and how we can emulate this for arbitrary use cases involving any type of display and interaction device. To achieve this, we do the following:

- We discuss related literature from engineering design and Human-Computer Interaction (HCI)
- We define what common types of inputs and input devices exist and what their important attributes are
- We evaluate these types and devices according to several input metrics
- We apply these insights to three 3D CAD use cases
- We describe possible avenues of research that could lead to better CAD input systems

State of the Art

The question of which 3D input devices are useful for modelling tasks predates the advent of current generation AR and VR. For example, Fiorentino et al. (2010) states that “while planar features modelling performs well with mouse and keyboard, Direct Modelling benefits from bi-manual interaction, six degrees of freedom input (6DOF) and 3D snapping.” They also already distinguish between use cases, where a VR system for conceptual design has different input requirements than mechanical CAD modelling on a 2D screen. For the latter task, they consider the effectiveness of different 6DOF haptic input devices for 3D rotations, and which input methods are most effective for them. In this case, devices are distinguished by aspects like the shape of their transfer function from device to display or how the device responds to muscle force being applied to it.

The literature also features many other input-system specific measures of accuracy. There are attempts to improve typing speed on keyboards (Jiang et al., 2020), studies about the ideal way to implement pointing tasks with VR controllers (Allgaier et al., 2022), accuracy of gaze-tracked inputs (Schuetz and Fiehler, 2022), and much more. Most of these papers focus on individual tasks. Overviews of how these systems fare over all task in multi-input device for multiple purposes are much rarer.

More direct quality metrics used in HCI research are the concepts of resolution and throughput (Bérard et al., 2011). Resolution describes the smallest displacement that an input device can measure, while throughput describes the average rate of information generated by a series of movements (Fitts, 1954).

In addition to these often well-researched common input devices, various dedicated experimental interfaces exist in the literature. Lemberski and Hemmerling (2010) turned an audio mixing device into a dedicated CAD controller, in which 10 faders control 10 parameters in the CAD modelling process. The advantage of a physical interface like a fader, is that their physical position reflects their absolute value. Sharma et al. (2011) created a multimodal system in which touch-based menu controls are replaced by voice commands, and the modelling is done with touch inputs. Specifically, this is about using voice and touch in the early design stages, when exact values are not necessarily that important. Research like this highlights that we may have been trained to think that precision is important above all, as the interfaces we currently use do not distinguish between early stage and late stage modelling precision. If all a keyboard can offer is precise numerical inputs, then precise numerical inputs is how we will approach modelling. Sharma et al. (2011) write that “it is commonly observed that designers prefer pen and paper for early stages of design”, which is most likely the case for this exact reason.

Parameters of CAD Input

While the literature features many precise quantitative measurements of precision and throughput, we want to propose broader, qualitative categories like: Can this device be used in the field? Are precise inputs even possible? Does using this device for prolonged periods get tiring? For this, we first need to establish a more systematic understanding of CAD inputs that stay relevant across many classes of interaction devices. This will aid us in establishing quality metrics that can be used as the basis for creating CAD interaction systems for arbitrary use cases and display setups.

Classification of Input Tasks

Fiorentino et al. (2010) “consider CAD modelling as a sequence of picking and navigation (rotation) actions”. More generally, they define four basic classes of CAD input tasks: *navigation and travel, selection, manipulation* and *system control*. We will base our analysis on this categorization, but will differentiate between spatial manipulations and numerical manipulations. *Spatial manipulations* involve the placement of points, drawing and alteration of lines and surfaces by direct spatial input. *Numeric manipulations* involve the input of numbers in relation to some part of the CAD model, for example by constraining the length of a wall to a specific precise length. Both of these interactions can have different levels or precision, depending on input devices. A mechanical slider like found on MIDI boards could potentially offer very precise inputs, compared to a completely virtual touch slider, as there is an inherent amount of smoothing and inference when dragging a finger over a touch surface.

Figure 1 shows the input tasks with their associated set of common operations. The input types are arranged in two classes—those that have a direct reference to the model being designed, and those that are used to put the software context into the proper camera position and program state to enable the next direct modelling steps.

Each task can be resolved in different way, depending on the input devices. Specifically, we will distinguish three classifications of input devices, which we will discuss in relation to selection, spatial manipulation and numeric manipulation tasks respectively.

Spatial Reference

Selection tasks happen at different resolutions, from object-level to the level of individual points on a geometry. The aspect that distinguishes them from each other more than anything however, is whether we select points

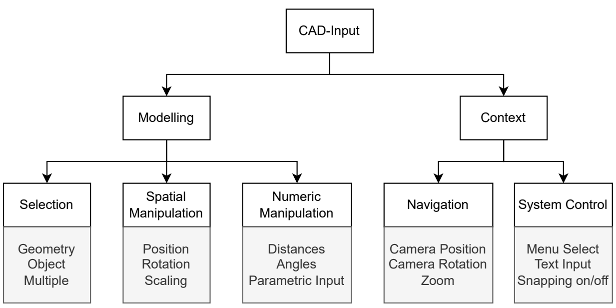


Figure 1: Hierarchical classification of CAD input tasks.

based on some projection to 2D (e. g. mouse with trackball), or whether the device can immediately reference 3D points independent of a display surface. Fig. 2 highlights this difference.

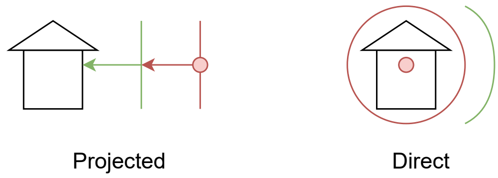


Figure 2: Green shows where the display is, red shows the interaction device, black denotes the object being selected.

Input Reference

Similar to selections, the way an input device does spatial manipulations will be most impacted by whether the device is creating a sensation of immediately interacting with the manipulated object, through a touch screen or in a midair 3D projection, or whether device movements happen on an unrelated surface or space and are somehow translated into the model. Figure 3 shows this distinction between translated and immediate inputs.

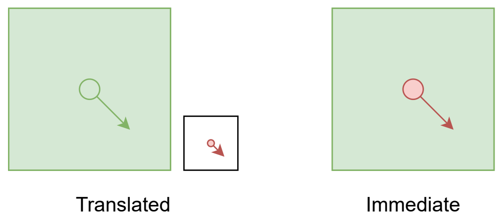


Figure 3: The green boxes represent a display device, the black box a separate zone of interaction. Red highlights the point where the input physically takes place.

Input Progression

Numeric inputs are distinguished by the basic fact of whether we are inputting the numerical value explicitly as a series of digits (usually to a high degree of accuracy), or whether we are providing a way to move as close as possible to it. The advantage of exact inputs is that the user can quickly set a-priori known values quickly—the disadvantage are case where the value is not known, like alignment cases. A system that allows for a continuous approach towards a value allows us to experiment with placements and sizes in a way that exact measurements do not. Though most relevant to numeric inputs, the capability to input numbers can have an impact on every kind of interaction, for example to select the scale at which a specific spatial operation should be resolved. Figure 4 highlights the process of these inputs over an imagined number line.

Input Devices

The main reference device for our analysis is the traditional combination of **mouse** and **keyboard**. The mouse usually acts as a combined selection ("Take *that*...") and spatial manipulation input ("...and put it *there*...") with reasonable accuracy, often supported by acceleration features. Meanwhile, the keyboard is used for numerical ma-

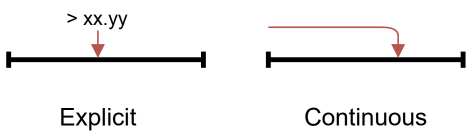


Figure 4: Numeric inputs are often collected by linear interfaces (e.g. slider). A device either moves from one side to the other (red line), or snaps to specific value on the line (red arrow)

nipulation ("...*there* being defined as *exactly*..."), but, often also provides alternative selection and manipulation key combinations. Individually and in tandem they also provide easy navigation and system control inputs. As stated in the introduction, this does not translate efficiently to 3D displays. This lack of good input devices is one of the main reasons, why we do not design in 3D[1]. Regardless of any other considerations, any alternative input system must be able to offer at least a degree of precision and throughput that gets reasonably close to what a mouse and keyboard can achieve. What is "reasonable" depends on the specific use case, as will be discussed later on in the paper.

The class of touch devices is one potential alternative interface technology. Here we use a pointing device (e. g. a stylus or our fingers) often on a surface (e. g. screen, touchpad, tablets) with different gestures (e. g. tap, one-finger slide vs. two-finger drag) to realize various inputs. The most useful distinction for our purpose here is simple **touch** vs. **multi-touch** with the finger modality, and then specific **touch devices** like a stylus, which can also be combined, like shown in Pfeuffer et al. (2021). The benefit of those inputs is that they are very intuitive and easy to learn. Traditionally those interfaces were 2D, but, nowadays are sometimes translated into the 3D world by devices like the **3D-stylus** by Jackson (2020). Sometimes, a 2D stylus is also used on a specific, external **drawing tablet**. When used like this, it is similar to a mouse, but, has additional sensors for pressure.

Then, there are devices that use more novel input systems with a touch-metaphor. This is the case for both the Meta Quest (menu controls often use hand movements as a touch-analog, like the swipe of the hand modelling a finger drag) and the Vision Pro, which both offer gesture control modes. In case of the Vision Pro, a combination of **gaze** and gestures enacts the same types of inputs like Apple touch devices, thus enabling cross-platform compatibility. The main distinction here is between **direct gestures**, which are done in direct contact with an object, and **indirect gestures**, which are done somewhere in space and area in some way related back to the object in focus. Apple (2024) makes this distinction explicit in their developer guidelines. Indirect gestures are those that work in tandem with gaze to implement common touch gestures from other platforms: Tap, Swipe, Drag, Touch (or pinch) and

[1]The lack of abstraction is another one. 2D plans actually enforce a simplification that is removing visual clutter and allowing users to focus on the relevant visual information. However, we know how to generate good, simple 2D views from 3D. So, it is technically a solved issue.

hold, Double tap, Zoom, Rotate. Direct gestures are a new addition specific to their AR platform and include the following: Touch, Touch and hold, Touch and drag, Double touch, Swipe, Pinch and drag two hands together or apart, Pinch and drag two hands in a circular motion.

A more specialized type of interaction hardware are controllers of any kind. These can include game controllers or even devices like MIDI controllers. These have in common that they combine multiple smaller input elements into a usually very ergonomic device. They feature devices like **buttons and triggers** (with just a few, sometimes analog buttons, as opposed to the many buttons on a keyboard), **analog sticks**, and **sliders and knobs**. These types of controllers are usually not tracked in space. Buttons and triggers are also usually part of other devices and are not deployed on their own. For the purposes of this paper, the class of analog sticks also encompasses any static 6DoF (Degrees of Freedom) controller, like 6DoF joysticks or 3D mouses.

A tracked extension of them are **motion controllers**, today mainly used for VR. These reduce the controller paradigm to what can comfortably be used in one hand, and add object and sometimes hand tracking to it. As such they model human hands, but other than direct hand gesture inputs they offer some translated inputs like analog sticks and triggers. Some motion controllers eschew additional controls and are instead focused on tracking and input accuracy—these types of controllers are equivalent to the aforementioned 3D stylus.

Some older VR controllers as well as controllers for room scale large display systems are only tracked in rotation to implement a laser pointer-like interaction. We will call this class of devices **3D pointing device**.

Finally, there is **Voice** control. Although this modality has fallen in and out of favor over the years, there could be great potential in using it in tandem with manual devices, specifically when these devices occupy both hands and can not by themselves cover every needed input.

Evaluation

In this section, we will take the previously established input types and devices, as well as their categorizations and evaluate them over some common metrics. This will yield a qualitative overview of which devices are appropriate for which use cases and how they can be combined.

Metrics of CAD Input

We first define some metrics to compare the different input devices in applicability to CAD design. Our metrics are based on the common quantitative metrics resolution and throughput defined for example in Bérard et al. (2011). *Resolution* is defined as the smallest possible displacement of an input device that can be attained reliably and quickly (i. e. with one short motion). *Throughput* describes, simply said, the average information an input device is capable of creating for the modelling task by a trained user. We distinguish in spatial throughput for spatial inputs, and numerical throughput for numerical inputs.

We redefine these metrics as *Precision* and *Capability* to make them qualitatively evaluateable and extend them by *Explorability* as follows:

Precision (Resolution) How precisely can we input the information into the model and how error-prone are inputs?

Capability (Throughput) How quickly can we get the needed information into the model and how much range do we have for the inputs?

Explorability Can we quickly try out variations while we input information into the model, for example by moving a control back and forth?

All apply to both spatial and numerical inputs, though they refer to slightly different aspects for both. For Selections, Navigation and System Control, we only consider a generalized *Effectiveness* metric. Explorability is a secondary metric that derives from both the throughput of a device and which input classes it falls into.

Apart from the direct input metrics, there are also metrics set by the context of use. Do we need a portable device? How many hands do we have free? Are we building robust systems for experts or trying to invite the general public to play around with modelling? Secondary metrics are:

Hands Is an interaction uni- and/or bi-manual?

Intuitive Are the interaction concepts intuitive?

Portable Can the input device quickly be used in the field?

Collaboration Does the input device allow for dynamic collaboration of multiple in one room, on the same reference display or projection?

Comfort Is the device ergonomic, and is using it not tiring even when used for hours on end?

Support How well is this type of device supported in common 3D CAD software today?

Factors like portability, software support can usually be easily determined based on experience with how these products are used in practice. Intuitivity and comfort are more complex psycho-physiological measures, but are still commonly studied metrics in the HCI literature (e.g. in Stern et al. (2008) and others referenced in Table 1).

Comparison of CAD Input Devices

With these types of inputs and metrics of quality established, we can consider which input devices are effective for which kind of inputs. A qualitative classification of this can be found in Table 1. We first compare all input devices based on the use case metrics and specify whether the device intuitivity, portability, collaborativity, comfort, and current support are low (▼), medium (◆), or high (▲). We then specify in which spatial reference, input reference and numeric input the device falls and evaluate the effectiveness of the device to execute selection, spatial inputs and numeric inputs. For the latter two we evaluate the precision, capability, and explorability individually and for the first one the overall effectiveness. We do the same for the effectiveness for navigation an system control. These evaluations are, where possible, based on results from the lit-

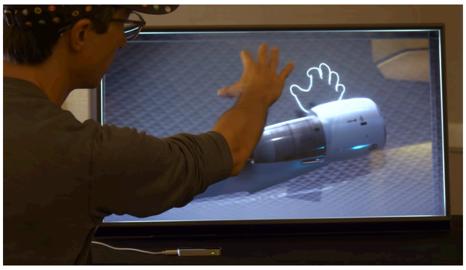


Figure 5: User interacting with a holographic display. Source: Looking Glass

erature. Example references are given at the end of the table. This table can be used to prioritize the quality measurements important for a specific use case, and quickly compare and identify appropriate input devices. Input devices can also be combined in any way until both hands are full, or if they are portable enough.

There are of course other possibilities than the ones shown in the table, like direct numerical inputs with the mouse by clicking on a virtual keyboard. However, we want to consider only inputs that are actually common in practice or could possibly become important as new display and interaction modalities are unlocked through AR and VR technologies. Many input methods allow for the use of virtual keyboard, but they will usually have lower precision (because of being error prone) than a physical keyboard, as well as much lower capability, and as such have never gotten a large amount of adoption.

Discussion

In order to show how to use this evaluation in practice, we will consider three input scenarios that are not currently common in CAD modelling, but could become more commonplace in light of recent technological advances.

Scenario 1: Single person, 3D Display

This use case is aimed at a trained user drawing up an entirely new model from scratch on a CAD workstation with a holographic or high-quality stereo display, capable of displaying models with 3D depth, or even fully volumetric displays as shown in Fig. 5. The user can see all input devices and there is no occlusion, as will be discussed in the second scenario.

The most important input types here are spatial, numeric and system control, and within the former two both precision and capability play a large role. As far as the secondary metrics go, the most important one is comfort, as expert users in such a system would be expected to work for hours uninterrupted.

According to Table 1, the following devices would be appropriate choices to fulfill all or some of the requirements:

- Mouse/Drawing Tablet (For translated inputs and system control)
- Keyboard (For numeric input and system control)
- 2D Touch Stylus (For immediate inputs)
- Sliders/Knobs (For continuous numeric inputs)
- Trigger/Buttons (For system control)

All of these devices can offer comfortable working conditions, depending on display-setup. For a touch stylus to be comfortable for hours, the display should be slanted beneath the user, something that may not be ideal when utilizing a 3D display. The issue with simply using mouse and keyboard is the fact that the display projects a 3D image. The 2D-based translated projection of the mouse will still work relatively well for selection tasks, however the translated spatial inputs are similarly projected into 2D, which will not be sufficient for a true 3D context. Instead, we need input with a high *comfort* level, a *direct* spatial reference and an *immediate* input reference. If we consider all options in Table 1, there is no combination that covers all three of these aspects perfectly. A 3D stylus would be the best option, but will not be comfortable for hours, even with for example an elbow resting on the table in front of the screen. In lieu, we could combine a mouse or drawing tablet for precision spatial inputs on the model in the current camera view, and the ability for the user to reach into the model with direct gestures whenever necessary. This way we achieve the direct and immediate inputs, without having to regularly grab a new input device.

To promote the gesture paradigm, we could include indirect gestures for system control, thus making the weak hand the "gesture" hand while the right hand is kept on the mouse. This only creates an issue in that it would be awkward to have a full keyboard in the same space in which the hand is frequently moving down onto the table for indirect gestures and up from the table for direct gestures. Here, we could instead use a special control board with sliders, knobs or buttons with the left hand, to do additional system control and to input continuous numeric values whenever necessary. If a direct numeric input is necessary, the user can use voice control to say the exact numbers needed.

Scenario 2: Single person, VR Glasses

In this input scenario we imagine an architect or civil engineer working on a CAD model that features drastic scale differences in individual parts. The user needs to be able to change camera position and scale within the world on the fly, in order to place themselves in front of a part, while also seeing as much of the surrounding context as possible. Because of the limited virtual view that a 3D display would offer, we need to utilize a VR headset to properly display the model context. This creates an issue with the input devices we used in Scenario 1: they would now be fully occluded by the virtual world. This means that the user cannot see their interaction with the input devices and could potentially lose their positioning after looking away from their focus point. Thus, all input devices that we use in this scenario should be tracked in 3D space, in order to place avatars of them into the virtual space. This is most feasible with any input device that has an immediate input reference and direct spatial reference, as well as any type of input that does not require a device at all:

- Gaze (Selection)
- VR Motion Controllers (For spatial and numerical

Table 1: Comparison of various input devices in their effectiveness for certain CAD editing tasks. For spatial and numerical inputs, effectiveness is split into the three quality measures precision/capability/explorability. We evaluate in: ▼- low, ♦- medium, ▲- high effectiveness.

	Use Case Metrics						Modelling Inputs						Context Inputs		
Input Devices	Hands	Intuitive	Portable	Collab.	Comfort	Support	Spatial Ref.	Select.	Input Ref.	Spatial	Input Prog.	Numeric	Nav.	Sys. Ctrl.	Example Ref.
2D Mouse	1	♦	▼	▼	▲	▲	Projected	▲	Translated	▲/▲/▲	Continuous	♦/▲/▲	♦	▲	Radhakrishnan et al. (2013), Besançon et al. (2017)
3D Mouse/Analog Stick	1	▼	▲	▼	▲	▲	-	-	Translated	♦/♦/▲	Continuous	▼/▲/▲	▲	▼	van Berkel et al. (2022)
Keyboard/Buttons	1 - 2	♦	▼	▼	▲	▲	-	-	-	-	Explicit	▲/▲/♦	▲	▲	Lepouras (2018)
Simple Touch	1	▲	▲	▲	♦	▲	Projected	♦	Immediate	▼/♦/▲	Continuous	▼/♦/▲	♦	♦	Besançon et al. (2017)
Multi-Touch	1 - 2	▲	▲	▲	♦	▲	Projected	♦	Immediate	♦/▲/▲	Continuous	♦/♦/▲	▲	♦	Radhakrishnan et al. (2013), Tuddenham et al. (2010)
2D Touch Stylus	1	▲	▲	♦	▲	♦	Projected	▲	Immediate	▲/▲/▲	Continuous	♦/♦/▲	♦	♦	Romat et al. (2021)
3D Stylus	1	▲	▲	♦	♦	▼	Direct	▲	Immediate	▲/▲/▲	Continuous	♦/♦/▲	♦	♦	Allgaier et al. (2022), Canavò et al. (2020)
Drawing Tablet	1	♦	▼	▼	▲	♦	Projected	▲	Translated	▲/▲/▲	Continuous	♦/▲/▲	♦	▲	Romat et al. (2021)
Gaze	0	▲	▲	▲	▲	▼	Projected	▲	Immediate	▼/♦/♦	Continuous	♦/▼/♦	▼	▼	Schuetz and Fiehler (2022), Blignaut (2017)
Indirect Gestures	1 - 2	♦	▲	▲	▲	▼	Projected	♦	Translated	▼/▼/▲	Continuous	♦/♦/▲	♦	▲	Lepouras (2018)
Direct Gestures	1 - 2	▲	▲	▲	▼	▼	Direct	▲	Immediate	♦/▲/▲	Continuous	♦/♦/▲	▲	♦	Allgaier et al. (2022)
Sliders/Knobs	1	▲	♦	▼	▲	♦	-	-	-	-	Continuous	▲/♦/▲	▼	▼	Tuddenham et al. (2010), van Berkel et al. (2022)
Trigger/Buttons	1	▲	▲	▼	▲	♦	-	-	-	-	-	-	▼	▲	Lepouras (2018)
3D Pointing Device	1	♦	▲	▲	▼	▼	Direct	▲	Immediate	▼/♦/▲	Continuous	♦/♦/▲	♦	♦	Mohr et al. (2019)
Voice	0	▲	▲	♦	♦	▼	-	-	-	-	Explicit	▲/♦/▼	▼	▲	Lepouras (2018)
Composite Devices															
VR Motion Controller	1	▲	▲	▲	▼	♦	Direct	▲	Immediate	♦/♦/▲	Continuous	♦/▲/▲	▲	♦	Allgaier et al. (2022), Canavò et al. (2020)
Gaze and Gesture	1 - 2	▲	▲	▲	▲	▼	Projected	▲	Immediate	▼/▲/▲	Continuous	♦/▲/▲	▲	♦	Lystbæk et al. (2022)

Table 2: A mapping of input tasks to predefined indirect gestures from Apples Vision Pro (Apple, 2024) and voice control.

Input	Gesture or device
Part Selection	Gaze and Tap
Navigation	Drag with one hand, Pinch and drag two hands in a circular motion
System Control	Gaze and Tap, Double Tap, Swipe
Moving and Altering Parts	Pinch and drag apart with two hands, Voice

Figure 6: VR headset and controllers in a CAD model. (Arch Virtual)

Figure 7: User interacting with a Apple Vision Pro. (Apple)

manipulation, navigation)
- 3D stylus (for precise spatial manipulation)
- Direct Gestures (Navigation)
- Indirect Gestures (For system control)
- Voice (For numeric inputs and system control)

The occlusion removes several interfaces that strongly depend on a good hand-eye-coordination like the mouse, keyboard or 2D stylus devices from our list. VR motion controllers or 3D styluses are better suited as they can be synchronized in the 3D model and users have a hand-eye-coordination feedback loop as shown in Fig. 6. One issue is that none of these devices offer high precision *and* a high comfort level. It is however unlikely that a user would wear VR glasses for a complex, multi-hour interaction process. Instead, this system is more likely to be deployed only once a high degree of spatial awareness is necessary. If longer working times are expected, we could in a first step either create more involved indirect gestures that also allow for numeric inputs, or for long-term work even create a custom tracked solution based on sliders, knobs and buttons.

Scenario 3: Collaborative, VR/AR Glasses

For this scenario, we imagine a group of designers doing design ideation around the same 3D model. They can twist and turn the model as well as highlight and switch out parts. The most important aspects here are selection, navigation and spatial manipulation. Hand-free interactions are more relevant than precision for this. While the input reference should be immediate, a direct spatial reference is not absolutely necessary. In this collaborative setting, users are expected to sit around a table in a relaxed manner, and will thus be relatively far from the model.

- Gaze (For selection)
- Indirect Gestures (For navigation and spatial manip.)
- Voice (for direct numeric input and system control)

Until recently, Microsoft's HoloLens display was the most prominent system that could implement this combination of inputs. However, display quality and depth of interaction was limited. Recently, the Vision Pro glasses by Apple have further developed this paradigm. What the Vision Pro offers us is a "Gaze and Gesture" input system as shown in Fig.7. Following Table 1, it is fit for comfortable collaboration, and can deliver reasonably well on selection, spatial manipulation and navigation tasks. We suggest a mapping of indirect gestures to common tasks in Table 2. The only issue is the precision of gestures in spatial manipulation. Placing parts in the model that is several meters away by hand gestures can be jittery and error prone. As such, these inputs should be transformed into step-wise numerical inputs with indirect gestures that mirror multi-touch gestures, and optional voice input for the highest precision.

Summary and Future Work

In this paper we review various classes of currently available input devices for CAD modelling applications. We identified commonalities and differences between these devices and categorized them according to multiple factors. Then, we established qualitative metrics for evaluating their expected performance for different input tasks. This being early stage research, the considerations were not validated quantitatively in a usability study. This is future work. We then applied these metrics to three use case scenarios that could profit from making use of novel display and interaction devices, discussing individual drawbacks of the device classes in context, and referencing specific actually existing devices where appropriate.

Establishing new input modalities is as much a process of breaking habits as it is of building effective devices. The issue that most new devices encounter, is that mouse and keyboard do most things very well, and specific 3D tasks reasonably well. A new device that does better at specific 3D modelling tasks will always have to compete with the broad usefulness and software support of paradigms like mouse and keyboard. Specific devices will be limited to specific software, a circumstance that can create enough inertia that even useful techniques do not end up in mainstream CAD software. However, in the near future we can expect a broader adoption of VR and MR devices—which has the capacity to break some of this technological inertia and thus introduce whole new device classes into the public mind. CAD software should prepare for this now, in order to exploit entirely novel possibilities offered by some of these technologies. There is potential for systems that make it easier to draw organic shapes for 3D printing applications, or inputs that natively include AI-assisted interactions, combining the precision of a manual stroke with the *intent* behind it.

References

Allgaier, M., Chheang, V., Saalfeld, P., Apilla, V., Huber, T., Huettl, F., Neyazi, B., Sandalcioglu, I. E., Hansen, C., Preim, B., et al. (2022). A comparison of input devices for precise interaction tasks in VR-based surgical planning and training. Computers in Biology and Medicine, 145:105429.

Apple (2024). Gesture specification - standard gestures. https://developer.apple.com/design/human-interface-guidelines/gestures. Accessed: 29.01.2024.

Benko, H. and Wigdor, D. (2010). Imprecision, inaccuracy, and frustration: The tale of touch input. Tabletops-Horizontal Interactive Displays, pages 249–275.

Bérard, F., Wang, G., and Cooperstock, J. R. (2011). On the limits of the human motor control precision: the search for a device's human resolution. In Int. Conf. Human-Computer Interaction.

Besançon, L., Issartel, P., Ammi, M., and Isenberg, T. (2017). Mouse, tactile, and tangible input for 3D manipulation. In CHI - Conf. on human factors in computing systems, pages 4727–4740.

Blignaut, P. (2017). Using smooth pursuit calibration for difficult-to-calibrate participants. J. Eye Movement Res., 10(4).

Cannavò, A., Calandra, D., Kehoe, A., and Lamberti, F. (2020). Evaluating consumer interaction interfaces for 3D sketching in virtual reality. In Int. Conf. on Arts, IT, Interactivity and Game Creation, pages 291–306.

Engelbart, D., English, W., and Huddart, B. (1965). Final report: Computer-aided display control. Technical report, Stanford Research Institute.

Fiorentino, M., Uva, A. E., Fabiano, M. D., and Monno, G. (2010). Improving bi-manual 3D input in CAD modelling by part rotation optimisation. Computer-Aided Design, 42(5):462–470.

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of experimental psychology, 47(6):381.

Jackson, B. (2020). Ovr stylus: Designing pen-based 3D input devices for virtual reality. In IEEE Conf. on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pages 13–18.

Jiang, X., Li, Y., Jokinen, J. P., Hirvola, V. B., Oulasvirta, A., and Ren, X. (2020). How we type: Eye and finger movement strategies in mobile typing. In CHI Conf. on human factors in computing systems, pages 1–14.

Lemberski, D. and Hemmerling, M. (2010). Mixer modeling–an intuitive design tool. In Conf. on Education in Computer Aided Architectural Design in Europe.

Lepouras, G. (2018). Comparing methods for numerical input in immersive virtual environments. Virtual Reality, 22(1):63–77.

Lystbæk, M. N., Rosenberg, P., Pfeuffer, K., Grønbæk, J. E., and Gellersen, H. (2022). Gaze-hand alignment: Combining eye gaze and mid-air pointing for interacting with menus in augmented reality. ACM on Human-Computer Interaction, 6(ETRA):1–18.

Mohr, P., Tatzgern, M., Langlotz, T., Lang, A., Schmalstieg, D., and Kalkofen, D. (2019). Trackcap: Enabling smartphones for 3D interaction on mobile head-mounted displays. In CHI - Conf. on Human Factors in Computing Systems, pages 1–11.

Pfeuffer, K., Dinc, A., Obernolte, J., Rivu, R., Abdrabou, Y., Shelter, F., Abdelrahman, Y., and Alt, F. (2021). Bi-3d: Bi-manual pen-and-touch interaction for 3d manipulation on tablets. In An. ACM Symp. on User Interface Software and Technology, pages 149–161.

Radhakrishnan, S., Lin, Y., Zeid, I., and Kamarthi, S. (2013). Finger-based multitouch interface for performing 3D CAD operations. Int. J. of Human-Computer Studies, 71(3):261–275.

Romat, H., Fender, A., Meier, M., and Holz, C. (2021). Flashpen: A high-fidelity and high-precision multi-surface pen for virtual reality. In 2021 IEEE Virtual Reality and 3D User Interfaces (VR), pages 306–315.

Schuetz, I. and Fiehler, K. (2022). Eye tracking in virtual reality: Vive pro eye spatial accuracy, precision, and calibration reliability. J. Eye Movement Res., 15(3).

Sharma, A., Madhvanath, S., Shekhawat, A., and Billinghurst, M. (2011). MozArt: a multimodal interface for conceptual 3D modeling. In Int. Conf. on multimodal interfaces, pages 307–310.

Stern, H. I., Wachs, J. P., and Edan, Y. (2008). Designing hand gesture vocabularies for natural interaction by combining psycho-physiological and recognition factors. Int. J. of Semantic Computing, 2(01):137–160.

Sutherland, I. E. (1964). Sketch pad a man-machine graphical communication system. In SHARE design automation workshop, pages 6–329.

Tuddenham, P., Kirk, D., and Izadi, S. (2010). Graspables revisited: multi-touch vs. tangible input for tabletop displays in acquisition and manipulation tasks. In SIGCHI - Conf. on human factors in computing systems.

van Berkel, N., Merritt, T., Bruun, A., and Skov, M. B. (2022). Tangible self-report devices: accuracy and resolution of participant input. In Int. Conf. on Tangible, Embedded, and Embodied Interaction, pages 1–14.

A FIRST EVALUATION OF THE SEAMLESS MARKERLESS AUGMENTED REALITY REGISTRATION SYSTEM SUPPORTING FACILITY MANAGEMENT

Leonardo Messi, Francesco Spegni, Massimo Vaccarini, Alessandra Corneli, and Leonardo Binni
Università Politecnica delle Marche (UNIVPM), Faculty of Engineering, DICEA Department, Ancona, Italy

Abstract

Augmented reality (AR), despite its great potential, still struggles to be widely used in real construction processes due to difficulties in registering holograms in mixed indoor-outdoor scenarios. Since a definitive technological solution for inside-out AR registration does not exist yet, a seamless markerless augmented reality registration system, integrating real-time kinematic positioning (RTK), inertial measurement units (IMU) technologies and image comparison based on convolutional neural networks (CNN), has been proposed by the authors. Experiment results have shown the need to continuously register AR at regular temporal interval within 1 seconds and/or 1 meter to achieve "fine-precision" positional accuracy (i.e., 0.10 m).

Introduction

The architecture, engineering, construction, and operations (AECO) sector, despite its reputation as one of the least digitized industries, is progressively embracing digital technologies to enhance various stages of a building's life cycle (Albahbah, Kıvrak, & Arslan, 2021). Among these stages, the operation and maintenance (O&M) phase in facility management (FM) incurs the most significant costs, accounting for 50–70% of annual facility operating costs and a whopping 85% of the overall lifecycle expense (Salman & Ahmad, 2023). As facilities become increasingly intricate, day-to-day tasks have become more demanding. This heightened complexity has propelled the construction industry pursuit of visualization technologies to facilitate better access to information for evaluation, communication, and collaboration.

Augmented reality (AR) technology emerges as a promising tool for streamlining O&M tasks, bridging the gap between facility managers and field workers, thereby minimizing delays and reducing operational expenses (Salman & Ahmad, 2023). AR technology revolutionizes human-computer interactions by establishing direct links between physical and digital realms. Its ability to overlay virtual objects onto the real world allows for the identification of concealed facilities and provides detailed maintenance guidance for field workers (Salman & Ahmad, 2023). However, integrating AR poses technical challenges. While display technology has made significant strides, achieving precise position tracking remains a major obstacle. Spatial registration, ensuring accurate alignment between virtual and real-world coordinates, stands as a critical aspect of AR functionality (Albahbah et al., 2021; Salman & Ahmad, 2023). Spatial registration methods can be categorized as "marker-based" or "markerless" (Albahbah et al., 2021; El Barhoumi, Hajji, Bouali, Ben Brahim, & Kharroubi, 2022; Salman & Ahmad, 2023). Marker-based approaches register virtual models using some kind of physical or hyperlink markers (El Barhoumi et al., 2022). These solutions can be potentially applied in mixed environments, but they generally require a preliminary survey, a manual setup and markers often generate aesthetic issues (Baek, Ha, & Kim, 2019). On the other hand, markerless methods, such as Global Navigation Satellite Systems (GNSS) or image-based localization, solve the previous gaps automating the AR registration process (Baek et al., 2019). However, the former cannot be applied indoor (Chen et al., 2019) and the latter requires a preliminary survey for collecting reference images (Baek et al., 2019; Messi, Spegni, Vaccarini, Corneli, & Binni, 2023) which restricts its use in large outdoor scenarios.

An innovative seamless inside-out and markerless localization system has been developed by the authors in (Messi et al., 2023), housing GNSS-based and image-based registration engines on an AR cloud platform, complemented by a switch engine managing their priorities. Since initial qualitative testing on a university campus has shown promising outcomes, AR registration performances need to be quantified in terms of drifts from an assumed benchmark. Hence, this paper serves as a follow-up to the one presented by the authors in (Messi et al., 2023) and aims to conduct initial assessments of the tracking engines embedded in the proposed system. These assessments provide a first evaluation of the capabilities of our seamless and markerless AR registration system (Messi et al., 2023). Existing studies report different registration accuracy thresholds, such as the "fine-precision" corresponding to 1°/0.10 m, and "high-precision" corresponding to 2°/0.25 m. The first one is considered as the minimum accuracy required for a good AR user experience in most settings (Sarlin et al., 2022; Sattler et al., 2017).

This paper progresses as follows. The Research questions section formalizes the objectives of the study. Then, the Methodology section presents the proposed system architecture and the evaluation method. The Experiments section describes the experiments carried out for the assessment of the drift and related results, followed by the Conclusions section that closes the paper.

Research questions

This study, since the seamless and markerless AR registration system presented by the authors has shown promising outcomes (Messi et al., 2023), carries out an accuracy assessment aimed to quantify its reliability. More in detail, the following research questions have been formalized:

RQ.1. What is the drift of the proposed system during FM operations?

RQ.2. What spatial/temporal AR registration frequency should be ensured to optimize the performance of the proposed system?

Methodology

This section presents the methodology applied by this study. To this purpose, the System architecture section will first recap the architecture of the markerless seamless AR system presented in (Messi et al., 2023), whereas the Evaluation method section describes methodology applied to assess drift issues related to the different registration engines embedded by the system.

System architecture

The architecture of the proposed system (Messi et al., 2023) is established on an AR cloud platform housing four essential elements: (i) the graph database, (ii) the GNSS-based AR registration engine, (iii) the image-based AR registration engine, and (iv) the switch engine (Figure 1). Although this system architecture has been defined for Microsoft HoloLens 2, it applies to all AR devices.

The AR cloud platform serves as a centralized resource for data processing, storage, and distribution. It enables precise mapping of virtual assets to real-world locations, crucial for integrating the virtual and physical realms. Images and point clouds are georeferenced and aligned within this platform using absolute world coordinates and accurate rotations. The graph database manages structured and unstructured data storage, allowing AR application accessibility. The GNSS-based AR registration engine utilizes RTK GNSS and IMU systems for open-space challenges, while the image-based AR registration engine employs convolutional neural networks (CNN) for indoor environments lacking reliable GNSS signals. The switch engine integrates the outdoor and indoor registration engines, facilitating uninterrupted AR experiences by dynamically switching between registration approaches.

The process for GNSS-based AR registration involves aligning the HoloLens 2 local frame to global coordinates using RTK measurements. The system computes the device position and orientation relative to geographical coordinates, enabling precise placement of building information models (BIMs), stored within the graph database. This process accounts for changing observer positions and altitudes, ensuring accurate object positioning and alignment with the local frame. Conversely, image-based AR registration relies on image comparison with survey data collected via cameras and LiDAR scanners. The process involves localizing the HoloLens 2 by comparing its current view with referenced images and related point clouds stored in the AR cloud platform (Sarlin, Cadena, Siegwart, & Dymczyk, 2018). CNNs, namely Hierarchical Feature Network (HF-Net) technology (Sarlin, Cadena, Siegwart, & Dymczyk, 2018), and the Perspective-n-Point (PnP) algorithm (OpenCV, 2023) are applied for accurate 6-degrees-of-freedom (6-DoF) localization. In this way, BIM models, stored within the graph database, can be overlapped to real world. The switch engine seamlessly integrates both registration approaches based on the scenario: RTK-only (mainly outdoor), RTK with images (mainly outdoor), and images only (mainly indoor). It identifies and triggers the appropriate registration engine based on available data, ensuring a continuous AR registration experience.

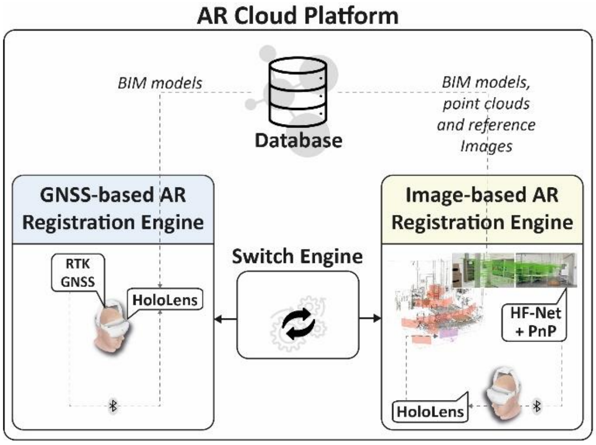


Figure 1: System architecture of the proposed system for seamless inside-out and markerless AR registration.

Evaluation method

For the purpose of a first evaluation of the reliability of the proposed markerless and seamless AR system, a methodology has been defined. It comprises a series of steps aimed at assessing drift values among the tracking engines of the system. More in detail, the methodology defined in this section aims to compare the HoloLens 2 tracking system with both the tracking engines of the system, namely the GNSS-based and the image-based engines. The evaluation steps are reported as follows:

1. **Tracking positions along a route:** this step consists in tracking the AR device position at predefined intervals while navigating a route that includes both outdoors and indoors scenarios. The AR device tracking is carried out using all the different tracking systems described above (i.e., HoloLens 2, GNSS-based, and image-based). The HoloLens 2 and GNSS-RTK tracking devices must be rigidly connected to each other.
2. **Registering HoloLens 2 reference system to the North direction:** this step aims to align the HoloLens 2 system to the North direction using RTK data, by computing the forward azimuth (Movable Type Scripts, 2023) from displacements.

3. **Aligning the image-based cloud system to the HoloLens 2 one:** by aligning the image-based cloud system to the HoloLens 2 one, this step ensures that their x, y, z axes correspond each other and are aligned to the North direction. This alignment is achieved using the position of three reference photos (Schönberger & Frahm, 2016) solely for the purpose of alignment.
4. **Data retrieval of the image-based tracking system:** this step involves calculating displacement vectors (i.e., "tvec" values (OpenCV, 2023; Schönberger & Frahm, 2016) and quaternions (i.e., "qvec" or "rvec" values (OpenCV, 2023; Schönberger & Frahm, 2016)) of the acquired images (i.e., their pose) with reference to the HoloLens 2 reference system.
5. **Data retrieval of the GNSS-RTK tracking system:** this step involves retrieving global coordinates from the GNSS-RTK system, transforming them into x, y, z coordinates using equirectangular projections (ArchGIS Esri, 2023), and computing the forward azimuth (Movable Type Scripts, 2023) of the tracking direction.
6. **Registering and assessing drift between HoloLens 2 and GNSS-RTK tracking:** this step involves calculating the discrepancy between the HoloLens 2 and the GNSS-RTK systems once aligned on the equirectangular projection.
7. **Registering and assessing drift between HoloLens 2 and image-based tracking:** similar to the previous one, this step involves computing the discrepancy between the HoloLens 2 and image-based systems once aligned.
8. **Evaluating drift at regular intervals:** repeating the discrepancy calculations at either specific time intervals (e.g., every 1, 2, 3 s) or distance covered (e.g., every 5, 10, 20 m) to evaluate drift values and assess the best registration frequency.

Experiments

The evaluation method introduced in the previous Evaluation method subsection has been applied to carry out a first reliability assessment of the proposed markerless and seamless AR system. Experiments have been executed assuming a FM use case based on a university campus as a case study. Specifically, the FM of the Digital Construction Capability Centre (DC3) Lab at the Università Politecnica delle Marche has been considered. The DC3 Lab, which covers an area as large as 240 m^2, is composed of a main open space, a changing room, an office, and a restroom. Within this context, the management of the electrical system and, in particular, of the internal electrical panel of the DC3 Lab, has been considered. During this activity, the technician in charge of FM operations spends time first locating the electrical panel. Then, in order to find the root cause of the problem, the technician may be asked to locate cablings associated to the panel, which extend externally to the building. These cablings can be accessed through manholes located on the road in front of the building, connecting it to the rest of the campus (Figure 2).

By testing the system proposed by the authors in (Messi et al., 2023) with the presented FM use case, its reliability in heterogeneous indoor/outdoor scenarios will be evaluated, considering the "fine-precision" positional accuracy threshold (i.e., 0.10 m) and "high-precision" one (i.e., 0.25 m) (Sarlin et al., 2022; Sattler et al., 2017).

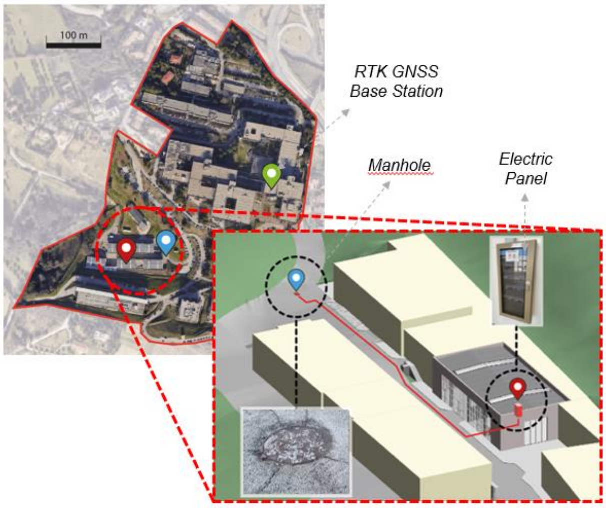


Figure 2: Aerial view of the university campus, assumed as case study, identifying the positions of the DC3 Lab, and the RTK base; The focus on the BIM model of the laboratory shows the positions of the indoor electrical panel and the outdoor manhole cover that are the objective of the simulated FM use case.

Experiments design and execution

In this study, we applied the methodology exposed in the previous Evaluation method subsection to the case study described at the beginning of Experiments section for assessing the outdoor scenario. The indoor scenario and, therefore, the evaluation of drift between the image-based and HoloLens 2 tracking systems will be addressed by future studies. Specifically, in this study, aiming to evaluate drifts between the GNSS-RTK and HoloLens 2 tracking systems (RQ.1) to understand the optimal registration frequency (RQ.2), we implemented points 1, 2, 5, 6, 8 of the proposed methodology (i.e., Evaluation method subsection). Regarding point 1, it should be noted that the conducted experiment simulates the maintenance activity as described at the beginning of Experiments section, following a path in a mixed indoor-outdoor scenario, using the HoloLens 2 device with a rigid add-on hosting the RTK receiver (Figure 3).

With these premises, the experiment was conducted in two parts. The first part involved comparing the GNSS-RTK and HoloLens 2 tracking systems in a mixed indoor-outdoor path, as indicated in point 1 of the methodology. Following point 6 of the methodology, the registration between the tracking systems was initially performed at a single point to assess the order of magnitude of drift over time and space. The second part of the experiment details the first part by comparing the same tracking systems and by following an only-outdoor path, executing multiple registrations at predefined time intervals. Specifically, registrations were executed each 20 s during the tracked simulation to assess the drift across several samples, aiming to deduce the minimum necessary frequency of tracking system registration while maintaining "fine-precision" or "high-precision" positional accuracy (Sarlin et al., 2022; Sattler et al., 2017).

Figure 3: View of the operator simulating FM activities using the HoloLens 2 with the RTK add-on.

Results and discussion

The results of the first part of the experiment, as described in the Experiments design and execution subsection, are presented in Figure 4 and Figure 5. Figure 4 displays the 2D layout of the DC3 laboratory and the trajectories recorded by both the GNSS-RTK tracking system (depicted in orange) and the HoloLens 2 (depicted in blue) during the simulation of FM tasks conducted by an operator. The size variation of the markers along the GNSS-RTK-traced path signifies positional accuracy, where larger markers correspond to decreased accuracy.

The registration of the two tracking systems occurred solely at point A, i.e. at the beginning of the route, and only once (Figure 4). Subsequent movements within the laboratory space (zone B) revealed notable inadequacies in the GNSS-RTK tracking system indoors, indicating a substantial decline in accuracy. Further observations highlight zone C, illustrating an analogous scenario to an urban-canyon setting, where the GNSS-RTK system exhibited reduced reliability, potentially favoring the application of the image-based registration system [1]. Upon assessing the external trajectory (zone D), where the RTK system has a centimeter-level accuracy as declared by manufacturers (ublox, 2023), the HoloLens 2 system unreliability outdoors became conspicuous. It noticeably diverged from the RTK data when venturing beyond a certain distance from the registration point, indicating a drift of around 1 m each 45 m traversed (Figure 5). Therefore, the first part of the experiment, while answering RQ.1, demonstrates the need for regular registrations of the tracking systems to contain drift.

In the second part of the experiment, whose results are detailed below and in Figure 6, the outdoor operator route mimics the one of zone D in Figure 4. Along this route, the HoloLens 2 tracking system was registered against the GNSS-RTK at 6 subsequent intervals, spaced 20 s apart, creating 6 subsets of the total path. Drift readings for each subset were registered 1 s apart (Figure 6). These recorded subsets are overlaid in Figure 6, displaying a relatively consistent trend. It is worth noting that only the initial 15 m of each subset are depicted in Figure 6, as values beyond this range prove less significant. In Figure 6, the "Subset4" (light green curve), registering the worst performance, shows that the drift remains below 0.10 m (i.e., "fine-precision" positional accuracy threshold) only within about 2 m of covered distance (around 1.5 s of tracking). Instead, on the overall, the trendline of all subsets (red dashed line) shows that the drift remains below 0.10 m within about 4 m of covered distance. Furthermore, the "Subset4" shows that the drift stays below 0.25 m (i.e., "high-precision" positional accuracy threshold) within around 7 m of covered distance (approximately 6 s of tracking). Instead, on the overall, the trendline of all subsets shows that the drift remains below 0.25 m within around 9 m travelled. This result further answers RQ.1.

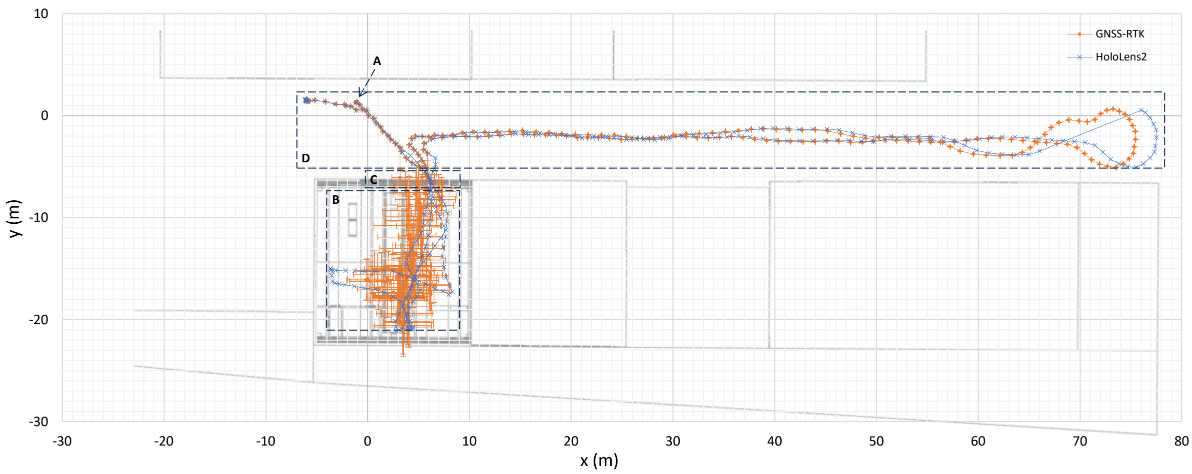


Figure 4: 2D view of the paths travelled by the operator and tracked by the GNSS-RTK (orange) and HoloLens 2 (blue) tracking systems. The grid has a mesh size of 1 m. Smaller orange markers along the GNSS-RTK route indicate higher accuracy of the samples.

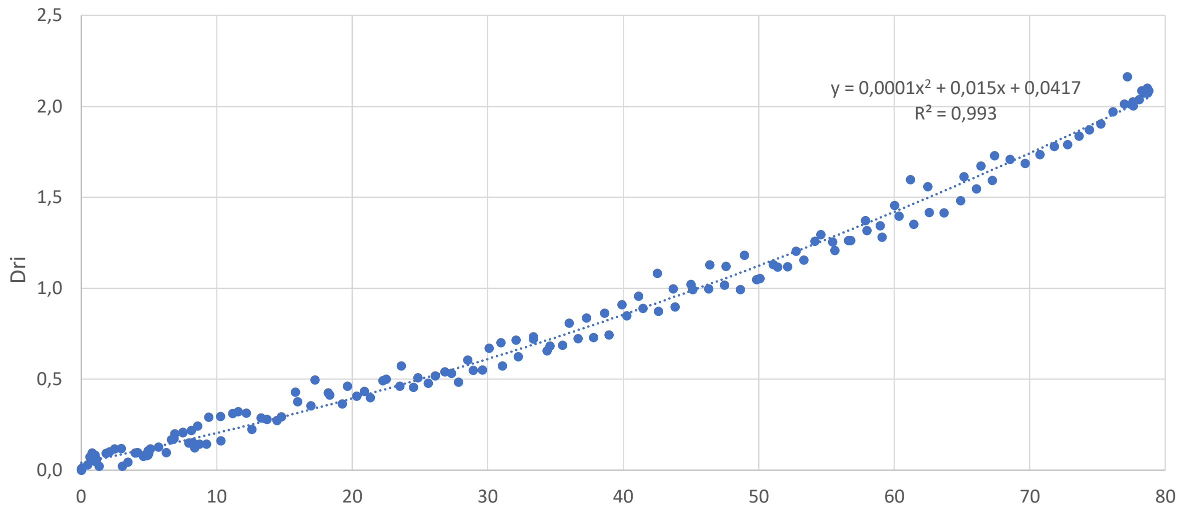


Figure 5: Outdoor assessment of HoloLens 2 drift against GNSS-RTK ground-truth localization (only zone D). The system is spatially registered only once at the beginning of the route.

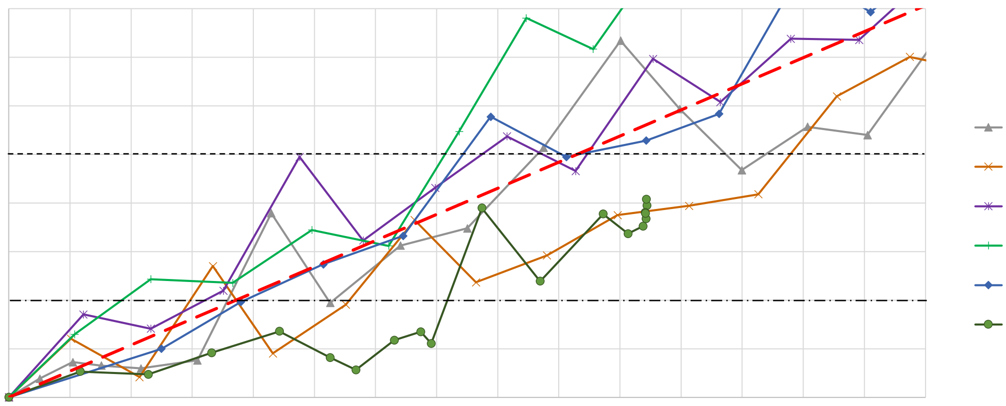


Figure 6: HoloLens 2 drift against GNSS-RTK ground-truth for each subset and trendline (red dashed line), related to the path traversed during the second experiment (only outdoor scenario – zone D). Horizontal dashed and dashed-dotted grey lines represent the "high-precision" and "fine-precision" positional accuracy thresholds at 0.25 m and 0.10 m respectively.

To sum up, for achieving in all outdoor circumstances a "fine-precision" positional accuracy, required for a good AR user experience (Sarlin et al., 2022; Sattler et al., 2017), the results underscore the need of aligning the HoloLens 2 tracking system with the GNSS-RTK approximately every 1.5 s or 2 m of travelled distance. Instead, to ensure the "fine-precision" positional accuracy on the overall, the results underscore the need of aligning the HoloLens 2 tracking system with the GNSS-RTK approximately every 4 m of travelled distance. This result answer RQ.2. It must be noted that, in this study, the RTK time latency was considered and the existing delay between the acquisition of GNSS-RTK samples and those from the HoloLens 2 was compensated.

Conclusions

Despite its immense potential, AR encounters barriers hindering its widespread implementation in real construction processes. One of the primary obstacles lies in the difficulty of aligning and sustaining the stability of holograms in mixed scenarios. The seamless integration of indoor-outdoor AR registration systems is imperative for effective built environment management.

The authors proposed a markerless augmented reality registration system that integrates RTK, IMU technologies, and image comparison based on CNN. Such system was qualitatively tested on a FM use case related to a university campus (Messi et al., 2023). This paper extends preliminary tests conducted on the proposed AR registration system. The aim of this study is quantifying drifts experienced within the different tracking systems embedded in the AR registration engines and determine the optimal registration frequency required for practicality in real FM scenarios. Experiment results have provided insights into the system performance, highlighting its capabilities, limitations, and areas of refinement. Such results can be summarized as follows:

- HoloLens 2 tracking system, as shown by the first part of the experiments, exhibited significant drift outdoors, reaching about 1 m per 45 m traversed

(Figure 4 and Figure 5). This result, while answering RQ.1, confirmed the need of periodic spatial registrations to contain drifts between the tracking systems.

- The second part of the outdoor experiments showed, in the worst subset, drifts below 0.10 m within around 1.5 s or 2 m of travelled distance and below 0.25 m within around 6 s and 7 m traversed (Figure 6). The overall trend of all subsets, instead, shows a drift below 0.10 m within about 4 m and below 0.25 m within about 9 m of travelled distance. These results extend the answer to RQ.1 and highlight the critical need of spatial registration of HoloLens 2 against GNSS-RTK.
- Finally, for achieving in all outdoor circumstances a "fine-precision" positional accuracy, required for a good AR user experience, aligning HoloLens 2 with GNSS-RTK data every 1.5 s or 2 m of travelled distance is crucial. Overall acceptable results can be achieved by aligning HoloLens 2 with GNSS-RTK data every 4 m of travelled distance. These results answer RQ.2.

Future works will focus on the assessment of the drift between image-based and HoloLens 2 tracking systems as explained in point 3, 4 and 7 of the presented methodology (Evaluation method section). Then, data fusion techniques should be applied to combine the different data sources and achieve better results in scenarios served by both the GNSS-RTK and image-based tracking systems.

Acknowledgements

This research has received funding from the project Vitality – Project Code ECS00000041, CUP I33C22001330007 - funded under the National Recovery and Resilience Plan (NRRP), Mission 4 Component 2 Investment 1.5 - 'Creation and strengthening of innovation ecosystems,' construction of 'territorial leaders in R&D' – Innovation Ecosystems - Project 'Innovation, digitalization and sustainability for the diffused economy in Central Italy – VITALITY' Call for tender No. 3277 of 30/12/2021, and Concession Decree No. 0001057.23-06-2022 of Italian Ministry of University funded by the European Union – NextGenerationEU.

References

Albahbah, M., Kıvrak, S., & Arslan, G. (2021). Application areas of augmented reality and virtual reality in construction project management: A scoping review. *Journal of Construction Engineering*, *4*, 151–172. https://doi.org/10.31462/jcemi.2021.03151172

ArchGIS Esri. (2023). Equirectangular. Retrieved December 22, 2023, from https://desktop.arcgis.com/en/arcmap/10.3/guide-books/map-projections/equirectangular.htm

Baek, F., Ha, I., & Kim, H. (2019). Augmented reality system for facility management using image-based indoor localization. *Automation in Construction*, *99*, 18–26. https://doi.org/10.1016/j.autcon.2018.11.034

Chen, K., Chen, W., Fellow, P., Li, C. T., Student, M., & Cheng, J. C. P. (2019). A BIM-based location aware AR collaborative framework for facility maintenance management. In *Journal of Information Technology in Construction (ITcon)* (Vol. 24). Retrieved from http://www.itcon.org/2019/19

El Barhoumi, N., Hajji, R., Bouali, Z., Ben Brahim, Y., & Kharroubi, A. (2022). Assessment of 3D Models Placement Methods in Augmented Reality. *Applied Sciences (Switzerland)*, *12*(20). https://doi.org/10.3390/app122010620

Messi, L., Spegni, F., Vaccarini, M., Corneli, A., & Binni, L. (2023). Seamless Indoor/Outdoor marker-less Augmented Reality registration supporting Facility Management operations. *23° International Conference on Construction Applications of Virtual Reality "Managing the Digital Transformation of Construction Industry."*

Movable Type Scripts. (2023). Calculate distance, bearing and more between Latitude/Longitude points. Retrieved December 23, 2023, from https://www.movable-type.co.uk/scripts/latlong.html

OpenCV. (2023). Perspective-n-Point (PnP) pose computation. Retrieved December 22, 2023, from https://docs.opencv.org/4.x/d5/d1f/calib3d_solvePnP.html

Salman, A., & Ahmad, W. (2023). Implementation of augmented reality and mixed reality applications for smart facilities management: a systematic review. *Smart and Sustainable Built Environment*. https://doi.org/10.1108/SASBE-11-2022-0254

Sarlin, P.-E., Cadena, C., Siegwart, R., & Dymczyk, M. (2018). *From Coarse to Fine: Robust Hierarchical Localization at Large Scale*. Retrieved from http://arxiv.org/abs/1812.03506

Sarlin, P.-E., Dusmanu, M., Schönberger, J. L., Speciale, P., Gruber, L., Larsson, V., … Pollefeys, M. (2022). *LaMAR: Benchmarking Localization and Mapping for Augmented Reality*. Retrieved from http://arxiv.org/abs/2210.10770

Sattler, T., Maddern, W., Toft, C., Torii, A., Hammarstrand, L., Stenborg, E., … Pajdla, T. (2017, July 27). *Benchmarking 6DOF Outdoor Visual Localization in Changing Conditions*. Retrieved from http://arxiv.org/abs/1707.09092

Schönberger, J. L., & Frahm, J.-M. (2016). *Structure-from-Motion Revisited*. Retrieved from https://github.com/colmap/colmap.

ublox. (2023). ZED-F9P module. Retrieved December 22, 2023, from https://www.u-blox.com/en/product/zed-f9p-module