# Generating metadata schema for data-driven smart buildings

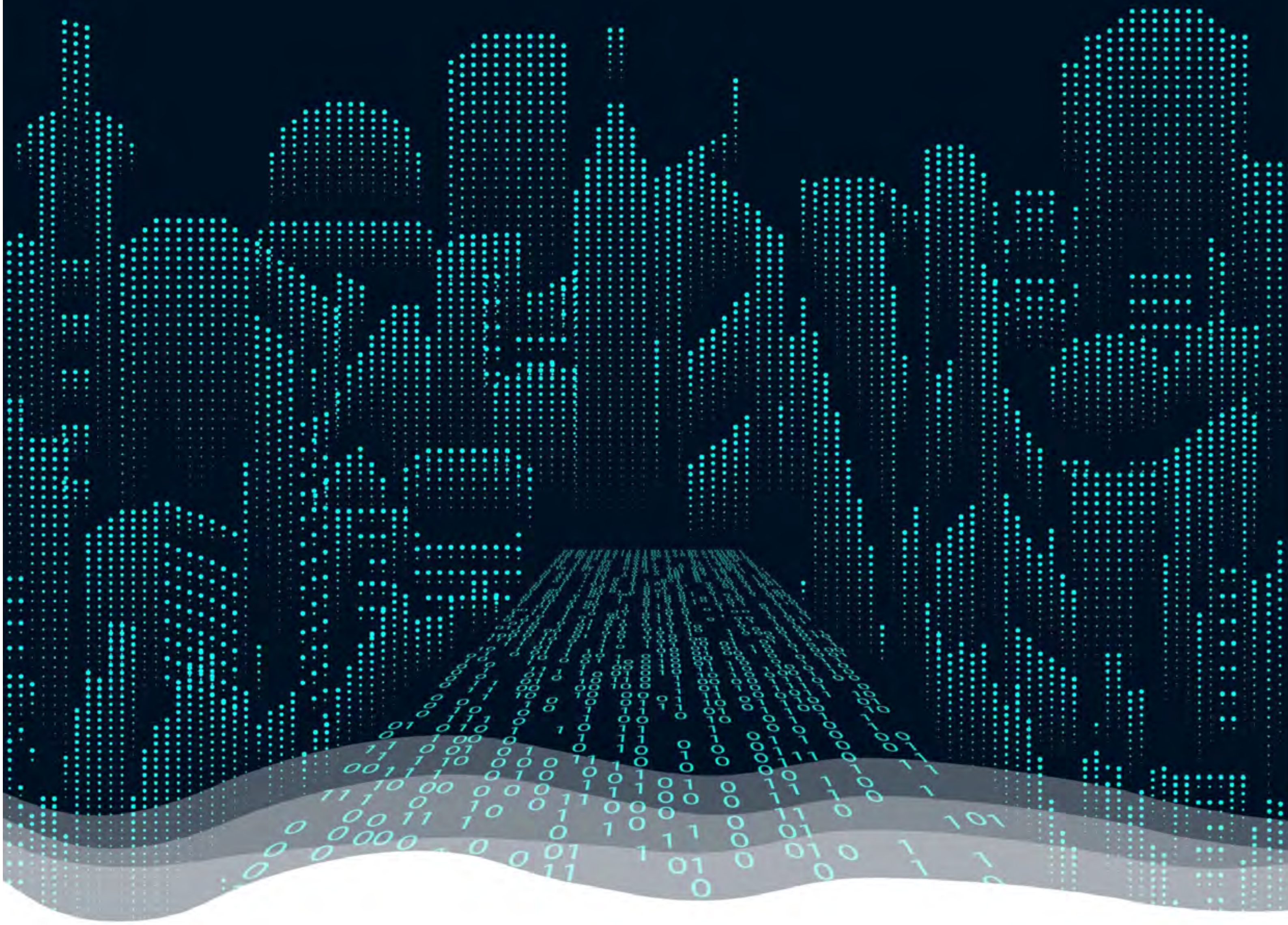
Chamari, L.; van der Weijden, J.A.M.; Boonstra, L.; Hoekstra, W.S.; Petrova, E.; Pauwels, P.

**Publication date**
2023

**Document Version**
Final published version

**Published in**
Proceedings of the 2nd Research Day on Digitalization of the Built Environment

# PROCEEDINGS OF THE 2nd RESEARCH DAY ON DIGITALIZATION OF THE BUILT ENVIRONMENT



**27-MARCH-2023**
**EINDHOVEN**
**THE NETHERLANDS**

**EDITED BY:**
**EKATERINA PETROVA**
**PIETER PAUWELS**

# 2nd 4TU/14UAS Research Day on Digitalization of the Built Environment

## Proceedings of the 2nd 4TU/14 UAS Research Day on Digitalization of the Built Environment

March 27, 2023

Eindhoven, the Netherlands

Edited by

**Dr. Ekaterina Petrova,** Eindhoven University of Technology

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

Published at Eindhoven University of Technology, Eindhoven, the Netherlands.

**To cite this publication:**

Petrova, E. & Pauwels, P. (2023). *Proceedings of the 2nd 4TU/14UAS Research Day on Digitalization of the Built Environment*, March 27, 2023, Eindhoven, the Netherlands.

A catalogue record is available from the Eindhoven University of Technology Library.

ISBN: 978-90-386-5717-2

# Generating metadata schema for data-driven smart buildings

*Lasitha Chamari, Eindhoven University of Technology, l.c.rathnayaka.mudiyanselage@tue.nl*
*Joep van der Weijden, Delft University of Technology, J.A.M.vanderWeijden@tudelft.nl*
*Lolke Boonstra, Delft University of Technology, L.Boonstra@tudelft.nl*
*Stefan Hoekstra, Delft University of Technology, W.S.Hoekstra@tudelft.nl*
*Ekaterina Petrova, Eindhoven University of Technology, e.petrova@tue.nl*
*Pieter Pauwels, Eindhoven University of Technology, p.pauwels@tue.nl*

***Abstract*** *– The integration of data from a variety of systems is critical for developing smart building applications. Design, construction, maintenance, energy management, and automation data are among few. For this integration, having a comprehensive metadata schema is crucial. Semantic web technologies and domain ontologies have been proposed as a means of modeling and linking domains and their relationships. This study proposes a semi-automatic metadata schema generator that combines an ontology database and a text search engine to generate a metadata schema. The proposed method is tested for a real-world building, and the resulting metadata schema is used to link timeseries data to the sensors and equipment in the building automation system.*

Semantic web and domain ontologies enable modeling and linking relationships between different domains. However, the process of generating a metadata schema for a specific building using ontologies cannot be generalized and need case by case approach. Previous attempts include point matching using occupant's inputs (Fürst et al., 2016), regular expressions to detect common patterns in metadata descriptors (Bhattacharya et al., 2015), utilize time-series values from Building Automation System (BAS) points to learn the mapping (Gao & Bergés, 2018), linguistic matching (Schumann et al., 2014), and using a combination of tags and time-series data (Balaji et al., 2015). However, not all buildings have accessible timeseries data storages, cannot use occupants as an input, or a combination of above. This study proposes a semi-automatic metadata schema generation process based on the metadata extracted from the BAS. The proposed method uses an ontology database and a text search engine to generate the scheme. This method is tested on a campus building and a metadata schema was generated successfully. Finally, timeseries data of the BAS is visualized using a web application using the generated scheme.

Table 1 presents a list of BAS object identifiers. The "Name" column contains a naming convention consisting of four elements, representing the building number (33), system number (201), control code (CV), and point type (V). Although the "Description" column partly describes the point's full details, it fails to convey that point in #1 pertains to AHU 201. As such, we chose to deconstruct the naming convention in the "Name" column to extract all metadata related to the point by using three additional mapping tables that detail the system number, control code, and point type. These three tables were initially in Dutch and were subsequently translated to English using the googletrans API. An example of this decomposition is illustrated in the "Description (EN)" column of Figure 2.

Table 1: Metadata extraction from BMS

| # | Item Reference | Object ID | Object Type | Name | Description (NL) | Description (EN) |
|---|---|---|---|---|---|---|
| 1 | XXX.FEC005.CLG-O | CLG-O | AO Mapper | (33) 201.CV-02V-- | Regelafsluiter koeler | Cooler control valve |
| 2 | XXX.FEC006.CLG-O | CLG-O | AO Mapper | (33) 202.CV-02V-- | Regelafsluiter koeler | Cooler control valve |
| 3 | XXX.SHWP1-FAULT | SHWP1-FAULT | BI Mapper | (33) 001.TP-01A-- | Transportpomp 1 storing | Transport pump 1 malfunction |

Each of the three mapping sets was matched with its corresponding Brick counterpart individually, using a text search engine[1]. The search engine was populated with Brick class names and their definitions extracted from the Brick ontology[2], which contain the necessary keywords to locate a match with the point names in the three mapping sets, as depicted in Figure 2. However, some object identifiers were not entirely recognized by the text search engine, expert human input was necessary. We then identified a logical pattern among the three identifiers in the subsequent phase. Using Brick relationships, we connected these three identifiers, as shown in Figure 2.

Out of the total of 2338 BAS object identifiers in the building, only 948 points had linked time series records. Of those 948 points, only 763 conformed to the logical naming structure and could be successfully incorporated into an RDF-based metadata schema using Brick classes and relationships. This scheme is used to link the BAS points and their time series data in the real time data stream and historical data storage. In conclusion, the metadata schema generation method was successful, and the resulting schema can be used to support data driven applications. Further expansion is possible to integrate with BIM and use real time monitoring applications. Figure 3 shows the Grafana[3] web application showing the historical data of the sensors connected to BAS. Using the graph generated greatly improves the efficiency of querying the historical data.

[1] https://docs.meilisearch.com/
[2] https://brickschema.org/ontology
[3] https://grafana.com/

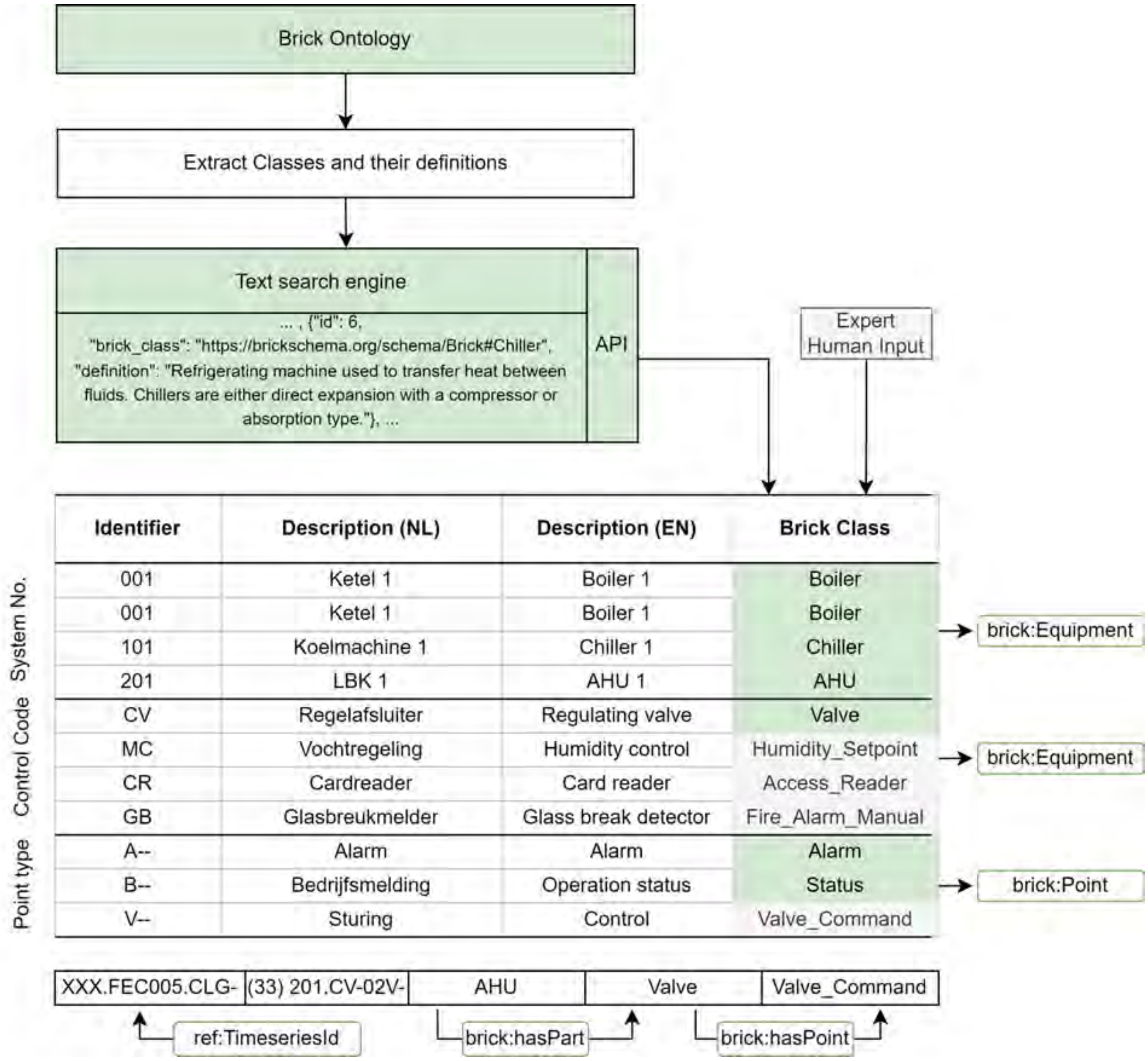



Figure 2 Workflow of mapping to Brick Ontology


Figure 3 Visualizing time series data using Grafana web application using the metadata schema.

## References

Balaji, B., Verma, C., Narayanaswamy, B., & Agarwal, Y. (2015). Zodiac: Organizing large deployment of sensors to create reusable applications for buildings. *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built*, 13–22. https://doi.org/10.1145/2821650.2821674

Bhattacharya, A. A., Hong, D., Culler, D., Ortiz, J., Whitehouse, K., & Wu, E. (2015). Automated metadata construction to support portable building applications. *BuildSys 2015 - Proceedings of the 2nd ACM International Conference on Embedded Systems for Energy-Efficient Built*, 3–12. https://doi.org/10.1145/2821650.2821667

Fürst, J., Chen, K., Katz, R. H., & Bonnet, P. (2016). Crowd-sourced BMS point matching and metadata maintenance with Babel. *2016 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops 2016*, 1-6. https://doi.org/10.1109/PERCOMW.2016.7457102

Gao, J., & Bergés, M. (2018). A large-scale evaluation of automated metadata inference approaches on sensors from air handling units. *Advanced Engineering Informatics*, *37*(February), 14–30. https://doi.org/10.1016/j.aei.2018.04.010

Schumann, A., Ploennigs, J., & Gorman, B. (2014). Towards automating the deployment of energy saving approaches in buildings. *BuildSys 2014 - Proceedings of the 1st ACM Conference on Embedded Systems for Energy-Efficient Buildings*, 164–167. https://doi.org/10.1145/2674061.2674081

## Acknowledgments

This project received funding from the Dutch Ministry of Economic Affairs and Climate Policy and Ministry of the Interior and Kingdom Relations under the MOOI program. We would also like to thank the data integration team from TU Delft for their support in providing with BAS data sets.