Enabling the Real-Time Integration of Big and Open Linked Data (BOLD) for decision-making

The design of an Extended Enterprise Architecture Framework

Thesis

by

W.A.W. Steinfort
4717457

in partial fulfilment of the requirements for the degree of

Master of Science
in Complex Systems Engineering and Management

at the Delft University of Technology,
to be defended publicly in September 2019.

Chair:	S. Cunningham,	TU Delft
Thesis committee:	A. Zuiderwijk-van Eijk,	TU Delft 1st supervisor
	S. Cunningham,	TU Delft 2nd supervisor
	B. Peters,	Gartner, Inc. 1st external
	D. Allessie	Gartner, Inc. 2nd external

This thesis is confidential and cannot be made public until September 2020.

This page was intentionally left blank.

Executive Summary

Research shows that CIOs of both public and private organizations are searching for new ways to support better outcomes of their data-driven decision-making. As linking external data with internal company data can result in interesting new insights and value, CIOs are increasingly developing IT strategies that involve the real-time use of external data. One of the innovations that is being adopted in real-time to enable better informed decision-making is external Big and Open Linked Data (BOLD). Combining internal company data silos with external BOLD in real-time helps in the creation of these real-time insights, such as new relationships, patterns, or more context to existing data that was not yet useful. These new insights can be used for better decision-making on a variety of problems and opportunities, such as strengthening customer relationship, better pricing, building a better product market fit, expanding products and services, and managing risk. However, the adoption of external BOLD addresses several adoption barriers that hinder the real-time integration of external BOLD with internal company data. The presence of adoption barriers show that there is a lack of alignment across the different levels of an organization and between the organizations concerned with the data integration. A common practice to align a new strategy and its implementation across the organization is Enterprise Architecture (EA). Countless Enterprise Architecture Frameworks (EAFs) are in existence, but none of them seems to focus on aligning an organization with its business ecosystem. This study shows that the foundation of EA is not capable of satisfying the external mindset of organizations and that an extended approach is required to align a new strategy between organizations. Therefore, this thesis answered the following main research question:

How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?

This thesis used a Design Science Approach to develop an extended EAF that enables the real-time integration of BOLD for decision-making. Throughout this thesis five sub-research questions are answered to derive to the answer of the main research question. First, two systematic literature reviews are conducted to explicate the research problem by identifying the adoption barriers of BOLD and the limitations of the foundation of EA. Secondly, explorative expert interviews and qualitative data analysis are conducted to define the requirements for an extended EAF that deals with the adoption barriers and enables the real-time integration of external BOLD with internal company data. Thirdly, several systems engineering techniques (Morphological chart, Best-of-class chart, and brainstorm sessions with creative methods) are used to transform the requirements into an extended EAF. Lastly, the extended EAF is evaluated on quality and suitability through a 'Plan Evaluation', which resulted in potential improvements of the design. The findings are synthesized for answering the main research question.

To explicate the misalignment across organizations and between organizations, the different barriers that hinder the real-time adoption of BOLD are identified. This resulted in a conceptual model of adoption barriers of BOLD related to information quality (data absence, data quality, a lack of meta-data), task complexity (task handling issues, lack of standards and different sources), technology (Lack of supporting infrastructure, legacy systems, data fragmentation), use & participation (Lack of knowledge, implementation issues, constraints, lack of acceptance), legislations (privacy, security issues, and permits and license issues) and governance (Policy issues, Scalability Issues and a Lack of Data Governance). If organizations want to use external BOLD for decision-making in real-time, no adoption barriers can be present. This study shows that the foundation of EA can align a new IT strategy and its implementation across the organization, but not of aligning a new strategy between organizations and its ecosystem. This means the practice can mitigate the adoption barriers that arise across an organization can be overcome, but not the adoption barriers that arise between organizations.

This thesis shows that a general transformation from a system of record to a system of engagement is required to adapt to this external mindset, which is key part of the paradigm shift that comes with an organization's digital transformation. Modern organizations should embrace the system of engagement, including the company's ecosystem as part of their Enterprise Architecture Framework. Hence, this thesis presents the 'Ecosystem Architecture Framework' that facilitates this transformation and has incorporated the required capabilities to mitigate the adoption barriers that arise between an organization and its ecosystem. The Ecosystem Architecture Framework provides guidance for organizations to implement an IT strategy that involves the real-time integration of external BOLD on an institutional, technological and transactional level. First, the framework

provides a proactive approach that initially deals with the potential adoption barriers. Secondly, the framework enables collaboration between the data users and BOLD providers by establishing strategic alliances with mutual incentives to mitigate the adoption barriers. Thirdly, the framework allows organizations to involve all stakeholders concerned with the BOLD integration in the design of the system. By developing a shared architecture vision and architectural requirements alignment between the organization and ecosystem is created. Fourthly, the framework enables organizations to initially link their architectures through a direct real-time connection. This study demonstrates that organizations need to invest in a hybrid integration platform (e.g. an API Management Platform) that facilitates that organizations can easily integrate and share integration services (e.g. APIs), different applications can be linked together, collaboration between applications is possible, and orchestration of the applications by a single interface. Fifthly, the framework looks beyond the technical implementation and ensures that an inter-organizational model developed to make the architecture working. The inter-organizational model consists of shared artifacts that make sure the same way of working is implemented within the business ecosystem. This can be achieved by developing uniform standards for data, meta-data, technical information about integration services; A common data model that ensures one data model, unified schema and semantics are used within the business ecosystem; An independent API policy that ensures that each external data provider can be internally linked to the architecture by conforming to this policy; and institutions between the organizations that enforce the same way of working to enable sustainability of the data. Lastly, the framework ensures that the architecture keeps running by adapting to changes of the environment. Therefore, it is important that organizations develop a governance mechanism that ensures the system is scalable, maintainable and sustainable. It is important that the inter-organizational model and shared artifacts of the ecosystem are managed and adapt to changes of the environment to keep the architecture running.

The outcomes of this research have both scientific and societal relevance. The former is achieved by adding the following contributions to science: new adoption barriers of BOLD; a critical assessment of the foundation of EA and EAFs; the required capabilities to mitigate the adoption barriers; and an approach that can be used to enable the real-time integration of BOLD with internal company data. The latter is achieved by providing a validated approach that can help to increase the success rate of data-driven strategies, which is relevant for national governments and organizations that want the data-driven strategies to succeed. The thesis also has some limitations. Firstly, the Ecosystem Architecture Framework is only validated trough a plan evaluation, which is a 'design on paper'. Secondly, the framework is dependent on the TOGAF-ADM and can only be used as design principles without TOGAF-ADM. Also, the report has become very lengthy, as the research covers many different research methods.

For future research, it is recommended to also validate the design through a 'product- and process evaluation' that investigate the implementation issues and effects in a real-world context. Secondly, it is recommended to do future research on additional capabilities to create an EAF that can be used independently. Thirdly, an in-depth study on the specific capabilities is recommended to investigate the different design patterns for incorporating the capabilities. Fourthly, it would be interesting to do research on engaging the Ecosystem Architecture Framework in different industries with different organizations part of different ecosystems and evaluate the differences.

<u>For illustration an example is used to explain the key findings:</u>
A flight comparison website supports data-driven decisions on the most suitable flight options. To improve the decisions, the website wants to adopt the external BOLD provided by a variety of airlines in real-time to advise customers the most suitable flights adapted to their needs. As external BOLD is very dynamic, it is important that the data that is used for the decision-making is in real-time. However, interoperability problems, a lack of standards, and data handling issues hinder the integration of the external BOLD in real-time. Therefore, it is important that these adoption barriers are mitigated to improve the decision-quality. To solve this problem, the flight comparison website can use the Ecosystem Architecture Framework to transform its system of record into a system of engagement in which a collaboration is established with the Airlines that provide BOLD with flight information in real-time. By involving the stakeholders concerned with the data-integration in the design of this system, architectural requirements are defined for the ecosystem architecture. The organization invests in an API Management Platform that integrates the APIs of all the different airlines shaped by an inter-organizational model to ensure a common way of working. Furthermore, the platform ensures that the architecture keeps running by adapting to changes of the environment.

Key words: Open Data; BOLD; Decision-making; Adoption; Barriers; Ecosystem Architecture Framework

List of Contents

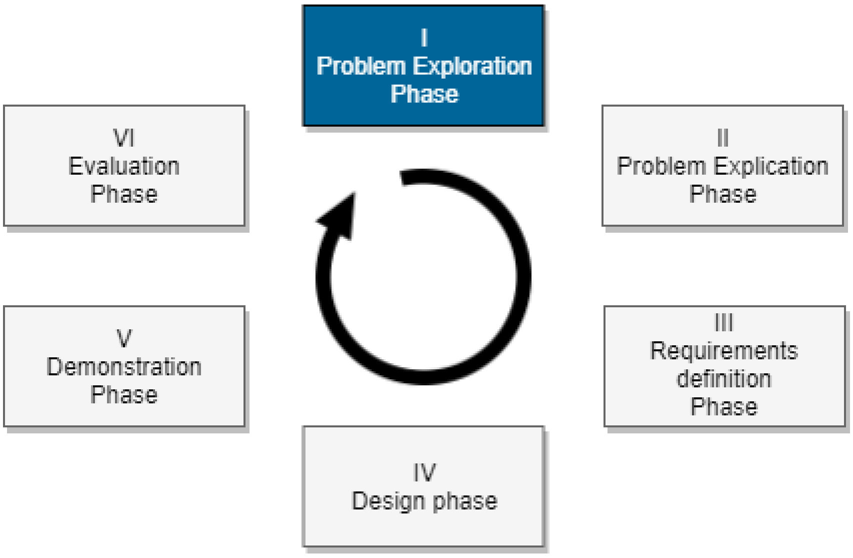
List of Figures

List of Tables

Terminology

Notion	Definition	Sources
BOLD	BOLD is the combination of (1) Big data, which typically are data with large data volume, a rapid velocity of data flows, and high variety in data types and sources, and (2) Open Data, which are the data that is opened by companies, governments, and individuals for free use, reuse and redistribution.	Lněnička and Komárková (2018)
Data-driven strategy	"A strategy that covers the data-driven business activities that need access to data, its analysis, and the tools needed to integrate the data analysis within the business activity. This data usage in business decision-making can enhance competitiveness through reduction of costs, increased added value, or any other parameter that can be measured against existing performance criteria."	Curry (2016)
Ecosystem	A business ecosystem is the network of organizations—including suppliers, distributors, customers, competitors, government agencies, and so on—involved in the delivery of a specific product or service through both competition and cooperation.	Investopedia (2019)
Ecosystem Architecture Framework	The design artifact that developed in this thesis. An Extension to Enterprise Architecture.	
Enterprise architecture	"A discipline for proactively and holistically leading enterprise responses to disruptive forces by identifying and analyzing the execution of change toward desired business vision and outcomes. EA delivers value by presenting business and IT leaders with signature-ready recommendations for adjusting policies and projects to achieve target business outcomes that capitalize on relevant business disruptions. EA is used to steer decision making toward the evolution of the future state architecture"	Gartner (2013)
Enterprise Architecture Framework	'An Enterprise Architecture Framework is a set of models principles, services, approaches, standards, design concepts, components, visualizations and configurations that guide the development of specific aspect architectures'	Schekkerman (2004)
Extended Enterprise Architecture	An Architecture Framework that is adapted with an extension to make the EAF more suitable to achieve its objectives.	Schekkerman (2004)
Open data	"A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike."	Open Data Handbook (n.d.)
Strategy	"Strategy is about being different. It means deliberately choosing a different set of activities to deliver a unique mix of value."	Porter (1996)

I. Problem Exploration Phase



1. Problem Exploration

This chapter presents the introduction, research problem, research objectives, research questions and the outline of this thesis. First, the research topic is introduced by discussing the current situation (section 1.1). Second, the research problem is presented by discussing the current state of literature, knowledge gaps, scientific and societal relevance (section 1.2). Third, the research approach is introduced (1.3). Finally, the outline of the thesis is discussed (section 1.4).

1.1. Introduction

CIOs of both public and private organizations are searching for new ways to support better outcomes of their data-driven decision-making (Wang & Lo, 2019). As linking external data with internal company data can result in interesting new insights and value, CIOs are increasingly developing IT strategies that involve the real-time use of external data (Wang & Lo, 2019). One of the innovations that is being adopted in real-time to enable better informed decision-making is external Big and Open Linked Data (BOLD) (Lněnička & Komárková, 2018). Combining internal company data silos with external BOLD in real-time helps in the creation of these real-time insights, such as new relationships between data subjects, patterns in datasets, or context to existing data that was not yet useful. These new insights can be used for better decision-making on a variety of problems and opportunities, such as strengthening customer relationship, better pricing, building a better product market fit, expanding products and services, and managing risk (Bayrak, 2015; Dutta & Bose, 2015; Saggi & Jain, 2018; Sheng, Amankwah-Amoah, & Wang, 2017; Wang & Lo, 2019).

Despite the effort of organizations, the adoption of external BOLD addresses multiple adoption barriers that hinder the real-time integration with internal company data, such as a wide variety in size, different sources and unstructured data (Lněnička & Komárková, 2018); different flexibility, interoperability and collaboration issues (M. Janssen, Estevez, & Janowski, 2014); and negative effects like violating privacy, possible misuse and misinterpretation of data (Zuiderwijk & Janssen, 2014). To overcome these barriers, a solution is required that enables adoption of BOLD in real-time. Figure 1 visualizes the current situation. The figure depicts that during the integration process several adoption barriers are addressed, which hinder the creation of real-time insights for decision-making.

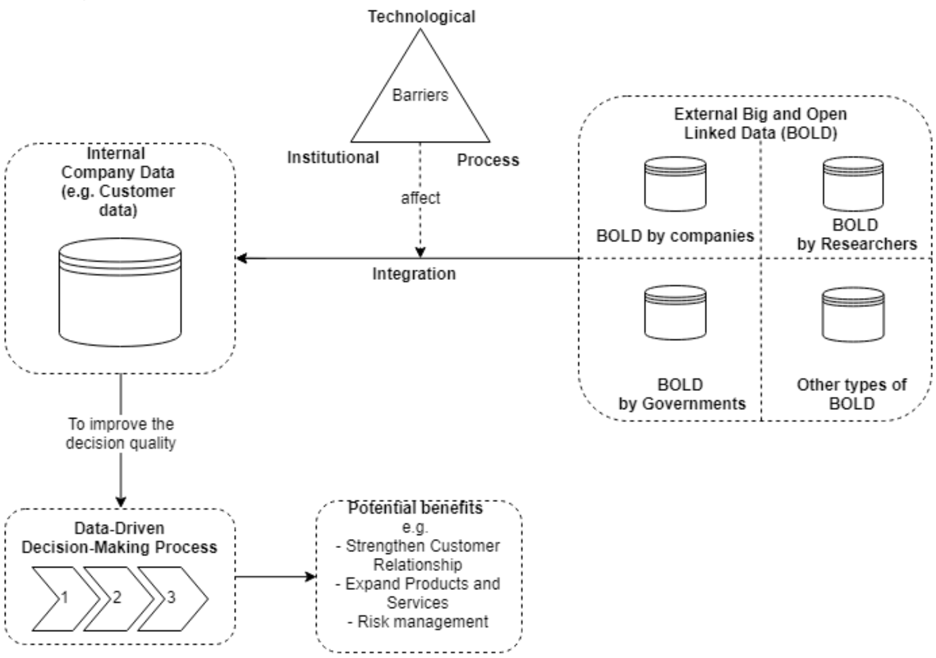


Figure 1: Current Situation

Example
A flight comparison website supports data-driven decisions on the most suitable flight options. To improve the decisions, the website wants to adopt the external BOLD provided by a variety of airlines in real-time to advise customers the most suitable flights adapted to their needs. As external BOLD is very dynamic, it is important that the data that is used for the decision-making is up-to-date. However, different interoperability, and task complexity barriers hinder the integration of the external BOLD in real-time. Therefore, it is important that the adoption barriers are mitigated to improve the decision-quality.

Organizations seem to have difficulties with taking advantage of this Big and Open Linked Data innovation, because organizations often do not have of the right technical and business capabilities to extract the data from the external data silos, process this data and use data analytics for drawing conclusions in real-time for decision-making (Gong & Janssen, 2017). Often an organization's existing ICT landscape is not flexible enough to acquire and implement a new IT strategy that involves the real-time adoption of external BOLD (Denstad & Bygstad, 2012; Gong & Janssen, 2017; Moreno Jr, de Souza Costa Neves Cavazotte, & de Oliveira Valente, 2009). Management-by-magazine seems to play a negative role, as the management of an organization decides to adopt a new technology or innovation without considering whether the rest of the organization can adapt to this new strategy (Gong & Janssen, 2017). The real-time integration of external BOLD does not only require alignment of business and IT across the internal organization, but also between the internal and external organization. The presence of adoption barriers show that the data users and data providers are not well-aligned with each other (Lněnička & Komárková, 2018). It is a big challenge for organizations to deal with these adoption barriers. Therefore, organizations are searching for an approach that helps to successfully implement a strategy that involves the real-time adoption of external BOLD and that mitigates the adoption barriers.

A common practice that is often used to implement a new IT strategy across the entire enterprise is Enterprise Architecture (EA). 'Enterprise Architecture' is generally used for aligning IT strategy and its implementation and to bridge the gap between the business and technology level of an enterprise (Rehman & Shamail, 2014). Furthermore, it describes the present and the future state of the business and can be used as a roadmap to transition from the present to the future state (Janssen, 2009). The real-time adoption of external BOLD requires the transition from a system that uses data and analytics within internal data silos towards a system in which internal and external data silos are combined to enable the creation of new useful insights and value (Lněnička & Komárková, 2018). Such a strategy requires (1) alignment across the business, data, application and technology level of the enterprise; and (2) alignment with the organizations that are involved with the real-time integration of BOLD. Nowadays, countless Enterprise Architecture Frameworks (EAFs) are in existence that allow for the integration of new technologies (Schekkerman, 2004). However, there is a lack of insight whether EAFs exist that can be used to create alignment between organizations. Moreover, the foundation of EA does not seem capable of aligning an organization with external organizations (Minoli, 2008; Schekkerman, 2004). It seems that an extension is required to the foundation of EA to enable the real-time adoption of external BOLD. Lněnička and Komárková (2018) state that limited EAFs exist that can be extended to satisfy the requirements of BOLD. Hence, this thesis aims to design an Extended EAF that guides organizations in the real-time integration of External BOLD with internal company data for decision-making.

1.2. Initial Literature Review: Research Problem

1.2.1. Systematic Literature Review Approach

To find the knowledge gaps in the current state of scientific literature, an initial literature review has been conducted on the research topics of 'data-driven decision-making', 'BOLD' and 'Enterprise Architecture'. The initial literature review is conducted to explore the potential research problems within this specific research area. The review is based on the guidelines defined by Levy and Ellis (2006), which consists of three steps: 1) Systematically collect the input publications for the literature review; 2) Systematically process the relevant publications; and 3) Systematically synthesize the outputs of the papers. The output of the initial literature review is a synthesis of the current state of scientific literature, the knowledge gaps of the current state of scientific literature and the research objectives of this thesis.

Step 1: Systematic Collection input publications

The first step focuses on collecting the 'input papers' of the initial literature review, which are all publications on a specific research area that are included in the systematic literature review (Levy & Ellis, 2006). In total, two academic databases are used to collect the publications: Scopus and Google Scholar. Scopus is chosen, because it enables to collect input papers from high-quality sources. Google Scholar is chosen, because it enables to collect input papers from the open web. This helps to identify the relevant articles that are not included yet in Scopus, but relevant for the development of the current state of scientific literature. The combination of these two academic databases enable a complete representation of the current state of scientific literature. The publications of the last 15 years are collected for the systematic review (2004-2019). The systematic collection only includes the input papers from journals, books and conference papers.

To identify the input papers used for the initial literature review, search queries representing different combinations of the search terms are used in the search engines. Table 1 presents the number of input papers that are collected during the first step of the initial literature review. In queries Q5 and Q6 'Big and Open Linked Data' is replaced for 'Open Data', because BOLD is a relatively younger research area and is a type of open data that resulted in a more results.

Table 1: Input papers for initial literature review

		Scopus			Google Scholar
ID	***Search term***	***Journals***	***Books***	***Conference papers***	***Input papers***
Q1	***"Data-driven decision-making"***	*363*	*36*	*36*	*1000+*
Q2	***"Big and Open Linked Data"***	*66*	*7*	*43*	*222*
Q3	***"Enterprise Architecture Framework"***	*820*	*121*	*820*	*1000+*
Q5	***Q1 AND "Open Data"***	*34*	*6*	*27*	*1000+*
Q6	***Q2 AND "Open Data"***	*5*	*1*	*13*	*140*

Step 2: Systematic Processing relevant publications

The second step focuses on processing the input papers and selecting the candidate papers that seem relevant to include in the overview of the current state of scientific literature (Levy & Ellis, 2006). First, the input papers are scanned on its title and/or abstract, and keywords to assess whether it is a candidate paper that is relevant to read. Secondly, the selected articles are scanned on the introduction and the conclusion to assess whether an article is indeed relevant to discuss in the initial literature review. Thirdly, the studies are thoroughly read, and all relevant information is discussed in the synthesis section.

Next to this procedure, also snowballing was used to find relevant articles. For this the guidelines by Wohlin (2014) were used. The author states that systematic reviews are becoming common and the use of snowballing as well. Therefore, a method is provided that discusses the steps from the start of the literature search until the final inclusion based on backward and forward snowballing. Both snowballing approaches are used. Backward snowballing starts with scanning the titles in the reference list, and then search for the citation in the article itself. Forward snowballing starts with looking at a citation and then look at the reference list. This helps in finding extra relevant publications that can be selected for the systematic review (Wohlin, 2014).

Step 3: Systematic Synthesis Current state of Scientific Literature

The third step focuses on synthesizing the most relevant studies to create an overview of the current state of literature. All relevant studies are summarized, analyzed and interpreted to discover knowledge gaps in the current literature. This section presents the current state of scientific knowledge in the research area with the most important findings; The knowledge gaps extracted from the review; and the scientific and societal relevance of the research problem. The key concepts used in this literature review are briefly explained in the Terminology section and in detail in Appendix A1: Definition of Key Constructs.

1.2.2. Current state of Scientific Knowledge in the Research Area

There is a shift towards data-driven decision-making

In the current society, the amount of data is exponentially increasing, just as the data held by organizations (Gong & Janssen, 2017). The European Commission (2014) states that we are going towards a thriving data-driven economy. This is evident from the initiatives by organizations to put more value out of their data. Organizations now have access to cheaper and more powerful computing platforms and modern analytics to enable real-time analytics on a variety of use cases (Bishop, 2018). According to Bishop (2018), these developments in technology and data convince organizations to put more focus on making evidence-based decision-making. ICT plays a more important role in the field of decision-making and is already integrated in different parts of the decision-making process of organizations (Pan, Jonoski, Castro-Gama, & Popescu, 2015). It is important for organizations to be data-driven and to link and exchange data through the whole enterprise or even beyond the boundaries of the enterprise (Curry, 2016; Vargas et al., 2016). Such data-driven strategies have become increasingly important for competitive differentiation (Bishop, 2018; Buchholtz, Bukowski, & Śniegocki, 2014). Organizations often perform data analysis within their locally kept data silos, which can create valuable new insights to enhance their decision-making. However, the decisions that are made are limited to the data that is available in the data silos. To enhance the decision quality, organizations often decide to strengthen their internal data sets with external data (Fernandes, O'Connor, & Weaver, 2012). Federating this internal data with data from other internal and external data silos can enable the creation of new useful insights and value(Gong & Janssen, 2017). A type of external data that increasingly adopted by organizations is Big and Open Linked Data (M. Janssen & Kuk, 2016).

The potential benefits of BOLD

BOLD is the combination of Big Data and Open Data. Big Data typically are data with large data volume, a rapid velocity of data flows, and high variety in data types and sources (Lněnička & Komárková, 2018). 'Open Data is data that is opened up by governments, organizations, and individuals to enable access for everybody to data without any pre-defined restrictions or conditions for use, reuse and redistribution' (M. Janssen, Matheus, & Zuiderwijk, 2015). With Linked Data is meant that the data is 'structured and machine readable so that it can be semantically queried' (Bizer & Heath, 2009). Figure 62 presents a classification of BOLD. BOLD brings several benefits which are attractive for organizations, among which cost efficiency, transparency, commoditization, interoperability, better informed decision making and opportunities for innovation (Janssen, Charalabidis, & Zuiderwijk, 2012). 'Information management and business analytics leaders in commercial businesses should embrace open-data practices to the same ends, deriving new organizational value in the process' (Gartner, 2015). It seems that organizations see the potential of BOLD, as CBS (2018) observed an explosive growth in the use of BOLD by organizations. Moreover, Wang and Lo (2019) observed the growth in the use of BOLD in organizations as well and discovered that the most influential factors for adoption were top management support, and competitive pressure. This confirms that organizations see the potential of integrating BOLD into the decision-making process to enhance their competitive position (Lněnička & Komárková, 2018; Wang & Lo, 2019). M. Janssen and Kuk (2016) state that, in despite of the already mentioned research, there is need for further examination why and how organizations can create values from BOLD and which approaches should be followed to integrate BOLD with internal company data.

Adoption Barriers of BOLD

Despite the effort of organizations, the adoption of BOLD does not immediately result in successful decisions. Organizations seem to have troubles with taking advantage of this big and open data innovation. Different primary studies are published about several barriers that hinder the successful adoption of BOLD. Zuiderwijk and Janssen (2014) state that open data cause negative effects like violating privacy, possible misuse and misinterpretation of data. These should be considered and be dealt with when integrating open data into the decision-making process. Janssen et al. (2012) describe that often a lack of data quality forms a problem. Research in open data has shown that quality rather than the quantity matters for service and digital innovation

(Kuk & Davies, 2011). Next to that, interoperability, flexibility and collaboration issues make the adoption of big data difficult (Gong & Janssen, 2017). Moreover, BOLD consists of a wide variety in size, different sources and unstructured data, which requires new capabilities to make integration of BOLD possible (M. Janssen et al., 2014; Lněnička & Komárková, 2018). For many organizations these capabilities are not readily available. Organizations have troubles with the integration of BOLD, because the processes and technology are not capable of extracting data from external data silos (databases), processing this information and using data analytics for drawing real-time conclusions for decision-making (Gong & Janssen, 2017). The real-time integration of external BOLD does not only require alignment of business and IT across the internal organization, but also between the internal and external organization. The presence of adoption barriers show that the data users and data providers are not well-aligned with each other (Lněnička & Komárková, 2018). It is a big challenge for organizations to deal with these adoption barriers. Therefore, organizations are searching for an approach that helps to successfully implement a strategy that involves the real-time adoption of external BOLD without the presence of adoption barriers (Lněnička & Komárková, 2018).

Lack of integration between Management and IT

New strategies for decision-making, such as the adoption of BOLD, are defined on management-level. Different authors state that the new strategies often cause difficulties in terms of implementation, because the IT-architecture does not allow it. This is often because the management team is not well-informed about the technical capabilities the organization has (Denstad & Bygstad, 2012; Moreno Jr et al., 2009). Gong and Janssen (2017) state that many strategies are driven by a kind of management-by-magazine approach, in which only the promised benefits of an innovation, such as BOLD, are considered without thinking about the required adaptation to the rest of the enterprise and external organizations concerned. Also, sometimes the technologies are not used in the way how the management had it in mind, which increases the chance of failure of a strategy (Jia, Wang, & Ge, 2018). Seen the different challenges to be dealt with during the integration of BOLD into the data-driven decision-making process, it is important that the business processes are well aligned with the new technology (Gerow, Thatcher, & Grover, 2015). Different literature is published about how the gap between management and IT can be bridged, but there is limited insight in how the internal and external enterprise can be better aligned (Huang & Hu, 2007).

Enterprise Architecture

In the literature, the problem of integration between business and technology is commonly known and different authors have suggested potential solutions for better alignment of these two domains. A common practice that is used to align strategy across the organization is Enterprise Architecture (EA) (Janssen, 2009; Rehman & Shamail, 2014). EA can be used as a discipline to guide an enterprise through change. 'Enterprise Architecture' is generally used for aligning strategy and its implementation and to bridge the gap between the business and technology level of an enterprise. Furthermore, it describes the present and the future state of the business and can be seen as the plan of transition from the present to the future state (Janssen, 2009). It provides design principles and a layered architecture that could enable the management of the key features of a district in separate layers (Jadda & Idrissi, 2015). Several Enterprise Architecture Frameworks (EAFs) are in existence that allow for the adoption of new innovations.

However, there is limited insight in what EAF provide guidance for extracting data from external sources, processing this information and using data analytics for drawing real-time conclusions for decision-making. This requires the move from data and analytics within data silos towards solutions in which internal and external data silos are federated to enable the real-time creation of new useful insights. Different authors propose EAFs that allow for the integration of similar innovations. Vargas et al. (2016) introduce the idea of inter-enterprise architecture. This type of enterprise architectures allows collaborative networks to integrate and coordinate different organizations. Lněnička and Komárková (2018) did research on how a Government Enterprise Architecture Framework to support the requirements of BOLD with the use of cloud computing should look like to deal with the many new information and communication technologies and the rapid increase of data. Gong and Janssen (2017) did research about the use of EA for the adoption of Big Data in governments and state that the discipline could help governments that are struggling with this activity and hypothesized that EAs can help governments to integrate big data with the rest of the current landscape. Although, each of these proposed EAFs have the right capabilities for parts of the research problem, it seems that no solution has all required capabilities needed to overcome the research problem.

There is limited insight in whether the foundation of EA is suitable enough to align an organization with external organizations (Minoli, 2008). It seems that an extension is required to the foundation of EA to enable the real-time adoption of external BOLD. Lněnička and Komárková (2018) state that limited EAFs exist that allow for revision and extension of existing EAFs regarding the requirements of BOLD. Hence, this thesis aims to design an Extended EAF that provides a method that enables organizations to build an EA that enables real-time data-driven decision-making using external BOLD.

1.2.2. Knowledge Gaps

Based on the initial literature review different knowledge gaps are addressed in the current state of literature:

I. Firstly, whereas a lot is published about the adoption barriers and potential benefits of BOLD, not much is written about how to overcome or mitigate the adoption barriers. M. Janssen and Kuk (2016) state that, despite the already conducted research (e.g. Lnenicka and Komarkova (2018); Wang and Lo (2019)) there is a need for further examination why and how organizations can create value from the adoption of external BOLD and which approaches should be followed to successfully adopt external BOLD and to integrate this with internal company data.

II. Secondly, there is limited insight whether the foundation of EA is suitable enough to implement a strategy that involves the adoption of Big and Open Linked Data (Gong & Janssen, 2017; Lněnička & Komárková, 2018). There is a need for a literature review that (1) identifies the Enterprise Architecture Capabilities that are required for the implementation of such a strategy that requires alignment across the enterprise and between enterprises and (2) identifies what EAF is most suitable to enable this.

III. Thirdly, authors recommend that an Extended Enterprise Architecture Framework is developed that is suitable for the requirements of BOLD (Lněnička, Máchová, Komárková, & Čermáková, 2017). Lněnička and Komárková (2018) state that limited EAFs exist that allow for revision and extension of existing EAFs regarding the requirements of BOLD.

1.2.3. Scientific Relevance

Gong and Janssen (2017) already examined how an EAF should look like for the adoption of BOLD; Lnenicka and Komarkova (2018) examined how an EAF should look like that can be used for the adoption of Big and Open Linked Data to enable cloud computing; and M. Janssen and Kuk (2016); M. Janssen et al. (2015) conducted research on the policies, practice and research on BOLD and show that it can bring a lot of insights in patterns and behaviors when it is combined with internal company data. Although, these studies are highly relevant there is a lack of research looking into the combination of internal company data with external BOLD and how this should be arranged at a technical and institutional level. Moreover, a lack of an EAF that provides guidance in the adoption of external BOLD and integrate this with internal company data in real-time to enable decision-making. This study addresses this knowledge gap by developing a validated EAF that guides both public and private organizations in the integration external BOLD with internal company data to enable decision-making in real-time (that forms parts of a solution).

1.2.4. Societal Relevance

The European Commission (2014) states that we are going towards a thriving data-driven economy, but that the implementation of such a data-driven strategy is difficult and one of the biggest challenges. In total, 85 percent of the organizations that have adopted a data-driven strategy fail (Bishop, 2018; Gartner, 2017). For national governments it is important that this percentage will be reduced, since it will have a negative impact on the economy of a country. As the number of data-driven organizations has been increasing for the past years, it is important that the adoption of such a strategy will be improved. It is important that not only a small percentage of the organizations are successful, but most of the companies are successful to stimulate the (local) economy (Marketwatch, 2017). From the perspective of organizations, it is important that a validated approach is developed that guides in the combination of external BOLD with internal company data to enable decision-making in real-time. If this becomes more feasible, it will have direct effect on the quality of the decision taken (Gartner, 2015).

1.3. Research Approach

1.3.1. Methodology

In this thesis, the Method Framework for Design Science by Johannesson and Perjons (2014) is used as a methodology for answering the main research question in a systematic, structured and logical way. The Method Framework for Design Science is based on the Design Science Research Cycle by Hevner (2007), which is explained in more detail later. It offers the possibility to do both literature research and practical research in the form of interviews and case studies and to use the research findings for the development and performance of design artifacts with the intention of improving its functioning. The translation of these research findings into a designed artifact forms an addition to the knowledge base (Hevner, 2007). Hevner (2007) states the main goal of design science research is achieving knowledge and understanding a specific problem domain by developing a designed artifact that can be used as a solution for problems that professionals encounter in the field. This thesis focuses on the design of an artifact that can help overcome the research problem.

The design cycle consists of an order of steps that allows for finding an answer to the research question. The Method Framework for Design Science by Johannesson and Perjons (2014) consists of five phases: Problem Explication Phase, Requirements Definition Phase, Design Phase, Demonstration Phase, Evaluation Phase, as displayed in Figure 2. In this thesis, an additional phase is added: The Problem Exploration Phase. The rationale for adding an additional phase is, because it allows to do an initial literature review to explore the research area and the research problem. Each of the phases corresponds to one of the sub-research questions. Table 2 presents a description of the main objective of each phase.

Table 2: Revised Method Framework for Design Science (Johannesson & Perjons, 2014)

Phase		Description
I.	*Problem Exploration Phase*	Aims to give an understanding of what the initial problem is.
II.	*Problem Explication Phase*	Aims to give an understanding of what causes the initial problem and gives an overview of all relevant information that is available that can be used to explicate the research problem.
III.	*Requirements Definition Phase*	Aims to discuss the research findings from the literature and interviews and translate the findings into a list of requirements for a design artefact that could help overcome the research problem.
IV.	*Design Phase*	Aims to build a design artefact with the requirements of the previous phase that could help overcome the research problem.
V.	*Demonstration Phase*	Aims demonstrate the design artefact during a field test for improvement and evaluation.
VI.	*Evaluation Phase*	Aims to provide an answer to the different research questions and Evaluation of the research problem

A design approach is chosen, because the research problem must be overcome by a solution that enables organizations with the real-time adoption of external BOLD for decision-making. The current foundation of EA does not provide an approach that guides organizations with the combination of external BOLD with internal company data in real-time. To enable this, it is important to first fully understand the problem and achieve the knowledge to design an artifact that facilitates this (Hevner, 2007; Johannesson & Perjons, 2014). Design Science research has a limitation, as it implies a study that focuses on shaping the existing world, while other types of research rather focus on describing and explaining the existing world (Iivari, 2007).

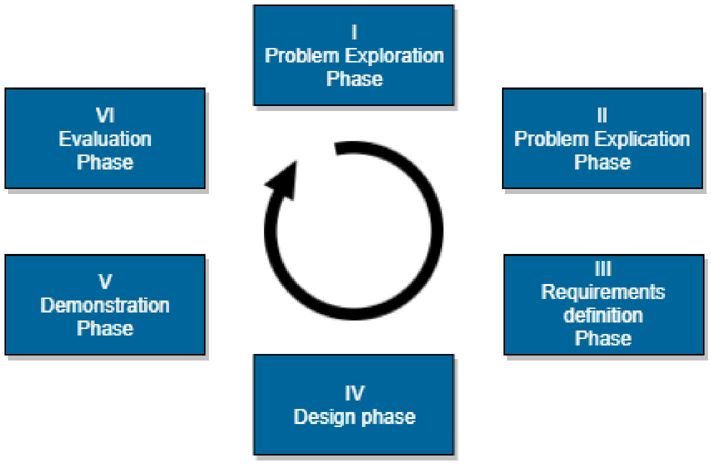


Figure 2: Method Framework for Design Science (Edited from Johannesson and Perjons (2014)).

1.3.2. Research questions

Based on the current state of scientific knowledge in the research area and the addressed knowledge gaps of the Research Problem section, the following main research question is defined:

> ***How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?***

With the following sub-research questions, which are consistent with phases of the Method Framework for Design Science (Figure 2), the main research question is answered:

Problem Exploration phase

The first phase is about identifying the initial research problem and the knowledge gaps. The outcome of this phase is used to define the research questions. The different sub-research questions are developed in such way that the outcome can be used to partly solve the research problem.

Problem Explication Phase

The problem explication Phase consists of two sub-research questions. The first sub research question is:

> ***RQ1****. Which adoption barriers affect the integration of external Big and Open Linked Data with internal company data?*

RQ1 aims to explicate which factors cause the research problem by investigating which different barriers cause difficulties for integrating external BOLD with internal company data to enable real-time decision-making. The outcome of this sub research question is an overview of the most common problems that cause the misalignment the between the data user and data providers of BOLD.

The second sub research question is:

> ***RQ2****. What are the capabilities from the current foundation of EA (Theories, Methods, Expertise, Experience, Meta-artefacts) that need to be incorporated in the Extended Enterprise Architecture Framework to mitigate the adoption barriers?*

RQ2 aims to collect all relevant information from the current literature available to create an overview of the foundation of Enterprise Architecture, its capabilities and an assessment of the most commonly used Enterprise Architecture Frameworks on these capabilities.

Requirements Definition Phase

The third sub research question is:

> ***RQ3****. What are the requirements for an Extended Enterprise Architecture Framework that enables real-time integration of external Big and Open Linked Data with internal company data for decision-making?*

RQ3 aims to discuss the most important findings of the literature reviews and explorative expert interviews and translates these findings into requirements for an extended Enterprise Architecture Framework.

Design Phase

The fourth sub research question is:

> ***RQ4**. What does an Extended Enterprise Architecture Framework that enables real-time integration of external Big and Open Linked Data with internal company data for decision-making look like?*

RQ4 aims to translate the list of collected requirements by the previous RQs into a suitable Extended Enterprise Architecture Framework that helps organizations to develop an EA enables the real-time adoption of BOLD.

Demonstration Phase

The fifth sub research question is:

> ***RQ5**. How can the quality and suitability of the Extended Enterprise Architecture Framework be improved?*

RQ5 aims to evaluate the design artefact that is developed by the previous RQ. Different experts are selected to evaluate the design artefact by means of a plan evaluation, where the design goals, design requirements and design specifications are evaluated on quality and suitability. The findings of the evaluative expert interviews are translated into potential improvements for the design.

Evaluation Phase

In the evaluation phase the main research question gets answered, by synthesizing all the findings of the different sub-research questions. This phase draws the conclusions and reflects on the thesis.

1.3.3. Design Science Research Cycle

The Method Framework for Design Science by Johannesson and Perjons (2014) is based on the Design Science Research Cycle by Hevner (2007), which consists of three different cycles: The Rigor Cycle, The Relevance Cycle and The Design Cycle. In Figure 3, the RQs are mapped in the Design Science Research Cycle.

(1) In the Rigor Cycle the foundations are discussed, which focuses on identifying the existing relevant theories, practical knowledge and Meta-artefacts that form the knowledge base to the research subject. It also helps to address what knowledge base additions are required to solve the research problem (Hevner, 2007).

(2) In the Relevance Cycle the environment is discussed, which focuses on identifying the relevant elements of the application domain: people, organizational systems, technical systems and problems & opportunities. Based on these findings, a list of requirements is defined which are required to build the design artefact and processes that can overcome the research problem. After the design phase, the design is tested in the application domain and gets evaluated (Hevner, 2007).

(3) In the Design Cycle, the requirements are translated into the design artefact and processes required overcome research problem. After the field testing the design artefact and processes are evaluated and improved (Hevner, 2007).

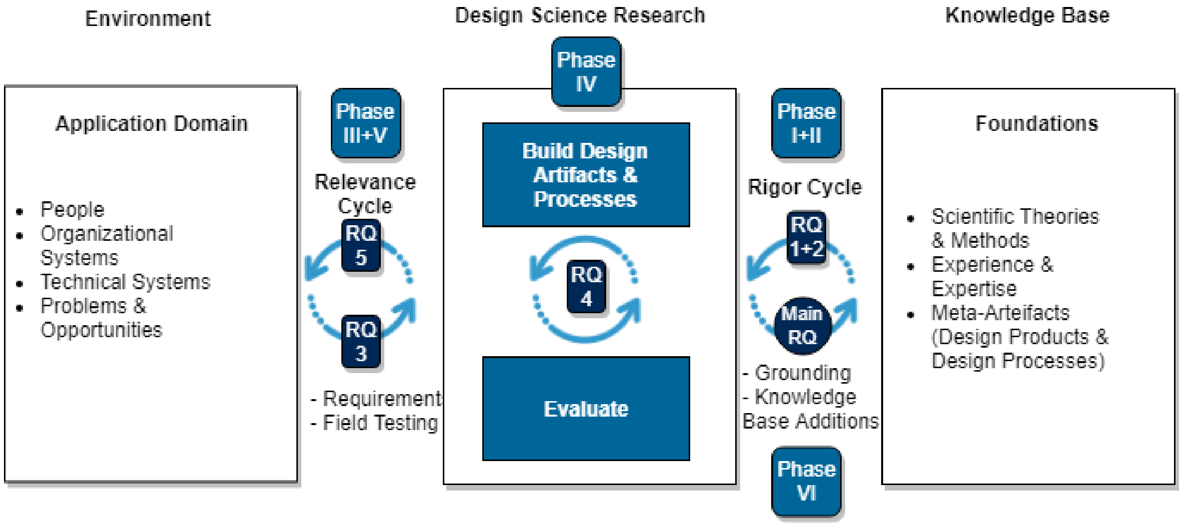


Figure 3: RQs plotted in the Design Science Research Cycle (Edited from (Hevner, 2007))

First, the knowledge base is defined by reviewing the existing scientific methods and theories, experience & expertise, and artefacts available. This review is used to identify the knowledge gaps. In this thesis the knowledge gaps are discussed in the 'research problem' section. The rigor cycle further focuses on exploring the initial problem and finding all relevant knowledge is used to explicate the problem. As depicted in Figure 3, sub research questions 1 and 2 relate to the rigor cycle. RQ1 explicates why BOLD providers and users are not well-aligned with each other, and RQ2 discusses the current available solutions to implement a new strategy across the enterprise and addresses the lacking capabilities through a design gap analysis.

Secondly, the knowledge gaps are discussed in the environment (application domain) through explorative expert interviews. The findings of the rigor cycle are used to develop the questions for the explorative expert interviews. The purpose of the relevance cycle is to collect the business needs from people, organizations and technology and use this for the definition of requirements for the design artefact. RQ3 discusses the design requirements for an Extended Enterprise Architecture Framework that satisfies the additional business needs.

Thirdly, the design requirements collected by the explorative expert interviews are used as input for the design cycle to build the design artefact and processes to implement the artefact. RQ4 represents the design of the Extended Enterprise Architecture Framework. As depicted in Figure 3, the design cycle is an iterative approach and the design artefact is evaluated through evaluated expert interviews. RQ5 represents the evaluation of the Extended Enterprise Architecture Framework in order to collect additional and revised requirements for the revision of the design artefact. Finally, the outcomes of the different RQs synthesized and used to answer to the Main RQ. The deliverable of the thesis is a design artefact that can be added to the knowledge base. It removes the knowledge gaps and proposes a solution to enable the real-time adoption of BOLD for decision-making.

1.3.4. Research Strategy

Figure 4 presents the research flow diagram that represents an overview of the research strategy employed for this thesis. The overview summarizes the required input data to answer a RQ, the research activity that requires that input data, and the expected output data (deliverables) that is achieved by performing the research activity per phase of The Method Framework for Design Science by Johannesson and Perjons (2014). Each phase consists of at least one chapter, and each chapter corresponds to one of the RQs (except for the introduction chapter). For each chapter, an overview of the expected deliverables is depicted. In the following section, the research activity of each phase is explained. **The specific approach that is used to perform each research is explained in detail in the concerned chapter itself.**

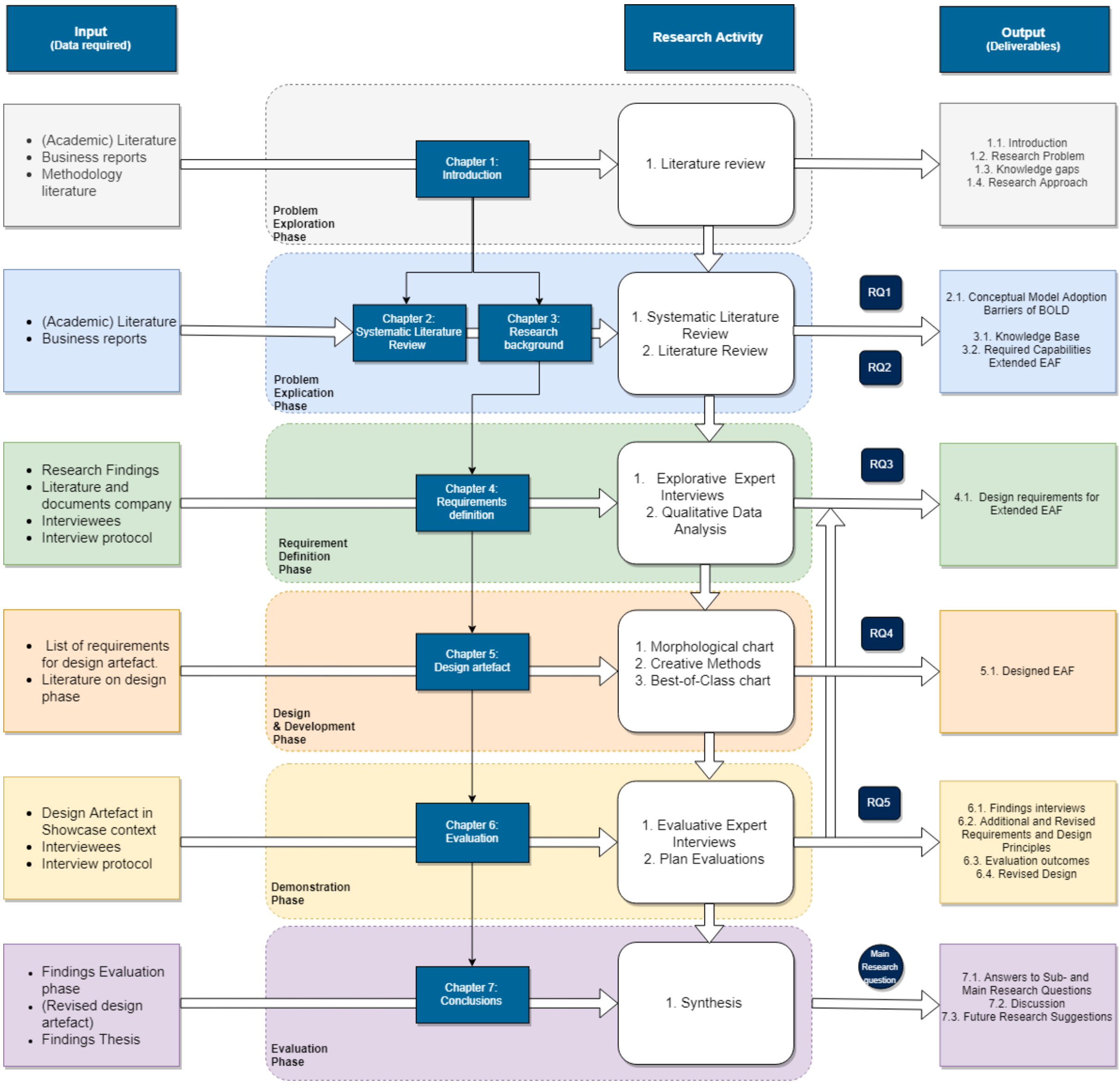


Figure 4: Research Flow Diagram

Phase I: Problem Exploration Phase

The problem exploration phase identifies the initial problem that the thesis addresses. The outcome of this phase are an overview of the current state of literature, the knowledge gaps and research questions that should remove the knowledge gaps. The following research methods and data collection methods are used to identify the latter:

Based on the guidelines by Levy and Ellis (2006) a big number of input papers are reviewed. During the literature review the research topics that are related to the problem are summarized, analyzed and interpreted, which helps to look at the findings from different perspectives in an objective way (Brereton, Kitchenham, Budgen, Turner, & Khalil, 2007). First all relevant literature was scanned at a surface level. This allows scanning as many different publications as possible related to the subject and helps to create a knowledge base about the underlying problem (Johannesson & Perjons, 2014). Thereafter, the selected articles are thoroughly reviewed. A limitation to this method is that there is a possibility that the search query does not cover all the relevant literature and is not accurate. In this case, all used search queries must be kept and saved. Also, selecting articles based on its abstract can cause lead to wrong interpretation. However, it makes you able to save time and handle a higher number of articles. The input for this method are business reports and academic articles. This will result in an overview of all relevant publications about BOLD and EA.

Phase II: Problem Explication Phase

The problem explication phase addresses the sub-research question **RQ1**. *'Which adoption barriers affect the integration of external Big and Open Linked Data in organizations with internal company data?'* and ***RQ2****. 'What are the capabilities from the current foundation of EA (Theories, Methods, Expertise, Experience, Meta-artefacts) that need to be incorporated in the Extended Enterprise Architecture Framework to mitigate the adoption barriers?'. The* following research methods and data collection methods are used to answer the research question:

For answering RQ1 and RQ2, a systematic literature review (SLR) is conducted. According to (Pettricrew & Roberts, 2006), '*a SLR aims to comprehensively identify all relevant studies to answer a particular question, and assesses the validity of each study taking this into account when reaching to conclusions*.' Therefore, a systematic literature review is conducted in order to create a conceptual model of the adoption barriers are encountered by organizations. 'A systematic review is a means of identifying, evaluating and interpreting all relevant research available to a particular topic.' This helps to look at the findings from different perspectives in an objective way (Brereton et al., 2007). A SLR is a secondary study that is systematically reviewing the primary studies about a certain topic (Kitchenham, 2004). The rationale for performing this type of review is that it enables summarizing existing evidence concerning the adoption barriers of the innovation of BOLD; helps to identify the gaps in the current research for further investigation; and it provides a background in order to position new research activities (Kitchenham, 2004). A systematic review is preferred in this thesis, because a lot is written about the topic and a single literature review about a topic that already has been investigated well is of little scientific value. A systematic review creates a synthesis existing work in such a way that delivers a complete summary of all the existing work in a report. The research findings of the SLR will be used to prepare the explorative interviews, where after the list of requirements will be defined. **For both RQs, the specific approach that is used for literature reviews is explained in detail in chapter 2 and 3.**

Phase III: Requirements Definition Phase

The requirements definition phase addresses the sub-research question ***RQ3*** *'What are the requirements for an Extended Enterprise Architecture Framework that guides organizations in the real-time integration of external Big and Open Linked Data with internal company data for decision-making?'.* The following research methods and data collection methods are used to answer the research question:

For answering RQ3, explorative expert interviews are conducted. In total 9 interviewees in different roles are interviewed. The research findings of RQ1 and RQ2 are used to develop the interview protocol and questions. The interviews are conducted in a semi-structured way in order to collect the design requirements in a structured way. A Limitation is that only a few persons can be interviewed, so it is important that the interviewees are selected based on specific requirements to ensure a complete set of requirements is extracted from the interviews. This is required to ensure a design artefact that enables to overcome the research problem. Second, after conducting the interviews qualitative data analysis is performed to summarize, analyze and interpret the interview results. **The specific approach that is used for the qualitative data analysis is explained in detail in chapter 4.**

Phase IV: Design Phase

The design phase addresses the sub-research question ***RQ4****. 'What does an Extended Enterprise Architecture Framework that guides organizations in the real-time integration of external Big and Open Linked Data with internal company data for decision-making look like?'.* The following research methods and data collection methods are used to answer the research question:

The design phase consists of three steps. In the first step a morphological chart is created to identify the design space for the extension for an EAF (Dym, Little, Orwin, & Spjut, 2009). In the second step, the design options for the design space are identified. The design options are means to fulfil the functional requirements that are collected during the explorative expert interviews. This is done through a brainstorm session and creative methods. In the final step, the actual design is made through another brainstorm session by selecting a combination of the design options that are identified in the second design step. A Limitation to this is, that there is not one perfect solution and a list of possible solutions will be provided. With a best-of-class method the best solution will be chosen and used for the Demonstration Phase (Dym et al., 2009; Johannesson & Perjons, 2014). **The specific approach that is used for the design is explained in detail in the chapter 5.**

Phase V: Demonstration Phase

The demonstration phase addresses the sub-research question ***RQ5.*** *'How can the quality and suitability of the Extended Enterprise Architecture Framework be improved?'.* The following research methods and data collection methods are used to answer the research question:

The Extended Enterprise Architecture Framework is designed through an iterative approach. In order to collect additional and revised requirements for an improved design, the design artefact was reviewed and evaluated by three different Enterprise Architects from different organizations. This chapter presents the evaluation phase of the design cycle by Johannesson and Perjons (2014). The Design Evaluation chapter is structured in three parts. First, the method of evaluation that was used to collect the potential improvements of the design is defined. The method of evaluation is based on the evaluation methodologies published by Verschuren and Hartog (2005). The second part of this chapter discusses the feedback of the expert evaluations that are conducted. The key findings of the interviews are discussed and transformed into potential improvements for the design. Finally, the key findings are translated into potential improvements for a new design and explained what improvements are processed. **The specific approach that is used for the evaluation is explained in detail in chapter 6.**

Phase VI: Evaluation Phase

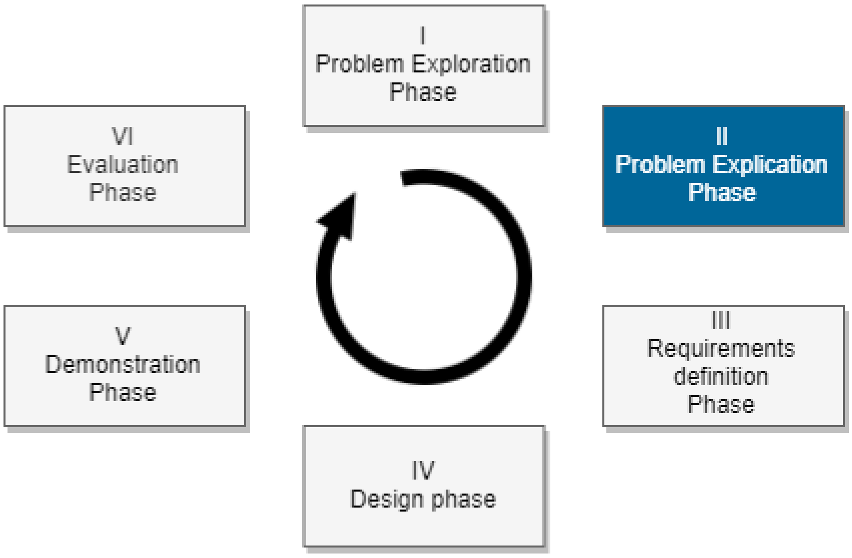
The evaluation phase addresses the main research question *'How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?'* with the following research methods and data collection methods:

For answering the main research question, a synthesis of all the research findings of the sub-research questions will be performed in order to give a substantiated answer. Furthermore, there will be reflected on the thesis, the contributions are discussed and recommendations for future research will be made.

1.4. Outline Thesis

The outline of this thesis can be found in Figure 4.

II. Problem Explication Phase



2. Literature Review: Adoption Barriers

In this chapter, the sub-research question 'RQ1. *Which adoption barriers affect the real-time integration of external Big and Open Linked Data with internal company data?'* is discussed. A systematic literature review is conducted to collect all the relevant adoption barriers that hinder the real-time integration of external BOLD for decision-making with internal company data. A variety of barriers affect the real-time integration with internal company data. Therefore, a conceptual model of the different categories of adoption barriers is developed to explicate the research problem. It provides an overview of the factors that cause mis-alignment between organizations regarding the integration of external BOLD. First, the research approach of the literature review is explained in detail (Section 2.1, 2.2. and 2.3). Secondly, the systematic literature review is discussed (section 2.4). Lastly, a conclusion is drawn, and the sub research question is answered (section 2.5).

2.1. Literature Review Overview

In order to solve the research problem, it is important to understand the different factors causing the research problem. As mentioned in the introduction, the real-time integration of BOLD with internal company data addresses a variety of adoption barriers. A lot is written about the adoption barriers of open data, big data, linked data, and BOLD. Therefore, a systematic literature review is conducted in order to create a conceptual model of the adoption barriers are encountered by organizations. 'A systematic review is a means of identifying, evaluating and interpreting all relevant research available to a particular topic.' It is a secondary study that is systematically reviewing the primary studies about a certain topic (Kitchenham, 2004). The rationale for performing this type of review is that it enables summarizing existing evidence concerning the adoption barriers of the innovation of BOLD; helps to identify the gaps in the current research for further investigation; and it provides a background in order to position new research activities (Kitchenham, 2004). A systematic review is preferred in this thesis, because a lot is written about the topic and a single literature review about a topic that already has been investigated well is of little scientific value. A systematic review creates a synthesis existing work in such a way that delivers a complete summary of all the existing work in a report.

The systematic literature review is conducted following the guidelines defined by Kitchenham et al. (2009). The review process involves activities in three core phases: The Planning Phase, the Conducting Phase and the Reporting Phase. Figure 5 presents an overview of the different activities that are associated in each phase of the Systematic Literature Review.

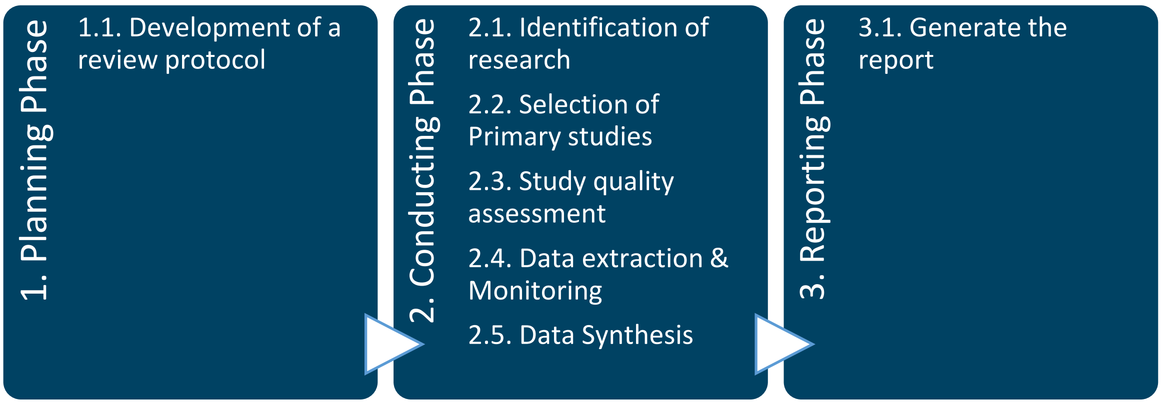


Figure 5: Systematic Literature Review Approach (Kitchenham, 2004)

2.2. The Planning Phase: Development of a Review Protocol

A Review Protocol is developed that is used to conduct a structured systematic review. A review protocol defines the methods that are used to perform the review and ensures the selection of the articles are not driven by researcher expectations. This is required to take away the chance for researcher bias (Kitchenham, 2004).

2.2.1. Background / The need for a Review

As discussed in the introduction, the rationale for the systematic review is that a lot is written about the adoption barriers of BOLD, but recently, no systematic review is conducted that summarizes the different adoption barriers of BOLD. This systematic review intends to provide a complete and up-to-date review. Figure 6 presents an overview of the number of publications related to the research topics. It is remarkable that the documents published are increasing every year. This shows that these topics are considered as scientifically relevant (As 2019 is not over yet it looks like the number of publications is decreasing, but in the coming months new papers can be published). It also shows that already a lot is written about the Adoption Barriers of BOLD and there is a need for a summary of all the relevant publications. This ensures that a complete conceptual model of the adoption barriers of BOLD can be developed, including the all the relevant primary studies. The articles used for the systematic literature review are primary studies that collected all the adoption barriers on open data, big data or BOLD.

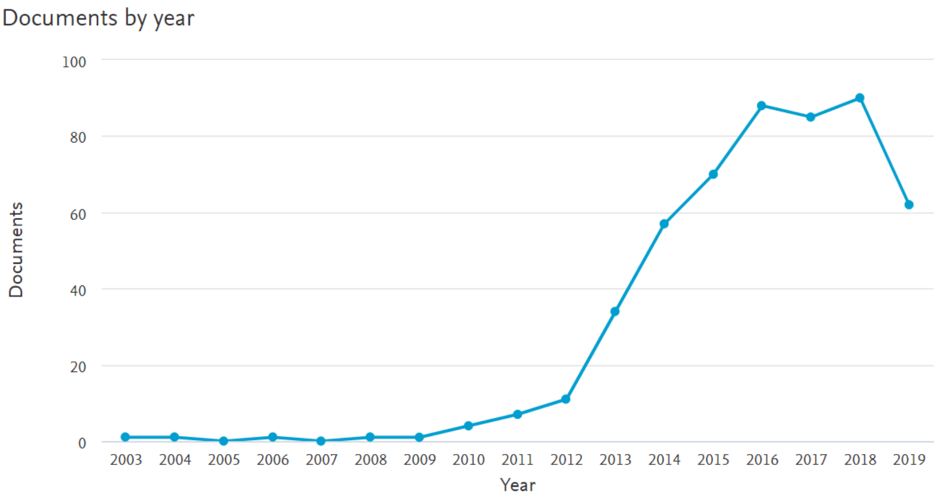


Figure 6: Publications related to the research topics of 'adoption barriers' and 'Big Data', 'Open Data' and 'BOLD'.

The first objective of this systematic literature review is to collect all relevant research that can be used to identify the adoption barriers for the adoption of Big and Open Linked Data. The second objective is to create a complete and up-to-date conceptual model of the adoption barriers, involving the most recent articles and the most cited articles. The third objective is that a systematic literature review is required to get a thorough understanding of the research problem and to give an elaborated explication of the research problem.

2.2.2. Research question

To perform a focused systematic review, a research question is defined that gives an answer to the first sub-question of this thesis. The output of the systematic review thoroughly explicates the research problem and is used to set up the explorative interviews for finding requirements for a solution. In total, one research question was defined:

RQ1. Which Adoption Barriers affect the real-time integration of Big and Open Linked Data with internal company data?

In order to prevent that the question is answered by providing a table with 100+ adoption barriers, the findings are summarized in a conceptual model that creates a mutual exclusive, collectively exhaustive overview of the different adoption barriers that organizations address.

2.2.3. Sources

To identify the primary studies that are currently available, different sources are used. In total, two academic search engines are used to identify the primary publications: Scopus and Google Scholar.
Scopus is chosen, because it delivers a comprehensive overview of the world's scientific research output across all disciplines. Next to that, it offers a platform that brings together the best quality, analytics and technology in one solution. Secondly, Google Scholar is chosen, because it enables to collect input papers from a wide variety of sources through the open web. This helps to identify the relevant articles that are not included yet in Scopus but could be relevant for the systematic review. The combination of these two academic databases enable a complete representation of the current state of scientific literature. The inclusion and exclusion criteria and quality criteria are discussed in section 2.2.4.

2.2.4. Inclusion/Exclusion criteria

For both academic search engines, the same inclusion/exclusion criteria are used. To get to the results the following search query is used:

TITLE-ABS-KEY ("Barriers" AND "Open Data" OR "Big Data" OR "Big and Open Linked Data") AND (LIMIT-TO (SUBJAREA , "COMP") OR LIMIT-TO (SUBJAREA , "ENGI"))

Table 3 presents an overview of the search results and the inclusion/exclusion criteria that are used to get to the number of publication considered to review.

Table 3: Search results

Source	Number of publications	Year	Subject area	Document Type	Language
Scopus	*514*	*2011-2019*	*All*	*Conference Paper; Journal Article; Book Chapter;*	*English*
Google Scholar	*1000+*	*2011-2019*	*All*	*Conference Paper; Journal Article; Book Chapter;*	*English*

2.2.5. Quality criteria

To ensure high quality of the systematic literature review, quality criteria are defined. Based on the quality criteria, primary studies are selected for the review. The following quality criterion is defined:

- The Systematic Literature Review must at least include publications from ten different authors;
- The Systematic Literature Review must at least include publications of the last seven years.

Firstly, these quality criteria are defined ensure that the review includes a representative number of authors with different views. This ensures that the review is not biased by the perspective a little number of authors. Secondly, it ensures that the review at least includes all the publications that are published since the big increase that is depicted in Figure 6.

2.2.6. Study Selection Procedure

First, a literature survey is conducted. A literature survey is useful to do a first assessment whether a primary study is useful or not. This includes scanning the articles on its title and abstract to assess whether the article could consist of relevant information or evidence for answering the research question. If the abstract and title seems relevant, the articles are selected for the systematic review. Secondly, the selected articles are scanned on the introduction and the conclusion to assess whether an article is indeed relevant for the systematic review. Otherwise, the selected article is not considered during the systematic review. Thirdly, the primary studies are thoroughly read, and all relevant information and evidence is collected and captured in an archive. Finally, the information and evidence are reviewed and synthesized into a conceptual model of adoption barriers.

Next to this procedure, also snowballing is used to find relevant articles. The guidelines by Wohlin (2014) are adhered to perform this. The author states that systematic reviews are becoming common and the use of snowballing as well. Therefore, a method is provided that discusses the steps from the start of the literature search until the final inclusion based on backward and forward snowballing. Both snowballing approaches are used. Backward snowballing starts with scanning the titles in the reference list, and then search for the citation in the article itself. Forward snowballing starts with looking at a citation and then look at the reference list. This helps in finding extra relevant publications that can be selected for the systematic review (Wohlin, 2014).

2.2.7. Synthesis extracted data

After the data is extracted from the selected articles used in the systematic review. A strategy for synthesizing the data is developed. First, all the relevant data is put in a database to create a clear overview of all the different adoption barriers that are collected. Secondly, the barriers are put into 5 pre-defined categories, based on the categories published by Janssen et al. (2012). After mapping all the adoption barriers into these categories (Information Quality Barriers; Task Complexity barriers; Technical barriers; Use & Participation Barriers; and

Legislation Barriers), the relationships between the adoption barriers are defined. Having an overview of the different relationships enables the development of sub clusters within the categories. After mapping all the adoption barriers in categories and sub clusters, the duplicate barriers are merged. This makes the value of an adoption barrier stronger, because it is recognized by multiple authors. The synthesis eventually resulted in the 'Conceptual model of adoption barriers of BOLD', depicted in Figure 7.

2.3. Step 2: The Conducting Phase

Following the research protocol, the review is conducted in 5 stages: 1. Identification of research 2. Selection of studies 3. Study quality assessment 4. Data extraction and monitoring progress 5. Data synthesis. The results are discussed in the reporting Phase.

2.4. Step 3: The Reporting Phase

2.4.1. Introduction

In chapter 1, the potential of adopting BOLD is discussed and it shows that a lot of value can be created but when putting it into practice multiple adoption barriers are addressed. A literature review is conducted to discover the adoption barriers of BOLD. In total, 87 barriers that influence the adoption of BOLD are extracted from the literature. Some of the barriers are recognized by many of the authors, and some barriers are only mentioned by a single author. Figure 7 presents a conceptual model of the adoption barriers that is developed to synthesize the findings of the review. The barriers are categorized in mutually exclusive and collectively exhausting clusters to give a more well-organized overview of the most common problems with the adoption of BOLD. The categories defined by Janssen et al. (2012) are adhered. Janssen et al. (2012) have stated that the adoption barriers of open data from a user perspective can be categorized into 5 types of barriers: Task Complexity Barriers; Use and Participation Barriers; Legislation Barriers; Information Quality Barriers; and Technology Barriers. In this thesis, for each category additional sub-clusters are defined that help to categorize the identified adoption barriers within a specific cluster. The barriers from the perspective of a data provider are not taken into consideration.

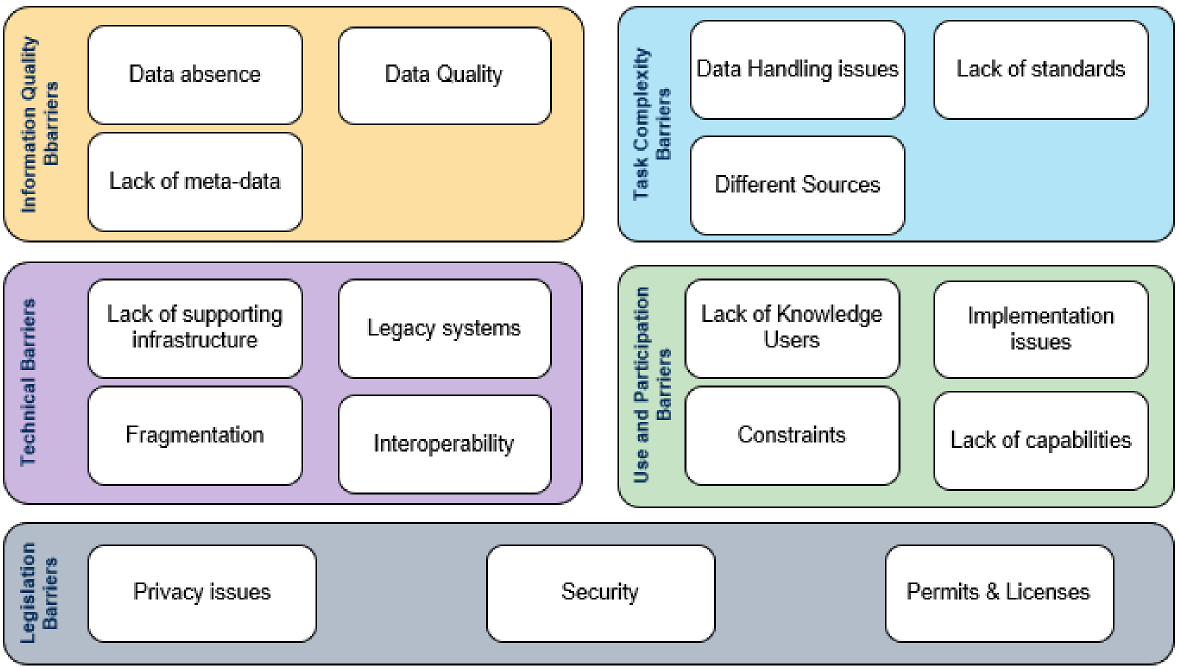


Figure 7: Conceptual model of the adoption barriers for BOLD.

2.4.2. Information Quality Barriers

The first cluster that is identified represents the 'Information Quality Barriers'. Information Quality is one of the most often identified problems for the adoption of BOLD. As where BOLD datasets consist of huge amounts of data and are relatively cheap, the datasets do not automatically guarantee quality. Three sub-clusters are identified: Data Absence; Data Quality and Lack of Meta-Data. Figure 8 depicts the addressed adoption barriers mapped into the sub-clusters.

The first sub-cluster consists of the adoption barriers that are related to data absence. In BOLD datasets, there is a chance that some data is lacking. Different authors describe that the open data-sets are often incomplete, and information is missing (Conradie & Choenni, 2014; Janssen et al., 2012; Mclaren, 2011; Saxena, 2018b; Vetrò et al., 2016).

The second sub-cluster consists of multiple adoption barriers that are related to the quality of the data. Often the accuracy of the information is lacking, which means that the datasets can consist of faults. Next to that, the data can be invalid or obsolete (Conradie & Choenni, 2014; Janssen et al., 2012). Furthermore, BOLD datasets often consist of duplicated data. The same data is captured in various forms, which causes a bias and negative effects on the performed data analysis (De Donato et al., 2018; Janssen et al., 2012). BOLD datasets often do consist of high-level quality data and can be outdated (De Donato et al., 2018; Donker, van Loenen, & Korthals Altes, 2017) Therefore, it is important that the BOLD datasets are implemented in real-time.

The third sub-cluster represents the adoption barriers that are related to the lack of meta-data. This means that the data often does not consist of an explanation of the meaning of the data (Janssen et al., 2012). Often data does not have any meta-data, or it is very poorly documented (Hossain, Dwivedi, & Rana, 2016; Vetrò et al., 2016; Zuiderwijk, Jeffery, & Janssen, 2012). Data without a meaning is useless. Besides that, there is often no (technical) information available about the data, which could give insights whether the data is already processed (Saxena, 2018a; Thorsby, Stowers, Wolslegel, & Tumbuan, 2017; Vetrò et al., 2016). Furthermore, often information about the quality of a data-set is lacking, which could be very helpful to assess whether to use a dataset for an analysis or not (Janssen et al., 2012).

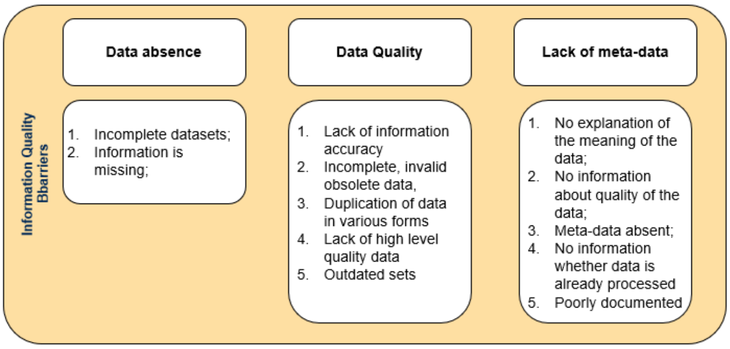


Figure 8: Information Quality Barriers.

2.4.3. Task Complexity Barriers

The second cluster is representing the 'Task Complexity Barriers', which include the barriers that make the adoption of BOLD in real-time difficult because of the complexity of the process. The barriers can be divided into three sub-clusters: Data Handling issues, Variety in Sources, Size and Types, and a Lack of Standards. Figure 9 depicts the addressed adoption barriers mapped into the sub-clusters.

The first sub-cluster represents the adoption barriers that are related to the Data Handling Issues that occur when integrating the data. Many organizations are struggling with the complexity of handling the collected data and integrate this in real-time into the applications used for decision making. For example, often an index is lacking, which makes combining and linking the datasets very difficult. (Conradie & Choenni, 2014; Janssen et al., 2012).

The second sub-cluster includes the barriers that are most-recognized and mentioned publications, namely 'Variety in sources, types and size' (Attard, Orlandi, Scerri, & Auer, 2015; Gerunov, 2017; Lněnička & Komárková, 2018). This makes combining the data very time-intensive, or even impossible without additional information.

Moreover, datasets can also be embedded in non-user-friendly files, such as PDFs, images, physical copies (Chatfield & Reddick, 2017; Corrêa, Paula, Correa, & Silva, 2017). Lastly, many organizations cannot deal with the unstructured state of the data (Lněnička & Komárková, 2018).

The third sub-cluster represents the barriers related to a lack of standards. A lack of standards makes the tasks way more complex. This happens in various forms. Firstly, the standards are or not well defined, or absent. When there are no standards for a well-defined format, a consequence will be that BOLD datasets will exists in a wide variety of types and formats. Also, the lack of meta-data standards make the standards hard to interpret and sometimes useless (Conradie & Choenni, 2014; Janssen et al., 2012). Lastly, data management without standards will in the end also cause task complexity. So, it is important that there will be defined clear standards, to make the way of working in the future less difficult (Conradie & Choenni, 2014; M. Janssen et al., 2014).

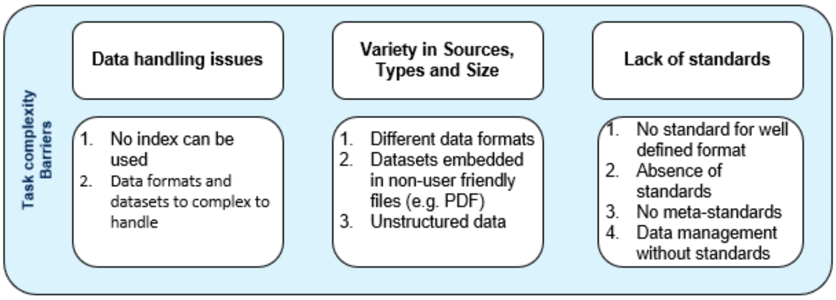


Figure 9: Task Complexity Barriers.

2.4.4. Technical Barriers

The third cluster is representing the technical barriers that include all the barriers relating to the required technology. These show that the need for a suitable technology landscape is indispensable to enable the real-time adoption of BOLD. The barriers of this cluster can be mapped into three sub-clusters: Lack of Supporting Infrastructure; Legacy Systems; and Interoperability Problems. Figure 10 depicts the addressed adoption barriers mapped into the sub-clusters.

The first sub-cluster represents the adoption barriers related to the lack of a supporting infrastructure. The lack of a supporting infrastructure is making the real-time adoption difficult. Different authors mention that a lack of tooling and technical resources are causing a poor infrastructure, such as supporting data sharing channels are forming a problem (Gong & Janssen, 2017; Huijboom & Van den Broek, 2011; Janssen et al., 2012). Besides that, often an integrated (open data) central portal or architecture is missing for the facilitation of BOLD integration (Conradie & Choenni, 2014; Janssen et al., 2012). Lastly, there is need for standard software for processing the open data (Janssen et al., 2012).

The second sub-cluster represents the presence of legacy systems. Organizations rather maintain and improve the current landscape of systems than opting for change. This complicates the process when adopting BOLD, since a reactive approach and the real-time integration are conflicting to each other. For the real-time adoption of BOLD, it is important the infrastructure can integrate ready-to-use data. This complicates the process if there are many legacy systems that should be made ready for this adoption. Relatively new organizations, that do not have large amounts of data yet will be way more flexible for the adoption (Janssen et al., 2012; Zuiderwijk, Janssen, Choenni, Meijer, & Alibaks, 2012).

The third sub-cluster represents the barriers related to Interoperability Problems, which is a major issue for real-time adoption. Different authors mention this as one of the critical factors of the adoption of BOLD (Gong & Janssen, 2017; Nugroho, Zuiderwijk, Janssen, & de Jong, 2015). It is important that systems are flexible enough to adopt data from external data silos. Hence, combining different systems seems to be a difficult challenge. Besides that, companies have accessibility problems to BOLD portals due to a variety of reasons, such as problems that arise with the user and access rights, or required registration(Corrêa et al., 2017; Saxena & Muhammad, 2018).

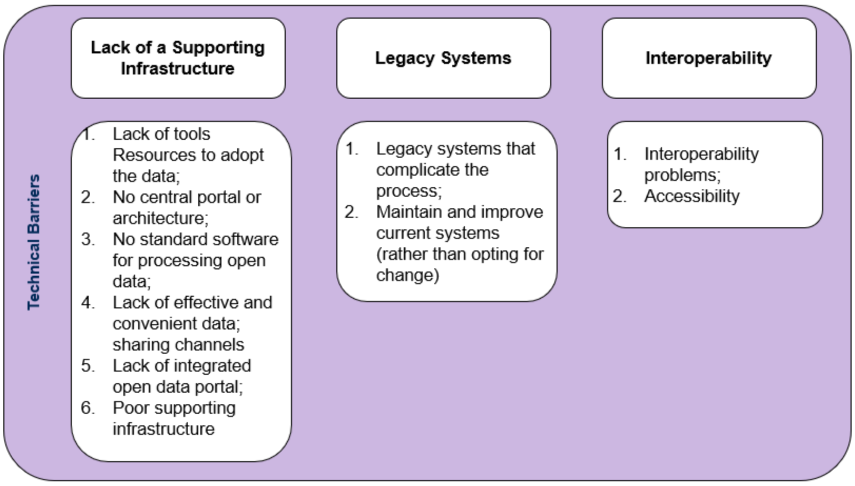


Figure 10: Technical Barriers.

2.4.5. Use and Participation Barriers

The fourth cluster represents the 'Use and Participation' barriers, which are categorized into four sub-clusters: Lack of Knowledge Users; Lack of Capabilities; Constraints; and Implementation Issues. This cluster includes the adoption barriers that are related to the problems that hinder the use and participation of the human user. This can have different Figure 11 depicts the addressed adoption barriers mapped into the sub-clusters.

The first sub-cluster includes the adoption barriers that are related to a lack of knowledge. Many organizations are not aware about the potentials, the issues and limitations the use of BOLD address (Janssen et al., 2012; Lee & Kwak, 2012). Janssen et al. (2012) state that using more complex and sophisticated data, such as BOLD, requires specific knowledge to be able to use it. Many organizations do not have this knowledge in-house, which forms a problem (Ohemeng & Ofosu-Adarkwa, 2015; Saxena & Muhammad, 2018). A lot of time is invested in knowledge on how the data can easily be used in software applications, where the data must be linked and combined (Janssen et al., 2012; Pennington & Cagnazzo, 2018). Also, organizations are lacking statistical knowledge about how to interpret and present the analysis with BOLD. Lacking this knowledge makes generating value out of the data more difficult (Lee & Kwak, 2012; Saxena, 2018b). Lastly, a lack of data literacy is making the adoption of BOLD less attractive for organizations (Lee & Kwak, 2012; Ohemeng & Ofosu-Adarkwa, 2015).

The second sub-cluster is representing the barriers related to the lack of capabilities that organizations should have to be able to adopt BOLD (Janssen et al., 2012). In many cases, organizations do not have the human and financial resources, expertise and implementing skills to adopt BOLD. Without these capabilities, organizations can generate less value. Moreover, organizations often do not have the technical and human capabilities to collect and extract the data, to process the data, to analyze the data, to interpret the data and to use the data (Lee & Kwak, 2012; Pennington & Cagnazzo, 2018). This is still a big challenge for organizations.

The third sub-cluster represents adoption barriers that are related to the different constraints to the adoption of BOLD. This can be either time constraints, financial constraints, technical constraints (De Donato et al., 2018; Janssen et al., 2012). Different authors mention that the extensive requirements, and terms and conditions often form a barrier for the adoption of BOLD (Huijboom & Van den Broek, 2011). For example, often registration is required before being able to download the data, which makes real-time integration of the data more difficult (Janssen et al., 2012). Besides that, the attempt of adopting the data often results in unforeseen implementation costs (De Donato et al., 2018; Janssen et al., 2012).

The final sub-cluster of the Use and Participation Barriers is representing the implementation barriers. The implementation issues are the most difficult barriers to overcome, because the human factor causes the problem. Firstly, a lack of user participation causes a problem for the use and re-use of BOLD. If people do not adopt BOLD according to the standards that are provided, the sustainability of the data decreases and the quality of the data will be influenced (Janssen et al., 2012; Lee & Kwak, 2012). Secondly, there can be competing interest and complicated relationship among the stakeholders involved with the BOLD data sharing (Janssen et al., 2012). This negatively influences the implementation of a strategy that includes the use of BOLD. Also, it is hard to define a clear strategy of how to use the data and use analytics to compete. Without a clear strategy, there will always be implementation issues.

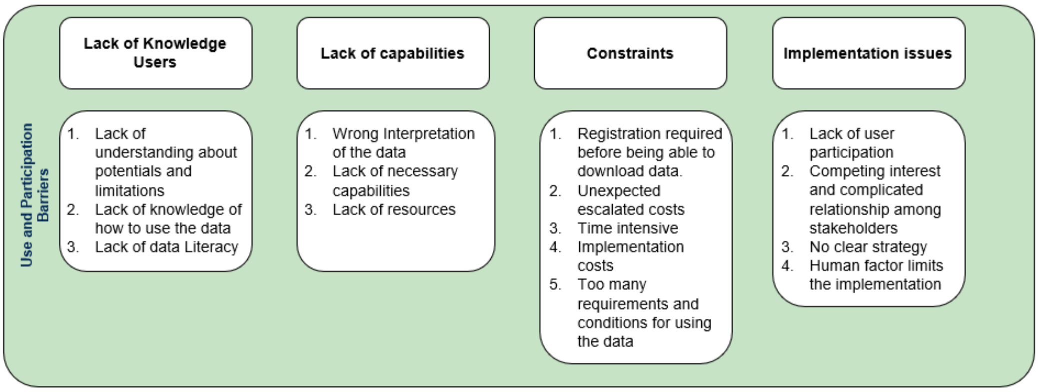


Figure 11: Use and Participation Barriers.

2.4.6. Legislation Barriers

The final cluster that is extracted from the systematic review represents the 'Legislation barriers', which includes all the legal barriers of the adoption of BOLD. The importance of the legislations depends of what sector an organization is part of, but the impact of the barriers can be enormous. This category consists of adoption barriers related to privacy, permits and licenses, and security hurdles.

The first sub-cluster represents the privacy-related adoption barriers. Firstly, the adoption of BOLD can violate the privacy of users by making sensitive customer data available after linking and combining the data (Guinney et al., 2017; Zuiderwijk & Janssen, 2014). The enactment of the General Data Protection Regulation (GDPR) makes this barrier even more relevant for organizations in the process of BOLD adoption (Janssen et al., 2012). Besides that, Zuiderwijk and Janssen (2014) state that data misuse is still also an issue that BOLD providers should be aware of. Unfortunately, open data also consists of a dark-side. Figure 12 depicts the addressed adoption barriers mapped into the sub-clusters.

The second sub-cluster represents the adoption barriers that are related to the issues with Permits and Licenses. First, Data providers that own the data, have copy right on the data and have licensing restrictions on the data must allow the use of the data first (Janssen et al., 2012; Mclaren, 2011). This forms a huge adoption barrier for organizations, because they do not want to violate legislations. Without the permissions, very high fees can be imposed. Many organizations cannot be at such a risk. Secondly, the data can have limitations attached to it. For example, some of the open data cannot be used for business goals without paying a fee (e.g. only for educational purposes). This makes the data less valuable for many organizations. Furthermore, some information or data can be confidential and can limit the data that can be used for decision-making (Janssen et al., 2012; Mclaren, 2011).

The final sub-cluster is representing the adoption barriers related to security issues. Organizations deem the security requirements as a problem, because providing data and securing the data at the same time is a big challenge (Guinney et al., 2017; Janssen et al., 2012). Furthermore, accessibility for the adoption of a dataset is often an issue. Accessing the data can become difficult because of data encryption. This is especially a problem within organizations that have confidential data.

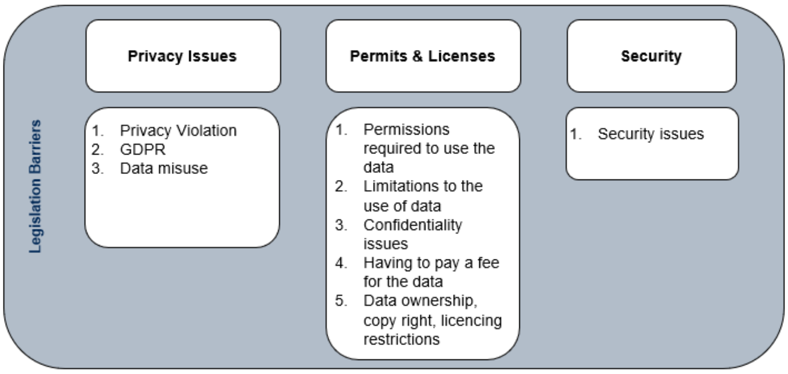


Figure 12: Legislation Barriers.

2.5. Summary Chapter 2 (Answer RQ1)

This chapter gives an answer to the first sub-research question. The first sub-research question ***'RQ1. Which adoption barriers affect the integration of external Big and Open Linked Data with internal company data?'*** identified all factors that hinder the adoption of BOLD. Through a systematic literature covering the primary studies on adoption barriers of BOLD published of the last 8 years, a conceptual model for adoption barriers of BOLD was developed. In total, five categories of adoption barriers are identified that the real-time integration of BOLD for data-driven decision-making difficult: Information Quality Barriers; Task Complexity Barriers; Technical Barriers; Use & Participation Barriers; and Legislation Barriers. Each category is divided into sub-clusters representing all addressed adoption barriers. Each identified adoption barrier was mapped into one of these sub clusters, which is presented in Figure 7. Many factors are influencing the adoption of BOLD, which makes the adoption of external BOLD in real-time very ambitious and difficult.

The rationale for creating a conceptual model of the different adoption barriers that organization encounter is to explicate the research problem. A problem can only be solved when the factors that cause the problem are known. The wide range of adoption barriers identified through the literature review and expert interviews show that organizations are struggling with the real-time adoption of BOLD. A wide range of factors is are causing the problem. The presence of adoption barriers show that organizations are not well aligned across and between organizations.

The key findings of RQ2 are:

I. Enriching internal data with external BOLD offers a lot of potential for real-time decision-making, but it addresses many challenges and adoption barriers.
II. Organizations are encountering a variety of adoption barriers during the different phases of the BOLD life cycle that makes the real-time adoption difficult.
III. The adoption barriers identified can be categorized into 5 overall categories, namely Information Quality Barriers; Task Complexity Barriers; Technical Barriers; Use and Participation Barriers; and Legislation Barriers.
IV. The Information Quality Barriers can be divided into three sub-clusters: Data absence; Data quality problems and Lack of meta-data.
V. The Task Complexity Barriers can be categorized into three clusters: Data Handling Issues; Lack of Standards and Different Sources.
VI. The Technical Barriers can be categorized into three different clusters: Lack of supporting infrastructure; Difficulties with Legacy Systems and Fragmentation of the data.
VII. Use and Participation Barriers can be clustered into the Lack of knowledge of Users, Implementation Issues, the existence of Constraints (Money, time, etc.), and the lack the right capabilities to adopt external BOLD in real-time.
VIII. Legislation barriers can be clustered into Privacy Issues, Permits and Licenses restrictions, and security issues.

3. Literature Review: Foundation of Enterprise Architecture

In this chapter, the second sub-research question '*RQ2. What are the capabilities from the current foundation of EA that need to be incorporated in the Extended Enterprise Architecture Framework to mitigate the adoption barriers?'* is discussed. Section 3.1. presents the literature review approach. Section 3.2. and 3.3. present the systematic literature review on the foundation of EA. Section 3.4. presents the EA capabilities that should be incorporated. Section 3.5 presents an assessment the most commonly used EAF for the selection of the most suitable EAF for extension.

3.1. Literature Review Approach

Enterprise Architecture is commonly used for the implementation of a new strategy across the Enterprise to ensure all levels of the enterprise are continuously aligned (Minoli, 2008). The systematic literature review on adoption barriers of external BOLD resulted in a conceptual model of the different adoption barriers that organizations encounter when adopting external BOLD. The presence of the adoption barriers show that there is a lack of alignment across organizations and between the organizations that are concerned with the integration of external BOLD. As discussed in the introduction, one of the knowledge gaps that are addressed is that there is limited insight whether the foundation of EA also can be used to align between organizations. In this chapter, the foundation of Enterprise Architecture is reviewed on its theories, methods, experiences and expertise, and existing meta-artifacts and there is critically reflected on its capabilities. The output of the literature review synthesizes the most relevant information and presents a list of EAF capabilities that should be incorporated by the Extended Enterprise Architecture Framework. The systematic literature review is conducted following the guidelines defined by Kitchenham et al. (2009). The review process involves activities in three core phases: 1) The Planning Phase, 2) the Conducting Phase and 3) the Reporting Phase (Figure 5).

3.1.1 Step 1: The Planning Phase: Development of a Review Protocol

The review protocol that is developed in Chapter 2 is followed to conduct the systematic literature review on the foundation of EA. This section discusses changes made to the research protocol to make it suitable for answering RQ2.

Research Question

To perform a focused systematic review, a research question is defined that gives an answer to the second sub-question of this thesis. The output of the systematic reviews thoroughly explicates the research problem and is used to set up the explorative interviews for finding requirements for a solution. In total, one research question was defined:

RQ2. What are the capabilities from the current foundation of EA (Theories, Methods, Experience and Expertise, and meta-artefacts) that need to be incorporated in the Extended Enterprise Architecture to mitigate the adoption barriers?

The following sub question is defined. To prevent reinventing the wheel, the Extended Enterprise Architecture is based on existing efforts. This sub question focuses on the selection of the most suitable EAF assessed on the extracted capabilities from the literature review.

RQ2a. What Enterprise Architecture Framework is most suitable for extension?

Inclusion/Exclusion criteria

In total, three different search engines are used to find relevant publications: Scopus, Google Scholar and IEEE. The rationale for choosing Google Scholar and Scopus is discussed in Chapter 2. In this literature review IEEE Access is added, because it provides open access to a variety of sources that publish architecture related papers. For all academic search engines, the same inclusion/exclusion criteria are used. To get to the results the following search query is used:

TITLE-ABS-KEY ("Enterprise Architecture" OR "Architecture Framework" AND capabilities OR "Success Factors") AND (LIMIT-TO (SRCTYPE , "p") OR LIMIT-TO (SRCTYPE , "j") OR LIMIT-TO (SRCTYPE , "k"))

Table 4 presents an overview of the search results and the inclusion/exclusion criteria that are used to get to the number of publications considered to review. Both primary and secondary studies are included for the systematic literature review.

Table 4: Search results SLR Foundation EA

Source	Number of publications	Year	Subject area	Document Type	Language
Scopus	*609*	*2000-2019*	*All*	*Conference Paper; Journal Article; Book Chapter;*	*English*
IEEE Access	*773*	*2000-2019*	*All*	*Conference Paper; Journal Article; Book Chapter;*	*English*
Google Scholar	*1000+*	*2000-2019*	*All*	*Conference Paper; Journal Article; Book Chapter;*	*English*

3.1.2 Step 2: The Conducting Phase

Following the research protocol, the review is conducted in 5 stages: 1. Identification of research 2. Selection of studies 3. Study quality assessment 4. Data extraction and monitoring progress 5. Data synthesis. The results are discussed in the reporting Phase.

3.1.3 Step 3: The Reporting Phase

After conducting the literature review, the collected data is synthesized and put in a report. The report consists of three steps (Figure 13). First, the foundation of EA is discussed. Secondly, the most relevant information is used to create a list of EA capabilities that should be incorporated by the Extended Enterprise Architecture Framework. Thirdly, the most suitable EAF for extension is selected.

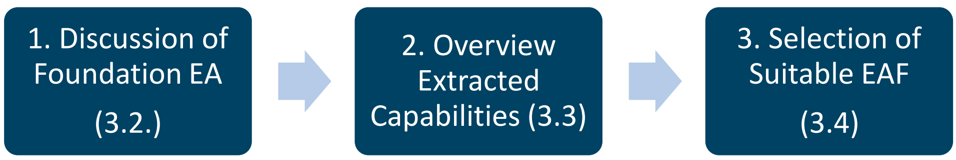


Figure 13: Structure Report

3.2. The Foundation of Enterprise Architecture

The data collected is synthesized by discussing the foundation of EA, including the capabilities extracted from existing theories, methods, experiences and expertise, and existing meta-artifacts.

3.2.1. Introduction to Enterprise Architecture

First, it is important to understand the foundation of EA and its context. In general, EA is used as a blueprint for the optimal deployment of the resources and capabilities to create an unified IT environment across the firm to support a new strategy (Minoli, 2008). Moreover, the goals are better alignment, standardization and the re-use of existing IT assets. The transition plan of an EA includes a map of IT assets and business processes, and a set of governance principles that help to express a new strategy through IT (Minoli, 2008). EA is therefore often seen as a management tool (Figure 14), that helps to ensure a strategy is aligned with the required products, processes, people and IT (Jonkers et al., 2006).

Figure 14: EA as a management tool (Jonkers et al., 2006).

A great number of EAFs are in existence that can be used for the development of an EA. An EAF provides an approach that can be followed by organizations for building a suitable EA. 'Architecture Frameworks make use of viewpoints to create views that represent the different perspectives of a system model' (Tang, Han, & Chen, 2004). In other words, each type of stakeholder has a different idea of how the system should look like, and it is important that each of these views are represented in the EA. According to the literature reviews by Minoli (2008) and Tang et al. (2004), most EAFs have 4 core views (Layers), which are Business Architecture, Information Architecture, Application Architecture and Technology (Infrastructure) Architecture. An explanation of the different views is depicted in Table 5.

Table 5: Main domains EAF

Layer	Focus
Business Architecture	A description of the structure and interaction between the business strategy, organization, functions, business processes, and information needs (Gartner, 2013).
Information/ Data Architecture	Point out where the core blocks of information are kept and how these information blocks can be addressed; And next to that also the sources of the information and the interrelations between this information blocks (Gartner, 2013).
Application Architecture	A description of the applications used in an organization, how these applications interact with each other and with its users (Gartner, 2013).
Infrastructure (Technology) Architecture	Includes the hardware, storage systems and networks of the company (Gartner, 2013).

Figure 15 presents a view of the macro environment of EA. First, external entities (e.g. customers, opportunities, regulators, technology, etc.) cause change in the business strategy of organizations. For example, the availability of BOLD is an opportunity that makes organizations develop a new strategy that is intended to improve decision-making. Organizations often do not have the right business and IT assets, and processes to implement a new strategy. So, the objective is to develop an IT infrastructure that enables, supports and facilitates this new business strategy. To this end, an organization usually has created an EA that provides a blueprint of the business, data, application, and technology architecture of the firm (Janssen, 2009; Minoli, 2008).

Organizations normally use industry mechanisms to develop the EA. Architecture Frameworks and models guide organizations through the design of a new EA and normally include: the IT industry techniques and methods to develop an EA; the architecture principles; and EA standards. Architecture development tools help to develop the EA. Next to the industry tools, also governance mechanisms ensure continuous performance of the EA, by means of different tools that help to identify the necessary adjustments to the EA. The outcomes of the industry mechanisms and governance mechanisms affect the development plan that results in the end state of the target architecture that is required to implement the new strategy.

Minoli (2008) explains that layered frameworks and models for EA are useful, because it ensures that the different architectures are non-overlapping. As mentioned in the Problem Exploration Phase, there are a great number of EAFs and models available. Organizations make use of different types of Enterprise Architecture Frameworks: Consortia-developed frameworks (e.g. TOGAF); Defense Industry Frameworks (e.g. DoDAF); Government Frameworks (e.g. FEAF); Open Source Frameworks (e.g. MEGAF); and Proprietary Frameworks (Gartner Enterprise Architecture).

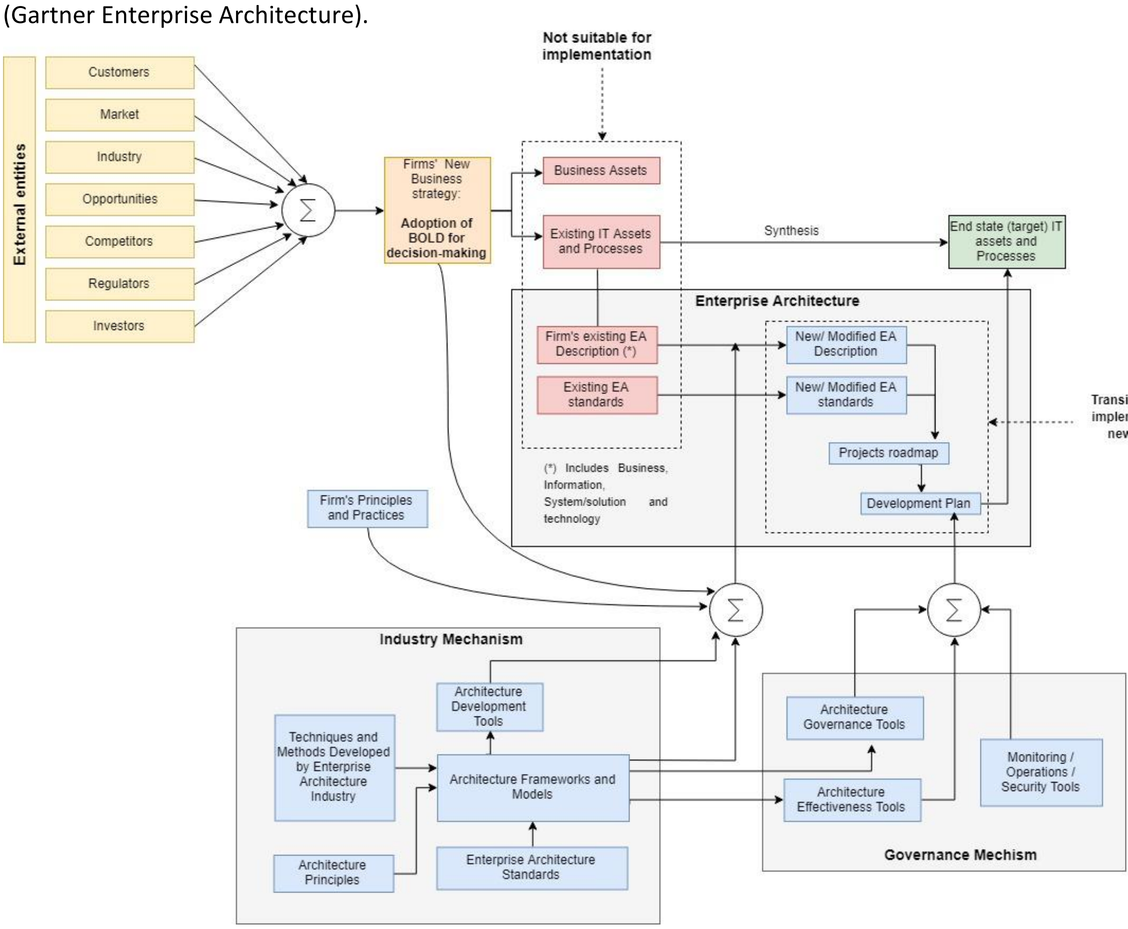


Figure 15: Macro Environment of Enterprise Architecture (Edited from (Minoli, 2008))

3.2.2. Architectural description

In this thesis the IEEE-Std-1471-2000, Recommended Practice for Architectural Description of Software-Intensive Systems is used as a standard for the description of the EA and EAF. The IEEE Computer Society has developed this standard as a recommended practice and organizations can decide whether to and how to incorporate these standards for the description of their EA. These standards are developed to define a direction for architectural thinking into standards; to create a framework and vocabulary that makes discussing architectural issues easier; and identify sound practices to establish a framework and vocabulary for software architecture concepts (Hilliard, 2000; Maier, Emery, & Hilliard, 2001). The IEEE-Std-1471-2000 focuses on complex systems where software is used to design, construct, deploy and evolve the system as a whole and the knowledge base is used to embrace the benefits of usability, flexibility, reliability, interoperability (Hilliard, 2000).

The IEEE-Std-1471-2000 provides a recommendation for how to describe an architecture. It is important to mention it is not a process, EAF, or methodology that should be followed to create an Enterprise Architecture. The architecture description is a set of artefacts to describe an architecture. The IEEE-Std-1471-2000 provides a conceptual framework that shows the relation of all relevant terms and concepts related to architectural thinking and helps to show the relationships between them (Figure 16). The role of this Conceptual Framework was to provide terminology in a research field where no common language existed (Maier et al., 2001). The conceptual framework helps Enterprise Architects with the establishment of Architectural descriptions in the context of stakeholders of the system, the life cycle of the system and the uses of Architectural description (Hilliard, 2000).

As depicted in Figure 16, when describing an Enterprise Architecture, the related and relevant stakeholders of the complex system are identified together with their concerns. According to Hilliard (2000) concerns form the basis for completeness of a system. If not all the concerns are covered, misalignment can be caused. Typical stakeholders are the Client, Owner, User, Operator, Architect, System Engineer, Developer, Designer, Builder, Maintainer, Service Provider, Vendor, Subcontractor, Planner (Hilliard, 2000). Dependent on the stakeholders involved in the project, an architectural description is organized by one or more views. 'A view is a representation of a whole system from the perspective of a set of concerns' (Maier et al., 2001). This can be compared with the building plan of a house, which could also be organized by multiple views, including the front plan, floor plan, plumbing plan, electrical plan, that each cover the concerns of different stakeholders involved in a business plan (Maier et al., 2001). A view consists of one or multiple models that represent an aspect of the complex system. In other words, views are modular and multiple modules can participate in views. A view is only related to one system and cannot be used across different systems. An Architectural Description should document every inconsistencies between the views (Hilliard, 2000).

A view conforms to exactly one viewpoint. Viewpoints capture descriptive frameworks captured from common architecture practices. According to Maier et al. (2001) 'viewpoints are the vehicles for writing, reusable, domain-specific architecture description standards'. A viewpoint includes the name of the viewpoint; the stakeholders it addresses; the associated concerns of these stakeholders; the viewpoint language, modeling techniques or analytical methods used; and the source that supports this viewpoint. Additionally, patterns or guidelines that support the creation of the view; consistency and completeness checks for well-alignment of the view; and evaluations and analyses methods for the models within the view (Hilliard, 2000; Maier et al., 2001). To make this more understandable, Table 6 represents an example.

Table 6: example of a viewpoint

Viewpoint name	**Business Context Viewpoint**
Stakeholders	Acquirer, architect, user
Concerns	- How to keep the system consistent? - How to make the system effective? - How to finance the system? - What process is required for successful implementation?
Viewpoint language	Block function diagram, Archimate
Source	Interview organization A

It is important to mention that viewpoints are not system specific, unlike the stakeholders and views. For this reason, viewpoints can be re-used for viewpoints descriptions. Therefore, the viewpoint library can be used for the re-use of relevant viewpoints supported by a citation, reference or other source. But it is also possible to establish a new viewpoint, based on interviews or other sources. As already mentioned, an EA is the blueprint of a system that is inhabits the environment and fulfills one or more missions. Change in the environment or missions may result in the need for a new or adjusted architecture and therefore a new architectural description. This is because change in environment may cause or create new concerns, which should be covered in the EA.

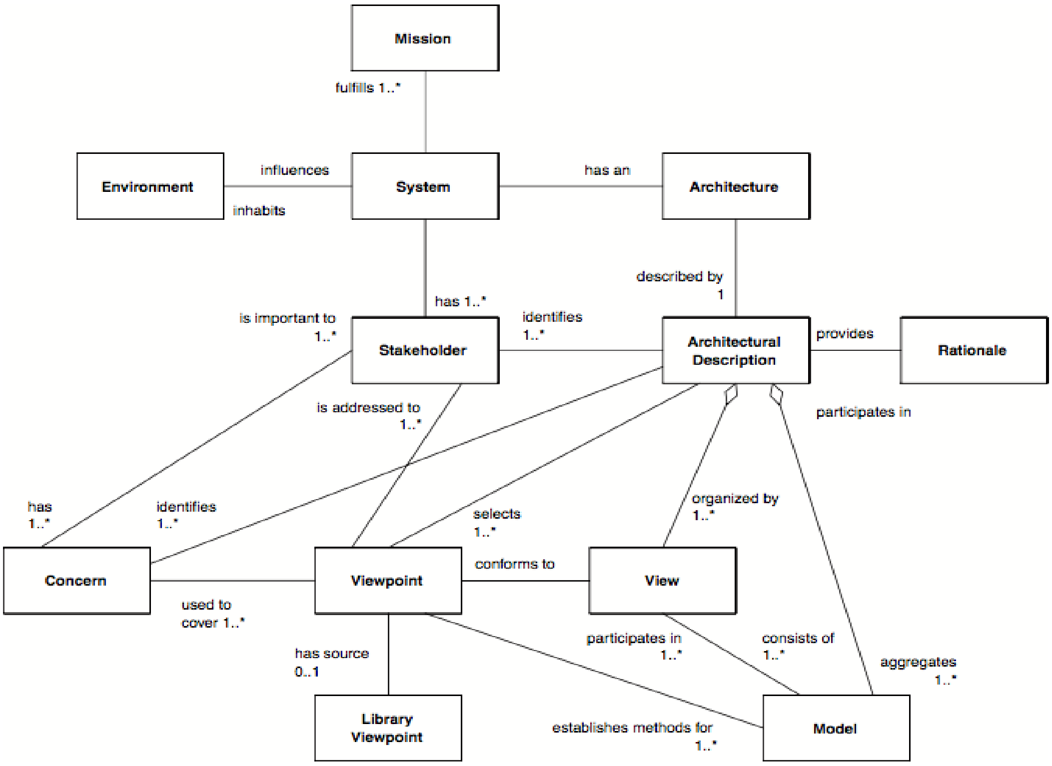


Figure 16: IEEE-Std-1471-2000. Conceptual Framework for Architectural Descriptions.

The IEEE-Std-1471-2000 conceptual framework for architectural description encompasses most EAFs. It makes the description possible for EAFs with one view (such as CMU), but also frameworks with multiple views, such as the Zachman Framework that is organized by 36 different views(points). Each established EAF is based on a selected set of techniques and methods developed by the industry, architecture principles and architecture standards. Hence, there are differences in the number of views and viewpoints between the existing EAFs.

3.2.3. Existing meta-artefacts: Most commonly Enterprise Architecture Frameworks

According to Hevner (2007), the foundation of a research topic also contains the existing meta-artefacts. Therefore, the most commonly used Enterprise Architecture Frameworks are analyzed and discussed, which the Zachman Framework; the Open Group Architecture Framework (TOGAF); and the Federal Enterprise Architecture Framework (FEAF) (Sessions, 2007; Urbaczewski & Mrdalj, 2006). For each EAF, a brief introduction is given, the key components of the EAF are described using the IEEE-Std-1471-2000 for description of Architectures (Figure 16) and, the EAF is assessed on the capabilities of EA.

Zachman Framework (ZIFA)

One of the first Frameworks for Enterprise Architecture is created by Zachman (1987). Zachman (1987) stated that the increased scope of design and the complexity level of information systems implementations required the use of a logical architecture. He introduced the Zachman Framework, which is 'a framework for Enterprise Architecture, which provides a formal and highly structured way of viewing and defining an enterprise'.

The basic purpose of this framework is 'that the same complex principles can be described for different purposes in different ways using different types of descriptions from 6 different perspectives' (Zachman, 1987). The framework allows you to describe any complex information system. The framework consists of two dimensions: Set of perspectives for describing complex enterprise systems; and a set of principles of the classical architecture that deals with 6 basic questions: 'What?' 'How?' 'Where?' 'Who?' 'When?' and 'What?' (Urbaczewski & Mrdalj, 2006). The architecture addresses 6 phases: Contextual, Conceptual, Logical, Physical, As Built and Functional. Each of these phases respectively correspond to the following Stakeholder roles: Planner, Owner, Designer, Builder, Subcontractor, and User. In total, the Zachman Framework results describing the 36 necessary categories to describe anything. The Zachman Framework does not provide guidance for the design of a system, but rather helps ensuring a complete system.

The Zachman Framework states that an entire Enterprise can be described by 36 different viewpoints that are associated with 36 deliverables. These deliverables are displayed in Figure 17 consisting of six deliverables per dimension. These deliverables are focused on inventory, process, network, organization, timing and motivation.

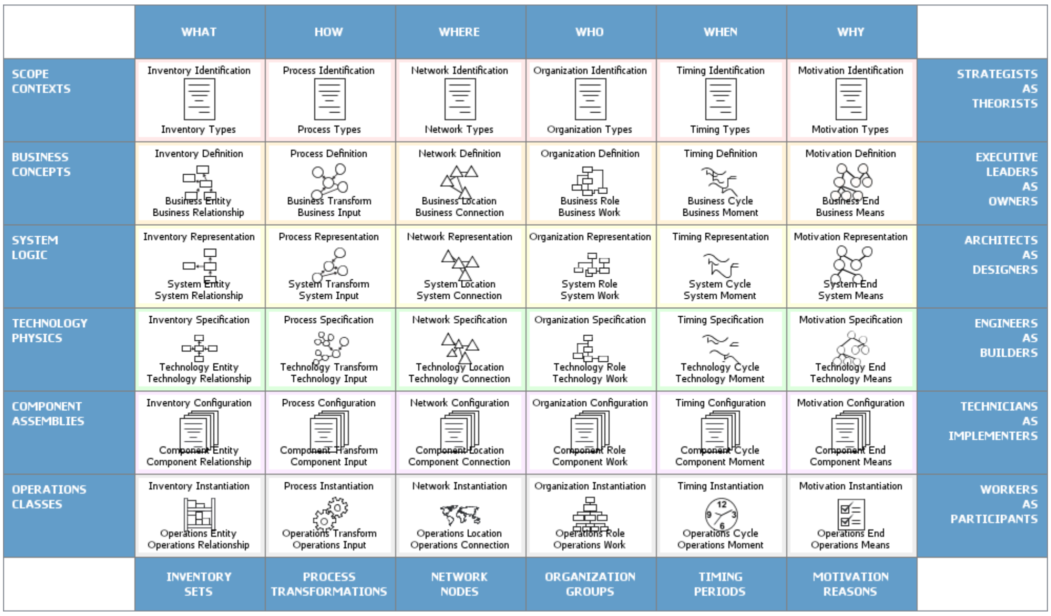


Figure 17: Zachman Framework (Zachman, 1987)

Each deliverable is a description of the system from the perspective of a stakeholder, also a so called 'viewpoint'. The Zachman Framework represents 6 different views, which are discussed in Table 7.

Table 7: View, Viewpoints Zachman Framework

View	Description	Viewpoints
Planner's view (Strategists/ Theorists)	Focus on the strategy and business objectives and goals. This view determines the scope in which the other views can work. It will provide the context in which the Owner's, Designer, Implementer, Subcontractor and Users view will be managed and derived.	Scope Contexts
Owner's view (Executives / owners)	This view describes what the organization must be capable and in which the information system must function.	Business Concepts
Designer's View (Architects)	This view describes how the information required will be satisfied by the system.	System Logic
Implementer's View (Engineers)	This view describes what solutions and technologies should be implemented to implement the system.	Technology Physics
Sub-constructor's View (Technicians)	This view describes what elements of systems are required and which elements need more attention before the system can be implemented.	Component Assemblies
User's View (Participants)	This view describes how the systems is used by the user in an operational environment.	Operation Classes

Components

According to Urbaczewski and Mrdalj (2006), the Zachman Framework consists of 4 different components: a Planning Component; an Analysis Component; a Design Component; and an Implementation Component of the complex system that should be developed to implement a new strategy. The maintenance Phase to keep the Enterprise Architecture up-to-date and capable of changes. Furthermore, the Zachman Framework does not provide guidance for the development and maintenance of a system in a dynamic environment, in which the architecture should be updated/maintained.

The Zachman Framework consists in total of 36 views (6 times 6), which makes this the EAF that has the most complete process. However, it does not provide any reference model guidance, practice model, maturity model, business focus, governance guidance, partitioning guidance, prescriptive catalog, vendor neutrality, information availability and time to value. Which makes it not useful to use it as a model for guidance for the development and maintenance of a EA. The real-time integration of external BOLD is a very dynamic process that require lots of maintenance, which makes the Zachman Framework less useful. However, it is still one of the most used EAFs, because it helps organizations making sure aspects of a complex system / of the enterprise are covered. The deliverables could be helpful to use in other EAFs, like TOGAF. Often, the Zachman Framework is used in combination with other EAFs.

Limitations[1] to the Zachman Framework (ZIFA) are (Sessions, 2007):

1. The ZIFA It is not clear how to unite the different deliverables, because each chain follows a different scheme.
2. The ZIFA does not provide guidelines for creating the new architecture.
3. It does not help with the decision-making in design choices.
4. Does not provide maintenance and on the current architecture and the target architecture

The Zachman Framework is perfect tool to determine the taxonomy of an enterprise (Sessions, 2007). However, it does not consider important components, such as maintenance, prescriptive architectures and it works best combined with a methodology that provides this guidance. Architectures are a continual process of creating, maintaining and leveraging the enterprise (Gartner, 2013).

The Open Group Architecture Framework (TOGAF)

The Open Group provided an EAF that can be used for the design, evaluation and building of architectures for Enterprises (Tang et al., 2004; Urbaczewski & Mrdalj, 2006). The Framework itself is a well-documented body of knowledge comprising a detailed method and supporting tools for developing enterprise Architectures. This method, the TOGAF Architecture Development Method (ADM), provides guidance during the process of developing an EA. TOGAF is designed by more than 300 architects from different world's leading organizations and is, therefore, based on general EA practice. Building and maintain an architecture is a process involving many stakeholders and decision processes. TOGAF facilitates this process, so organizations can build an Enterprise Architecture that is consistent, reflects the needs of stakeholders, employs best practice and considers the current requirements and the perceived future needs of the business. There are several common reasons why organizations are choosing for the TOGAF-ADM (Figure 18).

First, the EAF is vendor, tool and technology neutral using open standards. Secondly, the framework avoids that organizations must re-invent the wheel and allows organizations to build further on the framework. Thirdly, it ensures Business – IT alignment and is tailorable to meet a specific organization and industry needs. Fourthly, it is based on the best practices of architects from multiple market leading organizations. Fifthly, organizations can participate in the evolution of the framework. Sixth, it is free to use and widely adopted in the market.

[1] https://www.classes.cs.uchicago.edu/archive/2011/spring/51075-1/presentations/dray.pdf

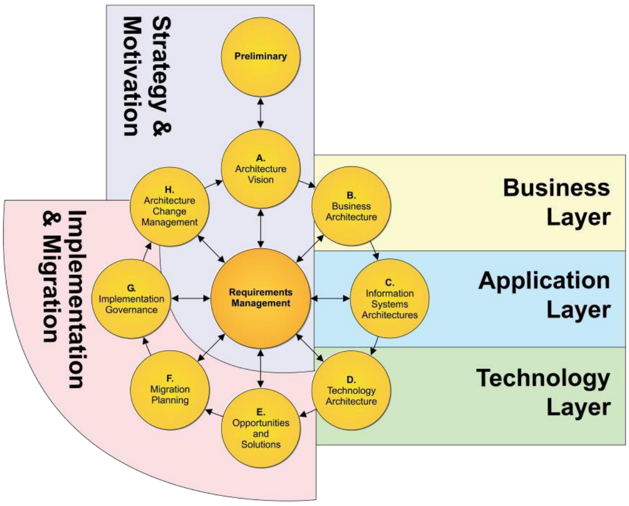


Figure 18: TOGAF-ADM

Components

The ADM consists of the following of 9+1 phases. First, the Preliminary Framework and principles are defined, which helps to define the base line architecture. This phase helps organizations to prepare the organization for a successful architecture project. (Tang et al., 2004). Secondly, the ADM-cycle starts by discussing the architecture vision (Phase A). The architecture vision sets the scope, constraints, and expectations for a TOGAF project. The architecture vision focuses on describing the current situation and the desired situation of the organization and it validates the business context. This is used to create a State of Architecture Work that is required for developing the different architectures (Group, 2019). Next, the Business Architecture is defined (Phase B). This is done by developing the baseline and target architectures and analyzing the gaps that should be filled by the target architecture. This is all documented and reviewed with the stakeholders. The same process is done for the Information Systems (consisting of Data architecture + Application architecture) and the Technology Architecture (Phase C+D) (Group, 2019).

After the architecture phases, the implementation and migration phases of the ADM-cycle are performed. In the Opportunities and Solutions phase, the initial implementation planning is done, which focuses on identifying the major implementation projects. If the transition is too much, there can be decided incremental architectures are required for the implementation. Also, the approach is determined how the transition will be realized. In phase F, the Migration Planning phase is performed. In this phase a concrete and detailed implementation and migration plan is developed, based on cost analysis, benefits, and risks identification. In the next phase (Phase G), the implementation and migration plan are implemented and architectural oversight for the implementation being done. This oversight ensures that each project that gets implemented is conform to the architecture (Group, 2019).

In the last phase, the Architecture Change Management Phase, continual monitoring and a change management process is provided that ensure the architecture responds to the needs of the enterprise. Any change in requirement scope, architecture vision, architecture domains, Planning, Implementation and Migration are monitored by the architecture change management and the governance decides if a transformation in architecture is required. In each phase, requirements are collected, stored and managed by Requirements Management, which ensures that every stage of the TOGAF project is based on and validates business requirements (Group, 2019).
TOGAF recommends a Business architecture view, Data architecture view, Application architecture view, Technology architecture view, System Engineering view, Enterprise Security view, Enterprise Quality view, Service view, and Enterprise manageability view. However, TOGAF allows for extension and modification by other Frameworks or Deliverables. It is usual to modify or extend the TOGAF-ADM to suit specific needs (Tang et al., 2004). TOGAF is a generic methodology, which can be used within a wide range of industries, geographies and sectors (Group, 2019).

Limitations to the methodology are that you are limited to four dimensions (Figure 19):

1. Breadth: Limited to the vision of entire company or parts of the company;
2. Depth: Enterprise level initiatives; Business domain level initiatives; Capability level initiatives; and smaller.
3. Time period: How much time organizations want to spend to implement.
4. Architecture domains: Limited to Business Architecture, Data Architecture, Application Architecture; and Technology Architecture.

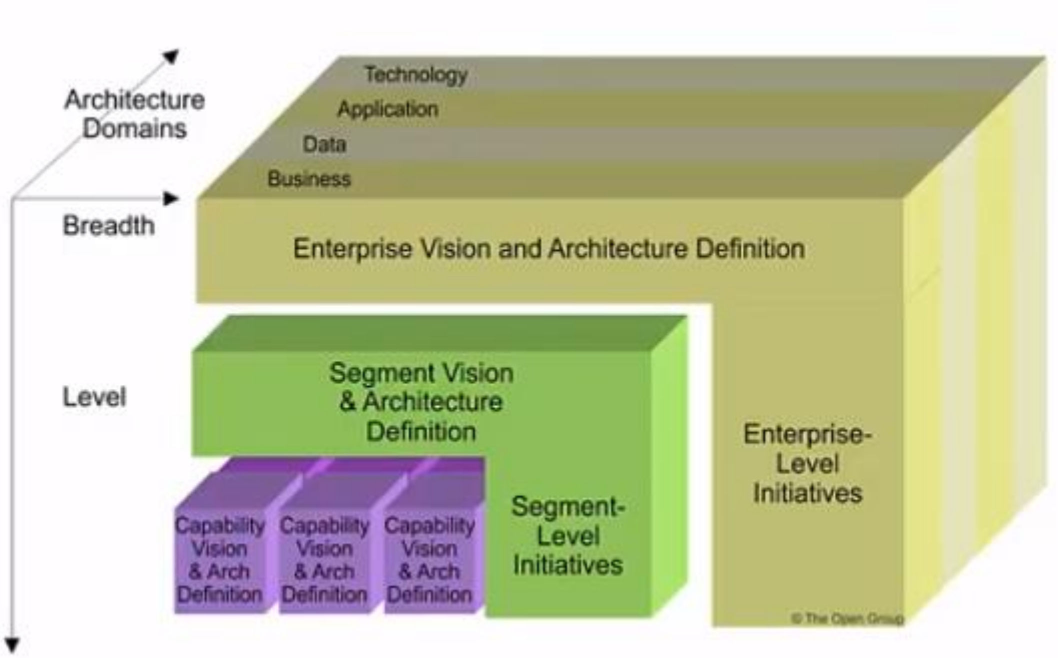


Figure 19: Integration levels EA (Group, 2019)

Federal Enterprise Architecture Framework (FEAF)

The Federal Enterprise Architecture Framework is developed by the U.S. Office Management and Budget and the U.S. Office of E-government and ICT to promote shared development of common US federal processes, interoperability and the sharing of information between Federal agencies and other entities of the government (Tang et al., 2004). The framework aims for improving design, use sharing and performance of federal information resources. The FEAF consists of four different levels. The first level concerns about the external stimulus and the strategic direction of enterprise and focus on the process of transforming the current architecture into a target architecture by following the architecture standards. The second level helps organizations with analyzing the business and design drivers of an architecture. The third level uses the business view, data view, application view, and technology view to model the target architecture. The last, fourth, level provides the methods to represent the different architecture layers (Tang et al., 2004; Urbaczewski & Mrdalj, 2006). The FEAF 'equips Federal agencies with a common language and framework to describe and analyze investments.

Components

The FEAF consists of 6 domains: Strategy, Business, Data, Applications, Infrastructure and Security. These domains are to develop to improve collaboration between different departments of the bureaucratic environment that otherwise would have been isolated. The focus is on utilizing the power of combining the 6 domains at each department within the environment. This can be achieved by using the methodology of FEAF consisting of five different steps. Step 1 focuses on identifying and validating. In this step all the stakeholders of the different departments are brought together to collaborate and develop requirements for the EA. Step 2 focuses on researching and leveraging. In this step the resources from the different departments are useful for other departments and can be utilized. Step 3 focuses on defining and planning. In this step the adjustments needed across the different domains are identified by defining the current situation and planning the future situation of the architecture. This results in a plan for transition. Step 4 focuses on investing and executing. This step the right investments to implement the plan for transition are made. The final step (step 5) is focusing on performing and measuring. This step monitors the performance and measures against metrics how the architecture is performing. The process is streamlined, and duplicated efforts are removed.

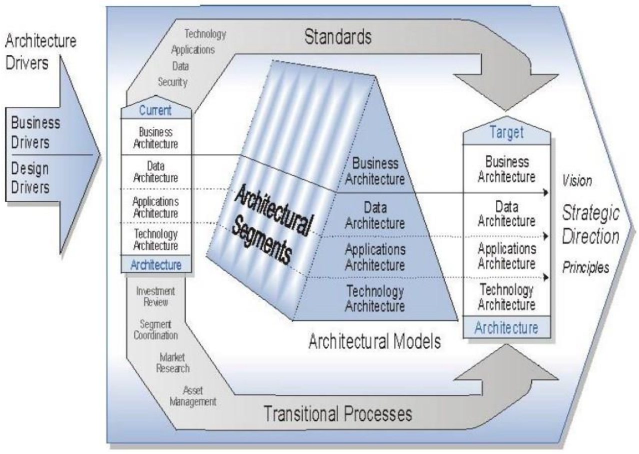


Figure 20: FEAF

3.3. Enterprise Architecture Capabilities

The second part of the report discusses the extracted EA capabilities from the literature. RQ2 focuses on identifying the capabilities of the current foundation of EA that should be incorporated in the design of the extended Enterprise Architecture Framework. To be able to explicate what additional EA capabilities are required to mitigate the adoption barriers, it is important to research what capabilities the existing foundation of EA has. Describing the foundation helps to define the knowledge that is lacking (Hevner, 2007).

In 2004, Schekkerman (2004) published a book on how to survive the jungle of Enterprise Architecture Frameworks. The author refers to the huge range of frameworks organizations can choose from when developing an Enterprise Architecture. To help organizations with the search for a suitable EAF, Schekkerman (2004) published a list of critical success factors that the ideal EA must have. This list of critical success factors is coherent with the existing knowledge base on EA, covering the scientific theories, methods, experience and expertise, and existing meta-artefacts. Based on this list of critical success factors an overview of the capabilities is developed that current EAFs can have.

Table 8 presents an overview of the capabilities that should be incorporated in the design of the Extended Enterprise Architecture Framework. For each capability the table presents an ID, description and the literature that substantiates the capability. This list of capabilities can be used to compare different EAFs with each other. In total 11 capabilities are extracted and should be incorporated in the design of the Extended Enterprise Architecture. Each capability is at least mentioned by one author.

Table 8: Capabilities Enterprise Architecture Frameworks (Edited from Schekkerman (2004))

ID	Capability	Description	Source
C1	Continuously align Business / IT	The framework creates and maintains a shared vision among the Business and IT perspectives of a business, by making sure the whole enterprise is aligned at all time.	Jonkers et al. (2006); Minoli (2008); Schekkerman (2004)
C2	Provide guidance for a process for transition towards the 'to-be' state of organization	The framework provides guidance for organizations to describe the current state of the organization and transforming the business towards the future state by means of a prescriptive architecture.	Janssen (2009); Jonkers et al. (2006); Minoli (2008); Schekkerman (2004)
C3	Lower the threshold for change	The framework builds agility of the enterprise by taking away the complexity barrier. This helps in less complex change of the organization.	Janssen et al. (2012); Schekkerman (2004)
C4	Enable/Increase flexibility of the organization	The framework makes it possible to form linkages with external partners and therefore increase the flexibility of an organization.	Schekkerman (2004)
C5	Develop a proactive organization	The framework helps organizations in meeting customer demands, driving innovation and overcoming competition.	Schekkerman (2004); Tang et al. (2004)

C6	Reduce risk and prepare for rapid, unplanned change	The framework helps to build an architecture that is able and prepared for unforeseen changes and the reduction of risk.	Schekkerman (2004)
C7	Avoid tension between business IT functions:	The framework removes and avoids friction between different viewpoints of the enterprise.	Minoli (2008); Schekkerman (2004)
C8	Create, unify and integrate business processes across enterprise	The framework helps organizations to create, unite and integrate business processes across the entire enterprise.	Jonkers et al. (2006); Matthes, Buckl, Leitel, and Schweda (2008); Schekkerman (2004)
C9	Unlock the power of information, unifying information silos.	The framework enables effective utilization of information among the enterprise and federating information silos for a variety of purposes.	Schekkerman (2004); Urbaczewski and Mrdalj (2006)
C10	Eliminate duplicate and overlapping technologies	The framework eliminates unnecessary applications and technologies.	Hilliard (2000); Schekkerman (2004)
C11	Reduce solution delivery time, development costs (maximize reuse):	The Framework ensures fast development time of the enterprise Architecture	Schekkerman (2004)

It is remarkable that the capabilities mainly enable alignment across the enterprise and not between an enterprise and its ecosystem, while a strategy that involves the real-time use of external BOLD concerns the business ecosystem. Figure 21 presents the different integration levels of EA and its limitations depicted in red. It shows that the current foundation is not capable of implementing ecosystem level initiatives based on ecosystem vision and architecture definition. The limitations are:

1. Breadth: Limited to the vision of entire company or parts of the company;
2. Depth: Enterprise level initiatives; Business domain level initiatives; Capability level initiatives; and smaller.
3. Time period: How much time organizations want to spend to implement.
4. Architecture domains: Limited to Business Architecture, Data Architecture, Application Architecture; and Technology Architecture.

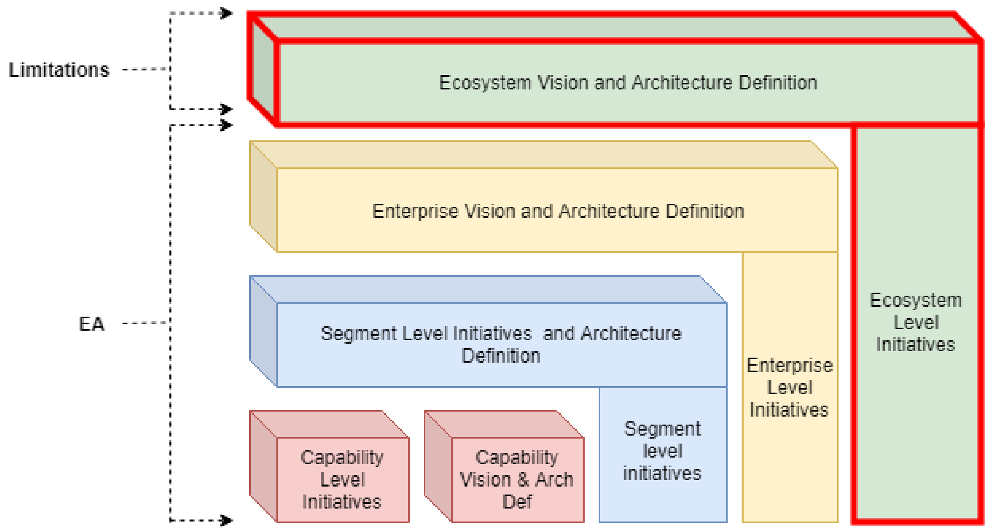


Figure 21: Limitation Integration levels of EA

In Chapter 4, the additional capabilities that are required to enable the real-time integration of external BOLD with internal company data are collected through explorative expert interviews. These additional capabilities should be built on the 11 capabilities from the current foundation of EA.

3.4 Selection of Suitable EAF for extension

The third part of the report is an assessment of the commonly used EAFs. The most suitable that should be extended to enable the real-time integration of external BOLD for decision-making is selected.

3.4.1. Selection Approach

The literature review on the foundation of EA shows that EA does not have the capabilities to align an organization and its ecosystem and, therefore, cannot deal with the adoption barriers of external BOLD. To enable the real-time integration of external BOLD with internal company data the foundation of EA must be extended with additional capabilities. To gain more insight what EAF is most suitable for extension, a method for comparison is developed. The purpose of the method is to assess the most commonly used EAFs and select an EAF that has most capabilities already incorporated. To avoid re-inventing the wheel, this EAF can be further build on to make it suitable for the real-time adoption of external BOLD and mitigating the adoption barriers.

'An Enterprise Architecture Framework is a set of models principles, services, approaches, standards, design concepts, components, visualizations and configurations that guide the development of specific aspect architectures' (Schekkerman, 2004). The benefit of an Enterprise Architecture Framework is that it makes available a design space in which different stakeholders can cooperate to solve a problem. Schekkerman (2004) published a book that helps architects with the search for a suitable EAF. The author describes that an organization can either choose or create an EAF, dependent on the business environment and the objectives of the organization. The creation of an EAF can be in various degrees: it can be the extension of an existing framework, it can be a combination of different frameworks or it can be a sophisticated framework from scratch. However, according to Schekkerman (2004) most of the time it is not necessary to start from scratch. It is better to avoid re-inventing the wheel and focus on extending an existing EAF where needed.

Hence, Schekkerman (2004) proposes a method with six steps in the creation of an EAF. Firstly, the business environment should be evaluated and understood well. Secondly, it is important that the goals and objectives the EAF should serve are defined. Thirdly, the currently existing frameworks should be assessed on suitability and one EAF is selected. In the fourth step, the selected EAF needs to be extended to the needs of organization. The fifth and sixth steps focus on testing, evaluating and refining the framework.

Based on this approach by Schekkerman (2004) the Revised EAF Selection Approach (R-ESA) was developed. In this approach the steps of the approach by Schekkerman (2004) are partly adopted. Only the steps for selecting a suitable base EAF are adopted (Figure 22):

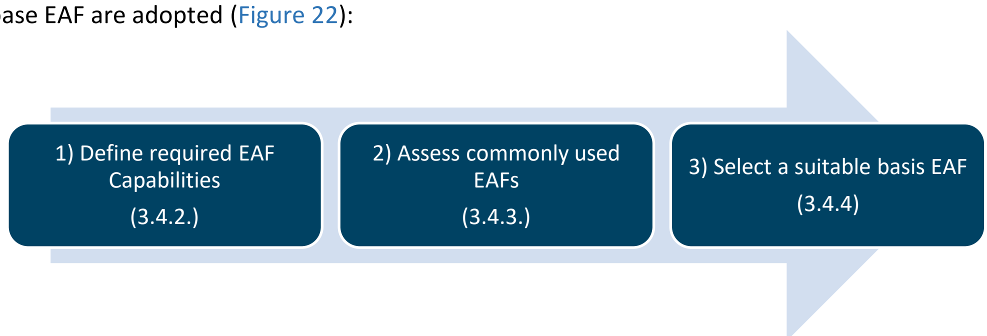


Figure 22: Revised EAF Selection Approach (R-ESA) (Edited from Schekkerman (2004))

Selection EAFs for analysis

As discussed in the literature review, a great number of EAFs are in existence which cannot be assessed each. Hence, there should be a rationale for the selection of appropriate EAFs that should be considered in the analysis. In this study, the most commonly used EAFs are selected. Table 9 presents the EAFs that are most commonly used by organizations to build their EA: Zachman Enterprise Architecture Framework (ZIFA), The Open Group Architecture Framework (TOGAF) and the Federal Enterprise Architecture Framework (FEAF) (Sessions, 2007; Urbaczewski & Mrdalj, 2006).

Table 9: Selected EAFs for analysis

ID	Enterprise Architecture Framework	Author / owner
1	Zachman Enterprise Architecture Framework (ZAF)	(Zachman, 1987)
2	The Open Group Architecture Framework (TOGAF)	The Open Group (Group, 2019)
3	Federal Enterprise Architecture Framework (FEAF);	The Federal Government of the United States

3.4.2. Step 1: Define Required Capabilities

Through a systematic review on the current theories, methods, expertise and experience, and meta-artefacts (which form the foundation) 11 capabilities are identified that should be incorporated in the design of the Extended EAF: Continuously aligning Business and IT; Providing a process for transition towards the 'to be' state of an organization; Removing the threshold for change; Enabling/Increasing Flexibility of the organization; Developing a proactive organization; Reducing risk and prepare for rapid, unplanned change; Avoiding tension between business and IT functions; Creating, Unifying and integrating business processes across the enterprise; Unifying information silos within the organizations; Eliminating duplicate technologies; and Reducing solution delivery time. Section 3.3 presents a description and sources of each of these capabilities (Table 8). The most commonly used EAFs are assessed on having these capabilities for the selection of the most suitable base EAF for extension.

3.4.3. Step 2: Assess most commonly used EAFs

In this step the three most commonly used EAFs are assessed on having the 11 identified capabilities. Each capability is given a score: Yes or No.

Zachman Framework

Table 10 presents the assessment of the Zachman Framework n the required capabilities. On the domain of taxonomy completeness, the Zachman-ADM is the most complete EAF. With a total of 36 viewpoints, each aspect is represented. However, the Zachman Framework does not provide guidance for transition, while that is the most important aspect on which the capabilities are based. Therefore, the Zachman does not have many of the capabilities that are required for the implementation of a strategy like the real-time integration of BOLD for decision-making. Nevertheless, the EAF could be useful as an addition to another framework or methodology that is more process focused but has less taxonomy completeness. Firstly, the Zachman Framework provides a proactive organization by describing all deliverables required to create a complete system. Secondly, the Zachman Framework helps to create, unify and integrate business processes across the enterprise. The business concept components helps organizations to identify the business processes required. Thirdly, the Zachman Framework helps to eliminate duplicate and overlapping technologies through the Technology Physics component.

Table 10: Assessment Zachman Framework

ID	Capability	Component	Score
C1	Continuously align Business / IT	N/A	No
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	N/A	No
C3	Remove the threshold for change	N/A	No
C4	Enable/Increase flexibility of the organization	N/A	No
C5	Develop a proactive organization	All viewpoints	Yes
C6	Reduce risk and prepare for rapid, unplanned change	N/A	No
C7	Avoid tension between business IT functions	N/A	No
C8	Create, unify and integrate business processes across enterprise	Business Concepts	Yes
C9	Unlock the power of information, unifying information silos.	N/A	No
C10	Eliminate duplicate and overlapping technologies	Technology Physics	Yes
C11	Reduce solution delivery time, development costs (maximize reuse)	N/A	No

TOGAF

Based on the information collected during the literature review about the TOGAF-ADM, an assessment is done on the capabilities the EAF has. Table 11 presents the assessment of the TOGAF-ADM. Each capability is covered by at least one of the components of the framework. To conclude, TOGAF has all capabilities of the current foundation of EA. Unfortunately, this literature review showed that the foundation of EA is limited on a couple of dimensions. The TOGAF does not have additional capabilities that to make the real-time data integration of external BOLD with internal company data for decision-making possible.

Table 11: Assessment TOGAF-ADM

ID	Capability	Component	Score
C1	Continuously align Business / IT	Architecture Vision Phase; Architecture Change Management.	Yes
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	Each phase TOGAF-ADM	Yes
C3	Remove the threshold for change	Architecture Change Management	Yes
C4	Enable/Increase flexibility of the organization	Application Architecture partly allows for it.	Yes
C5	Develop a proactive organization	Preliminary and Architecture vision	Yes
C6	Reduce risk and prepare for rapid, unplanned change	Architecture Change Management	Yes
C7	Avoid tension between business IT functions	Migration Planning; Implementation Governance	Yes
C8	Create, unify and integrate business processes across enterprise	Business Architecture; Migration Planning	Yes
C9	Unlock the power of information, unifying information silos	Information Architecture	Yes
C10	Eliminate duplicate and overlapping technologies	Technology Architecture	Yes
C11	Reduce solution delivery time, development costs (maximize reuse)	Opportunities and Solutions; Migration Planning	Yes

FEAF

Based on the information available about FEAF, an assessment is done on the capabilities the EAF has. Table 12 presents the assessment of the TOGAF-ADM. The FEAF scores relatively high on the different capabilities. However, it is too much focused on aligning within the own organizations and it also misses several capabilities: Remove the threshold for change; Enable/Increase flexibility of the organization; Reduce risk and prepare for rapid, unplanned change.

Table 12: Assessment FEAF

ID	Capability	Component	Score
C1	Continuously align Business / IT:	All steps	Yes
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	Step 3: Define & Plan	Yes
C3	Remove the threshold for change	N/A	No
C4	Enable/Increase flexibility of the organization	N/A	No
C5	Develop a proactive organization	Step 3: Define & Plan	Yes
C6	Reduce risk and prepare for rapid, unplanned change	N/A	No
C7	Avoid tension between business IT functions	Step 1: Identify and Validating	Yes
C8	Create, unify and integrate business processes across enterprise	Step 2: Research & Invest	Yes
C9	Unlock the power of information, unifying information silos across enterprise	Step 2: Research & Invest	Yes
C10	Eliminate duplicate and overlapping technologies	Step 5: Perform & Measure	Yes
C11	Reduce solution delivery time, development costs (maximize reuse)	Step 5: Perform & Measure	Yes

3.4.4. Step 3: Select a suitable Basis Enterprise Architecture Framework

Table 13 presents a comparison of the assessment of the different EAFs discussed in step 2. The outcome shows that the TOGAF-ADM is the most complete framework that can be used as a basis framework for extension. The Zachman Framework has its advantages and is a useful framework, but not for the purpose of this thesis. The TOGAF-ADM and FEAF are both suitable frameworks that can be used for extension, but FEAF is lacking some capabilities.

To avoid re-inventing the wheel, the basic framework will be adopted to incorporate the capabilities required to mitigate the adoption barriers of BOLD. However, this chapter also showed the limitations of the current foundation of EA and TOGAF. In the requirements definition chapter, it is important that the limitations of the TOGAF-ADM are removed by collecting requirements for this.

Table 13: Comparison of Assessment Traditional Enterprise Architecture Frameworks

	Ratings		
Criteria	Zachman	TOGAF	FEAF
C1	-	Yes	Yes
C2	-	Yes	Yes
C3	-	Yes	-
C4	-	Yes	-
C5	Yes	Yes	Yes
C6	-	Yes	-
C7	-	Yes	Yes
C8	Yes	Yes	Yes
C9	-	Yes	Yes
C10	Yes	Yes	Yes
C11	-	Yes	Yes

3.5. Summary Chapter 3 (Answer RQ2)

The second sub-research question '***RQ2. What are the capabilities from the current foundation of EA (Theories, Methods, Expertise, Experience and Meta-artefacts) that need to be incorporated in the Extended Enterprise Architecture Framework to mitigate the adoption barriers?'*** focused on two things: Identifying the capabilities of EA; and assessing whether these capabilities are suitable enough to mitigate the adoption barriers of BOLD.

Through a systematic review on the current theories, methods, expertise and experience, and meta-artefacts (which form the foundation) 11 capabilities that a successful EAF should have are identified: Continuously aligning Business and IT; Providing a process for transition towards the 'to be' state of an organization; Removing the threshold for change; Enabling/Increasing Flexibility of the organization; Developing a proactive organization; Reducing risk and prepare for rapid, unplanned change; Avoiding tension between business and IT functions; Creating, Unifying and integrating business processes across the enterprise; Unifying information silos within the organizations; Eliminating duplicate technologies; and Reducing solution delivery time. The rationale for this research question was to identify the knowledge base on EA, and to identify the limitations of the current foundation. These 11 capabilities should be at least incorporated by the design of the Extended Enterprise Architecture Framework to be able to mitigate the adoption barriers of BOLD. This will mitigate adoption barriers that cause misalignment across the different levels of an organization. The most commonly used EAFs are assessed on the possession of the capabilities. The TOGAF-ADM seemed the most suitable EAF for extension. The TOGAF-ADM can be used to aligning a new strategy across the enterprise.

However, the systematic literature review also identified limitations of the current foundation of EA. To enable the real-time integration of external BOLD with internal company data, no adoption barriers that cause misalignment with the BOLD providing organization can exist. Unfortunately, the foundation of EA has its limitations (Figure 23):

1. Breadth: limited to the entire company or parts of the company;
2. Depth: Enterprise level; Business domain level; Capability level; and smaller.
3. Time period: How much time organizations want to spend to implement.
4. Architecture domains: Limited to Business Architecture, Data Architecture, Application Architecture; and Technology Architecture.

To remove these limitations additional EA capabilities that are required. These capabilities must ensure an organization is continuously aligned with its ecosystem, that information silos between organizations can be unified and that an organization is flexible towards its ecosystem. Next to the 11 capabilities, an extension is required that has the capabilities to remove the limitations and to enable the real-time integration of BOLD with internal company data for decision-making.

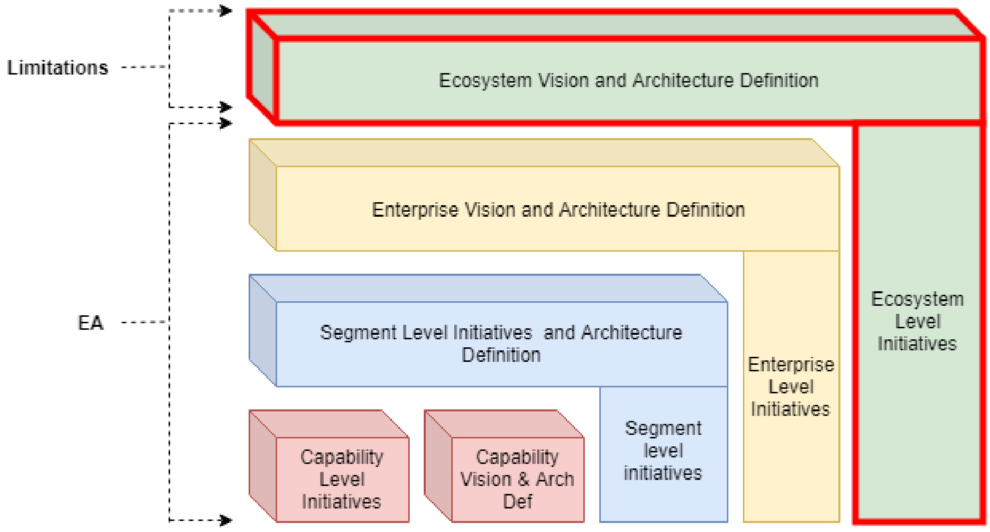


Figure 23: Limitations Integration levels

Key findings:

I. Countless Enterprise Architecture Frameworks are in existence.

II. Most commonly used Enterprise Architecture Frameworks are (1) The Zachman Framework; (2) TOGAF-ADM; and (3) Federal Enterprise Architecture Framework.

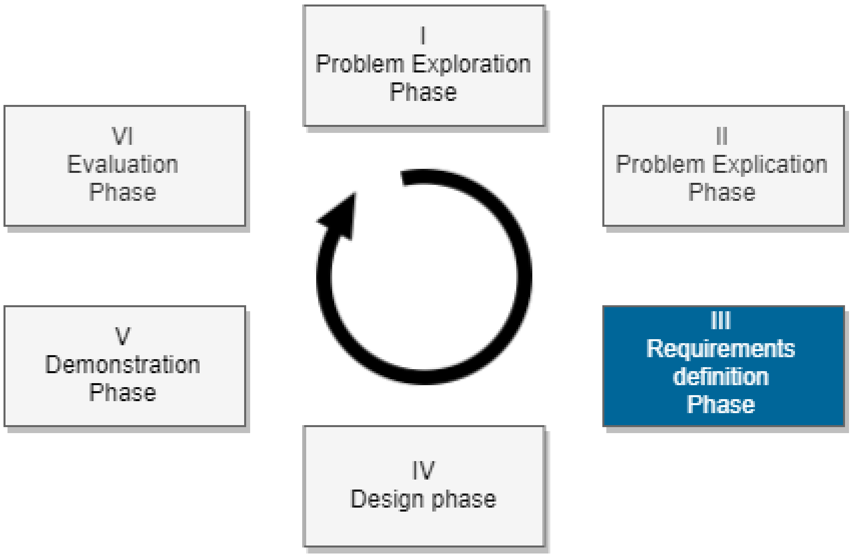
III. The current foundation of Enterprise Architecture is capable of: Continuously aligning Business and IT; Providing guidance for a process for transition towards the 'to be' state of an organization; Lower the threshold for change; Enable/Increase Flexibility of the organization; Develop a proactive organization; Reduce risk and prepare for rapid, unplanned change; Avoid tension between business and IT functions; Create, Unify and integrate business processes across the enterprise; Unify information silos within the organizations; Eliminate duplicate technologies; and Reduce solution delivery time.

IV. Enterprise Architecture is not suitable for the real-time integration of external data from data providers and needs to be extended.

V. TOGAF-ADM has most capabilities and is the most suitable Enterprise Architecture Framework for extension.

VI. By building on TOGAF-ADM, the capabilities required to mitigate the adoption barriers of BOLD across the organization are incorporated by the design of the Extended EAF.

III. Requirements Definition Phase



4. Requirements Definition

In this chapter, the requirements for the design of an Extended EAF that enables the real-time integration of external BOLD with internal company data are discussed. This chapter addresses *RQ3. 'What are the requirements for an Extended Enterprise Architecture Framework that enables the real-time integration of external Big and Open Linked with internal company data for decision-making?'*. The requirements are collected through both the literature reviews and the explorative expert interviews. Qualitative Data Analysis is performed to identify the key findings and to extract the requirements for the extended Enterprise Architecture Framework. Section 4.1. presents the Requirements Definition Overview. Section 4.2. presents field research findings on adoption barriers of BOLD. Section 4.3. presents the field research findings on the required capabilities. Section 4.4. presents a summary and answer of the RQ3.

4.1. Requirements Definition Overview

Figure 24 presents an overview of the Requirements Definition Phase. It presents how the literature Reviews are translated into requirements. This is done via three incremental steps. First, the literature reviews are used to set up the explorative interviews. Secondly, the explorative interview outcomes are used as input for the Qualitative Data Analysis, which is used for identifying the key findings. Thirdly, the key findings are translated into requirements for a suitable extension to the TOGAF-ADM.

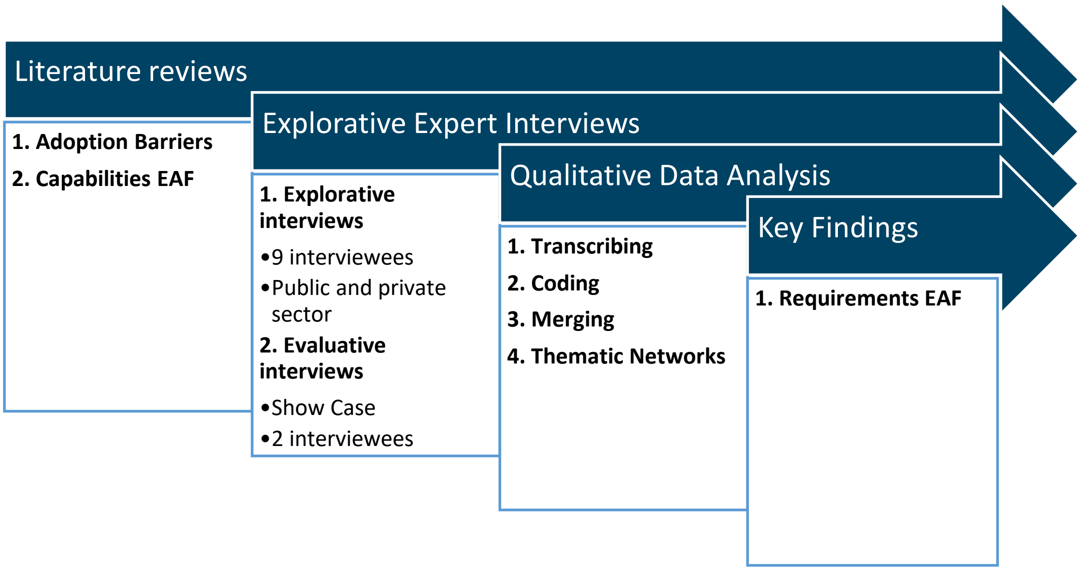


Figure 24: Overview Requirements Definition

The interview protocol, which can be found in Appendix B1: Explorative Expert Interview Protocol, is based on the literature reviews and is structured into four parts. In the first part, the potential of external BOLD for data-driven decision-making is discussed. Moreover, questions are asked about how the decision-making process based on external BOLD does look like. This has resulted in interesting examples of how companies use external BOLD to strengthen their internal company data and to base their decisions on. In the second part of the interview, the adoption barriers encountered by organizations during the (real-time) adoption of BOLD are discussed. The key findings of this part led to the revision and extension of the conceptual model of adoption barriers of external BOLD. In the third part, questions are asked about what organizations do in practice to overcome the identified adoption barriers. Finally, the ideal capabilities that an EAF should have to enable the real-time adoption of external BOLD for decision-making.

4.1.1. Semi-structured interviews

The explorative expert interviews are conducted in a semi-structured way, following the guidelines by Kallio, Pietilä, Johnson, and Kangasniemi (2016). The authors created a framework of the phases that are required to develop a semi-structured interview that helps to collect all the relevant information required needed to solve the research problem. The framework that guides development of a semi-structured interview includes five phases: '1) identifying the prerequisites for using semi-structured interviews; 2) retrieving and using previous knowledge; 3) formulating the preliminary semi-structured interview guide; 4) pilot testing the interview guide; and 5;) presenting the complete semi-structured interview guide' (Kallio et al., 2016). Following this framework resulted in the final version of the interview protocol, which can be found in Appendix B1: Explorative Expert Interview Protocol. In addition, the method proposed by Barriball and While (1994) was used to design the interview questions. The author discusses the measures that can help to deal with validity and reliability threads of semi-structured interviews, such as making sure an answer is collected for each question by each respondent; exploring attitudes, values and beliefs; evaluating the validity of the answer of the interviewees.

The findings of an article by Leech (2002) are used to determine the techniques of conducting the actual semi-structured interviews. Even when the protocol is very clear, each interview can proceed in a different way. Leech (2002) explains that using a combination of grand tour questions and floating prompts helps the interviewee to collect extra (or even all) relevant information and to push the conversation in the right way. Grand Tour questions are used to let the interviewees give a verbal overview of the subjects they know very well. This results in a good understanding of the most important subjects and gives context. Also, it creates understanding on what topics should be focused in the interview. The author came up with this technique, because interviewees often have useful answers. Even when the interview protocol does not work out as planned, the interviewees can give the answer to those questions. Prompts are used to discuss the things that are important to discuss, so they help to steer the conversation.

4.1.2. Qualitative Data Analysis: Transcribing, Coding and Thematic Networks

Qualitative Data Analysis is conducted to derive requirements for a suitable extension of the TOGAF-ADM. A General Inductive Approach is followed to get to the key findings of explorative expert interviews. The analytic strategies and questions of a General Inductive Approach is concerned with finding the core meanings evident in the text that are relevant to the research objectives of the thesis. The outcome of the analysis are the categories that are most relevant to the research objectives of the thesis. This will be presented by a description of the most important themes, which will be translated into requirements for a suitable extended EAF (Mihas, 2019; Thomas, 2006). To derive the most important themes and categories of the data, the interviews are transcribed, and a specific Coding Process for Inductive Analysis was used.

4.1.2.1. Transcribing

From each interview, an audio recording is made that is used to transcribe the interviews. The transcriptions are done in an accurate way. After the transcription is done, the transcription is sent back to the interviewees for validation. After validation, the transcriptions are used for the coding phase of the qualitative data analysis. The transcriptions are not put in the appendices due to confidentiality reasons but can be send upon request. The transcriptions are done in such way, the name and organization are anonymized. However, professional jargon can give an indication of what organization is interviewed. In total, 9 stakeholders in different roles are interviewed to ensure the collection of requirements for a complete, exhaustive design artefact (Johannesson & Perjons, 2014) (Table 14). The interviews are transcribed with Qualitative Data Analysis Software (Atlas.TI) to perform the General Inductive Analysis in a systematic way.

Table 14: List of Interviewees Explorative Expert Interviews

ID	Stakeholder Role	Description Tasks	Organization	# Years Experience
S1	BOLD Expert; Data analyst; Data User.	Conducting research about BOLD; Using External BOLD for Predictive Analysis used for decisions about setting up the infrastructure of one of world's biggest cities.	Technical University; One of World's Largest IT firms	7
S2	Enterprise Architect	Strategy making on the domain Information Provision; Deciding on what Information sources are used for data-driven decision-making, such as external BOLD and Open Data. Deciding on what Open Data is published and used. Focusing on the architectural decisions.	National Public Enforcement organization	30+
S3	Policy-maker	Strategy making on the domain Information Provision; Deciding on what Information sources are used for data-driven decision-making, such as External BOLD and Open Data. Deciding on what Open Data is published and used. Focusing on the policy-making.	National Public Enforcement organization	30+
S4	Enterprise Architect; Data Scientist; Open Data Expert	Mainly working on Open Data projects for International Governments. Using and Publishing Open Data; Using Open Data for Decision-Making.	One of world's biggest Consulting Firms	15+
S5	Decision-Maker	Decision-making and Policy making about Business Intelligence and Information quality. Deciding about what external open data sources are used for data-driven decision-making.	Regional Public Enforcement organization	20+
S6	CIO; Executive Partner	Advising CIOs with developing and implementing new IT strategies; Founder IoT-scale up that is disrupting the Dutch energy sector. Creating Open Data infrastructures. Using BOLD for data-driven Decision-making. Publishing external BOLD.	BOLD providing company; One of the biggest Dutch Energy Providing firms.	25+
S7	Data Scientist; Decision-maker.	Manager Data Scientist team; Developing Predictive models that can be used for decision-making by Insurance companies to determine damage assessments. Using external BOLD for this predictive models.	One of the largest Damage Assessment Organizations in Insurance Sector.	15+
S8	Decision-maker. Strategist.	Head decision-support. Making sure decision-quality of decisions is guaranteed. Leading decisions support about decisions were more than 1000 employees involved. Part of decisions are based on external BOLD.	World's biggest supplier for the semi-conductor industry.	20+
S9	Data Scientist	Creating Predictive models based on a variety of sources, such as external BOLD. These models are used for data-driven decision-making within the organization.	Regional Public Enforcement organization	20+

4.1.2.2. Coding

The transcriptions are coded in order to find the most important categories. To find these categories, the coding process by Thomas (2006) is used. Figure 25 displays the different steps that are taken. The first step covers the initial reading of the text to create an overall understanding of the content of the interview. In the next step, specific text segments related to the research objectives are identified and coded through open coding. After performing the coding for all the transcriptions, the codes are labeled through axial coding that documents the relationship among the different codes. This helps to determine what the connection is between the codes. Subsequently, a core variable is identified that can link segments into categories. In the next step, the transcripts are reread, and selective coding is done that helps to select any data that is related to the core variable/category that is identified. Finally, the overlap between the identified categories is reduced and redundant categories are deleted in order to create a thematic network that incorporates the most important categories (Thomas, 2006).

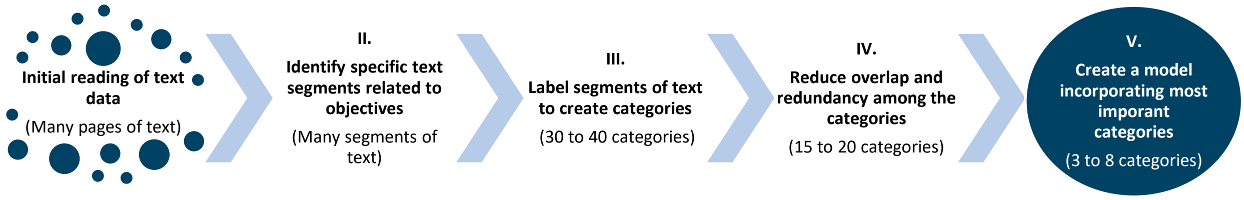


Figure 25: Coding Process Inductive (Thomas, 2006)

Following the coding process defined by Thomas (2006), the initial text is limited to a number of important categories. In Table 15, the quantity of variables per phase is displayed.

Table 15: Coding Phases: Quantity variables

Phase	*Initial reading*	*Text segments*	*Categories*	*Merged Categories I.*	*Merged Categories II.*
Adoption Barriers	17284 words	285	66	21	6
Ideal Capabilities EAF	17284 words	104	42	9	-

The transcriptions are coded into two code groups: Adoption Barriers; and Required Capabilities EAF (Table 16). The knowledge base that is created by means of the literature reviews is represented by these two elements. The insights from the environment are linked to these two elements, as these findings can form an addition to the knowledge base of those elements (Hevner, 2007). The insights are translated into requirements. After finalizing the coding process, thematic networks are created. This is explained in the following section.

Table 16: Overview of Code Groups

Element	Definition	Code Group
Adoption Barriers	The factors that hinder the successful adoption of BOLD into the decision-making process.	Adoption Barriers
Required Capabilites EAF	The capabilities that an EAF should have to overcome the adoption barriers of BOLD. Discusses the ideal situation according to the interviewees.	Capabilities EAF

4.1.2.3. Thematic Networks

Using the outcomes of the coding process, thematic networks are created. This represents the last step of the coding process by Thomas (2006). A thematic network is a qualitative data analysis technique for thematic analysis that results in a model incorporating the most important themes of each element of the literature review. In addition to the General Inductive Approach by Thomas (2006), a method to create thematic networks by Attride-Stirling (2001) is used. This method explains how to transform qualitative data into the key findings in a systematic way. The steps that are taken based on this method are: the merging of codes into more general variable; and the mapping of the findings into a thematic network of the most important variables.

In the thematic network, only the concepts that are mentioned by the interviewees are included. So, without the findings of the literature review. It is used to build on the knowledge base that is created through the literature reviews. In total two thematic networks are created: the Adoption Barrier Network and the Ideal Capabilities EAF Network. Both thematic networks are generated using the same coding process.

4.2. Adoption Barriers BOLD for Data-Driven Decision-Making

During the desk study, different adoption barriers are identified and a conceptual model of the most common adoption barriers of BOLD is created. First, a brief overview of the most important insights of the literature review are given, where after the insights of the interviews are discussed. Figure 27 presents the thematic network on adoption barriers that is generated from the qualitative data analysis performed on the transcriptions of the expert interviews.

4.2.1. Literature review insights

- Enriching internal data with external BOLD offers a lot of potential for real-time decision-making, but it addresses many challenges and adoption barriers.
- Organizations are encountering a variety of adoption barriers during the different phases of the BOLD life cycle that makes the real-time adoption difficult.
- The adoption barriers identified can be categorized into 5 overall categories, namely Information Quality Barriers; Task Complexity Barriers; Technical Barriers; Use and Participation Barriers; and Legislation Barriers.
- The Information Quality Barriers can be divided into three sub-clusters: Data absence; Data quality problems and Lack of meta-data.
- The Task Complexity Barriers can be categorized into three clusters: Data Handling Issues; Lack of Standards and Different Sources.
- The Technical Barriers can be categorized into three different clusters: Lack of supporting infrastructure; Difficulties with Legacy Systems and Fragmentation of the data.
- Use and Participation Barriers can be clustered into the Lack of knowledge of Users, Implementation Issues, the existence of Constraints (Money, time, etc.), and the lack the right capabilities to adopt external BOLD in real-time.

4.2.2. Explorative expert interview insights

During the Explorative Expert Interviews with interviewees in different roles from different public and private organizations, the interviewees are asked what adoption barriers they address in their organization when integrating BOLD for decision-making in real-time. In total, the interviewees addressed 63 adoption barriers. The adoption barriers mentioned have a lot of overlap with the conceptual model of adoption barriers of BOLD. This section discusses the key insights that are extracted from the interviews:

1. Information Quality Barriers

The interviewees mention different adoption barriers that are related to the Information Quality Barriers. The barriers that are mentioned can be categorized into the three clusters of the conceptual model of adoption barriers: Data Absence; Lack of Quality; and Lack of meta-data. More specific, most of the interviewees mention that often sources are missing data and that this data absence influences the information quality and therefore will have an impact on the decisions to be made. Secondly, sources do often lack meta-data: which can be information about the source, but also technical documentation of how APIs can be best accessed and implemented. One of the main reasons for this is that there are no meta-data standards that all BOLD providers use. Thirdly, a lack of data quality causes a big problem for organizations with the integration of BOLD. This can be caused by a variety of reasons, such as misinterpretation of the data.

2. Task Complexity Barriers

Also, adoption barriers are mentioned that are related to the Task Complexity Barriers. The different barriers mentioned can be divided into three different clusters: Lack of Standards; Data Handling Issues; and Variety in Sources, Types and Size. The barriers that are related to lack of standards are: Different Sources; and Different types of data. The mentioned adoption barriers related to Data Handling Issue barriers include Lack of semantics, Encrypted data sources, Real-time handling of data, Real-time communication of results, linking and contextualization in real-time and data visualization in real-time.

3. Technical Barriers

The interviewees also addressed technical barriers that consist of the issues with legacy systems that make the integration difficult. Also, a supporting infrastructure is often lacking, including the lack of tooling, technologies and a suitable architecture to make the integration of external data possible. Multiple interviewees mentioned that fragmentation of the data cause problems, because the data is separated in different systems across the enterprise. Lastly, problems with flexibility and interoperability and compatibility of the systems makes the real-time adoption of BOLD difficult.

4. Legislation Barriers

Fourth, Legal barriers are addressed by the interviewees. The interviewees did point out that permits and licenses, privacy issues, security form barriers for the adoption of BOLD for decision-making. One new legal barrier that is mentioned by different interviewees: often it would help if there is no legally accepted standard that organizations can use as a critical measure to compare their decisions with.

5. Use and Participation Barriers

Furthermore, the interviewed organizations address different Use & Participation Barriers, including constraints related to money and time, collaboration issues with partners and external other parties that provide BOLD, lack of acceptance of the new way of decision-making based on BOLD instead of the traditional way of working, and the lack of competences, such as knowledge and capabilities to make the real-time integration possible.

6. Governance Barriers

Finally, a new Adoption Barrier Category is identified: Governance Barriers. All interviewees are addressing Governance Barriers related to the adoption of BOLD. The governance barriers can be mapped into three clusters. Policy Issues; Scalability Problems; and a Lack of Data Governance. With governance is meant 'the way how organizations are managed and the systems for doing this'. In other words, the way how the data integration and exchange with external BOLD Providers is managed within and between organizations. Firstly, data governance is addressed as a barrier that make the real-time adoption. This entails the planning, oversight, and control over management of data and the use of data and data-related resources. Often there is no clear data-integration strategy, master data management and data governance that manage the quality, continuity, the integration of new and existing data sources. This reduces the quality of decision-making. One of the reasons for this is that companies do not have a moderator that is responsible for the data governance. Also, organizations address a lack of a good policy, implementation issues as reasons for problems with the real-time adoption of BOLD. Finally, the architecture of a company has often problems with the integration of BOLD by new data providers and with the exit of BOLD Providers. A governance mechanism that deals for this issues is often lacking.

4.2.3. Conclusion

Figure 26 presents the revised conceptual model of adoption barriers. These are the barriers the extended Enterprise Architecture Framework should deal with to enable the real-time integration of external BOLD with internal company data. Based on the key findings of the interviews, an additional category and some additional sub-clusters are added to the conceptual model (depicted in red).

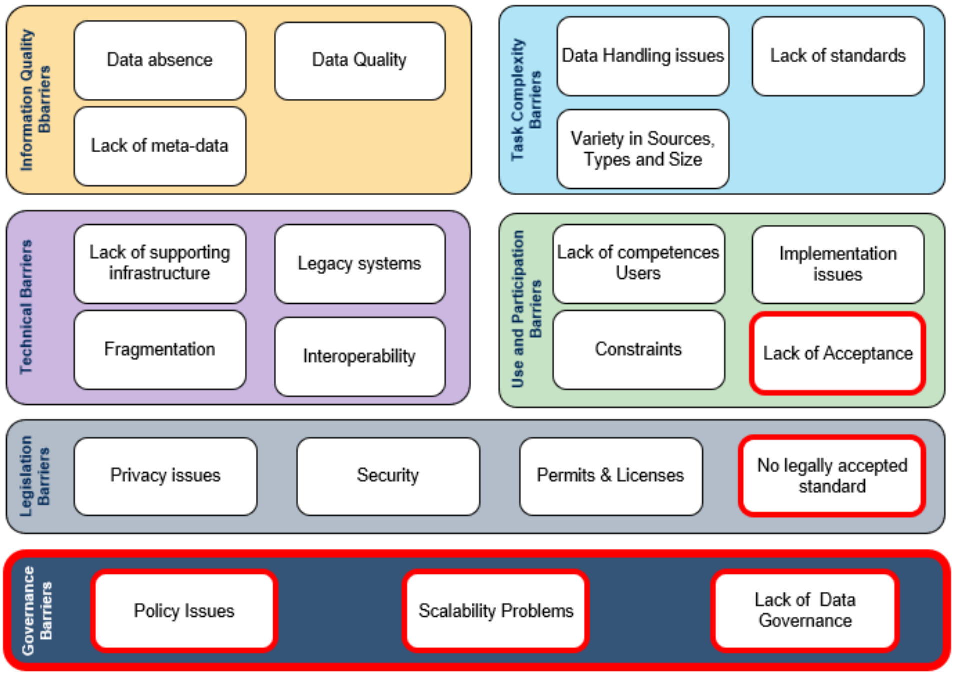


Figure 26: Revised Conceptual Model of Adoption Barriers

The qualitative data analysis resulted in the following key findings:

I. *Few organizations succeed in real-time integration of external BOLD with internal company data for decision-making. Most of the organizations are struggling with the adoption barriers that make the real-time use for decision-making difficult.*

II. *Organizations from both the Public Sector and the Private Sector are addressing adoption barriers with the integration of External Big and Open Linked Data for decision-making. Public organizations are especially struggling with the Legal Barriers, Use and Participation Barriers and the Information Quality Barriers, while Private Organizations are especially struggling with the Task Complexity Barriers, Information Quality Barriers and the Technical Barriers.*

III. *Organizations identify an additional adoption barrier category: 'Governance Barriers', consisting of the sub-clusters 'Policy Issues', 'Lack of Acceptance', and 'Scalability Problems'.*

IV. *Most of the adoption barriers are caused by the human factor.*

V. *Data integration from external to internal organizations is one of the biggest struggles of companies. The most successful organizations with BOLD integration are the companies with a clear integration strategy and are capable of sharing data with other organizations.*

The key findings and the revised conceptual model of adoption barriers of BOLD are translated into a requirement for a design artefact that must enable the real-time adoption of BOLD for decision-making.

Requirement 1: *The Enterprise Architecture Framework must* ***proactively deal*** *with the* ***Information Quality Barriers, Task Complexity Barriers, Technical Barriers, Use & Participation Barriers, Legislation Barriers, and Governance Barriers*** *that organizations address with the real-time integration of BOLD for decision-making.*

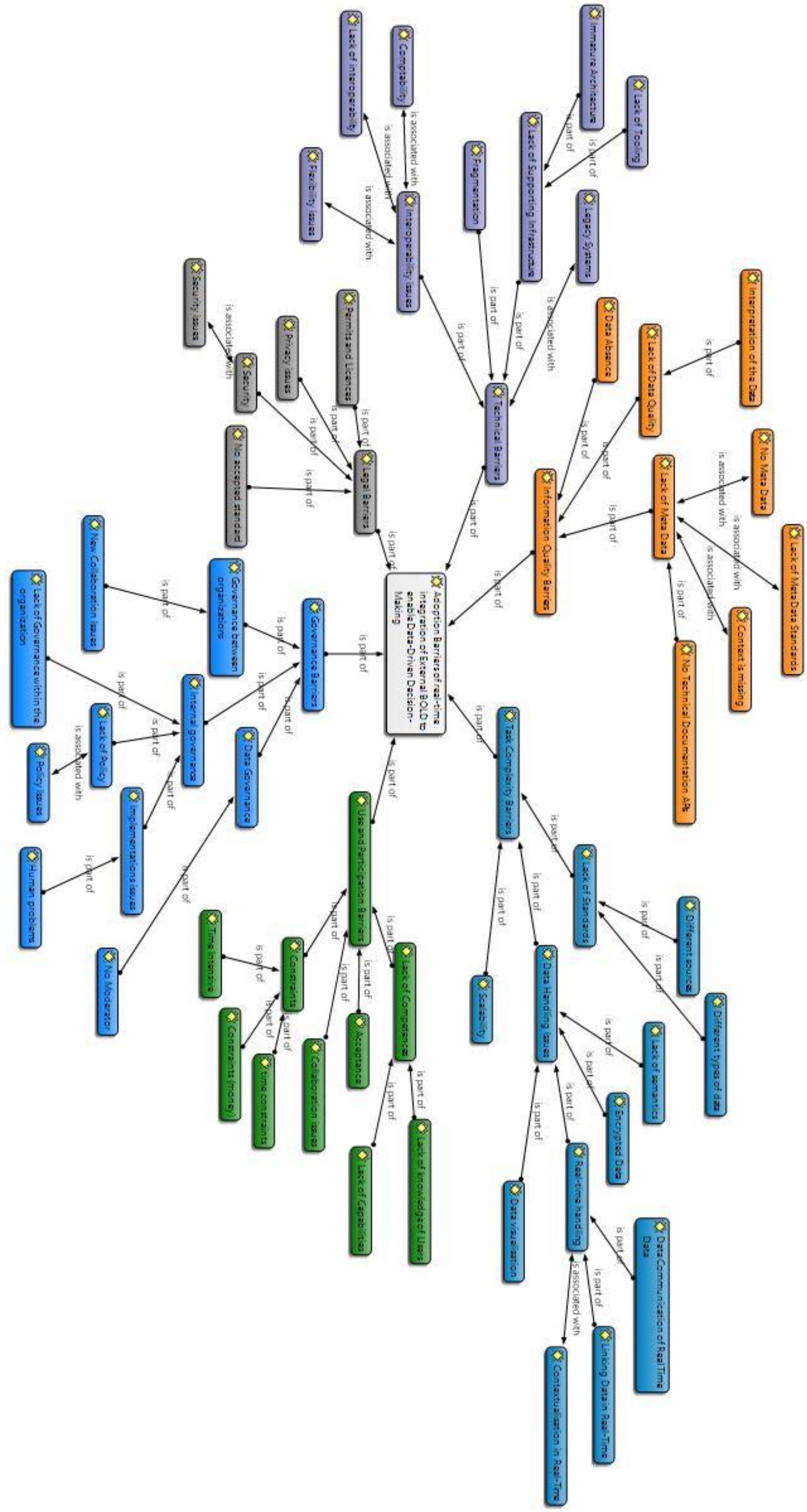


Figure 27: Thematic Network Adoption Barriers

4.3. Ideal Capabilities Enterprise Architecture Framework for enabling the real-time integration of external BOLD with Internal Company Data

4.3.1. Literature review insights

- Countless Enterprise Architecture Frameworks are in existence.
- Most commonly used Enterprise Architecture Frameworks are (1) The Zachman Framework; (2) TOGAF-ADM; and (3) Federal Enterprise Architecture Framework.
- The current foundation of Enterprise Architecture is capable of: Continuously aligning Business and IT; Providing guidance for a process for transition towards the 'to be' state of an organization; Lower the threshold for change; Enable/Increase Flexibility of the organization; Develop a proactive organization; Reduce risk and prepare for rapid, unplanned change; Avoid tension between business and IT functions; Create, Unify and integrate business processes across the enterprise; Unify information silos within the organizations; Eliminate duplicate technologies; and Reduce solution delivery time.
- Enterprise Architecture is not suitable for the real-time integration of external data from data providers and needs to be extended.
- TOGAF-ADM has the most capabilities and is the most suitable Enterprise Architecture Framework for extension.

4.3.2. Explorative expert interview insights

As the foundation of EA does not have the right capabilities to enable the real-time integration of external BOLD with internal company data for decision-making, additional required capabilities are collected during the explorative expert interviews. The ideal situation to prevent the adoption barriers from occurring is discussed. Through qualitative data analysis, different insights are extracted from the interviews. Figure 28 presents the thematic network that was developed for deriving the requirements. The key findings resulted in six different requirements. These are discussed in this section.

1. Proactive Approach

After discussing the adoption barriers with interviewees from different organizations, the ways to overcome the adoption barriers have been discussed. This has resulted in a discussion about the approach that should be followed to enable the real-time adoption of BOLD for decision-making. Apparently, a Reactive Approach and a Proactive Approach are possible to overcome the Adoption Barriers. Once a new external source is adopted, different adoption barriers can be encountered.

The Reactive approach focuses on overcoming the adoption barriers once they occur. For example, by investing in new tools and technologies that facilitate different processes among the data pipeline, such as filtration, transformation, cleaning and quality checks. The Proactive approach focuses on the development of a new approach to develop an architecture that is more suitable to integrate external sources. Moreover, a method that prevents the Adoption Barriers from occurring. However, most interviewees state that a Reactive Approach is adopted within their organization and that a reactive approach solves for maximum 80 percent of the adoption barriers. Due to all the extra steps required the real-time integration of BOLD for decision-making cannot be ensured. It only makes the integration more acceptable.

This resulted in the following key finding:

I. *A Proactive Approach should be taken to prevent the adoption barriers from occurring and is required to enable the real-time integration of external BOLD for decision-making.*

The first requirement is both based on the key findings of section 4.2. on the adoption barriers of BOLD, and on the key finding on a proactive approach:

Requirement 1: *The Enterprise Architecture Framework must **proactively deal** with the **Information Quality Barriers, Task Complexity Barriers, Technical Barriers, Use & Participation Barriers, Legislation Barriers, and Governance Barriers** that organizations address with the real-time integration of BOLD for decision-making.*

---

2. Collaboration between organizations

Adoption Barriers often occur because the data user and the data provider are not well-aligned with each other. One of the approaches to overcome an adoption barrier that is often mentioned by the interviewees is collaboration with the data-provider. When the data provider and data user collaborate closely with each other, the parties understand each other better, which prevents adoption barriers from occurring. This requires thorough cooperation of the external BOLD providers. Without cooperation of the data provider, it is difficult to understand how the adoption barrier can be removed.

Nonetheless, a data provider should have incentives to cooperate. Normally, cooperation goes in hand with reciprocity. In other words, a data provider is willing to adapt to the demand of the data user if it gets something for that in return. This can be different things, such as a financial compensation or data owned by the user. Without reciprocity, there are little incentives for data providers to adapt to the demand of the data users. Interviewee S5 demonstrates the importance of reciprocity:

"*Our organization would like to work more with external BOLD. The reason we do not work a lot with BOLD is because we encounter problems with the adoption. We would like to cooperate more with external parties to enable data-driven decision-making based on BOLD. But what you often see is that the providing parties would like to have something in exchange, so called: reciprocity"*
– Interviewee S5

Having mutual incentives often results in partnerships between organizations. A partnership is a long-term relation between organizations, in which both organizations offer value to each other. This results in continuous mutual benefits. In the case of the interviewees, collaboration between organizations often includes the sharing of data. The value of an organizations' own data increases, when it is combined with the data of the external organization. To utilize this value, companies are agreeing more often on partnerships to share data with each other. Hence, both parties then have incentives to create a common way of working to make it easier to integrate each other's data. To enable data analysis with real-time external data, the data integration must go smoothly in order to be able to make good decisions. Interviewee S6 explains the importance of data-sharing:

"(Big) data analysis is not solely done anymore within the internal organization (in contrast with five years ago). Organizations are opening up. Data from outside the internal organization is more valuable than data from the inside the internal organization. More organizations are creating assets that can be matched with the demand of other organizations. Organizations should share data assets with each other to create more value. Organizations would like to collaborate, but currently, do not yet share data with each other."
– Interviewee S6

Both organizations of the public and the private sector state that collaboration with stakeholders create trust, clarity, and takes away adoption barriers, such as information quality barriers. Public organizations are mainly struggling with information quality and legal barriers. The interviewees suggest that legal institutions should be involved to better deal with legal adoption barriers. Also, organizations should aim to commit to long-term relationships.

In the interviews, plenty of other examples are mentioned that confirm the same key findings:

I. *To enable real-time data-driven decision-making and prevent adoption barriers from occurring, an organization should collaborate more closely with its BOLD ecosystem.*
II. *There must be alignment between the business, information, application and technology layers of the internal organization; 2) and between the internal and external organizations that are part of the ecosystem.*

These key findings are translated into the following requirement:

Requirement 2: *The Enterprise Architecture Framework must* ***stimulate collaboration*** *between organizations, including partnerships, strategic coalitions, shared business models, mutual incentives and benefits and a clear integration strategy.*

3. Involvement Stakeholders Ecosystem

Another key finding related to collaboration between organizations, is the need for involvement of stakeholders of the business ecosystem for the development of the organization. Normally, when developing an Enterprise Architecture, all the stakeholders involved in the different views (Business, Data, Application and Technology architecture) have concerns about how the system should look like. These concerns are translated into viewpoints that represent the requirements for a view (architecture) (Hilliard, 2000). This ensures that the company is continuously aligned across the enterprise.

Yet, as continuous and real-time external data is becoming more important for organizations, it is important that the stakeholders concerned with the data-sharing process between organizations involved during the design of the Enterprise Architecture. This are the stakeholders that are part of the ecosystem of the organization, which have concerns that form a new viewpoint together. Taking this viewpoint into account will ensure better alignment between organizations. According to the interviewees, the involvement of stakeholders that are part of the ecosystem of an organization will result in better alignment and, therefore, less adoption barriers. Stakeholders that are part of the BOLD ecosystem of an organization are among others an ecosystem orchestrator, service providers, application providers, data producer, data publisher, data user, and data prosumer. Interviewee S4 explains the need for involvement of stakeholders during the design of a system:

"The data users should engage with the organizations providing the open data in order to understand how they can consume them. To give you a very practical example: I have been downloading XML from the tenders European daily (Database EU commission) where all the major procurements in Europe are entered. That is provided in XML, but it is only provided in bulk. So, I must download all bulks of procurement, while I'm only interested in IT procurement. This produces a practical problem and it takes a huge computing power to process the data, while it would have taken less if only the IT procurement could have been downloaded. Point is: You must think about the use cases that you are making the data available for. This is something that should be developed further."

– Interviewee S4

Another reason to involve stakeholders from the ecosystem is to ensure compliance. Different interviewees told that involving stakeholders concerned with privacy, transparency and secure integration of BOLD should be involved in order ensure the development of an architecture that overcomes the legal barriers. The interviewees gave diverse examples of situations in which involvement of parties from the ecosystem took away legal barriers for the adoption of BOLD. For example, Interviewee S2 and S3 explain that making agreements in advance with organizations, municipalities, legal and social authorities about a specific purpose can remove the adoption barriers, because it can result in a consent to develop a system that is used for data-driven decision-making. As depicted in this quote:

"In the public sector, the privacy and security barriers can be overcome if agreements are made in advance with data providers, governments, legal and social authorities if the purpose for using the data outweighs the legal barriers. For example, the tax authority has the right to adopt sensitive data in order to be able to link extra information that helps to identify tax evasion"

– Interviewee S2 & S3

This resulted in the following key findings:

I. *The Enterprise Architecture Framework must involve all stakeholders that are part BOLD Ecosystem, including Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, legal Institutions and Data Prosumer. This are the stakeholders that are concerned with the real-time integration of external BOLD for decision-making into account during the design of the Architecture to create better alignment of the Enterprise with its ecosystem.*

II. *Current Enterprise Architecture Frameworks are focusing too much on aligning strategy and technology of the internal organization and not on the alignment with the external organization.*

These key findings are translated into the following requirement:

***Requirement 3**: The Enterprise Architecture Framework must **involve all stakeholders from the BOLD ecosystem that are concerned** with the real-time integration external BOLD for decision-making **in the design of the Extended Enterprise Architecture to create better alignment of the enterprise with its ecosystem**, including Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, Data Prosumer and Legal Authorities.*

4. Direct Connection between Organizations

An interesting finding retrieved from the interviews is that organizations still think about integrating external BOLD traditionally. Many organizations are still focusing on integrating batches of data manually, while that is not suitable for data-driven decision-making in real-time. However, to integrate external real-time data streams into the decision-making process organizations need to start think differently. Organizations should make use of shared platforms in which the architectures of organizations are initially connected to each other. In this situation, organizations do not have to store the data locally anymore but can call the external data directly.

Interviewees S6 substantiates this:

"Organizations are still thinking traditionally without the use of shared platforms, while connecting with other organizations to share BOLD can create enormous new insights and value through combining the data silos and use this for methods, such as Artificial Intelligence. Continuous and Real-time data streams should be initially be integrated in this shared platforms, so that organizations can use this data in real-time."
– Interviewee S6

Interviewee S7 also explains that a change in the way how organizations think is required to enable a real-time connection between organizations:

*"Within ecosystems with partners, it could be useful to create more standardizing.
Yet, we are still too much focused on integrating the external data into the internal systems to ensure we have the data stored and it can be used for analyses. In the future, it is likely the data does not have to be stored locally anymore but can be called upon request in real-time via a webservice of an organization. This would make the life of many people easier. So that organizations are more focused on data-sharing and complying with standards to be able to request the data. On this domain, there is a lot to win."*
– **Interviewee S7**

Such a webservice would be available through APIs. An API is the abbreviation for Application Program Interface. An API is a set of routines, protocols, and tools for building software applications and allow programmers to access into another organization's programs or services easier. If each committed organization provides APIs for partners or even the public that use the same standards than the integration of external data would be way easier and possible in real-time. Different interviewees mentioned that APIs are the solution to the integration problem. Nonetheless, it is important that organizations than use the same standards, such as providing technical documentation about the API, to create compatibility between organizations. Interviewee S4 explains that APIs form opportunities and parts of a solution and to deal with the data integration problem:

"Lack of a supporting infrastructure is often a big problem. And organizations are doing a lot to improve this. Data integration is not a new thing, but the complexity is increasing as you increase the number of data sources and the different technical interfaces you are handling. Now you are implementing new technologies that can handle this technology like API brokers."
– Interviewee S4

Different interviewees think organizations should grow such a mindset to be able to overcome the adoption barriers of BOLD. Initially integrating BOLD data streams into shared platforms ensures ready-to-use data, which is required to enable real-time data-driven decision-making. Interviewee S1 explains the ideal situation in which this can be achieved:

In the ideal situation there should be one place where everyone can provide and retrieve data (e.g. open data portal). Some companies are already doing this, with the purpose of providing new solutions by means of a place where organizations publish this closed data that can be used by other organizations. (sustainability of data). If everyone provides open data, everything could be improved. Companies providing data to each other through API's, these can only be accessed because the companies are using open data standards + technical documentation. Otherwise other companies cannot access. Next to API's data can also be provided through other formats, such as XML, or through dashboards.

– Interviewee S1

Thus, an API is a mean for easier access to the services and applications of external parties. But there should also be an architecture that facilitates integrates and facilitates the use of APIs. Interviewee S6 explains that organizations will not make it in the future without adapting their end-state architecture focused on external organizations:

"How to involve in something we can use, and we can invest in something which works. Our organization now has an end-state architecture focused on the own enterprise. But we will never survive if we do not start thinking platform driven. Everything must be open in the future. Direct network effects. To enable this, we have to start thinking platform driven, monetizing APIs, data sharing mechanisms, suitable governance structures, new actors and technical functions are required"

– Interviewee S6

Yet, Enterprise Architecture Frameworks do not consider the architecture of the external data provider and the linkages between the Enterprise Architecture and the Architecture of the external parties. To facilitate this, it is important to create a layer that supports the collaboration of architectures between organizations. Enterprise Architecture focuses mainly on developing the business architecture, data architecture, application architecture and technology architecture of an enterprise. It also ensures that the linkages between the different architectures (business and IT) are aligned with each other to avoid tension between business functions, unify and integrate business processes, unifying information silos, and eliminate duplicate technologies.

But it does not align with the external organizations that are important to enable the real-time adoption of BOLD. There should be a shared platform that facilitates an environment in which data can be integrated and shared; different kinds of applications can be integrated and linked together (independent from platform, programming language or resource); collaboration between distributed and scattered applications; interoperability between different operating systems; security components that makes sure data is only shared with the right resources; and visual guidance by a single interactive user interface.

This resulted in the following key findings:

I. *Creating a shared platform that integrates APIs and can be accessed through APIs will enable real-time data-driven decision-making.*

II. *In the ideal situation, the internal organization is connected to the external organization to make the data integration and exchange easier to enable the creation of new insights and value and overcomes the adoption barriers of Big and Open Linked Data.*

III. *Continuous and Real-time data streams should be initially integrated into shared platforms to enable real-time decision-making.*

IV. *The Enterprise Architecture Framework helps to create an environment in which data can be integrated and shared; integrate different kinds of applications (independent from platform, programming language or resource) which can be linked together; collaboration between distributed and scattered applications; interoperability between different operating systems; security components that makes sure data is only shared with the right resources; orchestration by means of visual guidance by a single interactive user interface. (Hybrid Integration Platform*

These key findings are translated into the following requirement:

Requirement 4: *The Enterprise Architecture Framework must make the framework* ***enabling a direct real-time connection between data providers and data users.***

---

5. Uniformity in Ecosystem

Each interviewee mentioned that in the ideal situation, organizations must have the same way of working. According to the interviewees, the same way of working includes standards within ecosystems, between organizations and within an organization. This will lead to easier integration of external data. The interviewees also mentioned that this nearly impossible to arrange for every organization, but within the ecosystem of an organization with shared incentives it could be possible. Interviewee S5 and S9 substantiate this:

The ideal data integration occurs when the data is ready-to-use and not much time is needed for preparation and management. This can be achieved by following the same standards that are used in the partner organizations and even between and within the same departments of an organization. This leads to a uniform way of working.
– Interviewee S9

Create uniform standards. This will be difficult, because so much data is available already. Correcting will be very difficult, starting over would be a possibility.
– Interviewee S5

One interviewee refers to the common-pool resource theory by Nobel prize winner Eleonor Ostrom. Eleonor Ostrom states *"A common-pool resource typically consists of a core resource (e.g. water or fish), which defines the stock variable, while providing a limited quantity of extractable fringe units, which defines the flow variable. While the core resource is to be protected or nurtured in order to allow for its continuous exploitation, the fringe units can be harvested or consumed"* (Ostrom, 1990):

Is an architecture a common pool resource? Ostrom describes how actors should live with each other. She describes that even competitors/enemies should make agreements together to make sure a certain resource can be utilized, because they must survive. Without rules the resource will be exploited and not usable anymore and no one will survive. You can use this theory for data as well.
– ***Interviewee S6***

The interviewee states an analogy can be made with data. Without protection or nurturing, no continuous exploitation of data would be possible. Without a common way of working, common data model, common standards and common methods, the real-time sharing of data will become too difficult to handle. Therefore, using the theory by Ostrom (1990) the shared platform should be seen as an action arena in which game rules should be made to survive: Institutions between actors that ensure the continuous exploitation of data (core resource). If this will be done within ecosystems, the real-time integration of data for decision-making can be enabled.

With a common data model with meta-data standards, data will be more sustainable, and it allows for organizations to grow organically by connecting to more external organizations. In this case every organization of an ecosystem provides APIs that could be accessed easily, because they use the same way of working and provide technical documentation about the APIs. Interviewee S1 confirms this:

"Organizations are providing data to each other through APIs. APIs can only be accessed when companies are using open data standards + technical documentation. Otherwise other companies cannot access."
– Interviewee S1

The key insights derived from the qualitative data analysis are:

I. *The Enterprise Architecture Framework must ensure a uniform way of working within ecosystem of a company, including the common standards.*
II. *Collaboration with other organizations is stimulated by using the same way of working, the same standards within the ecosystem of the organization.*

This is translated into the following requirement:

***Requirement 5:** The Enterprise Architecture Framework must create an **uniform way of working within the organization and its ecosystem** by developing **Uniform Standards**; a **Common Data Model** (One data model; unified schema and semantics; building on existing efforts); **Uniform Meta-data standards** (technical documentation about APIs); **an Independent API Policy**; and **Institutions** (action rules) to shape/implement this way of working.*

6. Governance

Another important aspect retrieved from the interviews, is the need for governance. As discussed in the adoption barriers requirements definition section, the lack of governance cause problems with the adoption of BOLD. According to the interviewee's governance is needed on several domains. First, different interviewees state that the Enterprise Architecture that is connected to external organizations will have to grow organically. Hence, it is important for an organization to have an architecture that is flexible, scalable, maintainable and allows for new connections and linkages with the ecosystem of the organization. This must be facilitated by a governance mechanism. Interviewee S6 recommends to think about how this can be facilitated:

"You are going to conclude that architectures like this will grow. New parties will become part of the ecosystem, such as parties that are immature, parties that have a lot of data, parties that own infrastructures, parties that are customer owners. So, it is better to think: What are the guidelines/principles that help organizations with developing an architecture that can grow organically. For this you will make sure you do not focus on organization A, not organization B, but the focus on what should be developed in between them."
– Interviewee S6

The governance mechanism must govern the layer in which the companies are connected. It must focus on quality and continuity of the architecture, including API policy management, APIs and integration Services Catalog, Life Cycle Management, Meta-data management (documentation), Data quality. Without taking these aspects into account, the same problems will occur every time the something in the environment changes, such as a new ecosystem partner, a new API, a new legislation or new standards. Interviewee S7 explains the need for a governance mechaism

"When we are going to work with APIs and real-time data. Then it is needed that parties have full trust in each other. When you store the data locally first, it gives you time to check and assess the quality of the data, adjust and act. If you fully want to trust on the data, you need to think about partnerships with an additional governance mechanism that focuses on data quality and continuity."
*– **Interviewee S7***

The key insights derived from the qualitative Data analysis are:

I. Once an architecture is created through collaboration with the ecosystem. It is important that the architecture can keep running by being able to respond to changes in the environment.

This is translated into the following requirement:

***Requirement 6:** The Enterprise Architecture Framework must facilitate the development of an **architecture that can grow organically**, that is **scalable**, **maintainable** and **flexible towards** its **ecosystem**; ensures **continuity**; and allows for **new connections**.*

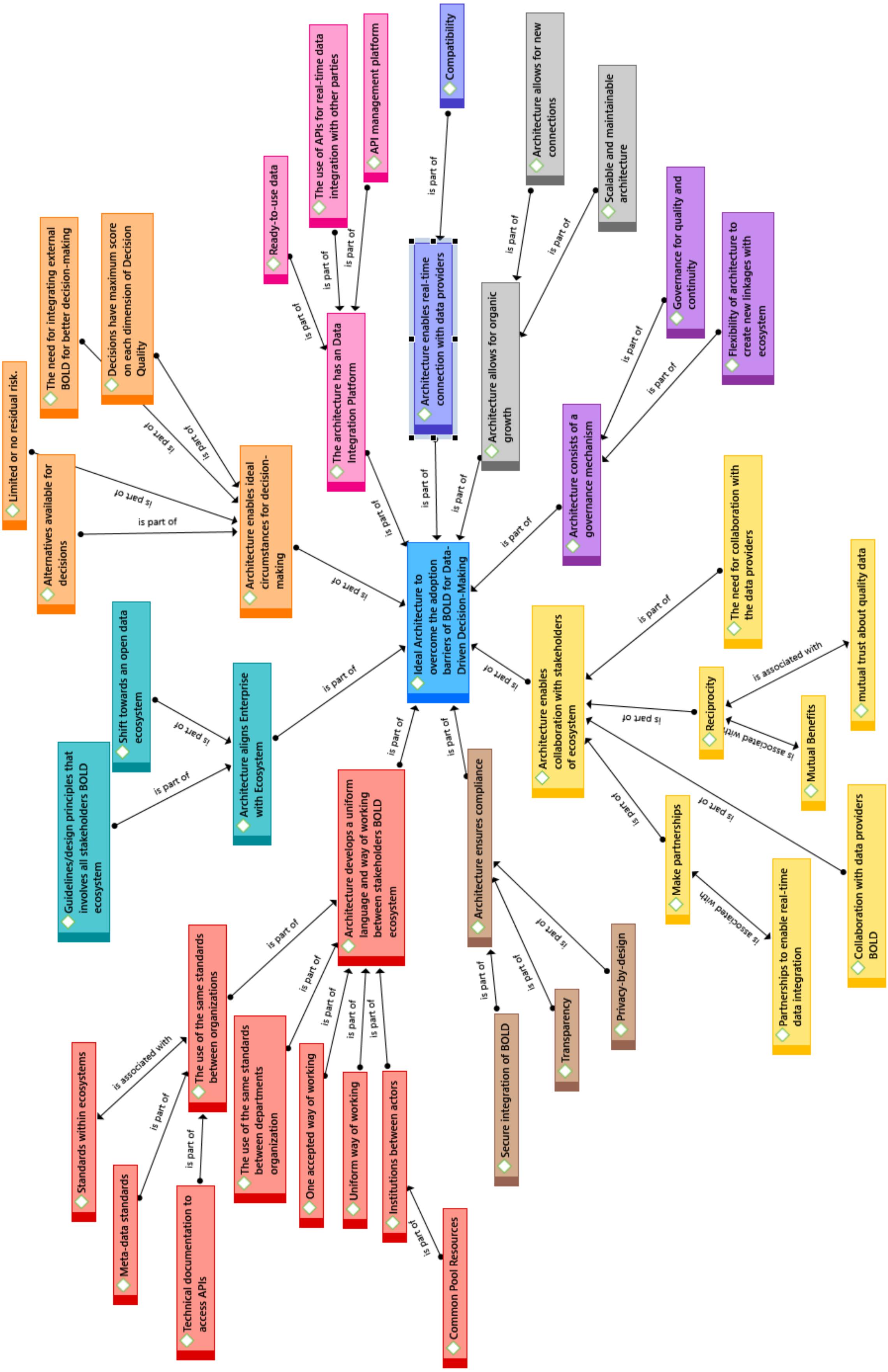


Figure 28: Thematic network Ideal Capabilities Extended Enterprise Architecture Framework

4.4. Summary Chapter 4 (Answer RQ3)

The third sub-research question ***'RQ3. What are the requirements for an Extended Enterprise Architecture Framework that enables real-time integration of external Big and Open Linked Data with internal company data for decision-making?'*** identified how the TOGAF-ADM must be extended in order to remove the limitations that hinder the adoption of BOLD. In total, 9 explorative interviews are conducted with different stakeholders of the research problem. Trough qualitative data analysis and thematic networks the transcriptions are analyzed, six requirements for an extended EAF are defined. This chapter presents how the requirements are derived from the explorative interviews through qualitative data analysis. The interviews resulted in key insights about the adoption barriers that are encountered by organizations and the required capabilities for an extended EAF that enables the real-time adoption of BOLD for decision-making. The key insights are summarized in Appendix C1: Overview Key Findings Explorative Expert Interviews. Based on these key insights six requirements for an Extended Enterprise Architecture Framework are defined. The key findings of qualitative data analysis are translated into requirements for the design. Table 17 presents the answer to the sub-research question of this chapter.

Table 17: Overview Requirements for Extended Enterprise Architecture Framework

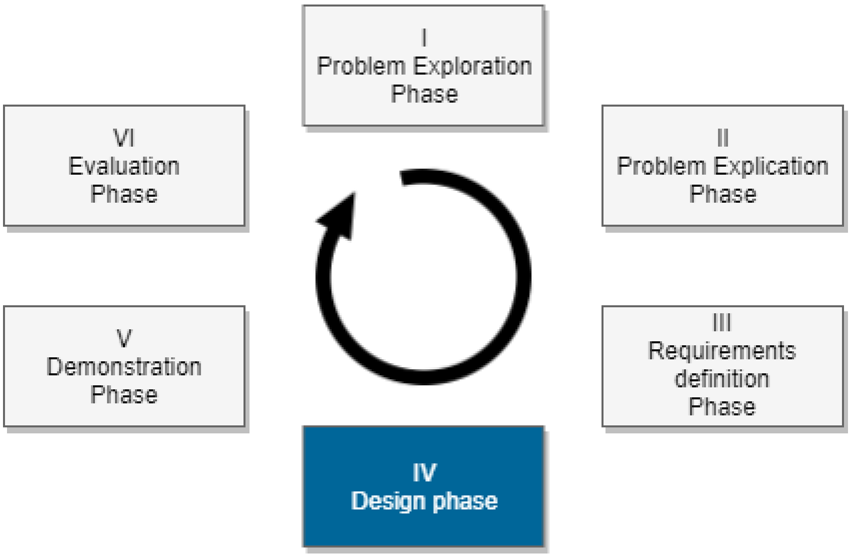
RQ ID	Definition
RQ1	The Enterprise Architecture Framework must **proactively** deal with the **Information Quality Barriers, Task Complexity Barriers, Technical Barriers, Use & Participation Barriers, Legislation Barriers,** and **Governance Barriers** that organizations address with the real-time integration of BOLD for decision-making.
RQ2	The Enterprise Architecture Framework must **stimulate collaboration** between organizations to create mutual commitment of the stakeholders, by developing a *shared vision, including partnerships, business models, mutual incentives / benefits, integration strategy.*
RQ3	The Enterprise Architecture Framework must **involve all stakeholders from the BOLD ecosystem that are concerned** with the real-time integration and sharing of external BOLD for decision-making system in the design of the Extended Enterprise Architecture **to create better alignment of the enterprise with its ecosystem**, including Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, Data Prosumer and Legal Authorities.
RQ4	The Enterprise Architecture Framework **must (build on TOGAF-ADM to)** make the framework capable of **enabling a direct real-time connection between data providers and data users.**
RQ5	The Enterprise Architecture Framework must **create a uniform way of working** within the organization and its ecosystem by developing **Uniform Standards**; a **Common Data Model** (One data model; unified schema and semantics; building on existing efforts); **Uniform Meta-data standards** (technical documentation about APIs); an **Independent API Policy**; and **Institutions between actors** (action rules) to shape/implement this way of working.
RQ6	The Enterprise Architecture Framework must facilitate the development of an **architecture that can grow organically**, that is **scalable**, **maintainable** and **flexible towards** its **ecosystem**; ensures **continuity**; and allows for **new connections**.

The requirements are translated into capabilities. Table 18 presents the capabilities that must be incorporated in the design of an Extended Enterprise Architecture Framework that enables the real-time integration of external BOLD with internal company data for decision-making. In Chapter 3, it is discussed that the TOGAF-ADM seems the most suitable EAF for extension (Sessions, 2007; Urbaczewski & Mrdalj, 2006). The TOGAF-ADM can be used for aligning across the enterprise, but not between enterprises. Incorporating these additional capabilities will result in a design artifact that can overcome the research problem. The capabilities are defined in a more general way to ensure an approach that can be used to implement different strategies that involve external parties from its ecosystem.

Table 18: Required Capabilities to enable real-time integration of external BOLD with Internal Company Data for decision-making

EA Capabilities		TOGAF
C1	Continuously align Business / IT:	Yes
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	Yes
C3	Remove the threshold for change	Yes
C4	Enable/Increase flexibility of the organization	Yes
C5	Develop a proactive organization	Yes
C6	Reduce risk and prepare for rapid, unplanned change	Yes
C7	Avoid tension between business IT functions in organization	Yes
C8	Create, unify and integrate business processes across enterprise	Yes
C9	Unlock the power of information, unifying information silos in organization	Yes
C10	Eliminate duplicate and overlapping technologies in organization	Yes
C11	Reduce solution delivery time, development costs (maximize reuse):	Yes
C12	Continuously align an Enterprise with its Ecosystem	-
C13	Stimulate collaboration between organizations	-
C14	Enable a direct real-time connection between data providers and data users.	-
C15	Enable/Increase Flexibility of Enterprise towards its Ecosystem. Develop an Enterprise that is scalable, maintainable, sustainable and can grow organically	-
C17	Develop uniformity within Ecosystem	-

IV. Design Phase



5. Design

This chapter discusses RQ4 translation of the design requirements into a design artifact that can overcome the research problem. First, the research approach is explained (Section 5.1). Secondly, the design phases are explained in detail (Section 5.2 and 5.3.). Thirdly, the final design is presented and explained in detail (section 5.4). Finally, the research question is answered (Section 5.5).

5.1. Overview Design Phase

As discussed in the Chapter 3, the current foundation of EA not suitable enough for implementing a strategy that involves the real-time adoption of BOLD for data-driven decision-making. The review addressed that 11 capabilities should be incorporated by the design of the extended EAF: Continuously align Business and IT; Provide a process for transition towards the 'to be' state of an organization; Remove the threshold for change; Enable/Increase Flexibility of the organization; Develop a proactive organization; Reduce risk and prepare for rapid, unplanned change; Avoid tension between business and IT functions; Create, Unify and integrate business processes across the enterprise; Unify information silos within the organizations; Eliminate duplicate technologies; and Reduce solution delivery time. In the literature review in Chapter 3, the most commonly used Enterprise Architecture Frameworks are assessed on having these capabilities. The literature review shows that these capabilities are incorporated in the TOGAF-ADM. Through conducted explorative expert interviews additional required capabilities for the real-time adoption of BOLD are identified. However, the TOGAF-ADM does not have the additional capabilities that are required. In the Requirements Definition Phase, these capabilities are translated into requirements for the design of an Extended EAF that can be used to enable the real-time integration of external BOLD with internal company data.

EA has been an accepted, valued method to build the enterprise for years. Hence, it is important that we do not try to re-invent the wheel but provide a method that can help improve the wheel to make it more suitable to function in the changing environment. In this chapter an extension is designed that builds on the successful foundation of Enterprise Architecture. The extension that is build can be used on the TOGAF-ADM, so the 11 capabilities extracted from the literature review are already incorporated in the design.

5.2.1. Overview Design steps

Figure 29 presents an overview of the Design Phase. The Design Phase consists of three steps. In the first step a morphological chart is created to identify the design space for the extension. In the second step, the design options for the design space are identified. The design options are means to fulfil the functional requirements that are collected during the explorative expert interviews. In the final step, the actual design is made by selecting a combination of the design options that are identified in the second design step. For each step, several research methods are used which is discussed in the following section.

Figure 29: Steps Design Phase

5.2.2. Step 1: Morphological chart

As mentioned before, a morphological chart is used to identify the different means that can be used to fulfill the different functional requirements (Dym et al., 2009). A morphological chart is used to construct and visualize the design space and the design options that are part of the design space. On the vertical axis of the morphological chart the functional requirements are displayed, and on the horizontal axis the means (solutions) are displayed. Combining a single means for each of the functional requirements will produce potential conceptual design solutions (Dym et al., 2009; Smith, 2007). This process can be repeated for several times to create a list of conceptual design solutions. In theory, all the possible combinations of means are possible. Only the feasible or most probable combinations are considered.

Before the design options are chosen, a reflection is done on the different building blocks discussing the importance of a building block, the key question the building block tries to answer, the function it must perform, and what from what perspective the building block is used. Figure 30 displays the outline of the morphological chart for the design the extension for the EAF.

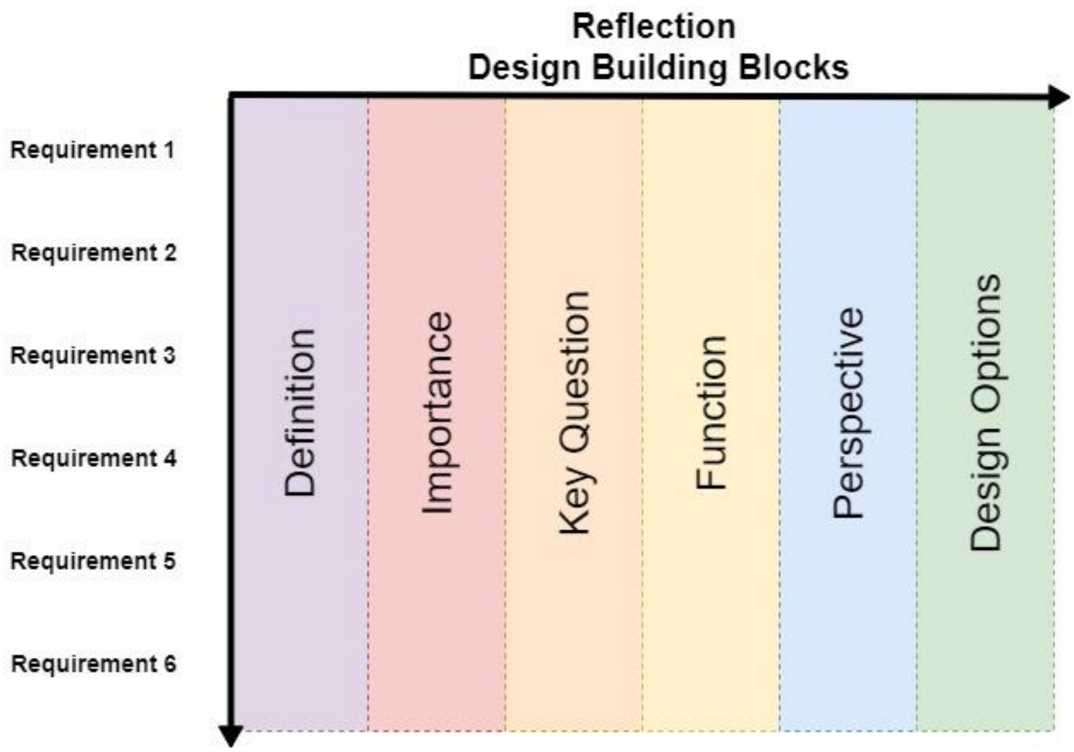


Figure 30: Morphological Chart Outline

5.2.3. Step 2: Brainstorm on Design Space

In the second step of the Design phase several brainstorm sessions are held. The first brainstorm session is held to identify the different design options that fit into the building blocks to fill the design space. For each of the requirements the reflection on the different design building blocks of the extended EAF are revisited and used as input for the identification of the most suitable design options. During the second brainstorm session three different conceptual design solutions that are considered the most suitable combinations to form the design artefact are selected. These three conceptual design solutions are assessed through a best-of-class chart (Dym et al., 2009; Smith, 2007).

The best-of-class chart has the objectives of the one axis and the conceptual design solutions on the other axis. The objectives are given a weight to make sure the most important objectives are represented by the selected conceptual design solution. Each conceptual design solution is given a score on each of the objectives. Finally, the weighted scores are summed up to determine what would be the most suitable solution. The weights are based on the times the requirement was mentioned during the explorative expert interviews. This solution is designed in the form of a framework.

5.2.3. Step 3: The Final Design

The third step of the Design Phase is about transforming the conceptual design solution into an Enterprise Architecture Framework extension that resolves the research problem. The design of an artefact is always an iterative process that delivers different versions. The different versions of the design artefact, the rationale for the implemented changes towards the newer versions, and the additional or adjusted requirements collected during evaluative interviews are discussed. In other words, the whole journey towards the final design is explained in detail.

5.2. Morphological Chart

Table 19 presents the reflection of the design building blocks that represent the requirements that are collected during the explorative interviews. Each building block is described by discussing the definition of the requirement it fulfils; its importance of being part of the design; the key question the building block should answer; the function the building block must fulfil; and the perspective on which the building block is based on.

As depicted in the Table 19, the design consists of six different building blocks. Each building block covers a part of the design artefact. Together the building blocks should form a design artefact that can be used to solve the research problem.

Table 19: Morphological Chart Reflection on the design building blocks

ID	Definition	Importance	Key Question	Function	Perspective
RQ1	*The EAF must Proactively deal with the adoption barriers*	*Adoption barriers should be prevented instead of reactively be overcome.*	*How to assess whether an EA extension is required to prevent the adoption barriers from occurring?*	*Assess the need for extending the Enterprise Architecture*	*Literature Review + Empirical*
RQ2	*The EAF must stimulate collaboration between organizations*	*Stimulants should be developed to increase the incentives for collaboration of the stakeholders.*	*How to create commitment of relevant stakeholders that are concerned with the adoption of external BOLD?*	*Getting the ecosystem committed*	*Literature Review + Empirical*
RQ3	*The EAF must involve all stakeholders that are concerned to create better alignment*	*Viewpoints of all concerned stakeholders must be developed to ensure a complete architecture*	*How should the system look like according to the committed stakeholders?*	*Getting the ecosystem involved*	*Literature Review + Empirical*
RQ4	*The EAF must enable a direct real-time connection between data providers and data users.*	*Architecture must be satisfying all viewpoints of concerned stakeholders to fulfill missing capabilities*	*How to build the Architecture that is required?*	*Getting the Architecture Running*	*Literature Review + Empirical*
RQ5	*The EAF must create a uniform way of working within the organization*	*Common way of working is required to makes sure the architecture is used in the right way and*	*What is required to make the architecture work?*	*Making the Architecture work*	*Literature Review + Empirical*
RQ6	*The EAF must facilitate the development of an architecture that can grow organically, that is scalable, maintainable, flexible towards ecosystem*	*Once the Architecture is running, it is important that the architecture is updated according to the needs of the environment.*	*What is required to keep the architecture running?*	*Keep the architecture running*	*Literature Review + Empirical*

5.3. Brainstorm sessions

In the previous section the morphological chart was filled by reflecting on the different building blocks that the design artefact should have. As depicted in Figure 30, the last step is defining potential design options that could represent the design building blocks. In total, two brainstorm sessions were done in order to identify the conceptual design solutions. In the first brainstorm session the potential design options for each design block are identified. This results in several means per design building block. In the second brainstorm session, different sets of design options are selected to create three conceptual design solutions that can solve the research problem. One of the conceptual design solution is selected through a best-of-class chart and is elaborated and explained in detail (Dym et al., 2009).

5.3.1. Brainstorm session 1: Identify the design options

Figure 31 presents the results of the first brainstorm session on the different design options (Appendix D1: Design Options). For each requirement, three means are identified that, respectively, represent three approaches: a descriptive approach, a prescriptive approach, and a guiding approach. For example, for the second requirement '*The EAF must stimulate collaboration between organizations*' three means are identified. The first mean 'Present an overview of different stimulants that increase the change for collaboration' is a descriptive approach, because it gives an overview of common ways to stimulate collaboration between organizations and let the organization choose one of the options. The second mean 'present guidelines that help organizations to write a proposal to collaborate' provides guidance for the process of writing a customized proposal that perfectly suites the situation of the organization. The third mean 'Prescribe a recommended approach to stimulate collaboration between organizations' is provides a recommendation that authoritatively puts forward an approach to stimulate collaboration between organizations.

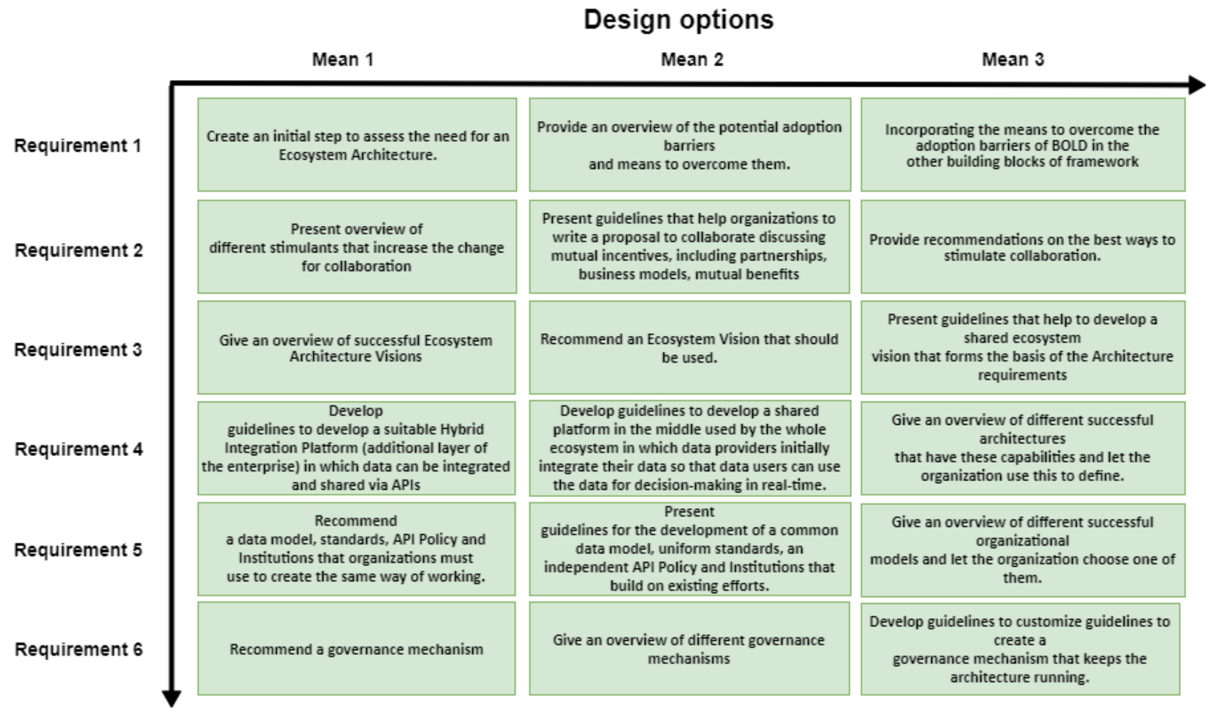


Figure 31: Design options Morphological chart

5.3.2. Brainstorm session 2: Select conceptual design solution

Figure 32 presents an overview of the three selected conceptual design solutions. Each color represents a set of means that together form a conceptual design solution. The green set of means represents the guiding approach, the yellow set of means represents the descriptive approach and the red set of means represents the prescriptive approach. The first conceptual design solution is the descriptive approach for building an extended enterprise architecture. In this solution (depicted in yellow), all the different means that use a descriptive approach are combined. The main approach to fulfil the requirement in each step is to give an overview of potential solutions which can be used by the user of the framework. The second conceptual design solution is the guiding approach for building an extended enterprise architecture. In this solution (depicted in green), all the means focus on providing guidelines to create a customized extended enterprise architecture that perfectly suites the situation of the organization. The final conceptual design solution is the prescriptive approach. In this solution (depicted in red), all the means authoritatively put forward a recommendation of how to fulfil the requirements in order to build an extended enterprise architecture.

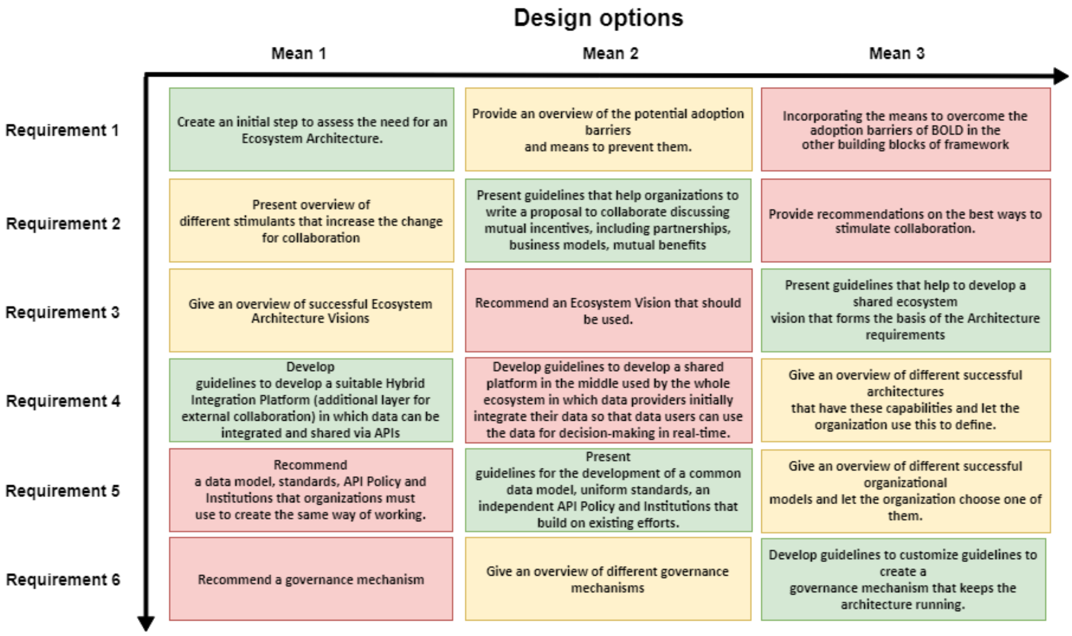


Figure 32: Conceptual Design Solutions

Dym et al. (2009) suggest an approach for choosing between different classes. In order to determine what conceptual design solution is most suitable to solve the research problem, a best-of-class chart has been used. Based on how often a requirement was mentioned during the explorative expert interviews, requirements are given weights to develop a representative score for each conceptual design solution. Table 20 presents the best of class chart. The scoring system is based on a ranking system with a range from 1 to 3, and in case of a draw the score is equally divided over the means. During the brainstorm session the means of each conceptual design solution were ranked and given a score. After calculating the weighted scores, the results show that the guiding approach is the most suitable conceptual design solution. Therefore, the guiding approach will be designed in the next section.

Table 20: Best-of-Class Chart

ID	Definition	# men-tioned	Weight	Descriptive Approach	Weighted Score	Guiding Approach	Weighted Score	Prescriptive Approach	Weighted Score
RQ1	*The EAF must Proactively deal with the adoption barriers*	*3*	*0.33*	*1*	*0.33*	*2.5*	*0.825*	*2.5*	*0.825*
RQ2	*The EAF must stimulate collaboration between organizations*	*5*	*0.55*	*1*	*0.55*	*3*	*1.65*	*2*	*1.1*
RQ3	*The EAF must involve all stakeholders that are concerned to create better alignment*	*7*	*0.77*	*2*	*1.54*	*3*	*2.31*	*1*	*0.77*
RQ4	*The EAF must enable a direct real-time connection between organizations.*	*9*	*1*	*1.5*	*1.5*	*3*	*3*	*1.5*	*1.5*
RQ5	*The EAF must create a uniform way of working within the organization*	*9*	*1*	*2*	*2*	*3*	*3*	*2*	*2*
RQ6	*The EAF must facilitate the development of an architecture that can grow organically, that is scalable, maintainable, flexible towards ecosystem*	*5*	*0.55*	*3*	*1.65*	*2*	*1.1*	*1*	*0.55*
Total Score					*7.57*		*11.885*		*6.745*

5.4. Incremental Design Steps

Based on the outcome of the best-of-class chart, the most suitable conceptual design solution is designed. This is done through an iterative approach in which the design artefact is tested through multiple evaluative expert interviews. During the evaluative expert interviews, the strong and weak points of the design artefact are collected and transformed into suggestions for potential improvements to the model. According to the design research cycle by (Hevner, 2007), the design has to be evaluated in order to collect additional requirements and to revise requirements. This results in a more valid design artefact. Different versions have been developed. Based on the Evaluation discussed in Chapter 6, the final design was adjusted. The steps taken to improve the final design are discussed in that section.

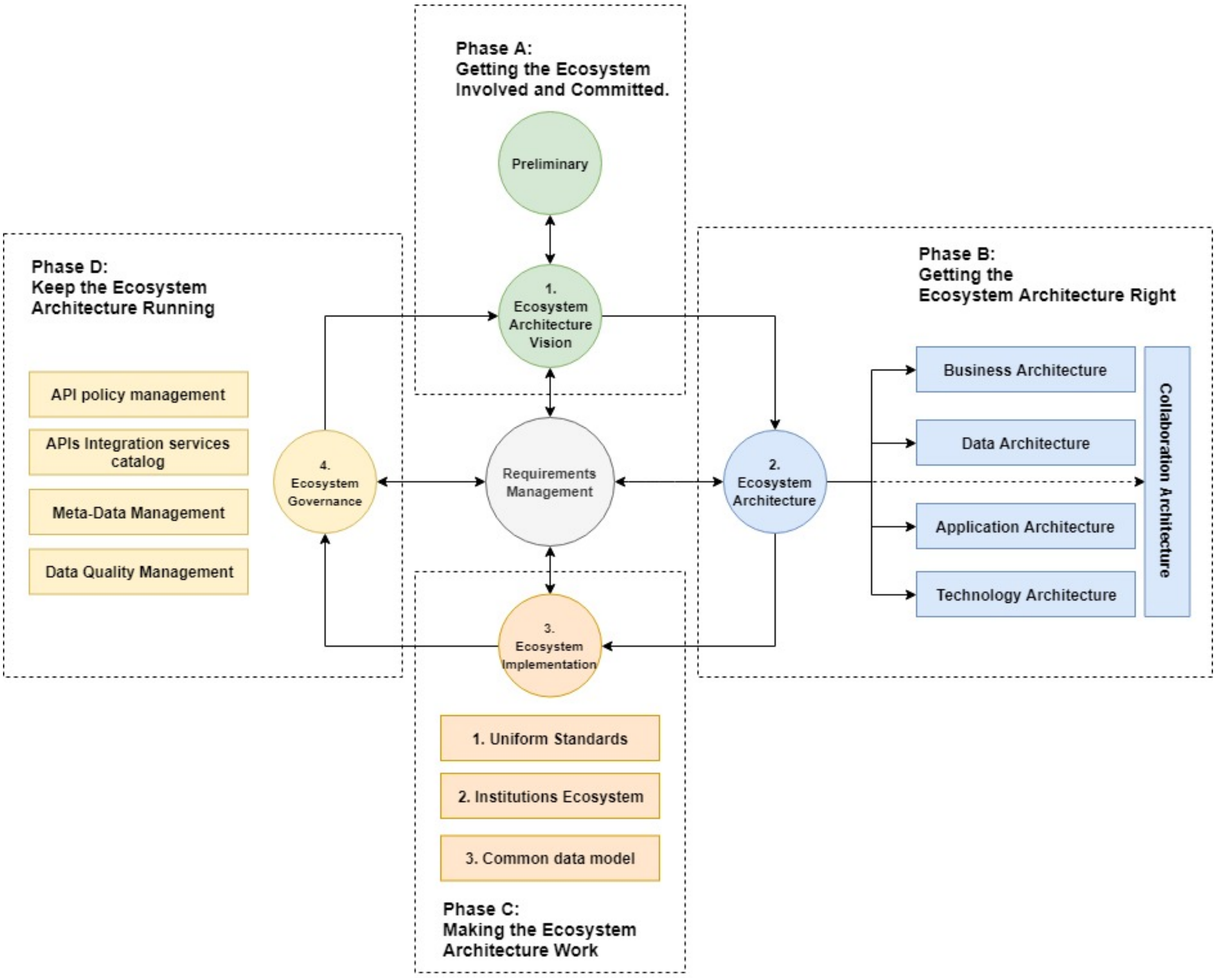


Figure 33: The Ecosystem Architecture Framework. Version 0.9

5.5. The Final Design: The Ecosystem Architecture Framework

The design phase has resulted in the Ecosystem Architecture Framework. The Ecosystem Architecture Framework can be used as an extension to the TOGAF-ADM that is discussed in Chapter 3. The rationale for building on the TOGAF-ADM is because the framework has incorporated all capabilities of the current foundation of EA but has its limitations regarding the requirements of integrating external BOLD with internal company data in real-time (Table 18). The Ecosystem Architecture Framework has incorporated the required capabilities to mitigate these limitations. As the implementation of this IT strategy does not only concern the business, data, application and technology architecture of the internal enterprise, the framework also aligns a new strategy and its implementation with the ecosystem of an Enterprise. Figure 34 presents the relationship of the Ecosystem Architecture with TOGAF. Figure 35 presents a high-level overview of the Ecosystem Architecture Framework, which is structured into four phases, depicted as Phase A, B C, and D.

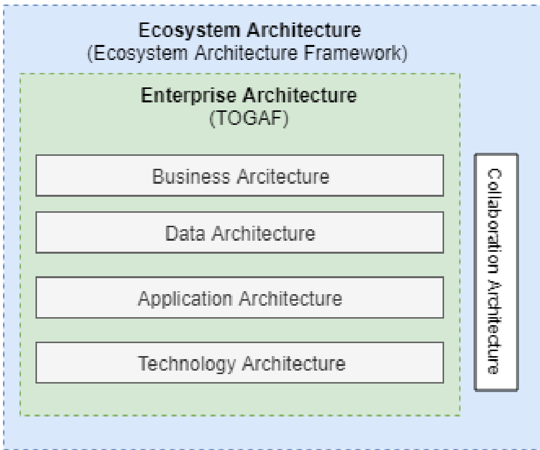


Figure 34: Macro overview Ecosystem Architecture

Phase A focuses on getting the stakeholders of an organizations' ecosystem involved and committed with the development of the Ecosystem Architecture. It starts with a preliminary phase, where a company identifies the need for the creation of an Ecosystem Architecture (Phase A.0). This results in a request for Ecosystem Architecture Work. Moreover, a proposal to external stakeholders to collaborate. The purpose of this phase is to create mutual incentives to mitigate the adoption barriers that arise that hinder the real-time integration of BOLD. In case the external stakeholders are willing to cooperate, a shared Ecosystem Architecture Vision is developed which the includes the shared business needs. Lastly, this vision is translated into architecture requirements for the Ecosystem Architecture (Phase A.1).

Phase B focuses on getting the Ecosystem Architecture right, by satisfying the Architecture Requirements through the creation of a suitable collaboration architecture. The collaboration architecture is a hybrid integration platform in which data can be integrated and shared, different applications can be linked together, makes possible collaboration between applications, stimulates interoperability between different operating systems and orchestrates the applications by a single interface. Several integration platforms are possible, but the Ecosystem Architecture Framework recommends an API platform to facilitate this (Phase B).

Phase C focuses on making the Ecosystem Architecture working by developing an inter-organizational model consisting of shared artefacts that ensure a common way of working is established: Uniform standards (e.g. Data standards and Meta-data standards); Institutions between organizations of the ecosystem; A Common Data Model used by all committed organizations; and an independent API policy. Having the same way of working ensures that the organizations are better aligned with each other and it is easier to develop APIs that can be accessed for the real-time integration of external BOLD through the Hybrid Integration Platform (Phase C).

Phase D focuses on keeping the Ecosystem Architecture running, by making sure the Collaboration Architecture and the inter-organizational model adapts to changes from the environment. It ensures that the Ecosystem Architecture optimally fulfils the business requirements and is aligned at all time. It ensures the Ecosystem partners are managed, so the architecture is updated when a new partner joins or leaves the strategic alliance; the API policy is updated according to the needs of the ecosystem; the APIs catalog is updated; the data quality and meta-data are managed; and the institutions are enforced (Phase D).

To create a better understanding of the Framework, the need and function of the Ecosystem Architecture Framework is explained by means of an example: *A flight comparison website supports data-driven decisions on the most suitable flight options. To improve the decisions, the website wants to adopt the external BOLD provided by a variety of airlines in real-time to advise customers the most suitable flights adapted to their needs. As external BOLD is very dynamic, it is important that the data that is used for the decision-making is in real-time. However, interoperability problems, a lack of standards, and data handling issues hinder the integration of the external BOLD in real-time. Therefore, it is important that these adoption barriers are mitigated to improve the decision-quality. To solve this problem, the flight comparison website can use the Ecosystem Architecture Framework to transform its system of record into a system of engagement in which a collaboration is established with the Airlines that provide BOLD with flight information in real-time. By involving the stakeholders concerned with the data-integration in the design of this system, architectural requirements are defined for the ecosystem architecture. The organization invests in an API Management Platform that integrates the APIs of all the different airlines shaped by an inter-organizational model to ensure a common way of working. Furthermore, the platform ensures that the architecture keeps running by adapting to changes of the environment.*

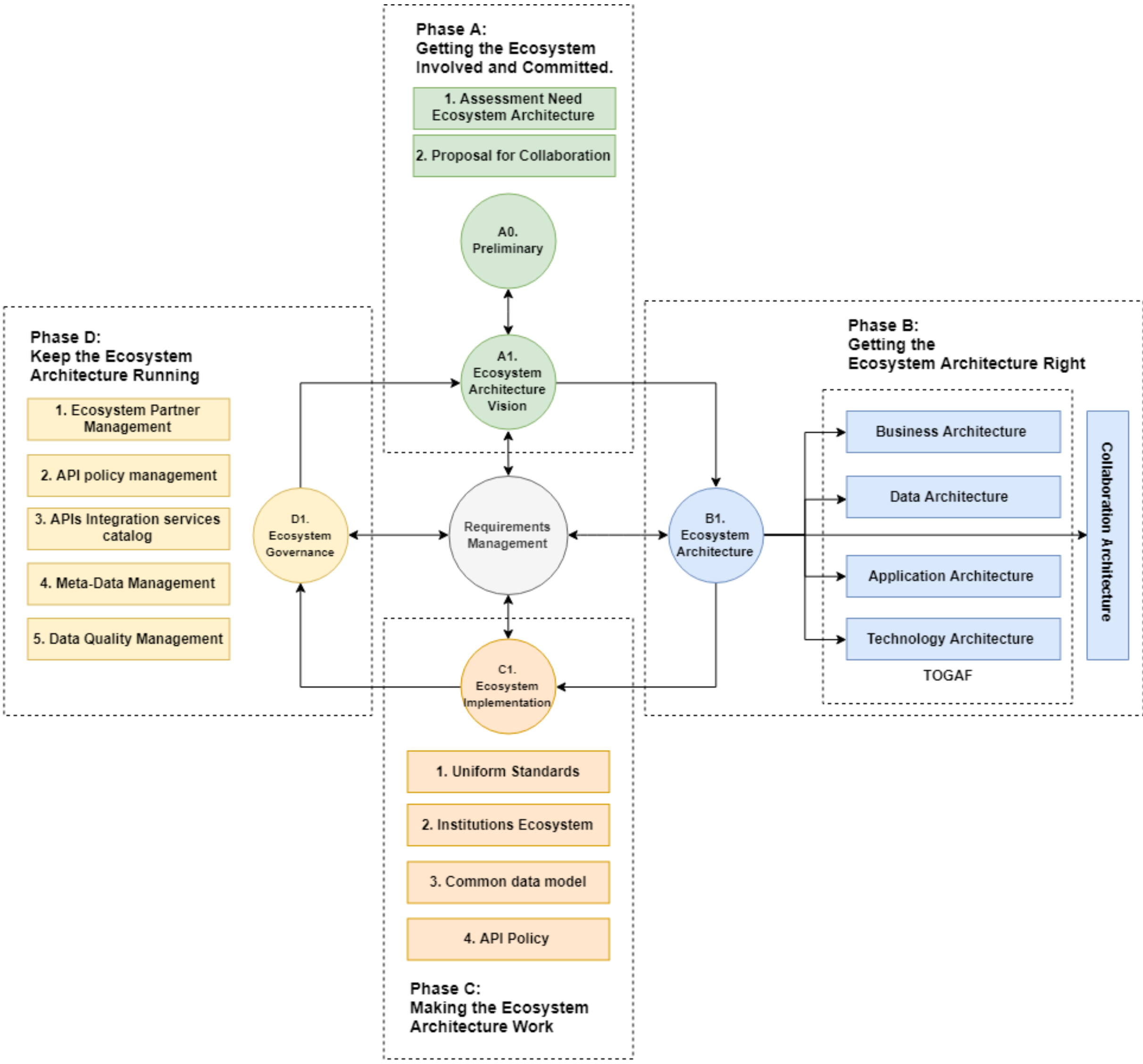


Figure 35: The Ecosystem Architecture Framework (v1.0)

5.5.1. Phase A: Getting the Ecosystem Involved and Committed

5.5.1.1. A0. Preliminary phase

The preliminary phase represents **RQ1** *'The Framework must Proactively deal with the Adoption Barriers of BOLD.'* and **RQ2** *'The framework must stimulate collaboration between organizations'*. In this phase, the Ecosystem Architecture Framework helps the organization to identify the need for collaboration with external parties to mitigate the (potential) adoption barriers that can arise during the integration of external BOLD. It helps the organization to assess the need for an Ecosystem Architecture and an inter-organizational model that facilitate that the architectures of data providers can be initially connected with the Enterprise Architecture of an organization to enable the real-time integration. However, both parties should be willing to cooperate to make this happen. In the preliminary phase, the stimulants for a successful collaboration are identified, such as a partnerships, shared business models, mutual incentives. This phase is mainly about getting the business ecosystem that is concerned with the real-time integration of external BOLD committed to the development of an Ecosystem Architecture in which the adoption barriers are mitigated. Furthermore, an assessment is performed of what architectural changes are required to make the collaboration work. The preliminary phase consists of three different core steps, which are explained in detail below. Figure 36 presents the guidelines of the preliminary phase of the Ecosystem Architecture Framework.

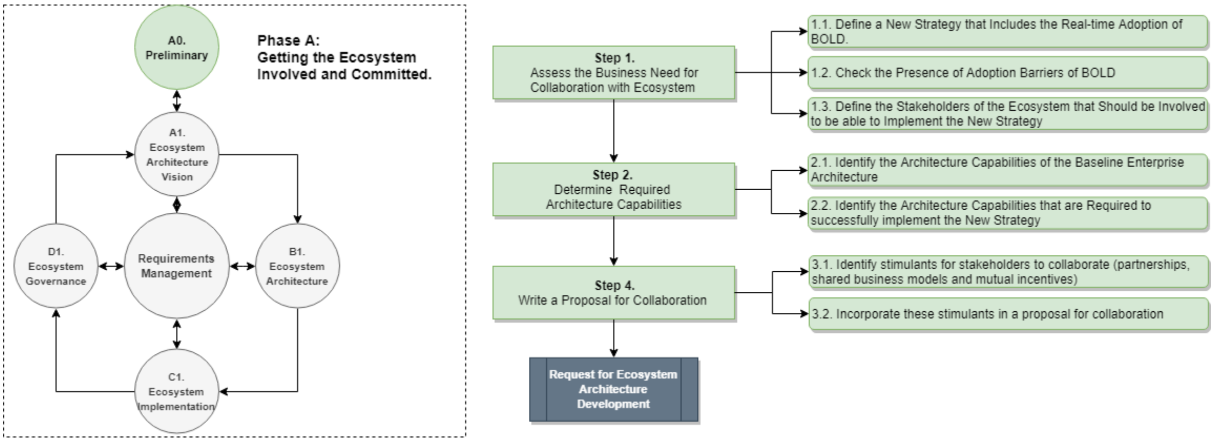


Figure 36: Guidelines Preliminary Phase (A0).

Step 1: Assess the Business need for Collaboration with Ecosystem

First, a new IT strategy is developed that involves the real-time use of external BOLD (step 1.1). Secondly, the adoption barriers are identified that can arise with the implementation of this new strategy (step 1.2). Thirdly, the stakeholders that should be collaborated with to mitigate these adoption barriers are identified (step 1.3). After these three steps are performed, the organization has a clear idea whether collaboration with the business ecosystem is required to implement the strategy.

Step 2: Determine Required Architecture Capabilities

To determine the architecture capabilities that the baseline architecture is lacking, the capabilities of the current and required architecture are identified (step 2.1 and 2.2). This helps to identify the additional architecture capabilities are required to implement the new strategy. This is done through a gap analysis. In the third step, the ability of transforming the organization and its Enterprise Architecture is assessed. First, the gap analysis is translated into business requirements and, secondly, there is assessed whether the organization is capable to fulfill these requirements.

Step 3: Write a proposal for collaboration

The third step focuses on creating a strategic coalition to mitigate the adoption barriers that can arise with the integration of external BOLD. In Step 1 and Step 2 the need for collaboration with stakeholders of the business ecosystem and the architecture capabilities to enable this are identified. This third step focuses on getting the stakeholders of the ecosystem on committed. Step 3.1. focuses on creating stimulants for the stakeholders to collaborate. It must be said that one party has more benefit of a collaboration than other party. Furthermore, each stakeholder as a certain degree of dominance in the negotiations of a collaboration. This step takes this into account and helps organizations to create strategic alliances through partnerships, shared business models and mutual incentives. This can be the sharing of data, a financial compensation or other incentives that stimulates the cooperation of the ecosystem partners. Step 3.2 focuses on writing a proposal that incorporates these stimulants to make it attractive for other organizations to establish a collaboration.

Output: Request for Ecosystem Architecture Development

The output of this initial phase is a request for Ecosystem Architecture Development. The request consists of a proposal for collaboration in the development of the Ecosystem Architecture of the organization. When the need for a collaboration with the ecosystem is identified, a request for Ecosystem Architecture Development is submitted. The proposal for collaboration is made to all the stakeholders that should be involved and committed. Once the stakeholders decide to cooperate, the actual Ecosystem Architecture Development Cycle starts (Phase A1.)

5.5.1.2. A1. Ecosystem Architecture Vision

In the Preliminary Phase, the need for an Ecosystem Architecture is identified and a request for collaboration is made to the stakeholders of the ecosystem to collaborate. Phase A1, the 'Ecosystem Architecture Vision Phase', represents **RQ3** '*The Enterprise Architecture Framework must involve all stakeholders of the business ecosystem that are concerned with the real-time integration of BOLD in the design of the Extended Enterprise Architecture to create better alignment of the enterprise with its ecosystem*'. In this phase, the stakeholders of the business ecosystem that have accepted the proposal for collaboration and architecture work are involved in the development process of the Ecosystem Architecture. Typical stakeholders are an Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, Data Prosumer and Legal Authorities. In order to create better alignment between the stakeholders that are concerned with the system, the concerns and requirements of each stakeholder are collected. Each stakeholder has an idea of how the system should look like. Hence, it is important that all the relevant stakeholders are involved in the design process. To ensure alignment between the enterprise and the ecosystem a shared Ecosystem Architecture Vision is developed. Developing this vision together ensures that the concerns of each stakeholder are represented in the architecture.

The Ecosystem Architecture Vision Phase is structured into three steps with associated guidelines, depicted in Figure 37. First, the Stakeholders of the Ecosystem are identified (Step 1). Secondly, the Shared Ecosystem Architecture Vision is developed (Step 2). Finally, the Ecosystem Architecture Vision is translated into architecture requirements (Step 3). These steps are explained in detail below.

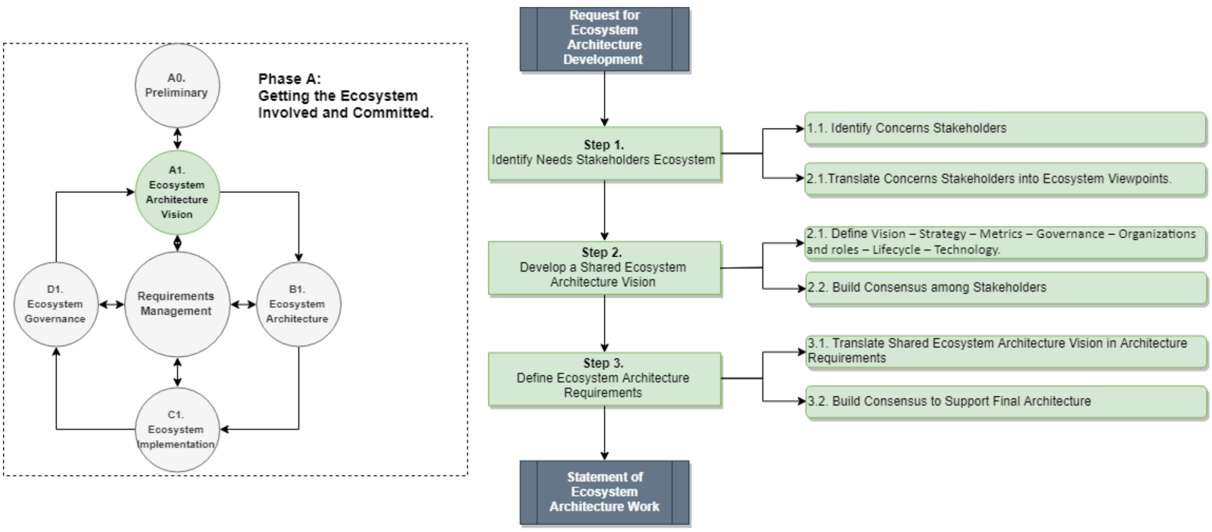


Figure 37: Ecosystem Architecture Vision Phase

Step 1: Identify Stakeholders of Ecosystem

In the first step, the needs of the different stakeholders that are concerned with the real-time integration of external BOLD are identified. As discussed in Figure 16, a system consists of different stakeholders that have concerns about how the system should look like. In order to create better alignment between the enterprise and its ecosystem, the concerns of each stakeholder are identified. Each stakeholder has an idea of how the system should look like. Hence, it is important that all relevant stakeholders are involved in the design process (Step 1.1). Secondly, the concerns that are related to each other are bundled together and translated into viewpoints.

Viewpoints are associated to views that form the architecture. Normally, an Enterprise Architecture Framework consists of four views: Business View, Data View, Application View, and the Technology View. In an Ecosystem Architecture, an additional view is represented. This represents the concerns of each external stakeholder to ensure the real-time integration is enabled and the potential adoption barriers are mitigated. 'The Collaboration View'. These views are derived from viewpoints that are based on the concerns of stakeholders of the system. This makes sure that the concerns of the stakeholders are represented in the system (step 1.2).

Step 2: Develop a Shared Ecosystem Vision

In the second step, the Shared Ecosystem Architecture Vision is developed. The Ecosystem Architecture Vision consists of seven building blocks: Vision – Strategy – Metrics – Governance – Organizations and roles – Lifecycle – Technology. These aspects must be discussed to ensure the development of a complete Architecture Vision (step 2.1). Each stakeholders has its own concerns about how these 7 building blocks should be filled in. It is important that ultimately, the Ecosystem Architecture Vision Represents the needs of all different stakeholders and a consensus is built among the different stakeholders (step 2.2).

Step 3: Define Architecture Requirements

In the third step, the architecture requirements for the Ecosystem Architecture are defined. First, Ecosystem Architecture Vision is translated into requirements for the Ecosystem Architecture. This are requirements for a system that ensures that the adoption barriers are mitigated, and the needs of the ecosystem partners are satisfied (step 3.1). It happens that the concerns of different stakeholders are conflicting with each other, therefore it is important that also consensus is built among the different stakeholders to support the final architecture. Otherwise, there is a chance new adoption barriers will arise. It is important these are mitigated before the system is developed (step 3.2).

Output: Statement of Ecosystem Architecture Work

The product that this phase eventually delivers is 'The Statement of Ecosystem Architecture Work'. In this statement, all the Ecosystem Architecture Requirements that must be satisfied by the architecture of the system are represented. This is used as a starting point to develop the target Ecosystem Architecture (Phase B).

5.5.2. Phase B: Getting the Ecosystem Architecture right.

Phase B focuses on getting the architecture right. In this phase the transformation from the Enterprise Architecture Framework into the Ecosystem Enterprise Architecture is performed. As mentioned, the Ecosystem Architecture Framework is an extension to the TOGAF-ADM, and therefore this phase only briefly explains the Business, Data, Application, and Technology architecture. These are incorporated in the TOGAF-ADM (Figure 34). This phase presents the guidelines for how the additional added layer, the 'collaboration architecture', should be developed.

5.5.2.1. B1. Ecosystem Architecture

Following the 'Statement of Ecosystem Architecture Work', the actual Ecosystem Architecture is developed. Figure 38 presents guidelines for the development of the target Ecosystem Architecture. The Ecosystem Architecture builds on the foundations of the Enterprise Architecture, which usually include a business architecture, data architecture, application architecture, and technology architecture that are representing the viewpoints of different sets of stakeholders. The Ecosystem consists of an extra view the 'Collaboration Architecture'. The TOGAF-ADM provides clear guidelines for the transformation of the baseline business, data, application, and technology architectures into the target architectures that are suitable for implementing a new strategy. The same techniques are adopted in the Ecosystem Architecture Framework, but the focus of this phase is mainly making sure the collaboration architecture is developed and that the other architectures are well-aligned with this layer. The collaboration architecture represents all the requirements of all stakeholders that are concerned with the real-time integration of external BOLD for decision-making. This phase represents **RQ4** '*The Enterprise Architecture Framework must (build on TOGAF-ADM to) make the framework capable of enabling a direct real-time connection between data providers and data users.*'

The collaboration layer represents a hybrid integration platform, which is an environment in which external data can be integrated and shared; different kinds of applications can be integrated and linked together (independent from platform, programming language and resource); collaboration between distributed and scattered applications is possible; interoperability between different operating systems is possible; security components ensure rightful data sharing and data use; and an orchestration mechanism provides visual guidance by a single interactive user interface. A hybrid integration platform is a technological agnostic concept that can be implemented through different integration patterns, which are discussed in more detail later.

The guidelines for the development of the Ecosystem Architecture are structured in four steps: developing the Baseline Architecture; exploring the Target Collaboration Architecture; developing the target Ecosystem Architecture; and performing the gap analysis. The different steps are explained in detail in the section below.

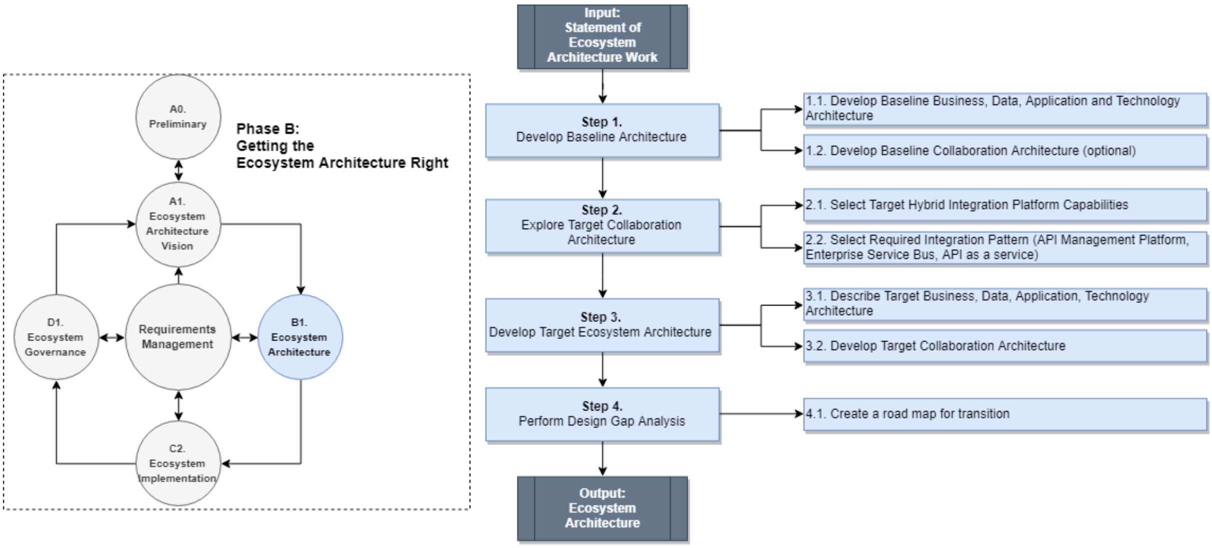


Figure 38: Guidelines Development Ecosystem Architecture

Step 1: Develop Baseline Architecture

The first step represents the development of the baseline architecture. The baseline architecture is representing the currently used system for the integration of the data, such as an 'Extract Transform Load' (ETL) system, Enterprise Service Bus or BGS approach (Figure 39). First, the baseline business, data, application and technology architecture are described. This is explained in detail in the TOGAF-ADM. In the Ecosystem Architecture Framework is only briefly mentioned (step 1.1). Secondly, the baseline Collaboration Architecture is developed. This step is optional as it is possible that the organization does not yet have a hybrid integration platform (step 1.2). A commonly used approach to develop the architecture is using Block Function Diagrams and ArchiMate. These are modeling techniques that help to represent the architecture of a system.

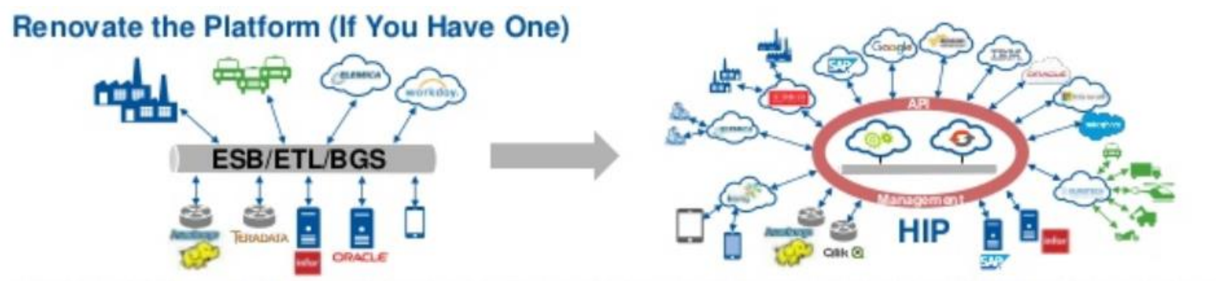


Figure 39: Example of Transformation (Gartner, 2018)

Step 2: Explore Target Collaboration Architecture

In the second step is focused on exploring how the target Collaboration Architecture should look like based on the Statement of Ecosystem Architecture Work. This statement represents the Architecture Requirements from the Ecosystem Architecture Vision. Based on this requirements the Hybrid Integration Platform (HIP) capabilities should be selected, such as an orchestration mechanism, security Mechanism, API Gateway, Publishing Tools, Developer portal, Reporting and Analytics, and Monetization. Gartner (2018) provided a framework of the potential capabilities that a HIP can incorporate (Figure 40). It is important that the organization explores what capabilities are required to mitigate the adoption barriers and to enable the real-time integration of external BOLD (step 2.1).

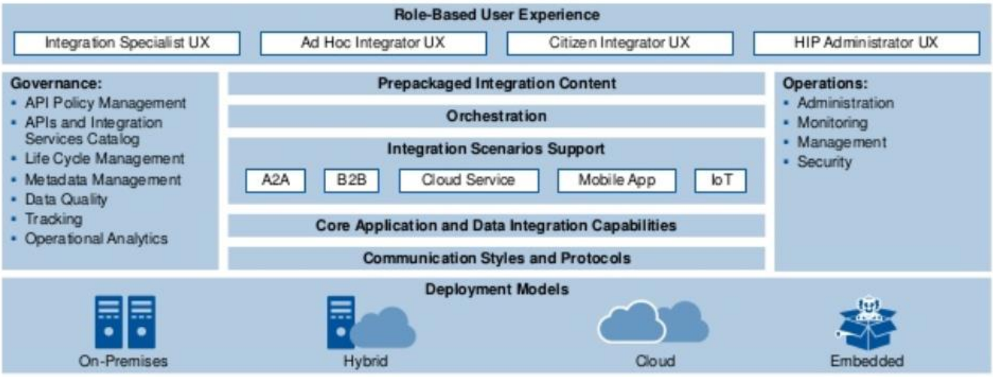


Figure 40: Hybrid Integration Platform Capabilities (Gartner, 2018)

Secondly, based on the required capabilities the organization should select a design pattern that has incorporated these capabilities. There are four design patterns that are most used by organizations (Gartner, 2018) (Figure 41):

I. **Classic, on-premises integration platforms** (ESBs, data integration tools, B2B gateway software): provide a rich set of core integration functionalities (such as transformation, mapping, adapters and orchestration), enable a range of integration patterns (such as application, data, process and B2B) and are perfectly suitable for an integration specialist constituency. However, they have limited support for cloud, mobile and IoT endpoints, are often unsuitable for a hybrid deployment model, and cannot really be used by ad hoc or citizen integrators

II. **API management platforms** implement key governance features and, by definition, manage access to any endpoint that exposes an API (whether on-premises, in the cloud, mobile or IoT). They usually support hybrid deployments and are a good fit for the requirements of ad hoc integrators and integration specialists. However, they provide only rudimentary core integration capabilities and aren't suitable for citizen integrators

III. **Integration platform as a service (iPaaS) offerings** provide an increasingly wide-ranging set of core integration capabilities, enable a broad variety of integration patterns, support hybrid deployment models and are, in most cases, effectively designed for ad hoc integrators. However, they have limited support for on-premises endpoints (for example, second-tier or industry-specific packaged applications) and, in most cases, support citizen integrators only partially

IV. **Integration software as a service (iSaaS) offerings** are wholly focused on citizen integrators and, to a lesser extent, ad hoc integrators. However, they typically provide only very basic core integration capabilities, don't support on-premises endpoints and, in most instances, don't enable on-premises or hybrid deployment models

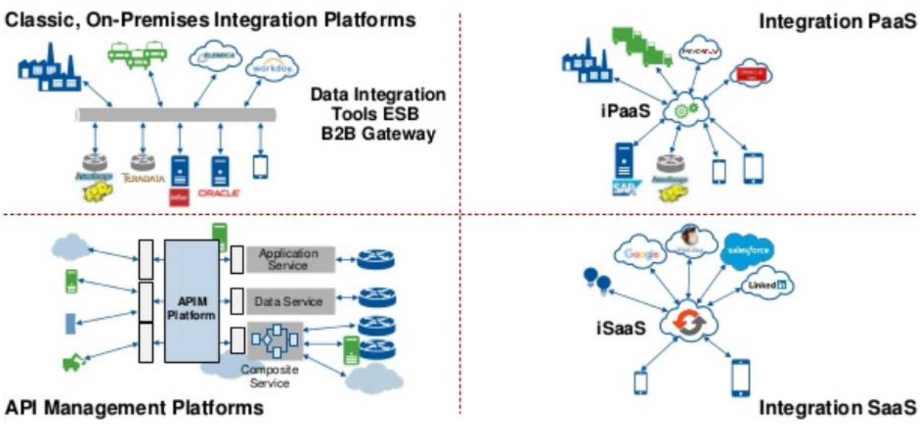


Figure 41: Hybrid Integration Platforms (Gartner, 2018)

Figure 42 presents an overview of the HIP capabilities and the design pattern that has these capabilities incorporated in its design. All four design patterns have advantages and disadvantages. It depends on the capabilities that an organization requires, which design pattern is chosen. It is not possible to combine design patterns. Hence, now no solution is available that combines all capabilities. For the integration of external BOLD, an API Management Platform seems most suitable and is recommended as it requires mainly governance and core application and data integration capabilities (Step 2.2).

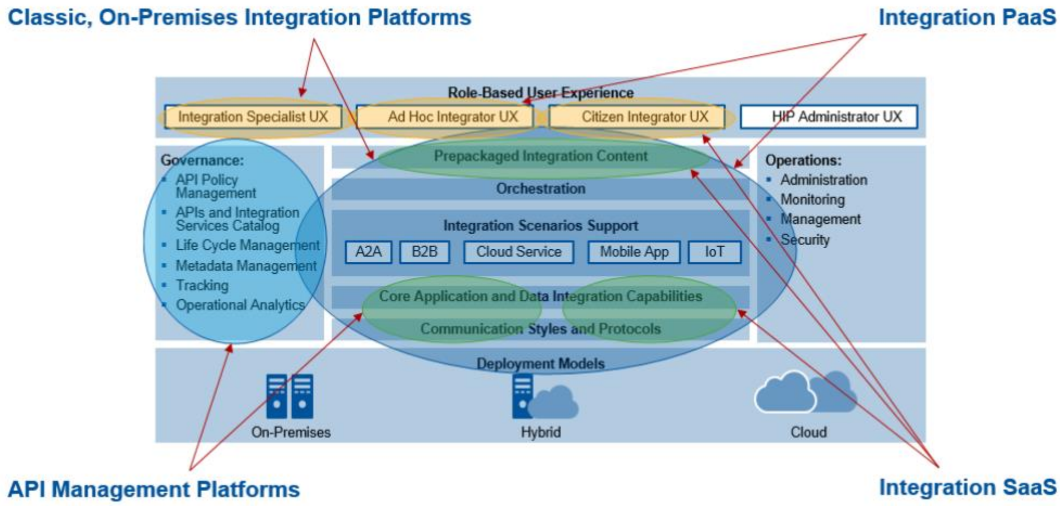


Figure 42: Design patterns and its HIP capabilities (Gartner, 2018)

Step 3: Develop Target Ecosystem Architecture

The third step represents the development of the Target Ecosystem Architecture Framework. First, the target business, data, application and technology architecture that are developed through the TOGAF-ADM are briefly described as the collaboration layer must be aligned with the collaboration architecture (step 3.1). Secondly, a target Collaboration Architecture is developed. The Target Collaboration Architecture incorporates the HIP design pattern that is selected that has the required HIP capabilities identified (step 3.2).

Step 4: Perform Design Gap Analysis

Comparing the baseline architecture and the target architecture with each other helps to identify the required transition of the organization. Step 4 represents the gap analysis that should be performed to identify the differences between the 'as is' situation and the 'to be' situation. These differences can be translated in a roadmap for transition with incremental steps that are required. This roadmap can be used by the organization to implement the transition and to develop a system that is capable of the real-time integration of external BOLD with internal company data.

Output: Ecosystem Architecture

The output is the target Ecosystem Architecture that is technically capable of integrating external BOLD with internal company data.

5.5.3. Phase C: Making the Ecosystem Architecture Work

In the previous two phases, a Shared Ecosystem Vision and a target Ecosystem Architecture are developed. However, as mentioned before the human factor is causing most of the adoption barriers of BOLD. This means that having a collaboration architecture that is technically capable of integrating external BOLD in real-time does not guarantee success. When people do not have a uniform way of working the Ecosystem Architecture, adoption barriers will be still present and not function optimal. Therefore, the third phase provides guidelines for making the Ecosystem Architecture work.

5.5.3.1. Ecosystem Implementation

In phase C1, guidelines are provided to create an inter-organizational model, consisting of shared artifacts that shape a common way of working. A common way of working ensures that the Target Ecosystem Architecture (Output of Phase B1) can function optimally and no adoption barriers will occur. If a design is not used in the way where it was designed for, new problems occur. For example, if you buy a Nespresso machine and you put Senseo pats in, the machine does not function optimal and it does not deliver the intended product. This phase represents **RQ5** '*The Enterprise Architecture Framework must create a uniform way of working within the organization and ecosystem'*. Through explorative interviews is derived that a uniform way of working can be enabled by developing the following shared artifacts: Uniform Standards; a Common Data Model (One data model; unified schema and semantics; building on existing efforts); Uniform Meta-data standards (technical documentation about APIs); An independent API Policy; and Institutions (action rules) to shape/implement this way of working.

Figure 43 presents guidelines for the creation of an inter-organizational model that ensures the Ecosystem Architecture is used in the right way. Phase B consists of four steps. First, the frameworks helps defining uniform standards that should be used by all ecosystem partners (Step 1). Secondly, it focuses on developing a common data model (Step 2). Thirdly, it provides guidance in the creation of an independent API policy that can be followed by any organization to develop APIs (Step 3). Finally, institutions are created between the actors to make sure data and the architecture stays sustainable (Step 4). The steps are explained in detail below.

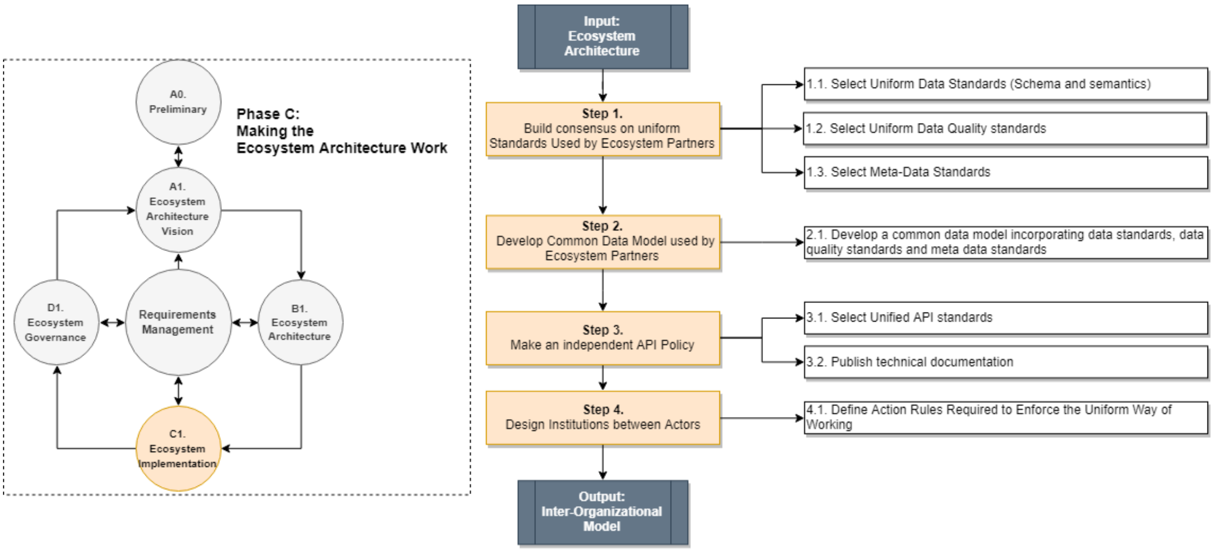


Figure 43: Ecosystem Implementation Guidelines

Step 1: Build consensus on uniform standards used by Ecosystem Partners

The first step represents the design of uniform standards used by the Ecosystem Partners. It consists of three guidelines that should be fulfilled. First, uniform data standards should be selected by the ecosystem. A wide variety of data standards exist, so it is smart to build on existing efforts instead of creating even more complexity. This step is focused on creating a consensus, which results in the selection of one data standard type, such as HL7 standards or CDISC standards. What all data standards have in common is that a specific schema and semantics are chosen. When these two aspects are different for each data provider, it is very difficult to integrate the data (step 1.1). Secondly, the ecosystem partners must agree on data quality standards to mitigate the adoption barriers related to information quality, which are completeness, uniqueness, timeliness, validity, accuracy and consistency (step 1.2). Thirdly, the ecosystem partners must agree on the meta-data standards. Meta-data is defined as 'data about data' and is used to understand the actual data. However, often meta-data is lacking, or no clear standards are used. Many meta-data standards exist, such as ISO, SGML and XML (step 1.3).

Step 2: Develop a Common Data Model

The second step focuses on the development of a common data model that has incorporated the data standards (schemes and semantics), data quality standards and meta-data standards selected in step 1. The common data model simplifies the process for the real-time integration of an external BOLD source, as the model enables consistency of data and its meaning across applications in which the data is stored. Without a common data model, every time a new external source is integrated custom implementation is required which is expensive and time extensive. Besides that, often repeated effort is required. A common data model in the ecosystem can resolve this and enable the real-time integration of external BOLD (step 2.1).

Step 3: Select an Independent API Policy

Different API policies exist. An API policy can be compared with a module that implements a specific function. Technically an API policy is an XML-formatted configuration file. It enables organizations to access an API easily. It is important that all ecosystem partners (data providers) offer APIs with an independent API policy which allows other organizations to easily access the API and integrate in in the hybrid integration platform (API management platform). An independent API policy consists of all the technical information required for organizations to easily integrate the API in their system. It also enables organizations to program behavior without writing any code. It consists of two core elements (API standards and technical documentation about the APIs) (step 3.1 and 3.2).

Step 4: Create Institutions between Actors

As discussed in the requirement definition phase, institutions are required to create data sustainability and to enforce a common way of working. Therefore, action rules should be designed that shape the common way of working and to prevent the presence of new adoption barriers. Action rules should be made to ensure the working of the different shared artifacts that are developed in step 1, 2 and 3 of this phase. Moreover, a business partner must comply with these institutions to be part of the strategic coalitions. An example can be that institutions can be that each data provider is obliged to offer APIs to its data applications together with an independent API policy that enables other organizations easily to integrate the APIs and access the data.

Output: Inter-organizational model

The output of this phase is an Inter-organizational model that consists of the shared artifacts developed in the four phases. It ensures that there is consensus about clear data standards, a common data model, independent APIs and institutions between actors to shape the uniform way of working. This is the key model that ensures that the technically capable Ecosystem Architecture Framework is used in the right way to create the ultimate benefit for all partners.

5.5.4. Phase D: Keep the Architecture Running

At this point, there is a consensus about how the system should look like; an architecture is developed that is technically capable of integrating external BOLD with internal company data; and an inter-organizational model is developed that ensures the architecture is used properly. However, a system that is used in a dynamic environment that is constantly changing requires updating. It is important that the shared artifacts are always adapting to the environment to ensure the Ecosystem Architecture Vision is represented at all time (including the metrics). Therefore, Phase D focuses on keeping the architecture running.

5.5.4.1. D1. Ecosystem Governance

Phase D1 focuses on creating a governance mechanism that keeps the Ecosystem Architecture running, but even more important keeps the Ecosystem Architecture work properly and aligned with the Ecosystem Architecture Vision. The Ecosystem Governance Phase deals with the **RQ6** *'The Framework must facilitate the development of an architecture that can grow organically, that is scalable, maintainable and flexible towards its ecosystem; ensures continuity; and allows for new connections*. On a high level, this is what the Ecosystem Governance Phase is about. The phase ensures that the shared artifacts and the Ecosystem Architecture adapt to changes of the environment. This is required to make sure that the architecture is compliant with the ecosystem architecture vision and delivers the business value that is demanded by the ecosystem. Changes of the environment can be caused by different factors related to the demand of customers, markets, industries, opportunities, competitors, regulators and investors**.** To make the architecture scalable, maintainable, flexible and continuous, a governance mechanism that manages all these aspects is developed in this phase. Figure 44 presents the guidelines for Ecosystem Governance Phase.

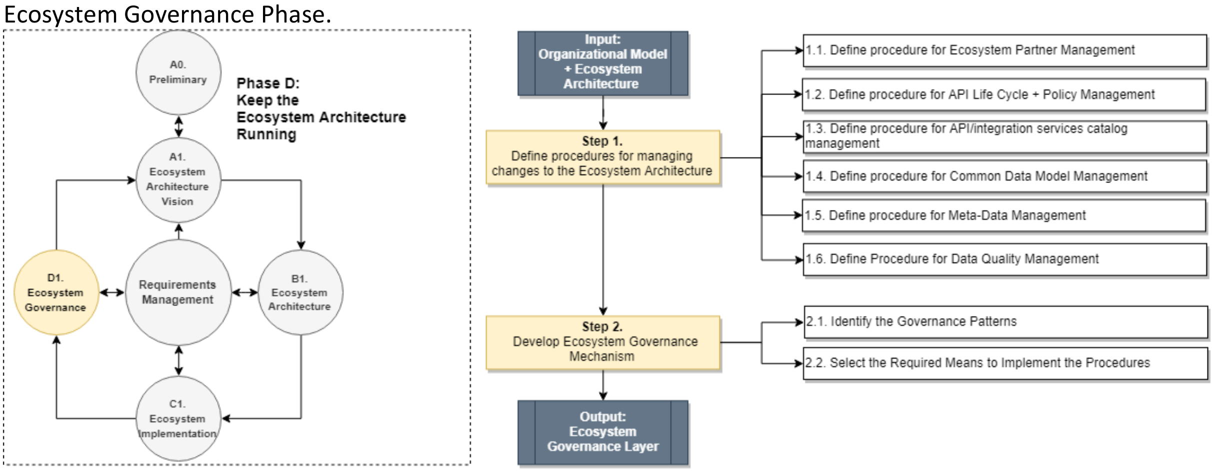


Figure 44: Guidelines for Ecosystem Governance

Step 1: Develop procedures for managing changes to the Ecosystem Architecture

To ensure that the architecture keeps working, procedures should be developed for how the Ecosystem Architecture should adapt to changes to the different shared artifacts. A procedure is a set of processes that should be performed to manage change. The governance mechanism ensures that the Ecosystem Architecture is always adapting to changes to the strategy, metrics, governance, organizations and roles, lifecycle, and technology. Changes can affect any of the following shared artifacts. First, the Ecosystem Architecture should be flexible for new connections with new ecosystem partners. Moreover, it should be possible for non-ecosystem partners to connect if they adopt the same inter-organizational model and integrate through the collaboration architecture. This should be facilitated through a developed procedure that deals with new or leaving Ecosystem Partners (step 1.1). Secondly, a procedure for changes to the independent API policy should be updated, so that an API can be integrated and accessed without problems. It is important that the procedures consider the life cycle management of APIs (step 1.2). Thirdly, the API catalog with descriptions should managed. Procedures must be developed that facilitate this. The API catalog provides an overview of all the APIs provided in the ecosystem for integration or sharing. Any changes in the API catalog and descriptions must be managed (step 1.3). Fourthly, any changes in the common data model and its incorporated meta-data standards, data quality standards and data standards (Schemas and semantics) should be managed. Procedures should be developed of how this is managed by the governance mechanism (step 1.4). Fifthly, meta-data should be managed by the governance mechanism. Metadata management is the administration of data that describes other data. It involves establishing policies and processes that ensure information can be integrated, accessed, shared, linked, analyzed and maintained to best effect across the organization. Procedures should be developed that govern this (step 1.5). lastly, Data Quality should be managed. Processes and policies should be developed to maintain and improve the data quality of the data that is shared within the ecosystem (step 1.6).

Step 2: Develop Ecosystem Governance Mechanism

To ensure the procedures with is associated processes and policies are managed, a governance mechanism is developed. First, the governance patterns are explored. According to Gartner (2018) and the expert that are interviewed there are different HIP governance structures exist: Centralized governance, Shared governance and Federated governance (Figure 45). Dependent of the situation of the organization a suitable HIP governance pattern should be selected. This thesis recommends the use of a federated HIP governance structure, as more than one ecosystem partner wants to integrate APIs of other Ecosystem partners (step 2.1). This means that the governance is done decentralized, and that one federated government body is developed to agree on things.

Secondly, the means are selected to perform the procedures that are defined in step 1 (step 2.2). One of the technologies that is used to manage all the different abovementioned is an API Management Platform. Therefore, this Ecosystem Architecture recommends selecting the API Management platform as it facilitates the governance of these aspects (Figure 42).

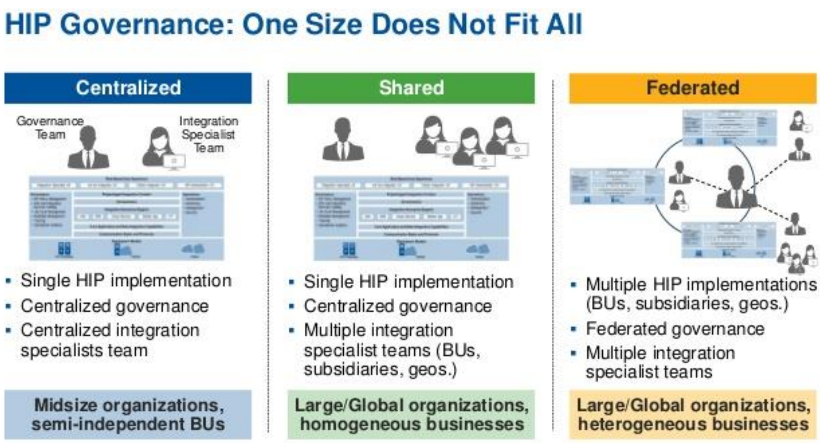


Figure 45: HIP Governance (Gartner, 2018)

5.6. Summary Chapter 5 (Answer RQ4)

The fourth sub-research question ***'RQ4. What does an Extended Enterprise Architecture Framework that enables the real-time integration of external Big and Open Linked Data with internal company data for decision-making look like?'*** focused on the development of a design artefact that satisfies the identified design requirements of RQ3. Through a morphological chart, brainstorm sessions, and a best-of-class chart the design artefact is developed based on the design requirements. The rationale for this research question is that it provides a new design artifact that overcome the research problem and it forms an addition to the current knowledge base on EA.

This chapter presents the design of the 'Ecosystem Architecture Framework', which is an Extended Enterprise Architecture that enables the real-time adoption of BOLD (Figure 35). It cannot only be used to align a new strategy and its implementation across the enterprise, but also between organizations. The Ecosystem Architecture Framework is structured in four different phases: Phase A: Getting the Ecosystem Involved and Committed; Phase B: Getting the Ecosystem Architecture right; Phase C: Making the Ecosystem Work; and Phase D: Keep the Architecture Running.

In Phase A, the need for the Ecosystem is assessed by checking the presence of adoption barriers of BOLD and identified strategic alliances are required to get the business ecosystem on board. This results in a proposal for an Ecosystem Architecture. Secondly, the committed organizations are involved during the design of the Ecosystem Architecture. All stakeholders concerned with the real-time adoption of BOLD are involved to create a shared Ecosystem Architecture Vision. This vision is translated into Architecture requirements.

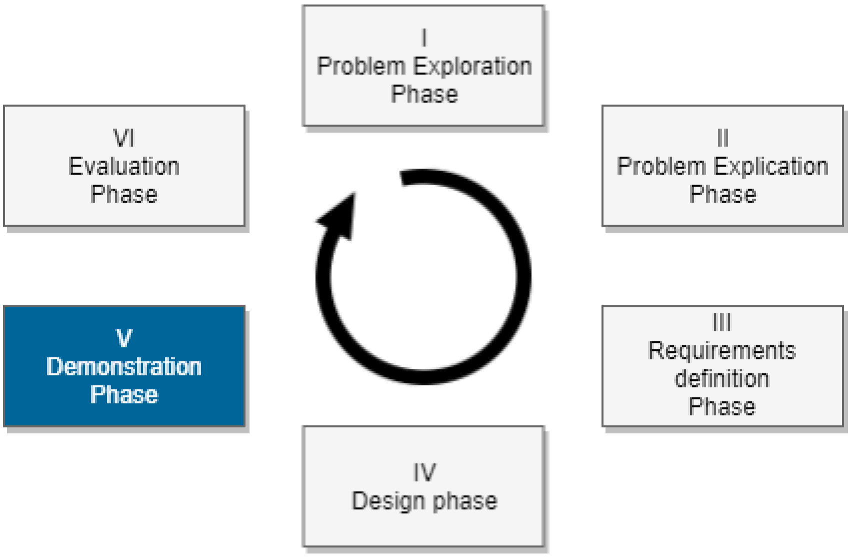
In Phase B, the Architecture Requirements are satisfied through the creation of a suitable collaboration architecture. The collaboration architecture is a hybrid integration platform in which data can be integrated and shared, different applications can be linked together, makes possible collaboration between applications, stimulates interoperability between different operating systems and orchestrates the applications by a single interface. Different types of hybrid integration platforms are possible, but the API Management Platform is recommended by the Ecosystem Architecture Framework.

Phase C helps to make the Architecture working by developing an inter-organizational model consisting of shared artefacts that ensure a common way of working is established: Uniform standards (e.g. Data standards and Meta-data standards); Institutions between organizations of the ecosystem; A Common Data Model used by all committed organizations; and an independent API policy. Having the same way of working ensures the organizations are better aligned with each other and it is easier to develop APIs that can be accessed for the real-time integration of external BOLD through the Hybrid Integration Platform.

Phase D ensures that the established way of working and Collaboration Architecture adapt to changes from the environment. It ensures that the Ecosystem Architecture optimally fulfils the business requirements and is aligned at all time. It ensures the Ecosystem partners are managed, so the architecture is updated when a new partner joins or leaves the strategic alliance; the API policy is updated according to the needs of the ecosystem; the APIs catalog is updated; the data quality and meta-data are managed; and the institutions are enforced. The governance mechanism ensures that the Ecosystem Architecture is always adapting to changes to the strategy, metrics, governance, organizations and roles, lifecycle, and technology.

The Ecosystem Architecture Framework forms parts of a solution to overcome the research problem and can be used on the TOGAF-ADM, which ensures the 11 capabilities identified in Chapter 3 are incorporated. The extension provides the additional 6 capabilities.

V. Demonstration Phase



6. Evaluation

This chapter presents the evaluation of the Ecosystem Architecture Framework that should overcome the research problem. This chapter answers the research question 'How can the Ecosystem Architecture Framework be evaluated?'. A design evaluation protocol is developed that provides a process to evaluate the design artefact in a structured and academically supported way. Three evaluative expert interviews are conducted in order to validate and evaluate the design. This chapter shows the potential improvements for the design that are transformed into revised and additional requirements that are more suitable to create a design that enables the real-time adoption of BOLD for decision-making. Section 6.1 presents an overview of the chapter. Section 6.2. presents a method for evaluation. Section 6.3. presents the feedback collected through the interviews. Section 6.4. presents the key findings of the interviews. Section 6.5. presents a summary and answer to the RQ of chapter 6.

6.1. Design Evaluation Overview

As discussed in design phase, the Ecosystem Architecture Framework is designed through an iterative approach. In order to collect additional and revised requirements for an improved design, the design artefact was reviewed and evaluated by three different Enterprise Architects from different organizations. This chapter presents the evaluation phase of the design cycle by Johannesson and Perjons (2014). The Design Evaluation chapter is structured in three parts. First, the method of evaluation that was used to collect the potential improvements of the design is defined. The method of evaluation is based on the evaluation methodologies published by Verschuren and Hartog (2005). The second part of this chapter discusses the feedback of the expert evaluations that are conducted. The key findings of the interviews are discussed and transformed into potential improvements for the design. Finally, the key findings are translated into potential improvements for a new design and explained what improvements are processed.

6.2. Method for Evaluation

6.2.1. Plan Evaluation

As mentioned, the evaluation method is based on the evaluation methodologies by Verschuren and Hartog (2005). The authors state that there are three methodologies for evaluation: Evaluation of the plan; Evaluation of the process; and Evaluation of the Product. The 'Plan Evaluation' is about assessing the quality and suitability of the design and whether the design meets all relevant design elements that are used as input for the design, such as the design requirements, the design assumptions and the design purpose. The 'Process Evaluation' is about the implementation of the design on a real-life case. This evaluation is in general focused on learning by doing, which results in the detection of shortcomings of the design process. The third Evaluation methodology 'Product Evaluation' focuses on evaluating the outcomes and effects of the design artefact after it is implemented. It helps to identify the consequences of the design (Verschuren & Hartog, 2005).

This thesis does not include the implementation of the design, but only focusses on the initial design of the Ecosystem Architecture Framework. The thesis only uses an example / showcase to show the working of the Ecosystem Architecture Framework. Hence, only the 'Plan Evaluation' methodology is used to empirically, logically and ethically evaluating of the quality and suitability of all design elements that are used to create the design, including the design objectives and goals, design requirements, and the design specifications (Verschuren & Hartog, 2005). This is reflected and represented in the design protocol is mainly based on the 'plan evaluation' methodology (Appendix E1: Evaluative Interview Protocol).

6.2.2. Design protocol

The design protocol is structured into 8 parts, each representing one of the design elements, including Design Purpose; Design Requirements; Overall Impression Design Artefact; Phase A of the design artefact; Phase B of the design artefact; Phase C of the design artefact; Phase D of the design artefact; and Overall Usability Design Artefact. Reflecting on the 'Plan Evaluation' methodology by Verschuren and Hartog (2005), the Design Purpose section is evaluating on the design objectives and goals; the Design requirement section is evaluating the design requirements that are used to create the design artefact; and the Overall Design section together with the specific Design Phases sections are evaluating the Design Specification that explains the design artefact in detail.

Figure 25 presents the expert evaluation form that was used to assess the design on the different design elements. Each design element was assessed on one or more evaluation criteria. The different criteria used for the evaluation are 'clarity, feasibility, completeness, structure and usability' to assess the quality and suitability of the design elements of the Ecosystem Architecture Framework. The chosen evaluation criteria are suitable for evaluating the quality and suitability of Enterprise Architecture Framework (Matthes et al., 2008). To assess the design elements on these criteria, statements are made about the design elements of the Ecosystem Architecture Framework. By means of a Likert scale the interviewee can give a score that expresses to what extent the expert agrees with the statement about the quality and suitability of the design elements. A Likert scale is a useful tool to compare the evaluations of different experts (Allen & Seaman, 2007). Also, an 'open question' is used in order to gather specific feedback and remarks on how the design can be improved.

Table 21: Expert Evaluation Form

Design Element	Evaluation Criteria	Statement	Score	General Feedback/ Remarks
1. Design Purpose	*Clarity*	*S1 'The purpose for the Design of the Ecosystem Architecture Framework to enable real-time integration of Big and Open Linked Data is clear'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
2. Design Requirements	*Clarity*	*S2 'The design requirements that are collected are clear.'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Feasibility	*S3. 'The design requirements that are collected are feasible.'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Completeness	*S4.' The design Requirements collected result in a design that fully solves the research problem.'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
3. Overall Impression Design Artefact	*Clarity*	*S5. 'The Ecosystem Architecture is clear and understandable.'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Completeness	*S6. 'The Ecosystem Architecture is complete and does not lacking important other phases.'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
4. Phase A: Getting the Ecosystem Committed and Involved	*Clarity*	*S7. 'The guidelines provided in phase A are understandable'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Structure	*S8. 'The guidelines of Phase A are structured in a logical way'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
5. Phase B: Getting the Ecosystem Architecture Right	*Clarity*	*S9. 'The guidelines provided in phase B are understandable'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Structure	*S10. 'The guidelines of Phase B are structured in a logical way'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
6. Phase C: Making the Architecture Work	*Clarity*	*S11. 'The guidelines provided in phase C are understandable'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Structure	*S12. 'The guidelines of Phase C are structured in a logical way'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
7. Phase D: Keep the architecture Running	*Clarity*	*S13. 'The guidelines provided in phase D are understandable'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
	Structure	*S14. 'The guidelines of Phase D are structured in a logical way'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	
8. Overall Usability Design Artefact	*Usability*	*S15. 'The Ecosystem Architecture Framework is easy to use'*	*'Strongly disagree – agree – neutral – agree – strongly agree'*	

6.2.3. Interviewees

In total, three different Enterprise Architects have evaluated the quality and suitability of the Ecosystem Architecture Framework. The rationale for choosing Enterprise Architects as evaluators is because they are the target end-users of the design artefact. Table 22 presents an overview of the stakeholders that have evaluated the design artefact. One of the interviewees was interviewed during the explorative expert interviews, which ensures validation whether the qualitative data analysis resulted in the correct design elements. The other two interviewees are enterprise architects that have a lot of knowledge about using Enterprise Architecture and Data architectures. The rationale for selecting two independent stakeholders for evaluation was in order to prevent goal reasoning. It is important that interviewees are exposed with the design artefact for the first time without having background information, because this will also be the case when the Ecosystem Architecture will be implemented in real-life.

The interviews were conducted physically and virtually. The Interview protocol was sent in advance to give the interviewee the chance to prepare for the interview. The evaluations resulted in interesting insights that are discussed in the next section.

Table 22: Evaluative Expert Interviewees

ID	Stakeholder Role	Organization	# Years Experience
S10	Enterprise Architect	One of the largest Consultancy firms	15
S11	Enterprise Architect	National Public Enforcement organization	30+
S12	Enterprise Architect	Consultancy Firm / Technical University	20

6.3. Feedback Expert Interviews

During the evaluative expert interviews, first an introduction was provided to the research subject. The interviewees already read the interview protocol in advance, which helped to discuss the important aspects right away. The interviewees already filled in the Evaluation form before the start of the interview. The interview itself was focused on gathering specific feedback on what the weak points of the design elements are and what the potential aspects for improvements are.

In this section the feedback on the different design elements is discussed and summarized. For each design element: the average score; some quotes that substantiate the score; and the key findings of the feedback are discussed. This section discussed the three important design elements defined by Verschuren and Hartog (2005): Design Objectives; Design Requirements; and Design Specifications.

6.3.1. Design Objectives

The first design element that is evaluated by the different interviewees is the 'Design Purpose'. This represents the 'Design Objectives' of to the 'Plan Evaluation' methodology by Verschuren and Hartog (2005). The knowledge gap was discussed, the problem was explicated, and the findings of the desk studies were discussed. Summarized, the Enterprise Architects agree that there is a need for an extension to the current foundation of Enterprise Architecture Frameworks to enable a strategy that uses external BOLD for real-time data-driven decision making. On the statement about the clarity of the design purpose *S1. 'The purpose for the design of the Ecosystem Architecture Framework to enable real-time integration of Big and Open Linked Data with internal company data for decision-making is clear'* the interviewees tend to agree (Figure 46).

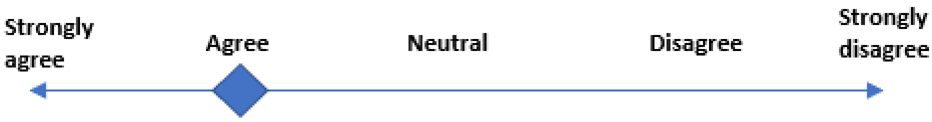


Figure 46: Average score 'Clarity' of design element 1. 'Design Purpose'

One of the interviewees states that the purpose for the design of the Ecosystem Architecture Framework can be described as following:

> *It looks like what you're describing ultimately comes down to general transformation from a system of record to a system of engagement and having your ecosystem reflect that. Modern organizations should embrace the system of engagement, this including the company's ecosystem as part of their enterprise information framework. It is a key part of the paradigm shift that comes with an organization's digital transformation. –* ***Interviewee S10***

The interviewee explains that the design purpose of the extended EAF reflects the need of organizations for guidance through this digital transformation. Organizations that keep working with a system of records are limited in their possibilities, while a system of engagement broadens the opportunities of an organizations for data integration.

Another interviewee states that this can be facilitated through Extended Enterprise Architecture Framework that provides guidance for the development of an Enterprise Architecture that aligns the needs of the ecosystem. The interviewee emphasizes that is not smart to use a platform in the middle approach (a joint architecture used by the whole ecosystem). These are complex adaptive systems that evolve based on triggers (e.g. legislation, new entrants and business models) and value streams and, therefore, bring many extra challenges to deal with. The interviewee substantiates this in the following quote:

> *After identifying the need for collaboration, is a shared architecture the mean that you need? I do not believe in an aligned architecture of the whole ecosystem. I believe in certain rules/standards that are used within the ecosystem and that you can classify this in the business/data/application/technology architecture from the perspective of the own Enterprise Architecture. So, you what you need is an Extended Enterprise Architecture Framework that is aligned with the ecosystem through an inter-organizational model –* ***Interviewee S12***

Although, the Ecosystem Architecture Framework provides guidance for the transition of a system of records to a system of engagement of the ecosystem, an interviewee states that it complicates the process for enterprise architects even more:

> *Be aware of the risk that TOGAF itself is already very complex and broad and that extending it can make it even more complex. Your extension on top of TOGAF makes it more complex for both understanding and applying. But in this framework, you need TOGAF, because without it your framework is like a loaded gun without bullets. Discuss that this is a challenge. –* ***Interviewee S12***

Besides this, the interviewees all agree the design purpose is highly relevant and has potential to mitigate the adoption barriers for integrating external BOLD with internal company data in real-time. In order to identify the potential improvements of the design, the collected feedback is summarized in the following key findings:

I. An extended Enterprise Architecture is needed that facilitates the transition from a system of records towards a system of engagement. This will mitigate the adoption barriers that will be addressed.

II. Be aware of the risk that TOGAF itself is already very complex and broad and that extending it can make it even more complex. Your extension on top of TOGAF makes it more complex for both understanding and applying.

III. You do not need a joint architecture with the ecosystem, but organizations need an Extended Enterprise Architecture Framework that is aligned with the ecosystem through an inter-organizational model.

IV. All interviewees state that the research problem is highly relevant and very interesting.

6.3.2. Design Requirements

The second design element that is evaluated by the different interviewees are the 'Design Requirements' discussing the eponymous element of the 'Plan Evaluation' methodology by Verschuren and Hartog (2005). During the interview the key findings of the explorative expert interviews and the literature review are discussed.

Subsequently, there is explained how the requirements are derived from the key findings. The 'Design Requirements' are evaluated on Clarity, Feasibility and Completeness.

Summarized, the Enterprise Architects all agree on the statement about the clarity of the design requirements (Figure 47); partly agree on the feasibility of the design requirements (Figure 48); and the opinions are divided about the completeness of the design requirements (Figure 49).

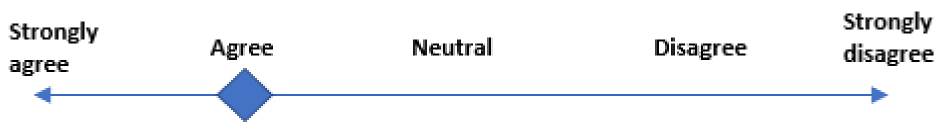


Figure 47: Average score 'Clarity' of design element 2. 'Design Requirements'

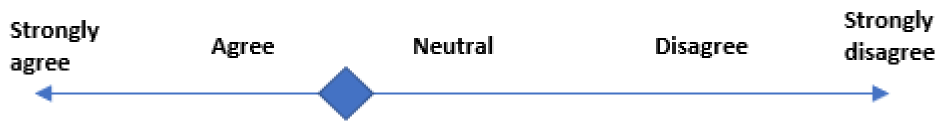


Figure 48: Average score 'Feasibility' of design element 2. 'Design Requirements'

Figure 49: Average score 'Completeness' of design element 2. 'Design Requirements'

One of the interviewees states that the requirements should make clearer that the requirements are focused on information ingestion of external data providers, and not also data-sharing. This requires additional requirements:

> *Make sure that the requirements reflect that a platform is required that facilitates the use of services in which third parties are providing ('brokering') information and the organization acts as a client. Make sure that the requirements reflect that. Try to scope this clearly, because the impact of the ecosystem may differ. **– Interviewee S10***

Another interviewee explains that some of the requirements are defined very strong, which makes it more difficult to satisfy the design requirements (less feasible). Interviewee S12 substantiates this:

> *Different requirements are not very feasible. For instance, it is impossible to get 'all' stakeholders involved. Be careful with how the requirements are defined. It is better to say: The Ecosystem Architecture Framework must make it possible for 'all' stakeholders to be involved. **– Interviewee S12***

About the completeness of the requirements, the opinions are divided. One interviewee agrees, one interviewee is neutral, while another interviewee disagrees. The interviewees state the requirements should better reflect it is only for an extension. Otherwise, more requirements should be collected that cover the design of an EAF from scratch.

> *There is a gap between the design requirements and the EAF design. You need more concepts, methods, tools, procedures, training to develop an effective EAF from scratch. Put more focus on the extension or collect more requirements. – **Interviewee S11***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. There is confusion whether the requirements are for an Enterprise Architecture Framework or for an Ecosystem Architecture Framework. The requirements are suitable for an Ecosystem Architecture or for an extension to the Enterprise Architecture Framework. However, the requirements are not suitable for building an entire Enterprise Architecture Framework from scratch.
II. Some of the requirements are defined too strong, which makes it less feasible.
III. The design requirements will result in a design that forms parts of a solution, but more requirements should be satisfied to ensure the whole design problem is removed. Collecting all the requirements is very difficult.
IV. The requirements should reflect better that the collaboration layer is a platform in which third parties integrate their services through e.g. APIs.

6.3.3. Design Specifications

The third design element that is evaluated by the different interviewees are the 'Design Specifications' discussing the overall design artefact and the specific phases of the design artefact. The design specifications is the third category of design elements that should be evaluated according to the 'Plan Evaluation' methodology by Verschuren and Hartog (2005). During the interview, the first version of the design was shown and explained in detail. Subsequently, the different phases and guidelines of are explained in detail. An explanation is given of what parts of the design are covering the requirements. The 'Design Specifications' are evaluated on Clarity, Structure and Completeness. This section discusses the evaluation of the overall design first, and subsequently, the evaluations of the different phases are discussed.

6.3.3.1. Overall Impression Design Artefact

In general, the interviewees consider the overall impression of the design artefact as clear and understandable. But the interviewees think that not 'all' required aspects to build an Ecosystem Architecture are included. Figure 50 and Figure 51 present the score of the overall design artefact.

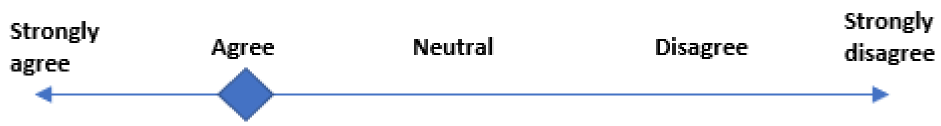


Figure 50: Average score of 'Clarity' of overall impression design artefact.

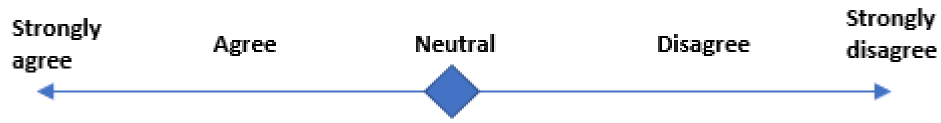


Figure 51: Average score of 'Completeness' of overall impression design artefact.

Interviewee S12 states that the Ecosystem Architecture Framework provides guidance for integrating external BOLD with internal company data concerns both a system of engagement and a system of record. The interviewee explains that the system of record cannot be treated in the same way as the system of record. He explains that you need to look at it as two circles:

First, you have the system of record is in the organization's circle of influence: The organization can exert complete control over the application landscape and data lifecycle.
Secondly, the system of engagement is in the circle of concern:
*The organization is concerned with the data but does not exert full control. – **Interviewee S12***

The overall design artifact must show how these different circles are represented in the framework, as both play a role for the real-time integration of external BOLD.

One of the interviewees thinks the overall design artifact looks complete, but explains it does not represent an important aspect:

*I miss how the security and privacy aspects are represented in the design artifact, as this is a major aspect that should be complied with when designing such Collaboration Architecture. – **Interviewee S11***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. Clarify that the Ecosystem Architecture Framework can be used to build an extension for the Enterprise Architecture to make it more suitable to align with a group of connected partners in the ecosystem, namely the Collaboration layer. Now, it looks it can be used to build a shared ecosystem architecture. However, that is too complex, and you can never be complete. Ecosystems are complex adaptive systems and hard to manage, govern and control when they are centralized. If you have 15 years of experience in practice you identify many more complexities that you cannot know by sitting behind the school desks.
II. Clarify how security and privacy aspects are guaranteed in the framework. This are important aspects to consider, especially for organizations of the public sector that work with Big Data Applications
III. Show how the Ecosystem Architecture Framework deals differently with systems that are in the circle of influence and with the circle of concern.
IV. Examples are needed, even for an experienced architect. Please note that even if you understand TOGAF, it is still difficult to 'just' add it to the EAF, both mentally and as a process.

6.3.3.2. Phase A: Getting the Ecosystem Involved and Committed

As depicted in Figure 52 and Figure 53, the interviewees agree with each other that phase A of the Ecosystem Architecture is clear and well-structured. Nevertheless, feedback was provided for potential improvements.

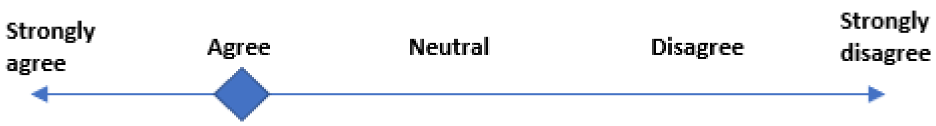


Figure 52: Average score of 'Clarity' of Phase A.

Figure 53: Average score of 'Structure' of Phase A.

Phase A.0: Preliminary Phase

For one of the interviewees it became not fully clear that Phase A is about creating a strategic alliance with the stakeholders/parties that are concerned with the integration of external BOLD. The interviewee explains that the Ecosystem Architecture Framework should achieve the following:

*You miss a phase in which strategic coalitions are build, for instance triggered by new legislation or business opportunities. You already mentioned it briefly as business architecture, but what in ecosystems is the most difficult is creating a business models for collaboration. You need a couple of parties that are required in an ecosystem, such as data users, data providers, data sources. These parties will only commit to the inter-organizational model (data model, data policies, standards, etc.) when there is a common business case, or it must be enforced by a new law. Without incentives or legal obligations, it will take years. Be aware of that. **– Interviewee S12***

Thus, it is important that this becomes clearer in the final design artifact. Another interviewee does see that this is incorporated in the Ecosystem Architecture, but creates awareness that there is also a chance that stakeholders do not want to collaborate:

> *Step 4 addresses the strategy of finding a way to commit other stakeholders from the ecosystem. However, it is only realistic to assume you cannot find commitment from all the stakeholders you need. My question is then: what do you do with them? Please refer to my earlier suggestion on third parties not in the 'circle of influence' needing a different approach.* ***– Interviewee S10***

Phase A.1: Ecosystem Architecture Vision

The interviewees liked the Ecosystem Architecture vision and have some suggestions for how the ecosystem architecture vision should be build:

> *The last step is focusing on finding consensus with the ecosystem of the company, including agree on terms, potential remuneration models, etc. So, it is important that the proposal for collaboration should become a sign contract in which the roles and responsibilities become clear and every stakeholders of the ecosystem knows what to do. At this stage, some high-level governance should already be envisioned. I would mention this is your detailed approach.* ***– Interviewee S10***

One interviewee refers to the 7 building blocks that are part of your strategic assessment: Vision – Strategy – Metrics – Governance – Organizations and roles – Lifecycle – Technology. These parts that should be discussed when developing the Architecture vision. Next to this the interviewee recommends revising Phase A to put more emphasis on specific guidelines.

> *The 'Document Ecosystem Architecture Vision' does not cover the importance of the sub-steps: Emphasize better the importance of building consensus and developing the architectural requirements. Each step is already focusing on documenting everything.* ***– Interviewee S10***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. Clarify in the Phase A.0 already focuses on forming a strategic coalition or collaboration.
II. Discuss what should be discussed in the architecture vision in more detail: Use the 7 building blocks for strategic assessment.
III. Make clear that the Initial Phase has two main objectives: Assess the need for an extended Enterprise Architecture and create commitment of ecosystem partners.
IV. Put more emphasis on the important guidelines, such as building consensus between the stakeholders and developing the architectural requirements.
V. The guidelines are a bit too high level. It is recommended to elaborate it a bit.

6.3.3.3. Phase B: Getting the Architecture right

As depicted in Figure 54 and Figure 55, The interviewees agree with each other that phase B of the Ecosystem Architecture is clear and well-structured. However, feedback was provided for potential improvements.

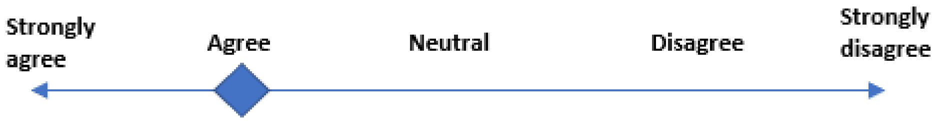


Figure 54: Average score of 'Clarity' of Phase B.

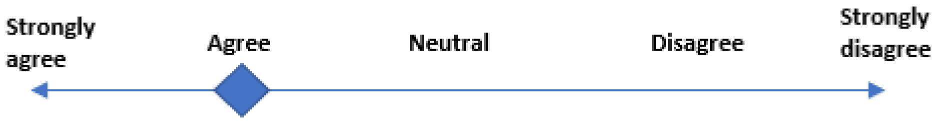


Figure 55: Average score of 'Structure' of Phase B.

One of the interviewees explains that integrating external BOLD into internal applications with company data is possible, but it should always go through a system that facilitates the whole enterprise, such as a data hub or a messaging system (e.g. Kafka). The interviewee explains that the collaboration layer should facilitates this:

> *I see where you want to go with this; essentially what you'll need is to do is extend your enterprise integration platform with Big and Open Linked Data and find a way to make this a seamless experience (so the applications can use the same way of working whether it concerns company data or ecosystem data) through a standardized enterprise architecture framework, a standardized interface (e.g. using URI's) and data model (such as pushed by ETSI-NGSI-LD for smart space projects).* ***– Interviewee S10***

The interviewees state that in general the steps make sense, but that it is recommended to explain everything in more detail. Also, the complexity can be decreased through changing the systems to giving design principles, rather than providing a fully complete method:

> *Complexity can be removed by giving design principles for building an Ecosystem Architecture instead of providing a 'complete' method to create an ecosystem architecture.* ***– Interviewee S12***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. Clarify that with ecosystem is meant that it is an extra layer added to the own Enterprise Architecture that makes it possible for external data providers to initially link their applications through integrations services such as APIs to enable real-time data integration. If you want to create a joint architecture, who is going to manage and govern the architecture. Ecosystems do not have governance.

II. Provide design principles instead of a complete EAF to ensure the framework does not conflict with completeness.

III. A bit too high level, e.g. what is a collaboration architecture? Why would parties collaborate?

6.3.3.4. Phase C: Making the Architecture work

Figure 56 and Figure 57 present the average score of the clarity and structure of Phase C of the design artefact. The interviewees agree with the statement that the design specifications of phase A are clear. About the structure of Phase C, the opinions are divided, while some interviewees agree and one of the interviewees tends to disagree.

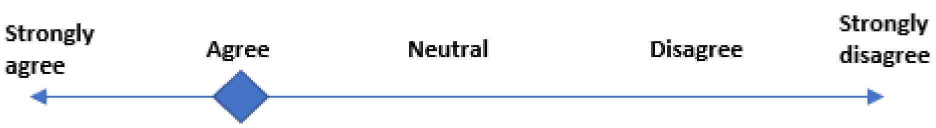


Figure 56: Average score 'Clarity' Phase C.

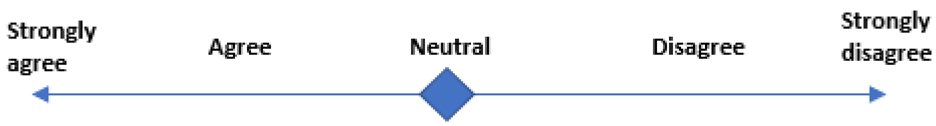


Figure 57: Average score 'Structure' Phase C.

Phase C is focusing on the creation of an inter-organizational model that strives for uniformity within the ecosystem. One of the interviewees responses to this:

> *Yes, one can strive for uniformity by agreeing on the use of the same standards or protocols. For example, information exchange. I think this is going too far although one can be ambitious. – **Interviewee S10***

Another interviewee thinks this is required to make the Architecture Work, and gives examples of inter-organizational artefacts that should be developed:

> *You must have shared or interorganizational artefacts: Whether it is a common data model; an API catalog with descriptions; or an administrative body. **– interviewee S12***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. The interviewees think 'the ecosystem implementation' phase is created from an interesting viewpoint and agree on the need for this phase.
II. Ensure that the inter-organizational artefacts that are required to make the architecture working are discussed.
III. Add where the API Policy, Policy descriptions, API stacks are created. Because now Phase D has API Policy Management, but in Phase C this is not described.
IV. Complex Adaptive Systems can be shaped with simple things, such as rules, collaboration agreements made with dominant players. Each ecosystem has dominant players: In this case this are the data source holders/ data providers, which can steer the way of working within the ecosystem.
V. The guidelines are a bit too high level. It is recommended to elaborate it a bit.

6.3.3.5. Phase D: Keeping the Architecture Running

The interviewees are satisfied with the clarity and the structure of Phase D (Figure 58; Figure 59). The interviewees have some interesting feedback to improve the design.

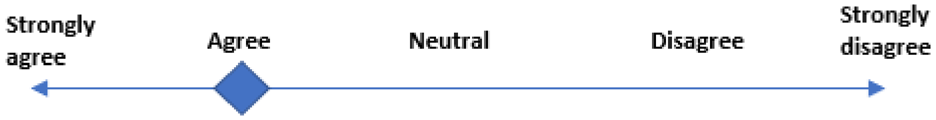


Figure 58: Average score of 'Clarity' of Phase D.

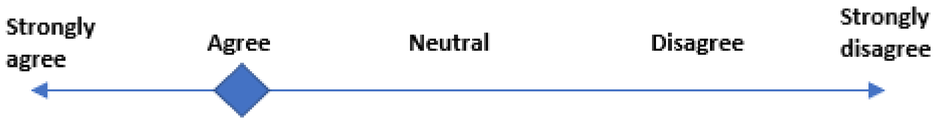


Figure 59: Average score of 'Structure' of Phase D.

Phase D focuses on keeping the architecture running and focuses on developing governance mechanisms that should ensure this. Interviewee S10 recommends investigating on what levels governance is required:

> *Try to look at exactly what to govern. Governance can exist at different levels and is tightly integrated with your organization, processes, etc. **– Interviewee S10***

The interviewees also give recommendations on how to govern:

> *You cannot create a central governance for the whole ecosystem, but you can create decentralized governance where you give advice as organization to other organizations – **Interviewee S12***

The interviewee gives as an example in the domain Healthcare: Many governance bodies exist, it is a network of governance bodies. How do these different governance bodies influence each other? Normally, through shared projects, shared domains or shared persons (part of both governance bodies). So, it is suggested to add a decentralized governance body that is part of the own organization that collaborates with the governance bodies of other organizations.

> *Discuss what decisions will be made? What are the shared artefacts? And what about legacy systems? How to create consensus? Change management within ecosystem is complex. I am happy with what you have until so far, because it shows you see the complexity of this.* ***– Interviewee S12***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. Clarify what type of governance is developed. Some ecosystems try to create centralized bodies, such as a community. However, this is difficult, because many decisions are made in all the decentralized bodies about the shared artefacts.
II. Discuss what decisions will be made? What are the shared artefacts? And what about legacy systems? How to create consensus? Change management within ecosystem is complex. I am happy with what you have until so far, because it shows you see the complexity of this.
III. How are you going to align the knowledge about the shared artefacts with the ecosystem?
IV. Discuss how to measure the metrics/KPIs. A bit too high level. What kind of KPIs do you suggest? What are the success indicators?

6.3.3.6. Usability

The last design element that is discussed is the usability of the design artefact. The interviewees are asked to response to the statement 'The Design Artefact is easy to use'. The interviewees responded divided to this statement. This resulted in a neutral average score.

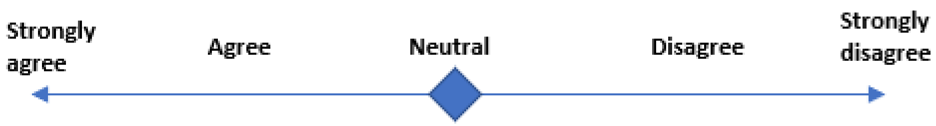


Figure 60: Average score on 'Usability'

> *I am missing the integration of these steps into TOGAF-ADM. How can I take this extension and integrate it into the ADM if I want to develop an ecosystem architecture? –* ***Interviewee S11***

In order to identify the potential improvements of the design, feedback is collected and summarized in the following overview:

I. You need to help users (which in your case are more business minded) to understand how to communicate, promote and use the EAF, for instance using examples/use cases. Extensive training will be required.
II. The TOGAF-ADM is already complex. An additional extension makes it even more complex. So, show how your extension can be built on the existing framework.

6.4. Key Findings

To conclude, relevant and useful feedback was collected throughout the evaluative expert interviews. For each design element, feedback and an average score is collected. The results show that the design artefact is mainly lacking on the feasibility of the design requirements, the completeness of the overall design artefact, the structure of the Ecosystem Implementation Phase, and the usability of the overall design artefact. The most important feedback is translated into potential improvements for the Ecosystem Architecture Framework, as depicted in Table 23.

Table 23: Potential Improvements Design Artefact

Potential Improvement	Explanation	Included in design?
1. Extended Enterprise Architecture	*- Clarify that the Ecosystem Architecture is an extension to another EAF, such as the Existing TOGAF, and cannot be used as an independent EAF, because in that case it is not complete enough. - Show how the extension can be used on the existing TOGAF-ADM. Give examples to make it more tangible. This will help Enterprise Architects to understand the Framework.*	*Yes. An additional section is included on usability where is explained how the extension can be added to the existing Enterprise Architecture Framework.*
2. Decentralized Ecosystem Architecture layer	*Clarify that the Framework does not build a centralized architecture for all ecosystem partners but that it creates a*	*Yes. An explanation has been added to the Design Specification of Phase B that makes clear that the collaboration layer is the extension to the existing business, data, application and technology layer. The external collaboration layer is decentralized and makes it easier for external data providers to initially link their APIs.*
3. Security & Privacy	*Put more focus on the security and privacy aspects of the collaboration. This should be considered in the first phase of the Ecosystem Architecture Framework.*	*Yes. Added an extra design principle that legal and security authorities should be involved during the design of the Ecosystem Architecture. It was already in the design, but in the new design there is more emphasis on the Security and Privacy aspects.*
4. Strategic Coalitions	*Clarify how strategic coalitions are developed to ensure the commitment of the stakeholders to collaborate. The most difficult part of creating an ecosystem architecture is to create shared business models.*	*Yes. The creation of strategic coalitions is more emphasized in the design. In Phase A0. Two main deliverables are developed. 1) Outcome of assessment that shows the need for Collaboration. 2) Proposal for a Strategic Coalition that represents shared incentives business models.*
5. Inter-organizational Model	*The inter-organizational model looks not fully complete yet. Make sure all required inter-organizational artefacts are considered, including API catalog; API descriptions; API policy; governance body.*	*Yes. Most of the inter-organizational artefacts to make the architecture work were already included. In the new design the suggestions of the interviewees were also include. Now Phase C has the deliverable: an inter-organizational model.*
6. Design principles for Ecosystem Architecture Framework	*To ensure the design artefact is more complete, consider giving design principles for each phase instead of providing steps that include everything. It is difficult to*	*Yes. The new design does not intend to create a full complete Ecosystem Architecture, but now only emphasized what should be at least considered when designing an Ecosystem Architecture Framework.*
7. Decentralized governance	*Ensure a decentralized body is developed that ensures that the inter-organizational artefacts are managed and updated.*	*Yes. This was already in the design. Additional explanation is added to make improve the clarity of Phase D.*
8. Relation TOGAF-ADM	*Show how the Ecosystem Enterprise Architecture be used on TOGAF*	*Partly.*
9. Circle of Influence / Circle of concerns	*How in the framework how an organization should deal with artifacts that are in the circle of influence and with the circle of concerns?*	*Not yet.*
10. Architecture Vision	*Explain in more detail what is discussed in the architecture vision.*	*Yes. 7 building blocks are discussed in the design.*
11. Discuss the power of different stakeholders.	*Some stakeholders, such as the BOLD providers have a more dominant role in the creation of the interorganizational artifacts and architecture vision.*	*Yes. This explained in more detail in the design phase.*

Most of the feedback that is collected during the evaluative expert interviews was already included in the design artefact. However, the evaluations show that the interviewees did not interpret the design artefact as it was meant to be. This is valuable information, as it shows that the design artefact is not clear enough without the background knowledge. The potential improvements are corrected in the final version of the design artefact (v1.0). Table 23 shows how the potential improvements are included in the final design artefact.

6.5. Summary Chapter 6 (Answer RQ5)

The fifth sub research question ***'RQ5. How can the quality and suitability of the Extended Enterprise Architecture Framework be improved?'*** covers the evaluation of the designed artefact. Based on the 'Plan Evaluation' methodology by Verschuren and Hartog (2005), a design protocol is created that enabled different Enterprise Architects to evaluate the design plan on quality and suitability. The design protocol ensures the design is evaluated on three aspects: the design objectives, design requirements, and design specifications. The Ecosystem Architecture Framework is demonstrated to three potential end-users (Enterprise Architects) to gain insights on how to improve the quality and suitability of the design. The rationale for this research question is that it is important the designed artefact demonstrated to the potential end-user and evaluation to collect additional and revised requirements. These can be used for the development of a suitable and high-quality design which can be implemented in the real-world.

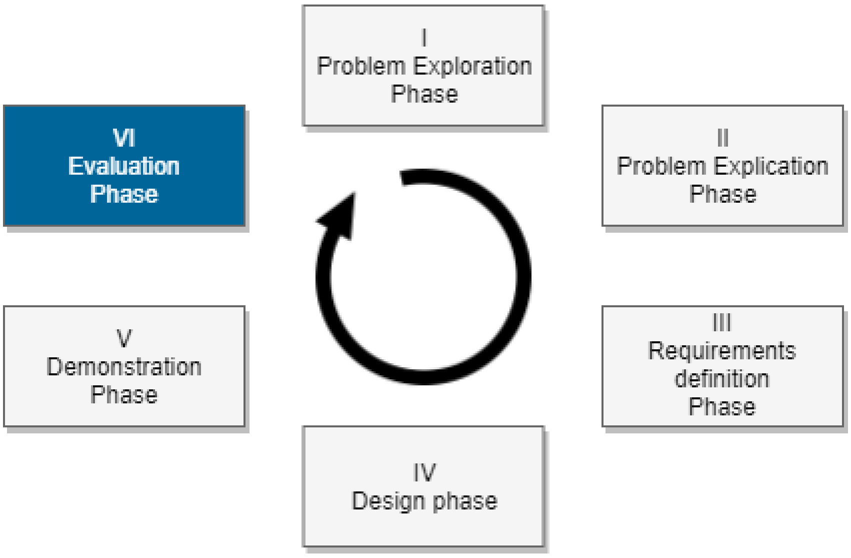
The findings of the evaluative expert interviews are translated into potential improvements for the design, which are depicted in Table 24. To improve the quality and suitability of the Ecosystem Architecture Framework these recommendations for improvement implemented. Potential improvements 1-7 and 11 are directly used to refine the design, which resulted in the final version of the design Figure 35. Three aspects are not improved yet: The relation with TOGAF-ADM; The role of the circle of influence and concern; and the building blocks of the architecture vision.

Table 24: Potential improvements design to improve suitability and quality

Potential Improvement		Explanation
1.	***Extended Enterprise Architecture***	*- Clarify that the Ecosystem Architecture is an extension to another EAF, such as the Existing TOGAF, and cannot be used as an independent EAF, because in that case it is not complete enough.* *- Show how the extension can be used on the existing TOGAF-ADM. Give examples to make it more tangible. This will help Enterprise Architects to understand the Framework.*
2.	***Decentralized Ecosystem Architecture layer***	*Clarify that the Framework does not build a centralized architecture for all ecosystem partners but that it creates a*
3.	***Security & Privacy***	*Put more focus on the security and privacy aspects of the collaboration. This should be considered in the first phase of the Ecosystem Architecture Framework.*
4.	***Strategic Coalitions***	*Clarify how strategic coalitions are developed to ensure the commitment of the stakeholders to collaborate. The most difficult part of creating an ecosystem architecture is to create shared business models.*
5.	***Inter-organizational Model***	*The inter-organizational model looks not fully complete yet. Make sure all required inter-organizational artefacts are considered, including API catalog; API descriptions; API policy; governance body.*
6.	***Design principles for Ecosystem Architecture Framework***	*To ensure the design artefact is more complete, consider giving design principles for each phase instead of providing steps that include everything. It is difficult to*
7.	***Decentralized governance***	*Ensure a decentralized body is developed that ensures that the inter-organizational artefacts are managed and updated.*
8.	***Relation TOGAF-ADM***	*Show how the Ecosystem Enterprise Architecture be used on TOGAF*
9.	***Circle of Influence / Circle of concerns***	*How in the framework how an organization should deal with artifacts that are in the circle of influence and with the circle of concerns?*
10.	***Architecture Vision***	*Explain in more detail what is discussed in the architecture vision.*
11.	***Discuss the power of different stakeholders.***	*Some stakeholders, such as the BOLD providers have a more dominant role in the creation of the interorganizational artifacts and architecture vision.*

The final version of the design these eight potential improvements are implemented, which has resulted in the Ecosystem Architecture Framework 1.0.

VI. Evaluation Phase



7. Conclusions

This chapter presents the conclusions, the reflection, and the recommendations of this thesis. First, the main research question '*How can organizations enable the real-time adoption of external Big and Open Linked Data with Internal Company Data for decision-making?'* is answered (section 7.1). Secondly, there is critically reflected on the research process, research choices and research outcomes (section 7.2). Thirdly, The contributions of this thesis are discussed (section 7.3). Fourthly, recommendations are provided on how the Ecosystem Architecture Framework can be improved, how it can be used in practice and what future research is required to strengthen the outcomes of the research (section 7.4). Lastly, a brief explanation is provided on how the study relates to the study program 'Complex Systems Engineering & Management (section 7.5).

7.1. Answering the Main Research Question

Throughout this thesis five sub-research questions are answered to derive to the answer of the main research question. To answer the main research question, it is important that the outcomes of the different sub-research questions are synthesized. The answers of the sub-research questions can be found at the end of each chapter. The main research question is:

How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?

CIOs of organizations see the potential of adopting external BOLD to strengthen their decisions and are increasingly developing IT strategies that take advantage of this innovation. This increase shows that there is a shift from an internal way of working towards a way of working in which the business ecosystem is involved. This new way of working goes in hand with the presence of adoption barriers, which indicate that there is a lack of alignment across and between organizations. Adoption barriers related to the information quality (data absence, data quality, a lack of meta-data) , task complexity (task handling issues, lack of standards, different sources), technology (Lack of supporting infrastructure, legacy systems, data fragmentation), use & participation (Lack of knowledge, implementation issues, constraints, lack of acceptance), legislations (privacy, security issues and license issues) and governance (Policy issues, Scalability Issues, Lack of Data Governance) hinder the real-time integration of BOLD. If organizations want to use external BOLD for decision-making in real-time, no adoption barriers can be present. The common approach to align a new strategy and its implementation across the enterprise is Enterprise Architecture. Yet, this study demonstrates that the external mindset of organizations demands alignment between organizations and that the current foundation of EA is not capable of satisfying this demand. The approach is limited in breadth, depth, time period and architecture domains.

To adapt to this external mindset a general transformation from a system of record to a system of engagement is required, which is key part of the paradigm shift that comes with an organization's digital transformation. Modern organizations should embrace the system of engagement, including the company's ecosystem as part of their Enterprise Architecture Framework. Hence, a new approach is required that creates alignment between an enterprise and its ecosystem. This study demonstrates that the foundation of EA must be extended with at least six design principles to enable the real-time integration of external BOLD with internal company data.

First, when developing an Enterprise Architecture organizations must embrace a proactive approach to mitigate the adoption barriers of BOLD. This study shows that most organizations embrace a reactive approach that includes the investment in new tooling, personnel or knowledge once an adoption barrier is addressed. However, it is recommended that organizations to embrace a proactive approach that initially deals with the potential adoption barriers during the development of the architecture. This will ensure in the development that is adapted to the adoption barriers that can occur.

Secondly, to ensure adoption barriers between organizations are mitigated the data providers and data users must collaborate with each other. This can result in agreements that benefit the adoption process and to identify potential (causes of) adoption barriers. Through collaboration, organizations can ensure that both parties are aligned with each other and a system can be created that enables the real-time integration of BOLD. However, cooperation must come from two sides, so it is important that there are mutual incentives to mitigate the adoption barriers. BOLD providers will have a more dominant role in the agreements that are made for the collaboration. Organizations must focus on creating incentives for the data provider to cooperate and to form strategic alliances.

Thirdly, organizations must involve the stakeholders concerned with the real-time adoption of BOLD in the design of the system that facilitates the data sharing. Stakeholders that should be involved are the Ecosystem Orchestrators, Service providers, Application providers, Data Producers, Data Publishers, Data Users, Data Prosumers and Legal Authorities. Each stakeholder has concerns about how the system should look like and what requirements the system should satisfy. When stakeholders related to potential adoption barriers are involved during the design there can be anticipated on prevention of these barriers. Creating a shared vision of how a system that facilitates the real-time data integration should look like stimulates the alignment between organizations. The shared vision can be translated into architectural requirements for the design of the system.

Fourthly, organizations must create an architecture that is initially linked to the architecture of the external data sources through a real-time and direct connection. With the current technology available, there is no need for storing the external data locally anymore. By using real-time integration services, such as APIs, the external data does can be called upon request in real-time to make data-driven decisions. This study demonstrates organizations need to invest in a hybrid integration platform that organizations can easily integrate and share integration services (e.g. APIs), different applications can be linked together, collaboration between applications is possible, and orchestration of the applications by a single interface. Different types of hybrid integration platforms are possible, but this thesis demonstrates that an API Management Platform is most suitable.

Fifthly, even when an organization is technically capable of integrating external integration services, such as APIs, inter-organizational and shared artefacts are required that make the system work as it is intended. This are the artefacts that make sure the same way of working is implemented within the business ecosystem. This can be achieved by developing uniform standards for data, meta-data, technical information about integration services; A common data model that ensures one data model, unified schema and semantics are used within the business ecosystem; An independent API policy that ensures that each external data provider can be internally linked to the architecture by conforming to this policy; and institutions between the organizations that enforce the same way of working to enable sustainability of the data.

Lastly, the system should be able to respond to changes of the environment. Therefore, it is important that organizations develop a governance mechanism that ensures the system is scalable, maintainable and sustainable. It is important that the inter-organizational and shared artefacts of the ecosystem are managed and adapt to changes of the environment. A governance mechanism is required to ensure the system is utilizing the maximum business value at all time.

Conforming to these design principles when developing a system that uses external BOLD in real-time for decision-making will ensure that there is alignment between organizations and adoption barriers are mitigated. In this thesis, the Ecosystem Architecture Framework is developed that satisfies the abovementioned design principles. The Ecosystem Architecture Framework provides guidelines for organizations to transform a system of record into a system of engagement. The Ecosystem Architecture Framework has incorporated the capabilities of the current foundation of EA and extends this with additional capabilities collected during field study. Whether the Ecosystem Architecture Framework is used to overcome the research problem or another approach that adheres to these design principles will, according to the outcomes of this study, result in successful integration of external BOLD with internal company data.

7.2. Reflection

This section presents a reflection of this thesis on three aspects. This section first reflects on the research process of this thesis. Secondly, there is reflected on the research choices that are made throughout the thesis. Lastly, the research outcomes of the thesis are reflected upon discussing the Ecosystem Architecture Framework and answer main research question.

7.2.1. Reflection on research process

As discussed in the evaluation session, Verschuren and Hartog (2005) developed different methodologies to evaluate on design-studies. The authors state that the study must be reflected upon four scientific quality criteria: The first criterion that should be reflected upon is the validity of a study; The second criterion is the reliability of a study; The third criterion is the researcher-independence; and the fourth criterion is the verifiability.

I. Validity

The first criterion 'Validity' represents the question whether the research process developed findings that correspond to reality (Verschuren & Hartog, 2005). This thesis collected data through several literature reviews, a systematic literature review, explorative expert interviews and evaluative expert interviews following the approaches proposed by well-known authors. For each research method, a variety of sources are used to ensure multiple perspectives were embraced in the research outcomes. A systematic literature review is a summary of the current state of literature. Nine interviewees from different organizations or departments in different stakeholder roles are interviewed to make sure multi perspective viewpoints are used to develop the design artefact. In chapter 6, the evaluation expert interviews are discussed that assessed whether the developed design artefact corresponds to reality by evaluation the quality and suitability of the design.

II. Reliability

The second criterion 'Reliability' represents the question whether other researchers using the same research approach would come to the same research outcomes (Verschuren & Hartog, 2005). In general, other researchers should come to the research findings using the same research approach consisting of the same research methods. However, it is important to reflect on a couple of aspects that influence the reliability.

First, it is important to mention that this study follows a design research approach, meaning the main deliverable is a designed artefact that forms a solution to the research problem. The researcher has chosen for a design approach, but it can be possible that another researcher would have chosen for another type of approach. Secondly, it is important to mention that the perspective of the researcher influences the research process. The researcher of this study is architecture minded and proposes a design artefact that has Enterprise Architects as end-users. Another researcher that is more economic minded could have decided for CIOs or other decision-makers as end-users. Thirdly, the interpretation of the researcher is an influencing factor. During the selection of the literature of the reviews and the development of the interview protocols, the researcher is having observer dependence. Even when the researcher tries to be as objective as possible, subjectivity will, unintendingly, always be influencing the research process. Lastly, during the brainstorm sessions of the design phase, the creative mind of the researcher is determinative. It is possible that another person with another creative mind would have developed another artefact.

Nevertheless, reliability is guaranteed as the steps that are taken during the research are clearly and comprehensively documented in this study. This makes sure that other researchers can use the same steps to be able to derive to the same conclusions.

III. Researcher-independence

The third criterion 'Researcher-Independence' represents the question whether the researcher was dependent to parties having incentives for a specific research outcome. In total, one person has conducted the research. A team of supervisors was involved during the research process helping with making decisions related to the research process. The interviews are conducting with interviewees in different interviewee roles from different organizations within the public and private sector ensure researcher-independence is guaranteed. Also, during the evaluative interviews, new objective interviewees are selected for this purpose.

IV. Verifiability

The fourth criterion 'Verifiability' represents the question whether it is possible to verify the correctness of the research (Verschuren & Hartog, 2005). Throughout this thesis, different research methods are conducted to get to the research outcomes. For each research method, each step taken is well-documented to ensure verifiability. First, an initial literature review based on the guidelines Levy and Ellis (2006) was performed to identify the knowledge gaps. Furthermore, two different systematic literature reviews are conducted using the approach by Kitchenham et al. (2009). These approaches consider the verifiability of the review by making sure different steps, such as the creation of a review protocol, inclusion and exclusion criteria, and search terms, are documented to enable other researcher to conduct the same research.

Secondly, the interview protocols of the explorative and evaluative interviews are well-documented and substantiated by literature that ensures verifiability. The transcriptions, the code-books and the thematic networks used for the qualitative data analysis can be sent upon request. This ensures the collection method of the qualitative data can be re-used to verify the outcomes. Lastly, each step of the design is documented through following a morphological chart and the best-of-class chart to come to the design. This can be reviewed by other researchers to see the design is derived from the design requirements.

7.2.2. Reflection on research choices

The second aspect the thesis that is reflected upon are the research choices that are made during the design process (Verschuren & Hartog, 2005). Throughout the research process several key research choices are made that have an impact on the research outcomes of this thesis. Each choice that is made earlier causes path dependency for later research decisions (Pierson, 2000). This means that earlier research choices affect the potential choices that must be made later. This section discusses the key research choices of this thesis.

I. The choice for a Design Science Approach

The first key research choice of this thesis was about the research approach of the study. In this thesis, the Method Framework for Design Science by Johannesson and Perjons (2014) is used as a methodology for answering the main research question. The research problem could have been studied through different research approaches. As the foundation of EA, including the existing design-artifacts, lacks the capabilities to overcome the research problem, it seemed logical to develop a new design artefact that can offer the required capabilities. The design science approach enabled the possibility to do both literature research on the adoption barriers of BOLD and capabilities of EA and field research in the form of interviews to collect the requirements for a design artefact that removes the knowledge gaps. This knowledge can be added to the knowledge base.

II. The choice for separate literature reviews

The second key research choice of this thesis was the choice for doing separate literature reviews. In this thesis two literature reviews are conducted to explicate the research problem. The first literature review explicates the reason for misalignment between organizations and across the enterprise when integrating external BOLD with internal company data. The second literature review explicated the limitations of the foundation of EA by identifying the capabilities EA that must be incorporated by the Extended Enterprise Architecture Framework. The latter discusses the cause of the research problem, while the literature review on adoption barriers discusses the consequences of not having the right capabilities. As both literature reviews address an answer to another sub question, the choice for separate literature reviews was made. The other option was to combine everything in one literature review, but this raised the risk of a lack of overview and complexity.

III. The choice for the approach of collecting the design requirements

The third key research choice of this thesis was the research choice on how the requirements for the design artefact should be collected. It was decided to collect the requirements through systems engineering theories. For the interviews, it was important that the right research questions and right interviewees were selected for the interviews. As discussed in this thesis, every stakeholder has its own concerns and viewpoints about how a design should look like. This also applies to the interviewee roles that are selected to collect the requirements of the research. For this reason, interviewees in four different interviewee roles were selected to ensure a complete set of requirements representing the concerns of very different interviewees. Also, an important research choice was to decide on who would be the end-user of the design artefact. This determined what interviewees were selected for the evaluation, as it is important that the design artefact satisfies the requirements of the end-users. However, as the requirements are collected through experts in different roles it is likely that the key findings would have been the same.

IV. The choice for developing the design

The fourth key research choice of this thesis was the design of an Extended Enterprise Architecture. In this thesis, was decided to follow systems engineering literature to develop the design. The design requirements were used for the morphological chart and best-of-class-chart following on the approach by Dym et al. (2009). Another approach could have been possible for developing the design, but the systems engineering approach fitted the other phases of the cycle.

V. The choice for the evaluation approach

The fifth key research choice of this thesis was creating the evaluation method. As a real-world implementation is not possible within the time available for this thesis, a plan evaluation was chosen in which the design purpose, design requirements, and design specifications are evaluated. Another option would have been through a case study, but this did not fit the time frame. Based on this evaluation it should be possible to test the design in the real-world but that requires an additional research.

7.2.3. Reflection on research outcomes

The third aspect on which the thesis is reflected are the research outcomes. The research process with its associated choices led to specific research outcomes. However, the research outcomes also have its limitations.

I. The Ecosystem Architecture Framework is only validated through a Plan Evaluation.

The design is only tested on 'paper' and not in real-world. This limits the research outcome, because it the Ecosystem Architecture Framework is only evaluated on design purpose, design requirements and design specifications. The evaluation show that the feasibility of the research could possibly an issue. For example, the interviewees state that 'involving all organizations concerned' in practice is hardly feasible. This kind of statements can only be validated if also a real-world experiment is done, however that is an already a study. Verschuren and Hartog (2005) explain that a valid design should also be evaluated through the 'Process Evaluation' and 'Product-evaluation', which entails implementing the Ecosystem Architecture Framework in real-life and the effects of the design artefact. This results in useful new insights

II. The Ecosystem Architecture Framework is dependent on TOGAF-ADM.

The Ecosystem Architecture Framework was designed as an extension to the TOGAF-ADM, as this EAF was selected as the most suitable commonly used EAF to build on. One of the interviewees stated that this makes the Ecosystem Architecture Framework dependent on TOGAF-ADM. The requirements collected are suitable for an extension to an existing EAF, and not suitable for a new EAF itself. It must be mentioned that the requirements without the use of a base EAF, are cannot fully cover the design of an independent extended EAF. However, the design artifact can also be design principles that a design should at least cover to overcome the research problem.

III. The Ecosystem Architecture Framework can be interpreted wrong.

The Ecosystem Architecture Framework can be interpreted wrong. It is meant as an Extended Enterprise Architecture Framework that increases the alignment between different organizations but is not meant as a collective architecture that is owned by the whole ecosystem. This addresses many additional challenges, such as responsibility on management and governance of the architecture. It only makes linking the architectures of different enterprises easier. Through a Collaboration Architecture and an inter-organizational model, the alignment is between organizations can also be achieved.

IV. The report has become very lengthy

This thesis has become more than hundred pages, which is quite long. To increase the reliability, verifiability and validity, many different design methods are used to come to the answer of the main research question. Unfortunately, this requires documentation of all the steps taken. This causes that the report became very long. Different sections are placed in the appendix to lower the number of pages. Also, a executive summary was developed to enable the reader to read the most important findings.

7.3. Scientific Contribution

Table 25 presents the new contributions to science of this thesis. The table shows the different aspects that are discussed throughout this thesis. It also shows what aspects are already covered by the current state of literature, and what aspects are added to the knowledge base. This thesis provides four Scientific Contributions:

I. Adding new Adoption Barriers of BOLD

Next to the adoption barriers collected through the systematic literature review, the adoption barriers that organization address when integrating external BOLD with internal company data in real-time are collected. Most of the adoption barriers collected are verifying the findings of the current literature, but also additional adoption barriers are found and used to revise the conceptual model of adoption barriers. One additional barrier category is identified: Governance Barriers. The governance barriers can be mapped into three clusters. Adoption barriers related to: Policy Issues; Scalability Problems; and a Lack of Data Governance. The governance barriers represent the barriers of how the data integration and exchange with external BOLD Providers is managed within and between organizations. Next to this, the lack of acceptance is a sub-cluster added to the use & participation barriers. Different properties of open data create the lack of acceptance by users, as it happens that this type of data is not yet accepted as an alternative to the traditional (data) sources for decisions-making.

II. Adding a critical assessment of the foundation of EA and the most commonly used Enterprise Architecture Frameworks

Through the literature review the capabilities of the foundation of EA that should be incorporated by the design are identified. This resulted in a list of 11 capabilities that a good EAF should have. The most commonly used EAFs are assessed on having these capabilities and the TOGAF-ADM is selected as the most suitable EAF for extension. This is among the first studies that assesses and compares the commonly used EAFs on having these capabilities. The research shows that the foundation of EA is not suitable enough for the implementation of a strategy that involves the adoption external BOLD for real-time decision-making and additional capabilities are required and should be extended.

III. Adding required capabilities for improving the foundation of EA

Through expert interviews the required capabilities to enable the implementation of a strategy that involves the real-time integration of external BOLD for decision-making are identified. This has resulted in six functional requirements satisfied in the design of an Extended EAF that has the capabilities to overcome the research problem. A contribution to the foundation of EA is This is the first study that shows that an extension to the foundation of EA is required to adapts to the importance of external data.

IV. Adding an approach that enables the real-time integration of BOLD with internal company data

Through a morphological chart, best-of-class chart, brainstorm sessions and evaluative expert interviews a new approach is developed that makes it possible to enable the real-time integration of BOLD with internal company data for decision-making. A validated Extended EAF has been developed that provides guidance for combing of internal company data with external BOLD for decision-making and how this should be arranged at a technical and institutional level.

Table 25: Contributions to science

Aspect	Current literature	Authors	This Thesis	Contributions to Science
Potential of BOLD	*Yes*	***e.g. Buchholtz et al. (2014); M. Janssen and Kuk (2016); Lněnička and Komárková (2018)***	*Yes*	
Adoption Barriers of BOLD	*Yes*	***e.g. Conradie and Choenni (2014); Janssen et al. (2012); Zuiderwijk and Janssen (2014)***	*Yes*	*Adding new adoption barriers of BOLD*
Capabilities of current foundation Enterprise Architecture	*Yes*	***e.g. Minoli (2008); Schekkerman (2004); Sessions (2007)***	*Yes*	
Assessment of most commonly used Enterprise Architecture Frameworks on having these capabilities.	*No*		*Yes*	*Adding a critical assessment of the foundation of EA and identification of most suitable EAF for extension*
Required capabilities for an Extension to EA to mitigate the adoption barriers of BOLD and enable the real-time integration of Big and Open Linked Data	*No*		*Yes*	*Adding required capabilities for improving the foundation of EA.*
Approach that enables the real-time integration of external BOLD with internal company data.	*No*		*Yes*	*Adding an approach that enables this: The Ecosystem Architecture Framework*

7.4. Recommendations

7.4.1. Recommendation for managers to use research in practice / society

We are going towards a thriving data-driven economy, in which the use of external data is becoming more important for organizations to use. More CIOs of organizations are developing strategies that involve the real-time integration of external data, such as BOLD, with internal data. Unfortunately, different adoption barriers that organization address show that there is mis-alignment between and across organizations. In total, 85 percent of the organizations that adopts a data-driven strategy fails (Bishop, 2018; Gartner, 2017). For both national governments and organizations itself it is important that the success rate of the implementation of such a strategy increases. Managers see the importance and potential of adopting external BOLD to strengthen their decisions with but encounter that the organization and ecosystem is often not ready for this innovation. To succeed in in the data-driven economy, organizations must make use of external BOLD. To enable this a transition from a system of records towards a system of engagement is required. Moreover, organizations must start thinking platform-driven, in which the architectures of data users and data providers can be initially linked and can be accessed through APIs.

This research provides an approach that can help Enterprise Architects with the development of an architecture that enables the combination internal company data with external BOLD in real-time and how this should be arranged at a technical and institutional level. This approach, the Ecosystem Architecture Framework, provides guidance to develop an architecture that not only focuses on the internal enterprise, but also on the external enterprise. Everything must be open, as the combination of data silos in real-time results in new (e.g. context, patterns and behaviors) which can be used to improve decisions. This research has implications for society and managers, as it shows that the foundation of EA seems not suitable anymore to adapt to the data-driven economy. The importance of adopting external data requires a renewed mindset of organizations, that should be reflected in the way of working of an organization. It is recommended for managers to use the Ecosystem Architecture Framework to ensure the organization adapts to the data-driven economy. This requires change of the mindset of the organization, which is facilitated by the guidelines of the Ecosystem Architecture Framework. It enables organizations to collaborate with data providers and to involve all the parties concerned with the real-time integration of BOLD with internal company data. It will ensure that the organization and its ecosystem will use the same way of working, so that the organizations can be initially linked together through a hybrid integration platform in which APIs are provided.

This new approach demands change and participations of organizations and its ecosystem. Therefore, it is important that mutual incentives are developed that ensure that all organizations of a business ecosystem are willing to cooperate. Only, if organizations are willing to cooperate the success rate of organizations will increase. This will contribute to a wealthy society.

7.4.2. Recommendation for Future Research

This section presents the recommendations for future research that can be made based on the evaluation, the conclusions, and the reflection.

This research identified six additional capabilities that should be incorporated by EAFs. In this thesis these six capabilities are discussed and embedded into the design of the Ecosystem Architecture Framework. During the interviews already interesting information about these capabilities are given, but it would be interesting to do further research on how to create these capabilities.

I. Future research would be investigating each capability in depth
II. Future research would be searching for additional capabilities

This research has validated the design through a plan-design by Verschuren and Hartog (2005). This has resulted in a list of potential improvements for the ecosystem architecture framework. The authors also suggest two other methodologies for evaluating a design: 'Process-Evaluation' and 'Product-Evaluation'. This would be interesting for future research. This will result in potential improvements to enrich the design to improve the artifact on completeness and feasibility.

I. Future research would be investigating the implementation of the Ecosystem Architecture Framework in the real-world to analyze the implications that will bring.
II. Future research would be analyzing the effects of the Ecosystem Architecture Framework in a real-world context.

This research does not focus on a specific sector or industry. The Ecosystem Architecture framework could be more successful in one industry than another. It would be interesting to do research on the implications of the new way of working in different industries.

I. Future research would be engaging the Ecosystem Architecture Framework in different industries with different organizations and ecosystems to identify and evaluate the differences.

This research concludes that organizations should collaborate, and external stakeholders should be involved during the design of the ecosystem architecture. This will have implications for maintain privacy and guarantying security of the internal systems and data.

I. Future research would be investigating how privacy and security can be maintained in an Ecosystem Architecture.

7.5. Link of this research with the CoSEM Master's Programme

The Complex Systems Engineering and Management (CoSEM) master's programme focusses on the design of technological innovations in complex socio-technical environments. This are environments in which multi-actor and technological complexities arise. The programme focuses on more than technology alone, but also on regulations, logistics, behavioral change and financial incentives. This thesis focuses on the complex environment of EA and External BOLD, in which multi-actor and technical complexities are present. The thesis looks beyond the design of a technical system that enables real-time integration of external BOLD and concentrates on what institutions and transactions (economics) are required to implement the system. This design is developed through a design science approach using different Systems Engineering theories and methods that are part of the CoSEM Curriculum. Among others, a morphological chart analysis, design space brainstorming sessions, and best-of-class chart are used to develop a design that overcomes the research problem. Also, systems engineering techniques are used to derive the design requirements for the solution.

Bibliography

Allen, I. E., & Seaman, C. A. (2007). Likert scales and data analyses. *Quality progress, 40*(7), 64-65.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly, 32*(4), 399-418.

Attride-Stirling, J. (2001). Thematic networks: an analytic tool for qualitative research. *Qualitative research, 1*(3), 385-405.

Barriball, K. L., & While, A. (1994). Collecting data using a semi-structured interview: a discussion paper. *Journal of Advanced Nursing-Institutional Subscription, 19*(2), 328-335.

Bayrak, T. (2015). A review of business analytics: a business enabler or another passing fad. *Procedia-Social and Behavioral Sciences, 195*, 230-239.

Bishop, S. (2018). *Using Data-Driven Decision-Making to Enhance Performance: A Practical Guide for Organizations.* University of Maryland University College,

Bizer, C., & Heath, T. (2009). Linked Data – The Story So Far. *International Journal on Semantic Web and Information Systems (IJSWIS), 5*(3).

Brereton, P., Kitchenham, B. A., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software, 80*(4), 571-583.

Buchholtz, S., Bukowski, M., & Śniegocki, A. (2014). Big and open data in Europe: A growth engine or a missed opportunity. *Warsaw Institute for Economic Studies Report Commissioned by demosEUROPA, 10*.

CBS. (2018). Gebruik Open Data stijgt fors.

Chatfield, A. T., & Reddick, C. G. (2017). A longitudinal cross-sector analysis of open data portal service capability: The case of Australian local governments. *Government Information Quarterly, 34*(2), 231-243.

Commission, T. E. (2014). Towards a thriving data-driven economy, Communication from the commission to the European Parliament, the council, the European economic and social Committee and the committee of the regions, Brussels.

Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly, 31*, S10-S17.

Corrêa, A. S., Paula, E. C. d., Correa, P. L. P., & Silva, F. S. C. d. (2017). Transparency and open government data: a wide national assessment of data openness in Brazilian local governments. *Transforming Government: People, Process and Policy, 11*(1), 58-78.

Curry, E. (2016). The big data value chain: definitions, concepts, and theoretical approaches. In *New horizons for a data-driven economy* (pp. 29-37): Springer, Cham.

De Donato, R., Ferretti, G., Marciano, A., Palmieri, G., Pirozzi, D., Scarano, V., & Vicidomini, L. (2018). *Agile production of high quality open data.* Paper presented at the Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age.

Denstad, H., & Bygstad, B. (2012). Managing the it Alignment Gap in Turbulent Times—An Inside View. *Journal of Information Technology Case and Application Research, 14*(2), 28-46.

Donker, F., van Loenen, B., & Korthals Altes, W. (2017). Societal cost-benefit analysis open data.

Dutta, D., & Bose, I. (2015). Managing a big data project: the case of ramco cements limited. *International Journal of Production Economics, 165*, 293-306.

Dym, C. L., Little, P., Orwin, E. J., & Spjut, E. (2009). *Engineering design: A project-based introduction*: John Wiley and sons.

Fernandes, L. M., O'Connor, M., & Weaver, V. (2012). Big data, bigger outcomes. *Journal of AHIMA, 83*(10), 38-43.

Gartner. (2013). Enterprise Architecture (EA). Retrieved from https://www.gartner.com/it-glossary/enterprise-architecture-ea/

Gartner. (2015). How to adopt Open Data for Business Data and Analytics – and why should? . Retrieved from https://www.gartner.com/document/3158117?ref=solrAll&refval=230530843&qid=7d7e5e18c7d688df934a78b9b

Gartner. (2017). Gartner Says Business Intelligence and Analytics Leaders Must Focus on Mindsets and Culture to Kick Start Advanced Analytics. . Retrieved from https://www.gartner.com/newsroom/id/3130017

Gartner. (2018). How to Implement a Truly Hybrid Integration Platform. Retrieved from https://www.gartner.com/document/code/350671?ref=ddisp&refval=350671

Gerow, J. E., Thatcher, J. B., & Grover, V. (2015). Six types of IT-business strategic alignment: an investigation of the constructs and their measurement. *European Journal of Information Systems, 24*(5), 465-491.

Gerunov, A. (2017). Understanding open data policy: evidence from Bulgaria. *International Journal of Public Administration, 40*(8), 649-657.

Gong, Y., & Janssen, M. (2017). *Enterprise Architectures for Supporting the Adoption of Big Data.* Paper presented at the Proceedings of the 18th Annual International Conference on Digital Government Research.

Group, O. (2019). TOGAF

Guinney, J., Wang, T., Laajala, T. D., Winner, K. K., Bare, J. C., Neto, E. C., . . . Pahikkala, T. (2017). Prediction of overall survival for patients with metastatic castration-resistant prostate cancer: development of a prognostic model through a crowdsourced challenge with open clinical trial data. *The Lancet Oncology, 18*(1), 132-142.

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of "big data" on cloud computing: Review and open research issues. *Information systems, 47*, 98-115.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian journal of information systems, 19*(2), 4.

Hilliard, R. (2000). Ieee-std-1471-2000 recommended practice for architectural description of software-intensive systems. *IEEE, http://standards. ieee. org, 12*(16-20), 2000.

Hossain, M. A., Dwivedi, Y. K., & Rana, N. P. (2016). State-of-the-art in open data research: Insights from existing literature and a research agenda. *Journal of Organizational Computing and Electronic Commerce, 26*(1-2), 14-40.

Huang, C. D., & Hu, Q. (2007). Achieving IT-business strategic alignment via enterprise-wide implementation of balanced scorecards. *Information Systems Management, 24*(2), 173-184.

Huijboom, N., & Van den Broek, T. (2011). Open data: an international comparison of strategies. *European journal of ePractice, 12*(1), 4-16.

Iivari, J. (2007). A paradigmatic analysis of information systems as a design science. *Scandinavian journal of information systems, 19*, 39.

Jadda, S., & Idrissi, M. A. J. (2015). *Strategic Alignment and Information System project portfolio optimization model.* Paper presented at the 2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA).

Janssen. (2009). Framing Enterprise Architecture: A metaframework for analyzing architectural efforts in organizations. *Coherency Management: Architecting the Enterprise for Alignment, Agility and Assurance*, 107-126.

Janssen, Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258-268.

Janssen, M., Estevez, E., & Janowski, T. (2014). Interoperability in big, open, and linked data-organizational maturity, capabilities, and data portfolios. *IEEE Computer, 47*(10), 44-49.

Janssen, M., & Kuk, G. (2016). Big and open linked data (BOLD) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce, 26*(1-2), 3-13.

Janssen, M., Matheus, R., & Zuiderwijk, A. (2015). *Big and open linked data (BOLD) to create smart cities and citizens: Insights from smart energy and mobility cases.* Paper presented at the International Conference on Electronic Government.

Jia, Y., Wang, N., & Ge, S. (2018). Business-IT Alignment Literature Review: A Bibliometric Analysis. *Information Resources Management Journal (IRMJ), 31*(3), 34-53.

Johannesson, P., & Perjons, E. (2014). *An introduction to design science*: Springer.

Jonkers, H., Lankhorst, M. M., ter Doest, H. W., Arbab, F., Bosma, H., & Wieringa, R. J. (2006). Enterprise architecture: Management tool and blueprint for the organisation. *Information systems frontiers, 8*(2), 63-66.

Kallio, H., Pietilä, A. M., Johnson, M., & Kangasniemi, M. (2016). Systematic methodological review: developing a framework for a qualitative semi-structured interview guide. *Journal of advanced nursing, 72*(12), 2954-2965.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University, 33*(2004), 1-26.

Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering–a systematic literature review. *Information and software technology, 51*(1), 7-15.

Kuk, G., & Davies, T. (2011). The roles of agency and artifacts in assembling open data complementarities.

Lee, G., & Kwak, Y. H. (2012). An open government maturity model for social media-based public engagement. *Government Information Quarterly, 29*(4), 492-503.

Leech, B. L. (2002). Asking questions: Techniques for semistructured interviews. *PS: Political Science & Politics, 35*(4), 665-668.
Levy, Y., & Ellis, T. J. (2006). A systems approach to conduct an effective literature review in support of information systems research. *Informing Science, 9*.
Lnenicka, M., & Komarkova, J. (2018). Big and open linked data analytics ecosystem: Theoretical background and essential elements. *Government Information Quarterly*.
Lněnička, M., & Komárková, J. (2018). *Big and open linked data analytics ecosystem: Theoretical background and essential elements*.
Lněnička, M., Máchová, R., Komárková, J., & Čermáková, I. (2017). Components of big data analytics for strategic management of enterprise architecture. *Research Gate*.
Maier, M. W., Emery, D., & Hilliard, R. (2001). Software architecture: Introducing IEEE standard 1471. *Computer*(4), 107-109.
Marketwatch. (2017). Americas most successful companies are killing the economy.
Matthes, F., Buckl, S., Leitel, J., & Schweda, C. M. (2008). *Enterprise architecture management tool survey 2008*: Techn. Univ. München.
Mclaren, R., & Waters, R. (2011). Governing location information in the UK. *The Cartographic Journal, 48(3), 172–178*.
Mihas, P. (2019). Qualitative data analysis. In *Oxford Research Encyclopedia of Education*.
Minoli, D. (2008). *Enterprise architecture A to Z: frameworks, business process modeling, SOA, and infrastructure technology*: Auerbach Publications.
Moreno Jr, V. d. A., de Souza Costa Neves Cavazotte, F., & de Oliveira Valente, D. (2009). Strategic Alignment and Its Antecedents: A critical analysis of constructs and relations in the international and Brazilian literature. *Journal of Global Information Technology Management, 12*(2), 33-60.
Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: lessons learned. *Transforming Government: People, Process and Policy, 9*(3), 286-308.
Ohemeng, F. L., & Ofosu-Adarkwa, K. (2015). One way traffic: The open data initiative project and the need for an effective demand side initiative in Ghana. *Government Information Quarterly, 32*(4), 419-428.
Ostrom, E. (1990). *Governing the commons: The evolution of institutions for collective action*: Cambridge university press.
Pan, Q., Jonoski, A., Castro-Gama, M. E., & Popescu, I. (2015). APPLICATION OF A WEB-BASED DECISION SUPPORT SYSTEM FOR WATER SUPPLY NETWORKS. *Environmental Engineering & Management Journal (EEMJ), 14*(9).
Pennington, D., & Cagnazzo, L. (2018). Connecting the silos: implementations and perceptions of linked data across European libraries. *Journal of Documentation*.
Pettricrew, M., & Roberts, H. (2006). Systematic reviews in social sciences. In: Oxford: Blackwell.
Pierson, P. (2000). Increasing returns, path dependence, and the study of politics. *American political science review, 94*(2), 251-267.
Porter, M. E. (1996). What is strategy. *Published November*.
Rehman, M., & Shamail, S. (2014). *Enterprise architecture and e-government projects in Punjab, Pakistan.* Paper presented at the Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance.
Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management, 54*(5), 758-790.
Saxena, S. (2018a). Drivers and barriers to re-use Open Government Data (OGD): a case study of open data initiative in Philippines. *Digital Policy, Regulation and Governance, 20*(4), 358-368.
Saxena, S. (2018b). Drivers and barriers towards re-using open government data (OGD): a case study of open data initiative in Oman. *foresight, 20*(2), 206-218.
Saxena, S., & Muhammad, I. (2018). Barriers to use open government data in private sector and NGOs in Pakistan. *Information Discovery and Delivery, 46*(1), 67-75.
Schekkerman, J. (2004). *How to survive in the jungle of enterprise architecture frameworks: Creating or choosing an enterprise architecture framework*: Trafford Publishing.
Sessions, R. (2007). A comparison of the top four enterprise-architecture methodologies. *Houston: ObjectWatch Inc*.
Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics, 191*, 97-112.
Smith, G. (2007). Morphological charts: a systematic exploration of qualitative design space.

Tang, A., Han, J., & Chen, P. (2004). *A comparative analysis of architecture frameworks.* Paper presented at the 11th Asia-Pacific software engineering conference.

Thomas, D. R. (2006). A general inductive approach for analyzing qualitative evaluation data. *American journal of evaluation, 27*(2), 237-246.

Thorsby, J., Stowers, G. N., Wolslegel, K., & Tumbuan, E. (2017). Understanding the content and features of open data portals in American cities. *Government Information Quarterly, 34*(1), 53-61.

Urbaczewski, L., & Mrdalj, S. (2006). A comparison of enterprise architecture frameworks. *Issues in Information Systems, 7*(2), 18-23.

Vargas, A., Boza, A., Patel, S., Patel, D., Cuenca, L., & Ortiz, A. (2016). Inter-enterprise architecture as a tool to empower decision-making in hierarchical collaborative production planning. *Data & Knowledge Engineering, 105*, 5-22.

Verschuren, P., & Hartog, R. (2005). Evaluation in design-oriented research. *Quality and Quantity, 39*(6), 733-762.

Vetrò, A., Canova, L., Torchiano, M., Minotas, C. O., Iemma, R., & Morando, F. (2016). Open data quality measurement framework: Definition and application to Open Government Data. *Government Information Quarterly, 33*(2), 325-337.

Wang, H.-J., & Lo, J. (2019). Factors Influencing the Adoption of Open Government Data at the Firm Level. *IEEE Transactions on Engineering Management*.

Wohlin, C. (2014). *Guidelines for snowballing in systematic literature studies and a replication in software engineering.* Paper presented at the Proceedings of the 18th international conference on evaluation and assessment in software engineering.

Zachman, J. A. (1987). A framework for information systems architecture. *IBM systems journal, 26*(3), 276-292.

Zuiderwijk, A., & Janssen, M. (2014). *The negative effects of open government data-investigating the dark side of open data.* Paper presented at the Proceedings of the 15th Annual International Conference on Digital Government Research.

Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012). Socio-technical Impediments of Open Data. *Electronic Journal of e-Government, 10*(2).

Zuiderwijk, A., Jeffery, K., & Janssen, M. (2012). The potential of metadata for linked open data and its value for users and publishers. *JeDEM-eJournal of eDemocracy and Open Government, 4*(2), 222-244.

Appendices

Appendix A1: Definition of Key Constructs

A1.1. Defining BOLD

As mentioned in the introduction, BOLD is the combination of Big Data and Open Data. Big Data typically are data with a large data volume, a rapid velocity of data flows, and high variety in data types and sources (Lněnička & Komárková, 2018). Open Data is data that is opened up by researchers, governments, organizations, and individuals to enable access for everybody to data without any pre-defined restrictions or conditions for use, reuse and redistribution (M. Janssen et al., 2015). With Linked is meant that the data is 'structured and machine readable that can be semantically queried' (Bizer & Heath, 2009). BOLD can take various forms and is gathered from a wide variety of sources, which can have different known and unknown qualities (M. Janssen & Kuk, 2016). In Figure 61, the relation of these different data categories of BOLD is displayed in a macro overview. In total, two data categories are related to BOLD: Big Data; and Open Data. As depicted in Figure 61, the combination of public Big Data and Open Business Data can be referred as BOLD.

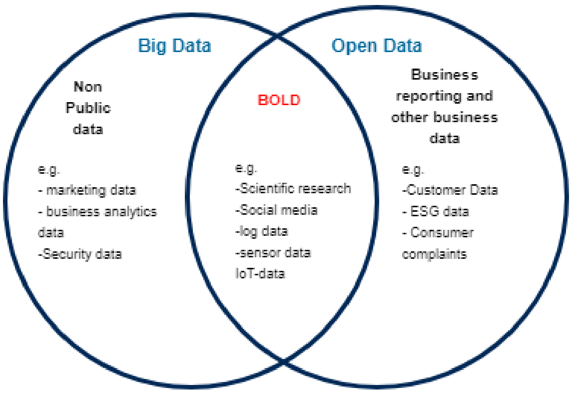


Figure 61: Macro overview of BOLD

A1.2. Classification of BOLD

Hashem et al. (2015) presents a classification of BOLD consisting of 5 classes: Data sources; Content format; Data format; Data staging; and Data processing (Figure 62Figure 62).

Data sources

BOLD can be retrieved from a variety of sources. Open data portals give users access to open data that is published by companies, governments, researchers and individuals. Data aggregators compile the data and information that is searched, gathered and presented in a report-based summarized format. Data markets are marketplaces where entities (e.g. companies) can sell their (customer) data to other companies. Social data is a data source that contains data that is extracted from social media applications. Also log, sensor and IoT data are sources that continuously are being collected by organizations.

Content format

The content format of the data can be structured, semi-structured and unstructured. Structured data is data that has been organized in a formatted database with columns and rows and relational keys. This is the easiest way to use the information for analyses. (e.g. Relational data). Semi-structured data is not in a relational database but do have organizational properties that make the analysis easier. These can be used to store them in in a relational database (e.g. XML data). Unstructured data is not organized and do not have organizational properties and does not fit in a relational database. Alternative platforms already exist to perform data analyses on unstructured data (e.g. JPEG, Word, PDF files.[2] Figure 63 presents that innovations result in more structured and open data (M. Janssen & Kuk, 2016). For example, by tagging faces in pictures, meta-data is created, which makes the data more structured.

[2] https://www.geeksforgeeks.org/difference-between-structured-semi-structured-and-unstructured-data/

Data stores

BOLD can be used in different data stores. NoSQL is a database that 'provides a mechanism for storage and retrieval of data that is modeled in means other than the tabular relations used in relational databases' [3], such as column stores and document stores. NewSQL is a type of relational database management system that provides scalability of NoSQL systems.

Data staging

There are different data stages, such as extracting data, cleaning the data, transforming the data, integration of the data, linking the different datasets and quality processes. Once these data staging processes are completed, the data can be loaded into a data target, such as data warehouses, data marts, data lakes.

Data processing

There are different ways how BOLD can be processed. BOLD can be processed as batches, that is, high volumes of data that is collected over a period of time. Another way of processing data is in real-time, which means that the input, process and output of the data is continuously done. This means that is processed in a very small period (e.g. radar systems, customer services).[4] Stream processing is almost instantly analyze data that is being used.

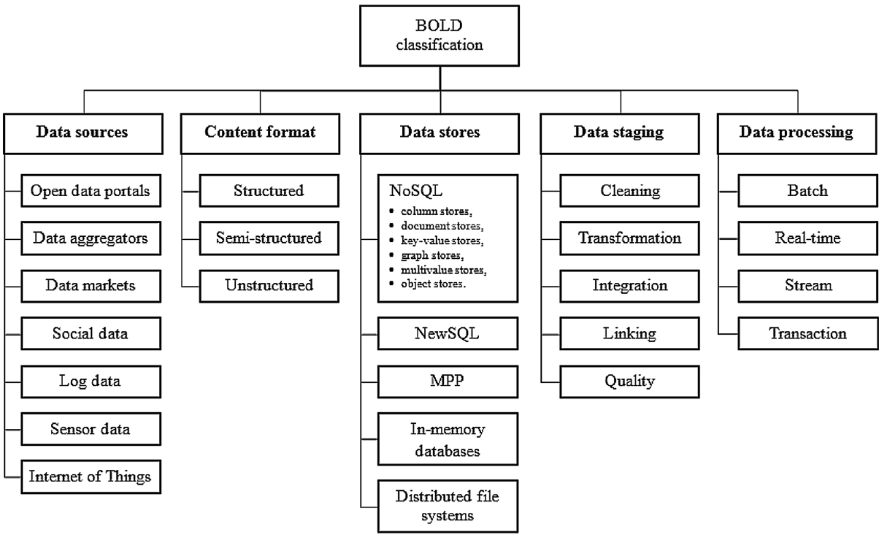


Figure 62: A classification of BOLD (Hashem et al., 2015)

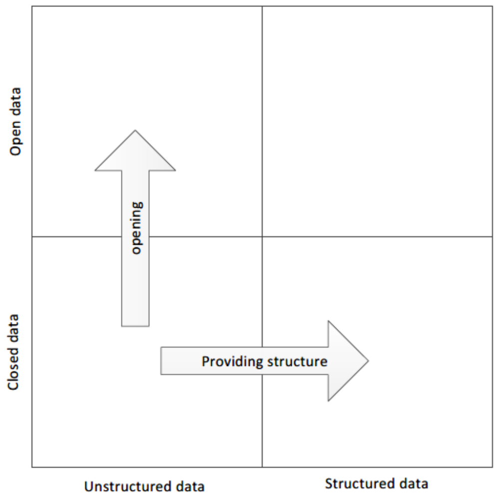


Figure 63: A classification of BOLD (M. Janssen & Kuk, 2016)

[3] http://nosql-database.org/

[4] https://www.datasciencecentral.com/profiles/blogs/batch-vs-real-time-data-processing

A1.3. BOLD Life Cycle

According to Lněnička and Komárková (2018), there are 5 phases in the BOLD lifecycle from the data sources unto the decision-making processes. Based on a literature review on the topic of BOLD, the authors created a BOLD lifecycle with 5 phases with its associated activities (figure 11).

First, the 'Data Collection and Acquisition' phase includes activities that are focusing on what data sources will be used and the collection and extraction process. This phase includes the generation of BOLD, selecting and filtering the data sources, gathering the data and the extraction and loading of the data. Important components are interfaces, data types, selection and transfer of the data. Second, the 'Data Management and Preparation' phase is about cleaning and preparing the data, pre-processing and validation of the data, and transforming the format of the data to ensure it can collaborate with each other. The third phase is about 'Storing and Archiving' the data, which includes storing and securing the data, transporting and transferring the data to right locations, searching and finding processes of the data and archiving and curating the data. Fourth, the 'Data processing and analysis' phase is about performing data analysis, modelling and simulating the data for useful insights, and predicting and optimizing BOLD for improved insights. In the next phase, the 'Data Visualization' phase the output of the analyses are visualized and interaction (?), where after the visualizations are evaluated and interpreted. Finally, the outcomes will be exported and reported. The last phase before the decision-making processes is the 'Data publication, sharing and reuse' phase, which focusses on legislation, defining standards, interlinking and connecting data, mashups, collaboration, requesting and suggesting and sharing and spreading the BOLD.

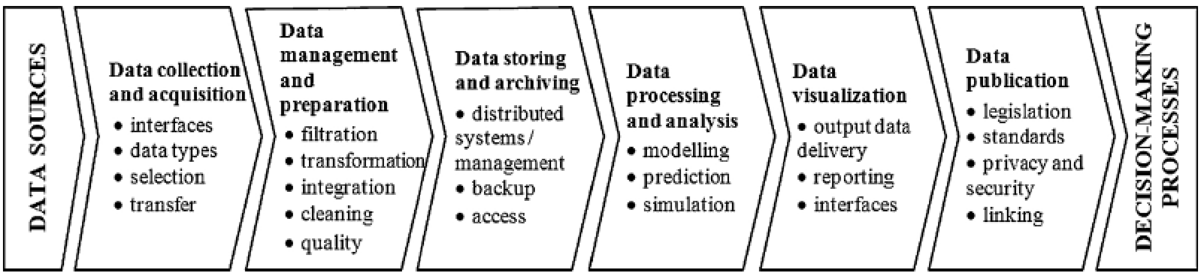


Figure 64: OLD life cycle (Lněnička & Komárková, 2018).

A1.4. Open data sharing

The concept of open data is used in different forms, so it is important to define what is considered open data. Officially open data is defined as 'data or content that can be accessed by anyone and is free to use, re-use and re-distribute.[5] But in practice, 'Open data means that the publication of data in an open manner will exist on a spectrum of open access'. Gartner (2015) made an overview of the spectrum of open data accessibility with on the one side of the spectrum fully and unconditionally open data and on the other side publicly available data as part of the business model, as displayed in Figure 65

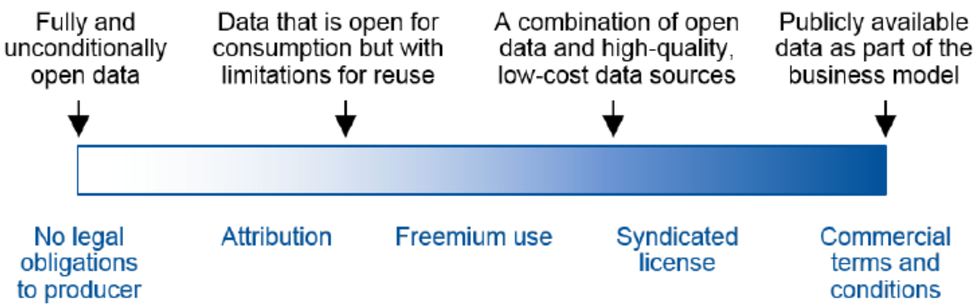


Figure 65: Spectrum of Open Data Accessibility (Gartner, 2015)

. Open data can be published internally between different departments within a company, between business partners, and publicly accessible for anyone (McKinsey, 2019). In Figure 66, these open data dimensions are depicted. There are different dimensions of Open Data. First, Open Data can be shared within different business and IT units of an organization. In this case, data is opened to unify data silos within the internal organization.

[5] http://opendatahandbook.org/guide/en/what-is-open-data/

Secondly, data can be opened for partners of the organization. In this case a collaboration agreement is made for data-sharing between different companies. A business model can be part of this collaboration. Thirdly, open data can be shared publicly. This means that the data can be accessed by anyone, for free or for a license.

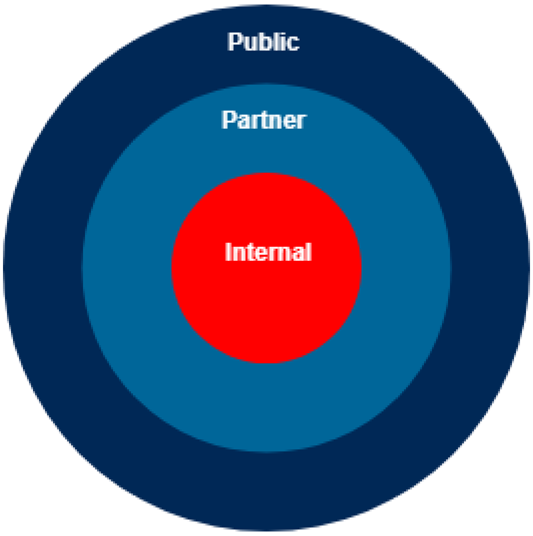


Figure 66: Data integration dimensions BOLD

Appendix B1: Explorative Expert Interview Protocol

In this appendix the Explorative Expert interview protocol is discussed. The interview protocol discusses the approach of the interview. An interview protocol consists of the instruction for the interviewer including the opening statements, the key questions to be asked, possible follow up questions and transition messages for the interviewer. Also, there should be space in the interview for a reflection of the interview by the interviewee and space for the interviewer to record comments. The purpose of the protocol is to make sure that the questions are well understood by the interviewee and the interview is conducted in a structured way. Each question aims to result in valuable research findings that can contribute to the thesis. The guidelines by Johanesson & Perjons (2014) were used to set up the interview protocol. As displayed in appendix 1, different interview roles are interviewed: Business Analyst, Data Analyst, Enterprise Architect, Decision-maker and BOLD-experts.

Introduction of the thesis subject:

Chief Information Officers (CIOs) of companies are adopting external Big and Open Linked Data to strengthen their real-time data-driven decision-making. The adoption of external BOLD could support companies with their decisions relating to a variety of problems and opportunities, such as strengthening customer relationship and creating a better product-market fit. My thesis intends to build an Enterprise Architecture Framework that should help companies deal with the most common challenges of the real-time integration of BOLD for decision-making.

During the interview I aim to explore the problem, discuss (additional) causes of this problem (adoption barriers), and collect requirements for possible solutions for this problem. I'm especially curious about what companies/organizations are doing in practice to overcome the adoption barriers of BOLD and how the data-driven decision-making process of these companies look like.

I would like to emphasize that the interview is voluntary and anonymous, the interviews can be stopped at any moment, and no names of interviewees and companies will be mentioned in the thesis. I would like to make audio recordings of the interviews, so I get the chance to code and transcribe the interviews for qualitative data analysis. I will send the interview results after the interview, so you will have the chance to review and comment on how I interpreted things. I would like you ask you to sign the informed consent form if you agree with the questions. A data management plan has been developed to comply with the General Data Protection Regulation (GDPR). I will publish only the code-book and not the transcripts, and the thesis can be accessed via the Academic Education Repository.

The interview will consist of four parts: Background; Use of BOLD for Data-Driven Decision-Making; Adoption barriers of BOLD; and Overcoming Adoption Barriers. The interview will approximately take 1 hour.

I. Background

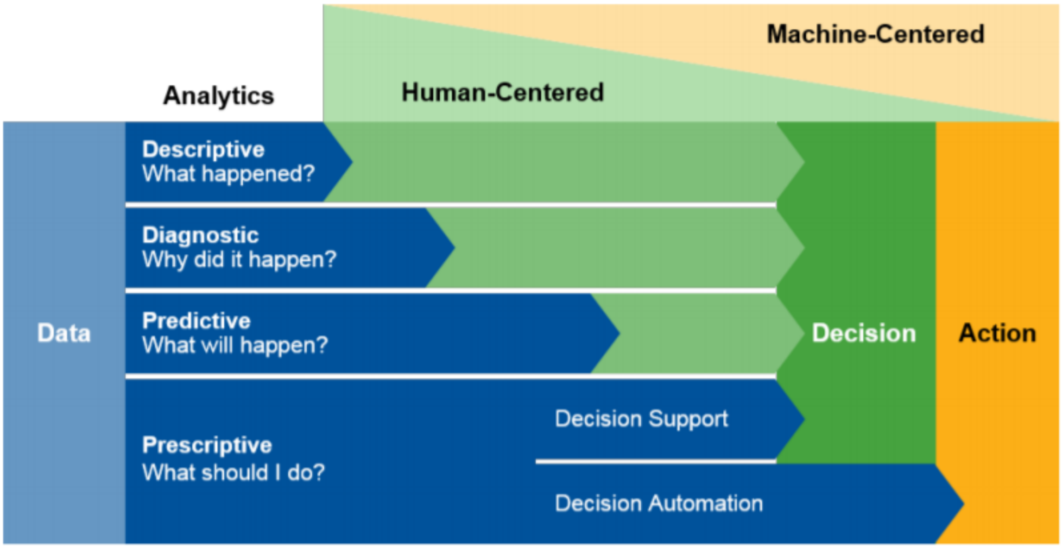
First, some questions will be asked about your background.

1) *Which organization do you work for?*
2) *What role do you have within your organization?*
3) *How many years of experience do you have in your current role?*

II. Definition of Key Constructs

In this thesis, I use the following definition of data-driven decision-making: 'Data-driven decision- making is about performing data analytics (either descriptive, diagnostic, predictive) and use the findings for decision support or even decision automation'. In this thesis there is mainly focused on the combination of predictive and prescriptive data analytics.

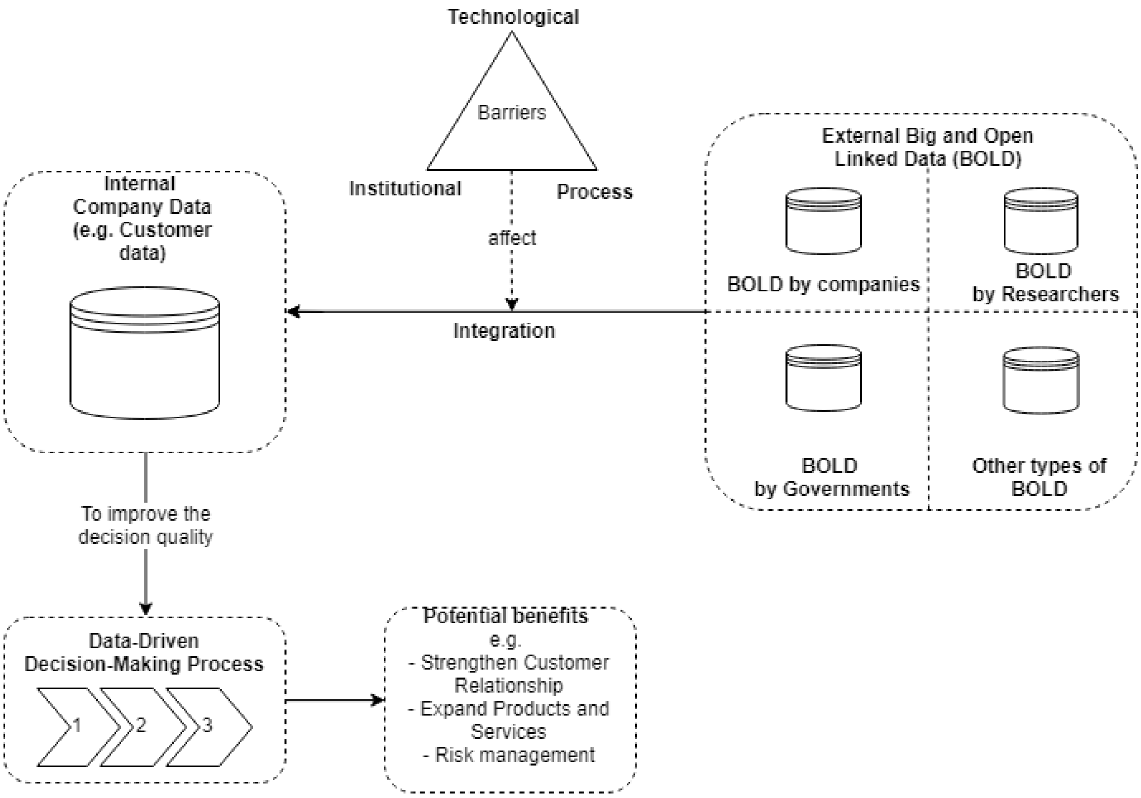
Figure 4. The Four Analytic Capabilities



4) *Does my definition of data-driven decision-making match your definition? Please explain.*
5) *Do you have experience with data-driven decision-making?*
6) *Data-driven decision-making can be made on various quality levels. What does your organization consider high-quality decision-making (fast, well-informed, etc.)?*
7) *Are you familiar with the concept of Big and Open Linked Data (BOLD)?*

III. Adoption of BOLD / Big and Open Data

Interviewer explains what external BOLD and about companies adopting BOLD for their decision-making process.



These questions discuss the experiences of companies (or personal perception) with the adoption of BOLD for data-driven decision-making.

8) Does your organization have ever used external BOLD for data-driven decision-making? (or personal perspective)

If yes:

- *a. Did you encounter problems with the adoption of BOLD?*
- *b. Did you integrate the data in real-time?*
- *c. What does a data-driven decision-making process in your organization look like?*
- *d. Did the integration of external BOLD with company-owned data result in better decision-making?*
- *e. Can you give an example of a data-driven decision?*

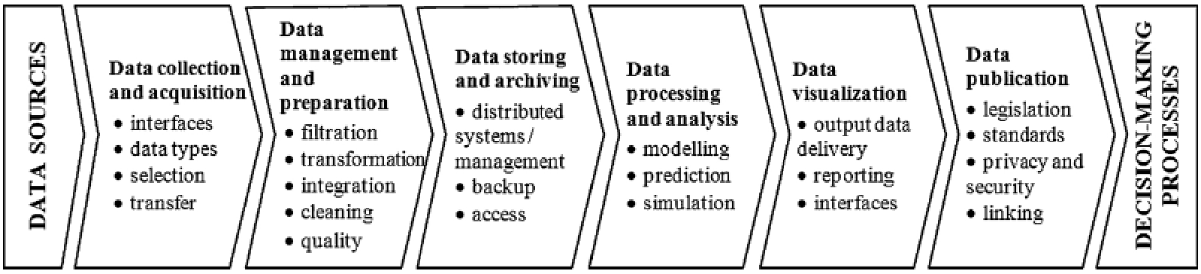
If no:

A1) What does a data-driven decision-making process in your organization look like?

A2) How do you measure the quality of such a decision?

IV. Requirements for an Enterprise Architecture Framework

In the literature review conducted the following data pipeline was identified. This figure represents the BOLD lifecycle from data sources unto the decision-making process. I'm curious if you recognize the different stages from data adoption up to and including the decision-making processes.



Source: (Lnenicka & Komarkova, 2018)

9) Does your organization encounter the same stages as the figure showed? Do you have additional stages/activities?

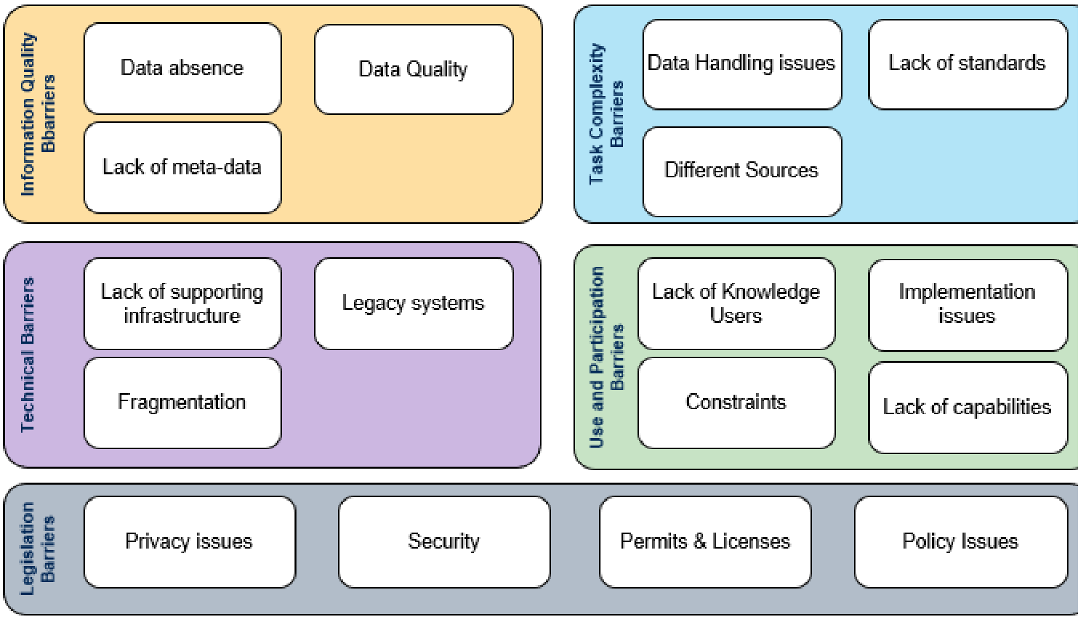
10) What applications / technologies does your organization use to enable these stages?

V. Adoption Barriers

These questions are used for the validation of the problem by stakeholders (So, testing the problem in the environment). Based on a literature review the most common adoption barriers for external BOLD were identified and divided into clusters and sub-clusters.

11) Do you encounter problems with the adoption of BOLD for data-driven decision-making? Which adoption barriers?

This is the Conceptual Model of the Adoption Barriers of BOLD I just mentioned. 6 categories have been identified and the barriers were placed in sub-clusters. Also, I will give you a handout with examples of barriers that have been mapped in each sub-cluster.



13) Can you choose the five most critical barriers for your organization?

14) Which of these barriers have the biggest influence on data-driven decision-making?

VI. Good/best practices to overcome adoption barriers

The following questions aim to collect the common and best practices that are used by companies to overcome the adoption barriers. Next to that, the questions are focused on the collection of requirements. I'm curious what you did to overcome the problems you encountered to successfully integrate BOLD.

Example questions:
15) What does your organization do to deal with the information quality barriers of BOLD? (Data absence, Data quality, and a lack of metadata)
16) What does your organization do to deal with the task complexity barriers of BOLD? (Data handling Issues, Lack of capabilities, Interoperability/flexibility/collaboration issues)
17) What does your organization do to deal with the technical barriers of BOLD? (Lack of supporting infrastructure, Lack of standards, Fragmentation, and legacy systems)
18) What does your organization do to deal with the use and participation barriers of BOLD? (Lack of knowledge users, implementation issues, constraints, Lack of capabilities)
19) What does your organization do to deal with to deal with the legislation barriers of BOLD (Privacy issues, security, Permits & Licenses, Policy issues)?

Thank you for giving the interview. I will listen to the audio recordings and I will use the recordings for coding and transcribing, and qualitative data analysis. The interview results will be shared by e-mail.

Appendix C1: Overview Key Findings Explorative Expert Interviews

Table 26: Overview key findings of the explorative interviews

Category		Key Finding
0.	Adoption Barriers	**Key Finding 1:** Few organizations succeed in real-time integration of external BOLD for decision-making. Most of the organizations are struggling with the adoption barriers that makes the real-time use for decision-making difficult **Key finding 2:** Organizations from both the Public Sector and the Private Sector are addressing adoption barriers with the integration of External Big and Open Linked Data for decision-making. Public organizations are especially struggling with the Legal Barriers, Use and Participation Barriers and the Information Quality Barriers, while Private Organizations are especially struggling with the Task Complexity Barriers, Information Quality Barriers and the Technical Barriers. ***Key finding 3**: Organizations identify an additional adoption barrier category: 'Governance Barriers', consisting of the sub-clusters 'Policy Issues', 'Lack of Acceptance', and 'Scalability Problems'.* ***Key Finding 4:** Most of the adoption barriers are caused by the human factor.* ***Key finding 5:** Data integration from external to internal organizations is one of the biggest struggles of companies. The most successful organizations with BOLD integration are the companies with a clear integration strategy and are capable of sharing data with other organizations.*
1.	Proactive Approach	***Key Finding 6:** A Proactive Approach should be taken to prevent the adoption barriers from occurring and is required to enable the real-time integration of external BOLD for decision-making.*
2.	Collaboration between organizations	***Key Finding 7:** To enable real-time data-driven decision-making and prevent adoption barriers from occurring, an organization should collaborate more closely with its BOLD ecosystem.*
3.	Involvement BOLD Ecosystem	***Key finding 8:** The Enterprise Architecture Framework must involve all stakeholders that are part BOLD Ecosystem, including Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, legal Institutions and Data Prosumer. This are the stakeholders that are concerned with the real-time integration of external BOLD for decision-making into account during the design of the Architecture to create better alignment of the Enterprise with its ecosystem.* ***Key Finding 9:** Current Enterprise Architecture Frameworks are focusing too much on aligning strategy and technology of the internal organization and not on the alignment with the external organization.*
4.	Real-time connection between organizations	***Key finding 10:** Creating a shared platform that can be accessed through APIs will enable real-time data-driven decision-making.* ***Key finding 11:** In the ideal situation, the internal organization is connected to the external organization to make the data integration and exchange easier to enable the creation of new insights and value and overcomes the adoption barriers of Big and Open Linked Data.* ***Key Finding 12:** Continuous and Real-time data streams should be initially integrated into shared platforms to enable real-time decision-making* ***Key Finding 13:** The Enterprise Architecture Framework helps to create an environment in which data can be integrated and shared; integrate different kinds of applications (independent from platform, programming language or resource) which can be linked together; collaboration between distributed and scattered applications; interoperability between different operating systems; security components that makes sure data is only shared with the right resources; orchestration by means of visual guidance by a single interactive user interface. (Hybrid Integration Platform*
5.	Uniformity	***Key finding 13:** The Enterprise Architecture Framework must ensure a uniform way of working within ecosystem of a company, including the common standards.* ***Key finding 14:** Collaboration with other organizations is stimulated by using the same way of working, the same standards within the ecosystem of the organization.*
6.	Governance	***Key finding 15**: Once an architecture is created through collaboration with the ecosystem. It is important that the architecture can keep running by being able to respond to changes in the environment.*

Appendix D1: Design Options

RQ ID	Definition	Design options
RQ1	The Enterprise Architecture Framework must **proactively** deal with the **Information Quality Barriers, Task Complexity Barriers, Technical Barriers, Use & Participation Barriers, Legislation Barriers,** and **Governance Barriers** that organizations address with the real-time integration of BOLD for decision-making.	1. **Create an initial step to assess the need for an Ecosystem Architecture.** 2. Provide an overview of the potential adoption barriers and means to overcome them. 3. Incorporating the means to overcome the adoption barriers of BOLD in the other building blocks of framework
RQ2	The Enterprise Architecture Framework must **stimulate collaboration** between organizations to create mutual commitment of the stakeholders.	1. Present overview of different stimulants that increase the change for collaboration 2. **Present guidelines that help organizations to write a proposal to collaborate discussing mutual incentives, including partnerships, business models, mutual benefits** 3. Provide recommendations on the best ways to stimulate collaboration.
RQ3	The Enterprise Architecture Framework must **involve all stakeholders from the BOLD ecosystem that are concerned** with the real-time integration and sharing of external BOLD for decision-making system in the design of the ecosystem architecture **to create better alignment of the enterprise with its ecosystem.**	**1. Present guidelines that help to develop a shared ecosystem vision that forms the basis of the Architecture requirements** 2. Recommend an Ecosystem Vision that should be used. 3. Give an overview of successful Ecosystem Architecture Visions
RQ4	The Enterprise Architecture Framework **must (build on TOGAF-ADM to)** make the framework capable of **continuously aligning an Enterprise with its Ecosystem, unifying information silos within and between organizations, enabling real-time connection between data providers and data users, increasing the flexibility of an enterprise towards its ecosystem.**	1. **Develop guidelines to develop a suitable Hybrid Integration Platform (additional layer of the enterprise) in which data can be integrated and shared via APIs** 2. Develop guidelines to develop a shared platform used by the whole ecosystem in which data providers initially integrate their data so that data users can use the data for decision-making in real-time. 3. Give an overview of different successful architectures that have these capabilities and let the organization use this to define
RQ5	The Enterprise Architecture Framework must create an **uniform way of working within the organization and ecosystem** by developing Uniform Standards; a **Common Data Model** (One data model; unified schema and semantics; building on existing efforts); **Uniform Meta-data standards** (e.g. technical documentation about APIs); and **Institutions** (action rules) to shape/implement this way of working.	1. Recommend a data model, standards, API Policy and Institutions that organizations must use to create the same way of working. 2. **Present guidelines for the development of a common data model, uniform standards, an independent API Policy and Institutions that build on existing efforts.** 3. Give an overview of different successful organizational models and let the organization choose one of them.
RQ6	The Enterprise Architecture Framework must facilitate the development of an **architecture that can grow organically**, that is **scalable**, **maintainable** and **flexible towards** its **ecosystem**; ensures **continuity**; and allows for **new connections**. (By providing a governance mechanism that facilitates all of this)	**1. Develop guidelines to customize guidelines to create a governance mechanism that keeps the architecture running.** 2. Give an overview of different governance mechanisms 3. Recommend a governance mechanism

Appendix E1: Evaluative Interview Protocol

Enabling the real-time adoption of Big and Open Linked Data for decision-making through the 'Ecosystem Architecture Framework'

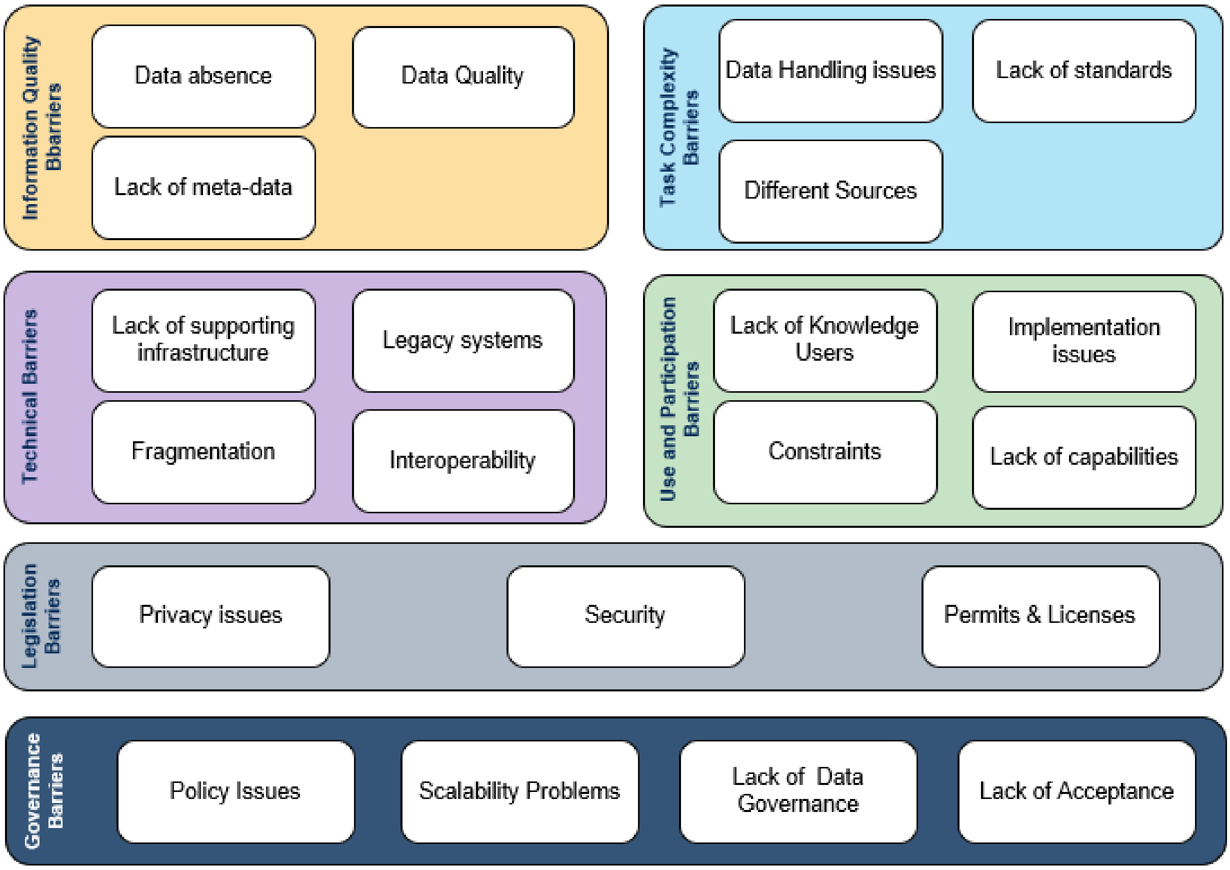
0. Evaluation Criteria / Statements

The evaluation form below can be used to assess the Design Artefact. The purpose of the evaluation interview is to improve the current version of the Design Artefact by collecting additional or revised requirements to develop a revised version of the Design Artefact. The Evaluation form asks questions about 8 design elements, which all correspond to a section of this evaluation document.

Design Element	Evaluation Criteria	Statement	Score	General Feedback/ Remarks
1. Design Purpose	*Clarity*	*S1 'The purpose for the Design of the Ecosystem Architecture Framework to enable real-time integration of Big and Open Linked Data is clear'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
2. Design Requirements	*Clarity*	*S2 'The design requirements that are collected are clear.'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Feasibility	*S3. 'The design requirements that are collected are feasible.'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Completeness	*S4.' The design Requirements collected result in a design that fully solves the research problem.'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
3. Overall Impression Design Artefact	*Clarity*	*S5. 'The Ecosystem Architecture is clear and understandable.'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Completeness	*S6. 'The Ecosystem Architecture is complete and does not lacking important other phases.'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
4. Phase A: Getting the Ecosystem Committed and Involved	*Clarity*	*S7. 'The guidelines provided in phase A are understandable'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Structure	*S8. 'The guidelines of Phase A are structured in a logical way'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
5. Phase B: Getting the Ecosystem Architecture Right	*Clarity*	*S9. 'The guidelines provided in phase B are understandable'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Structure	*S10. 'The guidelines of Phase B are structured in a logical way'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
6. Phase C: Making the Architecture Work	*Clarity*	*S11. 'The guidelines provided in phase C are understandable'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Structure	*S12. 'The guidelines of Phase C are structured in a logical way'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
7. Phase D: Keep the architecture Running	*Clarity*	*S13. 'The guidelines provided in phase D are understandable'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
	Structure	*S14. 'The guidelines of Phase D are structured in a logical way'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	
8. Overall Usability Design Artefact	*Usability*	*S15. 'The Ecosystem Architecture Framework is easy to use'*	*'strongly agree – agree – neutral – disagree – strongly disagree'*	

1. Design Purpose

Chief Information Officers (CIOs) of organizations are adopting external Big and Open Linked Data (BOLD) to strengthen their real-time data-driven decision-making. The adoption of external BOLD could support organizations with their decisions relating to a variety of problems and opportunities, such as strengthening customer relationship and creating a better product-market fit. Unfortunately, the adoption of BOLD addresses several adoption barriers. During a literature review and explorative expert interviews, a conceptual model of the different adoption barriers was developed, depicted in the figure below. The adoption barriers are categorized and mapped in clusters to create a complete overview of the problems that occur.



The goal of this thesis is to build an Enterprise Architecture Framework that deals with the adoption barriers of BOLD to enable real-time decision-making. During the explorative interviews and Literature Review (1) the ideal situation was discussed to prevent the adoption barriers from occurring and (2) what capabilities an Enterprise Architecture Framework should have to enable real-time data-driven decision-making with BOLD. This has resulted in the list of Enterprise Architecture Framework Capabilities that is depicted in the table below. The most commonly used Enterprise Architecture Frameworks (EAFs) were assessed on the possession of these capabilities. None of the Enterprise Architecture Frameworks have all the capabilities. This resulted in a design gap, which is highlighted in red. TOGAF seems the most suitable EAF for extension.

EA Capabilities		TOGAF
C1	Continuously align Business / IT:	++
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	++
C3	Remove the threshold for change	++
C4	Enable/Increase flexibility of the organization	+

C5	Develop a proactive organization	++
C6	Reduce risk and prepare for rapid, unplanned change	++
C7	Avoid tension between business IT functions in organization	++
C8	Create, unify and integrate business processes across enterprise	++
C9	Unlock the power of information, unifying information silos in organization	++
C10	Eliminate duplicate and overlapping technologies in organization	++
C11	Reduce solution delivery time, development costs (maximize reuse):	++
C12	Continuously align Enterprise with its Ecosystem	-
C13	Unlock the power of information by unifying information silos between organizations	-
C14	Enable a real-time connection between data providers and data users.	-
C15	Enable/Increase Flexibility of Architecture towards ecosystem.	-
C16	Develop an Enterprise that can grow organically	-
C17	Develop uniform rules, standards, principles, models between organization and ecosystem	-

Based on this design gap analysis, there was concluded that the foundation of Enterprise Architecture Framework is no longer suitable for strategies, such as the real-time integration of BOLD for data-driven decision-making. A design gap was identified of EAF capabilities that an EAF should have to overcome the research problem.

2. Design Requirements

The key findings of the Literature Review and Explorative interviews are translated into requirements for the design of the new Enterprise Architecture Framework, which are displayed in the table below (This is worked out in detail, but too long to put in this document. If you are interested, the chapter can be sent upon request).

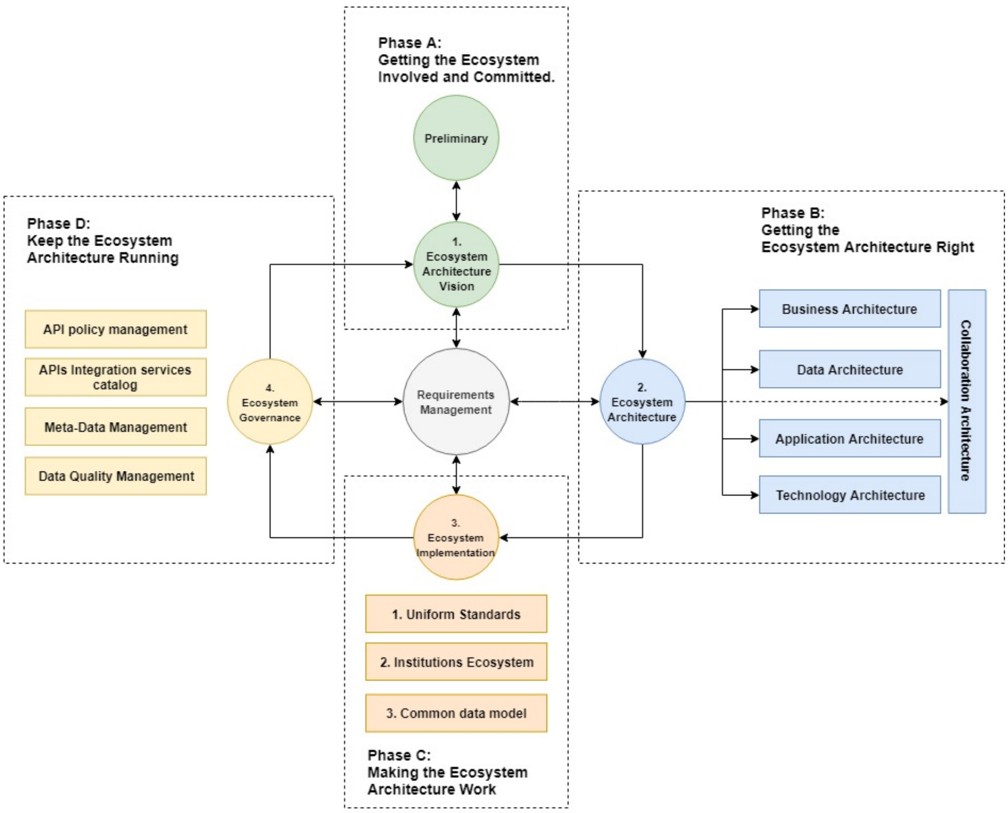
RQ ID	Definition	Included in Design (next section)
RQ1	The Enterprise Architecture Framework must **proactively** deal with the **Information Quality Barriers, Task Complexity Barriers, Technical Barriers, Use & Participation Barriers, Legislation Barriers,** and **Governance Barriers** that organizations address with the real-time integration of BOLD for decision-making.	All Phases
RQ2	The Enterprise Architecture Framework must **stimulate collaboration** between organizations to create mutual commitment of the stakeholders, by developing a *shared vision, including partnerships, business models, mutual incentives / benefits, integration strategy, etc.*	Phase A
RQ3	The Enterprise Architecture Framework must **involve all stakeholders from the BOLD ecosystem that are concerned** with the real-time integration and sharing of external BOLD for decision-making system in the design of the ecosystem architecture **to create better alignment of the enterprise with its ecosystem**, including Ecosystem Orchestrator, Service providers, Application provider, Data Producer, Data Publisher, Data User, Data Prosumer and Legal Authorities.	Phase A
RQ4	The Enterprise Architecture Framework **must (build on TOGAF-ADM to)** make the framework capable of **continuously aligning an Enterprise with its Ecosystem, unifying information silos within and between organizations, enabling real-time connection between data providers and data users, increasing the flexibility of an enterprise towards its ecosystem.**	Phase B

RQ5	The Enterprise Architecture Framework helps to create an **environment** in which **data can be integrated and shared**; **integrate different kinds of applications** (independent from platform, programming language or resource) which can be linked together; **collaboration between distributed and scattered applications**; **interoperability** between different operating systems; **security** components that makes sure data is only shared with the right resources; **orchestration** by means of visual guidance by a single interactive user interface. (**Hybrid Integration Platform**)	Phase B
RQ6	The Enterprise Architecture Framework must create an **uniform way of working within the organization and ecosystem** by developing Uniform Standards; a **Common Data Model** (One data model; unified schema and semantics; building on existing efforts); **Uniform Meta-data standards** (e.g. technical documentation about APIs); and **Institutions** (action rules) to shape/implement this way of working.	Phase C
RQ7	The Enterprise Architecture Framework must facilitate the development of an **architecture that can grow organically**, that is **scalable**, **maintainable** and **flexible towards** its **ecosystem**; ensures **continuity**; and allows for **new connections**. (By providing a governance mechanism that facilitates all of this)	Phase D

3. The Design Artefact

Based on the requirements, the Ecosystem Architecture Framework (Eco-AF) has been developed. The Ecosystem Architecture Framework is an extended Enterprise Architecture approach that builds further on the foundations of Enterprise Architecture. To avoid reinventing the wheel it can be seen as an extension to the commonly used TOGAF-ADM. The Ecosystem Architecture Framework provides guidance for organizations to implement strategies that involve the real-time adoption of external Big and Open Linked Data. It does not only help with the implementation of a strategy within an Enterprise, but also aligns the strategy with the ecosystem of an Enterprise.

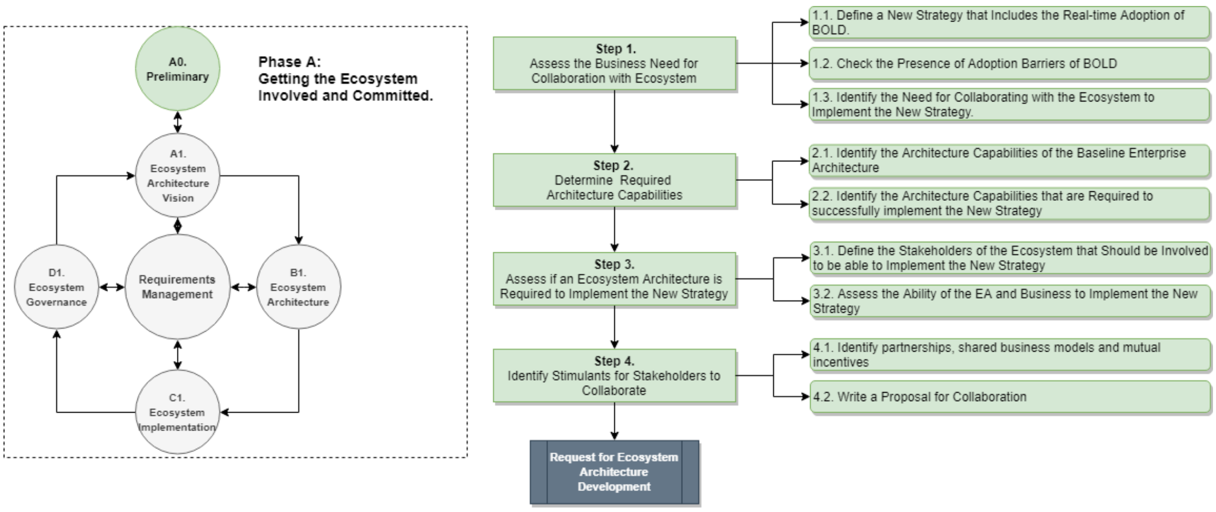
The figure below presents a high-level overview of the Ecosystem Architecture Framework that is the output of the Design Phase. The Ecosystem Architecture Framework (EsAF) is structured into four phases, displayed as Phase A, B C, and D.



4. Phase A: Getting the Ecosystem Involved and Committed

A0. Preliminary phase

The preliminary phase represents the first requirement 'The Framework must Proactively deal with the Adoption Barriers of BOLD.' In this phase, an organization notices that a strategy that includes the use of external BOLD for data-driven decision-making addresses multiple adoption barriers. The Ecosystem Architecture Framework helps the organization to identify the business need for collaboration with data providers within the ecosystem to prevent the adoption barriers from occurring. It helps the organization to assess the need for an Ecosystem Architecture in which the architectures of data providers are initially connected with the Enterprise Architecture of the organization to enable the real-time adoption. Secondly, the requirement 'The Framework must stimulate collaboration between data providers and data users.' In this phase, different stimulants for a successful collaboration are identified, such as a partnerships, shared business models, mutual incentives. This phase is mainly about creating awareness for the mutual and commitment for the shared development of an Ecosystem Architecture. Furthermore, an assessment is performed of what architectural changes are required to make the collaboration work. The preliminary phase consists of three different core steps, which are explained in detail below. The figure below presents the guidelines of the preliminary phase of the Ecosystem Architecture Framework.

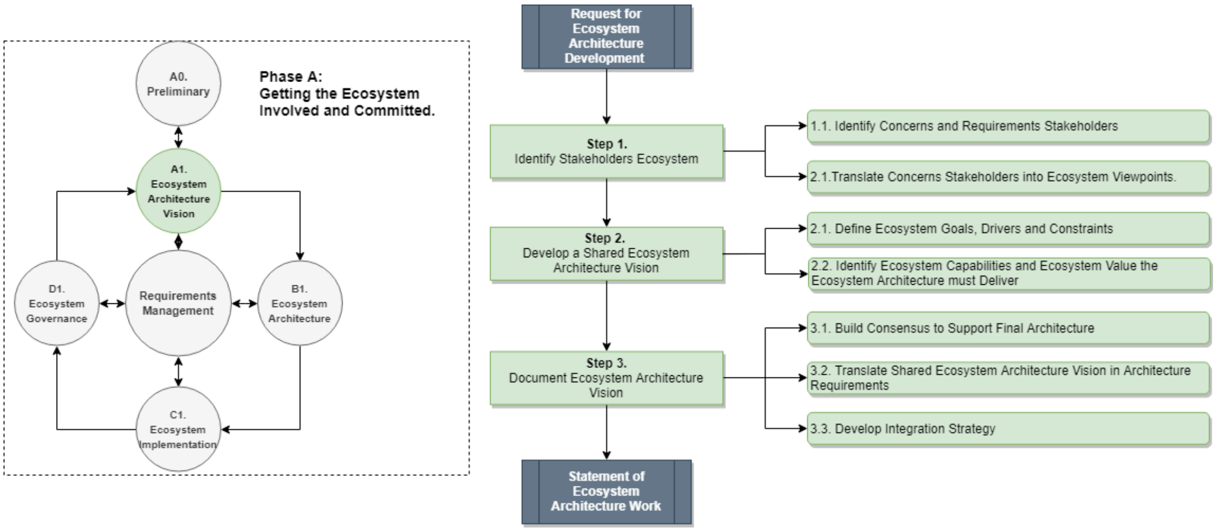


A1. Ecosystem Architecture Vision

In the Preliminary Phase, the need for an Ecosystem Architecture was identified and a request was made to stakeholders of the ecosystem to collaborate. The Ecosystem Architecture Vision Phase represents the requirement 'The Framework must involve the stakeholders the business ecosystem that are concerned with the system to create better alignment between the enterprise and the ecosystem.' In this phase, the stakeholders of the business ecosystem are involved in the development process of the Ecosystem Architecture. In order to create better alignment between the stakeholders that are concerned with the system, the concerns and requirements of each stakeholder are collected. Each stakeholder has an idea of how the system should look like. Hence, it is important that all the stakeholders are involved in and committed to the design process.

To ensure alignment between the enterprise and the ecosystem a shared ecosystem architecture vision is create. Developing this vision together ensures that the concerns of each stakeholders are represented in the architecture. A vision consists of the goals of the ecosystem, the drivers and constraints for an ecosystem architecture, and the capabilities and business value the architecture should deliver. Finally, the Ecosystem Architecture Vision is documented by building consensus about the final architecture, translating the vision into architecture requirements, and the development of an integration strategy (Part of Viewpoint?).

The Ecosystem Architecture Vision is structured into three steps, presented in the figure below. First, the Stakeholders of the Ecosystem are identified. Secondly, the Ecosystem Architecture Vision is developed. Finally, the Ecosystem Architecture Vision is documented and translated into architecture requirements. These steps are explained in detail below.



5. Phase B: Getting the Ecosystem Architecture Right

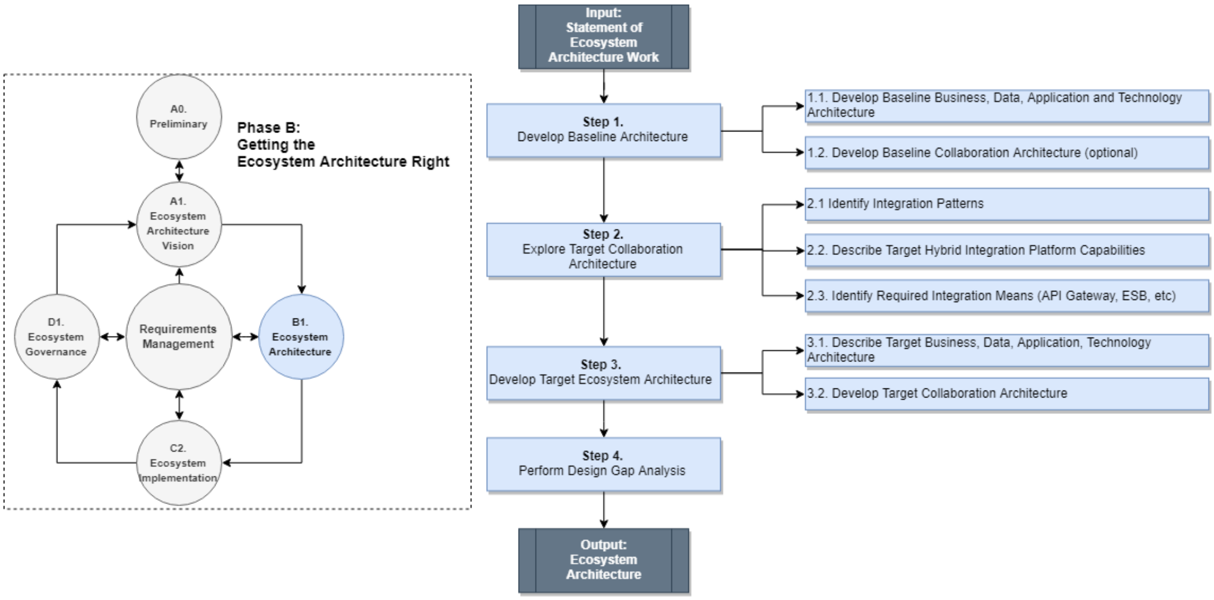
B1. Ecosystem Architecture

Building on the 'Statement of Ecosystem Architecture Work', the Ecosystem Architecture is developed in phase B1. The figure below presents guidelines for the development of the target Ecosystem Architecture.
The Ecosystem Architecture builds on the foundations of Enterprise Architecture, which normally includes a business architecture, data architecture, application architecture, and technology architecture that are representing the viewpoints of different sets of stakeholders. The Ecosystem consists of an extra view the 'Collaboration Architecture'.

As mentioned earlier, this framework is meant as an extension to the existing and commonly used TOGAF-ADM. The TOGAF-ADM provides clear guidelines for the transformation of the baseline business, data, application, and technology architectures into the target architectures that are suitable for implementing a new strategy. The same techniques are adopted in the Ecosystem Architecture Framework, but the focus of this phase is mainly making sure the collaboration architecture is developed and that the other architectures are well-aligned with this layer.

The collaboration architecture is based on the viewpoints of the stakeholders that are part of the ecosystem and represents an integration platform. This is an environment in which external data can be integrated and shared; different kinds of applications can be integrated and linked together (independent from platform, programming language and resource); collaboration between distributed and scattered applications is possible; interoperability between different operating systems is possible; security components ensure rightful data sharing and data use; and an orchestration mechanism provides visual guidance by a single interactive user interface.

The guidelines of for the development of the Ecosystem Architecture are structured in four steps: the Development Baseline Architecture; The exploration of the Target Collaboration Architecture; The development of the target Ecosystem Architecture; and performing the gap analysis. The different steps are explained in detail in the section below.

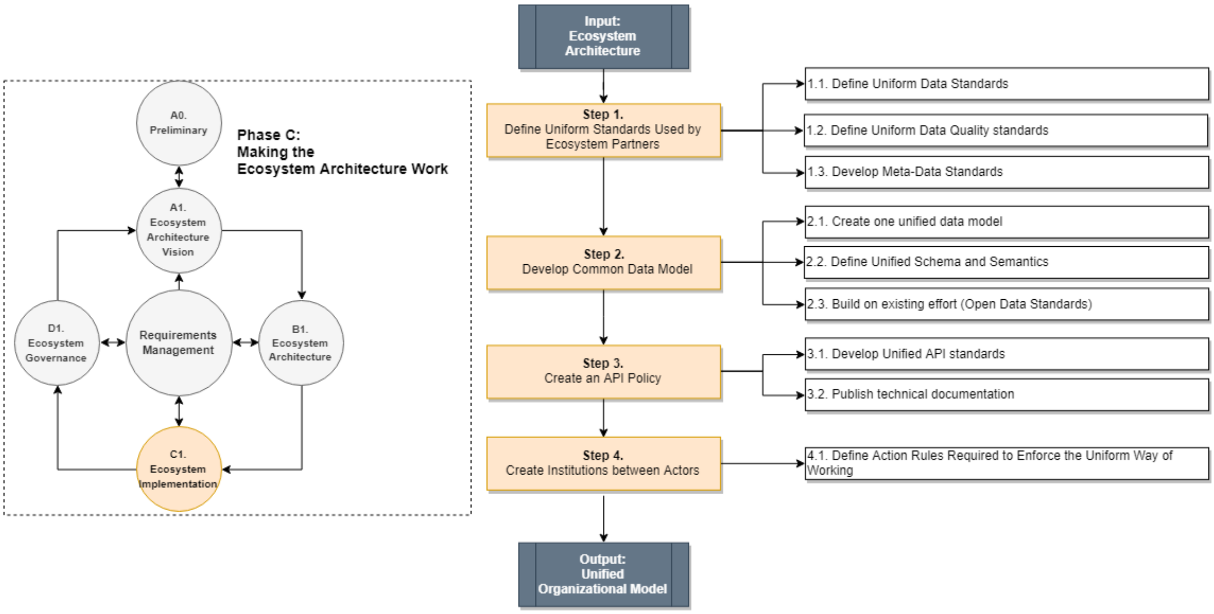


6. Phase C: Making the Ecosystem Architecture Work

C1. Ecosystem Implementation

In phase C1, guidelines are provided to create a unified organizational model that implements a common way of working to ensure the Target Ecosystem Architecture (Output of Phase B1) can function optimally and no adoption barriers will occur. If a design is not used in the way where it was designed for, new problems occur. For example, if you buy a Nespresso machine and you put Senseo pats in, the machine does not function optimal and it does not deliver the intended product. This phase represents the requirement 'The Enterprise Architecture Framework must create a uniform way of working within the organization and ecosystem'. During the explorative interviews was discovered that this uniform way of working can be ensured by developing Uniform Standards; a Common Data Model (One data model; unified schema and semantics; building on existing efforts); Uniform Meta-data standards (technical documentation about APIs); An independent API Policy; and Institutions (action rules) to shape/implement this way of working.

The figure below presents an guidelines for the creation of an organizational model that ensures the Target Ecosystem Architecture is used in the right way. It makes sure that the architecture works. The overview consists of four steps. First, the frameworks help defining uniform standards that should be used by all ecosystem partners. Secondly, it focuses on developing a common data model. Thirdly, it provides guidance in the creation of an independent API policy that can be followed by any organization to develop APIs. Finally, institutions are created between the actors to make sure data and the architecture stays sustainable. The steps are explained in detail below.



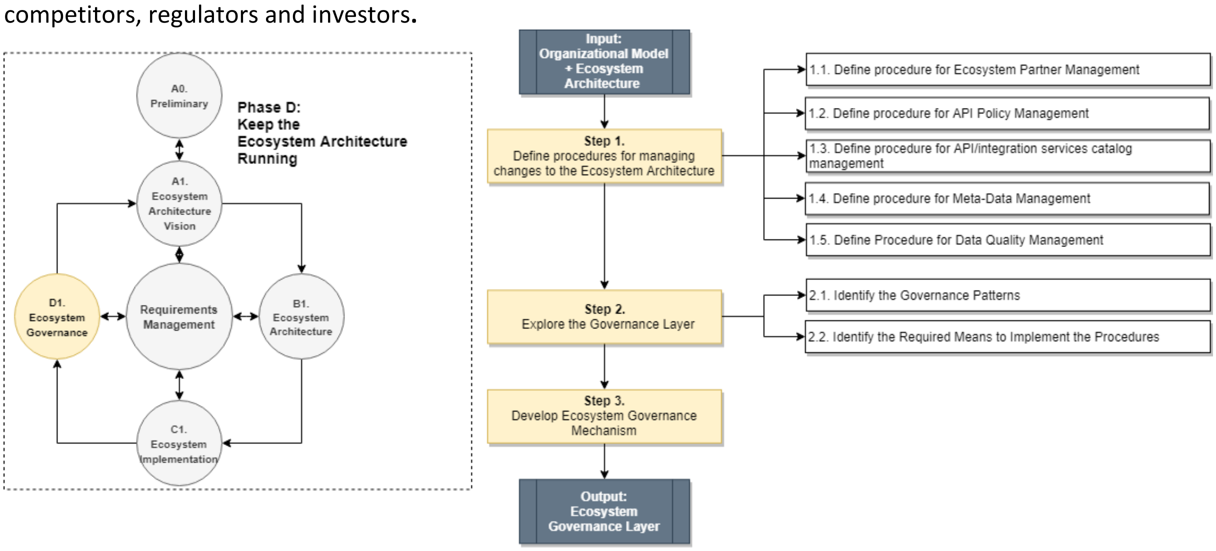
7. Phase D: Keep the Ecosystem Architecture Running

<u>D1. Ecosystem Governance</u>

Phase D1 focuses on creating a governance mechanism that keeps the Ecosystem Architecture running, but even more important keeps the Ecosystem Architecture work properly and aligned with the Ecosystem Architecture Vision. The Ecosystem Governance Phase deals with the requirement 'The Framework must facilitate the development of an architecture that can grow organically, that is scalable, maintainable and flexible towards its ecosystem; ensures continuity; and allows for new connections. Secondly, it fulfils the requirement 'The Enterprise Architecture Framework provide a governance mechanism of the Ecosystem Architecture, including API policy management; APIs and other integration services catalog management; meta-data management; and data quality'.

On a high level, this is what the Ecosystem Governance Phase is about. Any changes in the aforementioned aspects have a direct impact on the Ecosystem Architecture and should be updated to make sure that the architecture is compliant and delivers the business value that is required. To make the architecture scalable, maintainable, flexible and continuous, the Governance phase creates a governance mechanism that manages all these aspects. The figure below presents guidelines for Ecosystem Governance.

This phase has builds further on foundation of Enterprise Architecture, that Ensures that the Enterprise Architecture responds to the need for change. It provides governance guidelines a dynamic environment, consisting of different factors related to the demand of customers, markets, industries, opportunities, competitors, regulators and investors**.**



Enabling the real-time integration of Big Open Linked Data (BOLD) for decision-making

Warner Steinfort – TU Delft 4717457

Faculty of Technology, Policy and Management – Delft University of Technology, 2628BX Delft, NL

Abstract

Research shows that an increasing number of CIOs of organizations decide to adopt external Big and Open Linked Data (BOLD) in real-time to support their decisions. The adoption of external BOLD offers a lot of potential for the support of decisions on a variety of problems and opportunities, However, the adoption of external BOLD addresses several adoption barriers that hinder the real-time adoption. The presence of adoption barriers indicate that there is a lack of alignment between organizations. A common approach to align a new strategy and its implementation across the enterprise is Enterprise Architecture (EA). However, the foundation of EA does not seem capable of creating alignment between organizations. Through a systematic literature review and explorative expert interviews (1) the different adoption barriers that hinder the real-time adoption of BOLD are identified; and (2) the EA capabilities that must be incorporated in an Extended Enterprise Architecture Framework to overcome these adoption barriers are identified. This paper provides design principles for an EAF should have to successfully enable the real-time integration of external BOLD in real-time for decision-making. For future research, it is recommended to do an in-depth study on the specific capabilities; and to investigate the different design patterns for incorporating the capabilities.

Key words: Open Data; BOLD; Decision-making; Adoption; Barriers; Ecosystem Architecture Framework

1. Introduction

CIOs of both public and private organizations are searching for new ways to support better outcomes of their data-driven decision-making (Wang & Lo, 2019). As linking external data with internal company data can result in interesting new insights and value, CIOs are increasingly developing IT strategies that involve the real-time use of external data (Wang & Lo, 2019). One of the innovations that is being adopted in real-time to enable better informed decision-making is external Big and Open Linked Data (BOLD) (Lněnička & Komárková, 2018). Combining internal company data silos with external BOLD in real-time helps in the creation of these real-time insights, such as new relationships between data subjects, patterns in datasets, or context to existing data that was not yet useful. These new insights can be used for better decision-making on a variety of problems and opportunities, such as strengthening customer relationship, better pricing, building a better product market fit, expanding products and services, and managing risk (Bayrak, 2015; Dutta & Bose, 2015; Saggi & Jain, 2018; Sheng et al., 2017; Wang & Lo, 2019).

Despite the effort of organizations, the adoption of external BOLD addresses multiple adoption barriers that hinder the real-time integration with internal company data, such as a wide variety in size, different sources and unstructured data (Lněnička & Komárková, 2018); different flexibility, interoperability and collaboration issues (M. Janssen et al., 2014); and negative effects like violating privacy, possible misuse and misinterpretation of data (Zuiderwijk & Janssen, 2014). To overcome these barriers, a solution is required that enables adoption of BOLD in real-time. Figure 67Figure 1 visualizes the current situation. The figure depicts that during the integration process several adoption barriers are addressed, which hinder the creation of real-time insights for decision-making.

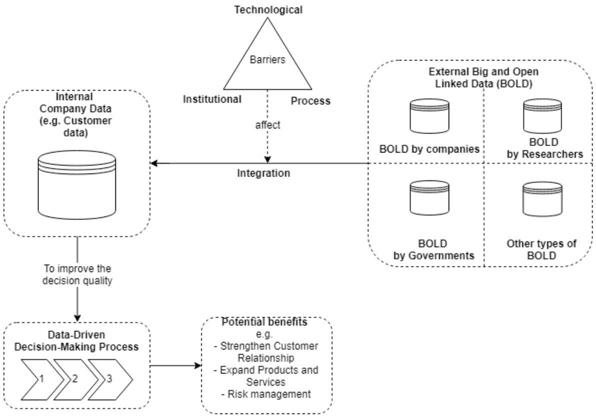


Figure 67: Current Situation

Organizations seem to have difficulties with taking advantage of this Big and Open Linked Data innovation, because organizations often do not have of the right technical and business capabilities to extract the data from the external data silos, process this data and use data analytics for drawing conclusions in real-time for decision-making (Gong & Janssen, 2017). Often an organization's existing ICT landscape is not flexible enough to acquire and implement a new IT strategy that involves the real-time adoption of external BOLD (Denstad & Bygstad, 2012; Gong & Janssen, 2017; Moreno Jr et al., 2009). Management-by-magazine seems to play a negative role, as the management of an organization decides to adopt a new technology or innovation without considering whether the rest of the organization can adapt to this new strategy (Gong & Janssen, 2017). The real-time integration of external BOLD does not only require alignment of business and IT across the internal organization, but also between the internal and external organization. The presence of adoption barriers show that the data users and data providers are not well-aligned with each other (Lněnička & Komárková, 2018). It is a big challenge for organizations to deal with these adoption barriers. Therefore, organizations are searching for an approach that helps to successfully implement a strategy that involves the real-time adoption of external BOLD and that mitigates the adoption barriers.

A common practice that is often used to implement a new IT strategy across the entire enterprise is Enterprise Architecture (EA). 'Enterprise Architecture' is generally used for aligning IT strategy and its implementation and to bridge the gap between the business and technology level of an enterprise (Rehman & Shamail, 2014). Furthermore, it describes the present and the future state of the business and can be used as a roadmap to transition from the present to the future state (Janssen, 2009). The real-time adoption of external BOLD requires the transition from a system that uses data and analytics within internal data silos towards a system in which internal and external data silos are combined to enable the creation of new useful insights and value (Lněnička & Komárková, 2018). Such a strategy requires (1) alignment across the business, data, application and technology level of the enterprise; and (2) alignment with the organizations that are involved with the real-time integration of BOLD. Nowadays, countless Enterprise Architecture Frameworks (EAFs) are in existence that allow for the integration of new technologies (Schekkerman, 2004). However, there is a lack of insight whether EAFs exist that can be used to create alignment between organizations. Moreover, the foundation of EA does not seem capable of aligning an organization with external organizations (Minoli, 2008; Schekkerman, 2004). It seems that an extension is required to the foundation of EA to enable the real-time adoption of external BOLD. Lněnička and Komárková (2018) state that limited EAFs exist that can be extended to satisfy the requirements of BOLD. Hence, this paper explores the EA capabilities that must be incorporated in the design of an EAF to enable the real-time integration of BOLD. This paper is answering the following research question:

How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?

This paper is structured into five parts. The first section discusses the research problem and the research question. The second section discusses the research approach taken to be able to answer the research question. The third section discusses the research background, including a discussion of the literature reviews and explorative expert interviews on the adoption barriers of BOLD and the foundation of EA. Also, it discusses the requirements that must be fulfilled to solve the research problem. The last section draws the conclusions of the paper and provides an answer to the research question. Also, recommendations for future research are proposed.

2. Research Approach

This paper uses the design approach by Johannesson and Perjons (2014) to develop design principles for an extended enterprise architecture framework enables the real-time adoption of BOLD for decision-making. By following the guidelines by Kitchenham (2004), a systematic literature review on the adoption barriers of BOLD is conducted in order to explicate the problem. Secondly, a literature review and explorative interviews are conducted on what EA capabilities must be incorporated by the design of an extended EAF to mitigate the adoption barriers of external BOLD and enable the real-time integration.

3. Results

3.1. Adoption Barriers of Big and Open Linked Data

Figure 68 presents the conceptual model of the adoption barriers for the real-time integration of BOLD. Through of a systematic literature review and explorative expert interviews the adoption barriers that hinder the real-time adoption of BOLD are identified. This are the factors that create a lack of alignment between organizations. The research shows that organizations encounter six categories of adoption barriers: Information Quality Barriers; Task Complexity Barriers; Technical Barriers; Use & Participation Barriers; Legislation Barriers; and Governance Barriers (Conradie & Choenni, 2014; Janssen et al., 2012; Saxena & Muhammad, 2018). Each category can be divided into sub-clusters representing all addressed adoption barriers. Each identified adoption barrier was mapped into one of these sub clusters.

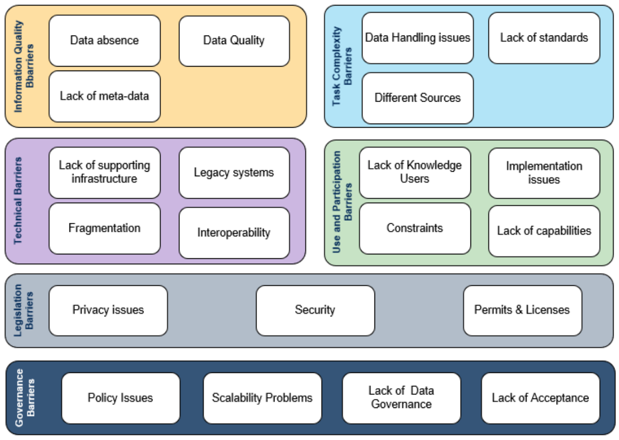


Figure 68: Conceptual Model for the adoption Barriers of BOLD

The rationale for creating a conceptual model of the different adoption barriers that organization encounter, is to explicate the research problem. A problem can only be solved when the factors causing the problem are known. The wide range of adoption barriers identified through the literature review and expert interviews show that organizations are struggling with the real-time adoption of BOLD. A wide range of factors is causing the problem. The presence of adoption barriers show that organizations are not well aligned across and between organizations.

3.2. Foundation of Enterprise Architecture

The foundation of EA is assessed on its ability to implement a new strategy that involves the real-time adoption of BOLD for decision-making (Table 27). The Literature review and explorative expert interviews show that the current theories and methods, expertise and experience, and meta-artefacts related to EA are missing the capabilities to enable the real-time adoption of BOLD (Schekkerman, 2004). This shows that the knowledge base and the current foundation EA must be extended.

The explorative expert interviews show that additional EAF capabilities are required to enable the real-time adoption of BOLD: An Architecture Framework Extension must proactively deal with the adoption Barriers; stimulate collaboration between organizations; involve stakeholders concerned with the adoption of BOLD; enable a real-time direct connection between organizations; ensure uniformity within the ecosystem; and ensure the architecture is scalable, maintainable and sustainable. An Extended Enterprise Architecture Framework that fulfills these requirements can enable the real-time adoption of BOLD.

Table 27: Design Gap Analysis Capabilities EA

EA Capabilities		TOGAF
C1	Continuously align Business / IT:	Yes
C2	Provide a process for transition from the 'as-is' towards the 'to-be' state of organization	Yes
C3	Remove the threshold for change	Yes
C4	Enable/Increase flexibility of the organization	Yes
C5	Develop a proactive organization	Yes
C6	Reduce risk and prepare for rapid, unplanned change	Yes
C7	Avoid tension between business IT functions in organization	Yes
C8	Create, unify and integrate business processes across enterprise	Yes
C9	Unlock the power of information, unifying information silos in organization	Yes
C10	Eliminate duplicate and overlapping technologies in organization	Yes
C11	Reduce solution delivery time, development costs (maximize reuse):	Yes
C12	Continuously align an Enterprise with its Ecosystem	-
C13	Stimulate collaboration between organizations	-
C14	Enable a direct real-time connection between data providers and data users.	-
C15	Enable/Increase Flexibility of Enterprise towards its Ecosystem. Develop an Enterprise that is scalable, maintainable, sustainable and can grow organically	-
C17	Develop uniformity within Ecosystem	-

The most commonly used EAFs are assessed on the possession of these capabilities. The TOGAF-ADM seems the most suitable EAF for extension (Sessions, 2007; Urbaczewski & Mrdalj, 2006). The TOGAF-ADM can be used for aligning across the enterprise, but not between enterprises. There is built on existing efforts to avoid re-inventing the wheel.

4. Conclusions

The main research question of this paper was:

How can organizations enable the real-time integration of external Big and Open Linked Data with internal company data for decision-making?

CIOs of organizations see the potential of adopting external BOLD to strengthen their decisions and are increasingly developing IT strategies that take advantage of this innovation. This increase shows that there is a shift from an internal way of working towards a way of working in which the business ecosystem is involved. This new way of working goes in hand with the presence of adoption barriers, which indicate that there is a lack of alignment across and between organizations. Adoption barriers related to the information quality (data absence, data quality, a lack of meta-data) , task complexity (task handling issues, lack of standards, different sources), technology (Lack of supporting infrastructure, legacy systems, data fragmentation), use & participation (Lack of knowledge, implementation issues, constraints, lack of acceptance), legislations (privacy, security issues and license issues) and governance (Policy issues, Scalability Issues, Lack of Data Governance) hinder the real-time integration of BOLD. If organizations want to use external BOLD for decision-making in real-time, no adoption barriers can be present. The common approach to align a new strategy and its implementation across the enterprise is Enterprise Architecture. Yet, this study demonstrates that the external mindset of organizations demands alignment between organizations and that the current foundation of EA is not capable of satisfying this demand. The approach is limited in breadth, depth, time period and architecture domains.

To adapt to this external mindset a general transformation from a system of record to a system of engagement is required, which is key part of the paradigm shift that comes with an organization's digital transformation. Modern organizations should embrace the system of engagement, including the company's ecosystem as part of their Enterprise Architecture Framework. Hence, a new approach is required that creates alignment between an enterprise and its ecosystem. This study demonstrates that the foundation of EA must be extended with at least six design principles to enable the real-time integration of external BOLD with internal company data.

First, when developing an Enterprise Architecture organizations must embrace a proactive approach to mitigate the adoption barriers of BOLD. This study shows that most organizations embrace a reactive approach that includes the investment in new tooling, personnel or knowledge once an adoption barrier is addressed. However, it is recommended that organizations to embrace a proactive approach that initially deals with the potential adoption barriers during the development of the architecture. This will ensure in the development that is adapted to the adoption barriers that can occur.

Secondly, to ensure adoption barriers between organizations are mitigated the data providers and data users must collaborate with each other. This can result in agreements that benefit the adoption process and to identify potential (causes of) adoption barriers. Through collaboration, organizations can ensure that both parties are aligned with each other and a system can be created that enables the real-time integration of BOLD. However, cooperation must come from two sides, so it is important that there are mutual incentives to mitigate the adoption barriers. BOLD providers will have a more dominant role in the agreements that are made for the collaboration. Organizations must focus on creating incentives for the data provider to cooperate and to form strategic alliances.

Thirdly, organizations must involve the stakeholders concerned with the real-time adoption of BOLD in the design of the system that facilitates the data sharing. Stakeholders that should be involved are the Ecosystem Orchestrators, Service providers, Application providers, Data Producers, Data Publishers, Data Users, Data Prosumers and Legal Authorities. Each stakeholder has concerns about how the system should look like and what requirements the system should satisfy. When stakeholders related to potential adoption barriers are involved during the design there can be anticipated on prevention of these barriers. Creating a shared vision of how a system that facilitates the real-time data integration should look like stimulates the alignment between organizations. The shared vision can be translated into architectural requirements for the design of the system.

Fourthly, organizations must create an architecture that is initially linked to the architecture of the external data sources through a real-time and direct connection. With the current technology available, there is no need for storing the external data locally anymore. By using real-time integration services, such as APIs, the external data does can be called upon request in real-time to make data-driven decisions. This study demonstrates organizations need to invest in a hybrid integration platform that organizations can easily integrate and share integration services (e.g. APIs), different applications can be linked together, collaboration between applications is possible, and orchestration of the applications by a single interface. Different types of hybrid integration platforms are possible, but this paper demonstrates that an API Management Platform is most suitable.

Fifthly, even when an organization is technically capable of integrating external integration services, such as APIs, inter-organizational and shared artefacts are required that make the system work as it is intended. This are the artefacts that make sure the same way of working is implemented within the business ecosystem. This can be achieved by developing uniform standards for data, meta-data, technical information about integration services; A common data model that ensures one data model, unified schema and semantics are used within the business ecosystem; An independent API policy that ensures that each external data provider can be internally linked to the architecture by conforming to this policy; and institutions between the organizations that enforce the same way of working to enable sustainability of the data.

Lastly, the system should be able to respond to changes of the environment. Therefore, it is important that organizations develop a governance mechanism that ensures the system is scalable, maintainable and sustainable. It is important that the inter-organizational and shared artefacts of the ecosystem are managed and adapt to changes of the environment. A governance mechanism is required to ensure the system is utilizing the maximum business value at all time.

Conforming to these design principles when developing a system that uses external BOLD in real-time for decision-making will ensure that there is alignment between organizations and adoption barriers are mitigated. For future research, it is recommended to do an in-depth study on the specific capabilities to investigate the different design patterns for incorporating the capabilities.

5. References

Bayrak, T. (2015). A review of business analytics: a business enabler or another passing fad. *Procedia-Social and Behavioral Sciences, 195*, 230-239.

Conradie, P., & Choenni, S. (2014). On the barriers for local government releasing open data. *Government Information Quarterly, 31*, S10-S17.

Denstad, H., & Bygstad, B. (2012). Managing the it Alignment Gap in Turbulent Times—An Inside View. *Journal of Information Technology Case and Application Research, 14*(2), 28-46.

Dutta, D., & Bose, I. (2015). Managing a big data project: the case of ramco cements limited. *International Journal of Production Economics, 165*, 293-306.

Dym, C. L., Little, P., Orwin, E. J., & Spjut, E. (2009). *Engineering design: A project-based introduction*: John Wiley and sons.

Gong, Y., & Janssen, M. (2017). *Enterprise Architectures for Supporting the Adoption of Big Data.* Paper presented at the Proceedings of the 18th Annual International Conference on Digital Government Research.

Janssen. (2009). Framing Enterprise Architecture: A metaframework for analyzing architectural efforts in organizations. *Coherency Management: Architecting the Enterprise for Alignment, Agility and Assurance*, 107-126.

Janssen, Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers and myths of open data and open government. *Information Systems Management, 29*(4), 258-268.

Janssen, M., Estevez, E., & Janowski, T. (2014). Interoperability in big, open, and linked data-organizational maturity, capabilities, and data portfolios. *IEEE Computer, 47*(10), 44-49.

Johannesson, P., & Perjons, E. (2014). *An introduction to design science*: Springer.

Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University, 33*(2004), 1-26.

Lněnička, M., & Komárková, J. (2018). *Big and open linked data analytics ecosystem: Theoretical background and essential elements*.

Minoli, D. (2008). *Enterprise architecture A to Z: frameworks, business process modeling, SOA, and infrastructure technology*: Auerbach Publications.

Moreno Jr, V. d. A., de Souza Costa Neves Cavazotte, F., & de Oliveira Valente, D. (2009). Strategic Alignment and Its Antecedents: A critical analysis of constructs and relations in the international and Brazilian literature. *Journal of Global Information Technology Management, 12*(2), 33-60.

Rehman, M., & Shamail, S. (2014). *Enterprise architecture and e-government projects in Punjab, Pakistan.* Paper presented at the Proceedings of the 8th International Conference on Theory and Practice of Electronic Governance.

Saggi, M. K., & Jain, S. (2018). A survey towards an integration of big data analytics to big insights for value-creation. *Information Processing & Management, 54*(5), 758-790.

Saxena, S., & Muhammad, I. (2018). Barriers to use open government data in private sector and NGOs in Pakistan. *Information Discovery and Delivery, 46*(1), 67-75.

Schekkerman, J. (2004). *How to survive in the jungle of enterprise architecture frameworks: Creating or choosing an enterprise architecture framework*: Trafford Publishing.

Sessions, R. (2007). A comparison of the top four enterprise-architecture methodologies. *Houston: ObjectWatch Inc*.

Sheng, J., Amankwah-Amoah, J., & Wang, X. (2017). A multidisciplinary perspective of big data in management research. *International Journal of Production Economics, 191*, 97-112.

Urbaczewski, L., & Mrdalj, S. (2006). A comparison of enterprise architecture frameworks. *Issues in Information Systems, 7*(2), 18-23.

Verschuren, P., & Hartog, R. (2005). Evaluation in design-oriented research. *Quality and Quantity, 39*(6), 733-762.

Wang, H.-J., & Lo, J. (2019). Factors Influencing the Adoption of Open Government Data at the Firm Level. *IEEE Transactions on Engineering Management*.

Zuiderwijk, A., & Janssen, M. (2014). *The negative effects of open government data-investigating the dark side of open data.* Paper presented at the Proceedings of the 15th Annual International Conference on Digital Government Research.