Enterprise Architecture: AI-Driven Capability Mapping for the Agile Strategic Processes

Matti Murtomäki

Haaga-Helia University of Applied Sciences

Degree Programme in Business Technologies

Information Services and Systems

Master's Thesis

2025

Abstract

Author(s) Matti Murtomäki
Degree Master of Business Administration
Report/thesis title Enterprise Architecture: AI-Driven Capability Mapping for the Agile Strategic Processes
Number of pages and appendix pages 79 + 14

Rapidly changing business environments and technological developments challenge organizations' ability to sustain business and remain competitive in today's business world. Enterprise architecture (EA) is an essential part of organizations' ability to manage and develop their business environment. EA must respond to the challenge of digitalization and business environment changes by adapting to become agile and dynamic according to the strategic requirements of the business.

The aim of this thesis was to research and design an AI-based system and an EA process model that integrates and creates a comprehensive, agile, and dynamic way to develop and utilize the case company's capability maps, focusing on their role in enhancing the development of EA. The objective was to develop a solution that accelerates and enhances the customer's capability management as well as explores AI-driven automation opportunities and challenges in EA.

The literature review consists of the capability management in an organization's EA and strategy. The review specifically focuses on the Open Group's TOGAF EA methodology and practices to manage the capabilities. To support the implementation and use of the identified TOGAF-based solution, an AI-based Retrieval-Augmented Generation (RAG) system was defined that supported the goals of the company.

The study was conducted from September 2024 to January 2025, applying qualitative research methodologies, following a practical case study approach supported by real-life use case scenarios. To develop a TOGAF-based process model that integrates into the AI-based solution, the Design Science Research (DSR) methodology was employed. Data collection and analysis were conducted through iterative thematic interviews, with immediate analysis of data against literature review and requirements to refine the developed artifacts.

Summarizing the results and conclusions, after identifying the AI-based system, the following high-level architecture was successfully designed along with the TOGAF-based EA process model. The results shown high potential of integrity between the system and the process model. However, the system design is a large-scale and technically demanding, requiring a lot of knowledge and experience in both EA and AI. Additional use cases are expected to enhance execution and productivity to meet the requirements. The process model enables the implementation of a unified, dynamic, and iterative system that guides EA, as well as the entire organization, toward becoming a digital enterprise.

Keywords
Enterprise architecture, artificial intelligence, TOGAF, process model, Retrieval-Augmented Generation, operational efficiency

Table of contents

1 Introduction

Driven by technological advancements and rapidly evolving market conditions, business environments are transforming in unprecedented ways. Organizations must continuously adapt, innovate, and align their operations to maintain competitiveness and achieve strategic objectives. According to the World Intellectual Property Organization (2019, 13), AI is increasingly driving key developments in business and technology, with applications spanning from medical diagnosis to autonomous vehicles and advanced manufacturing.

EA provides an essential framework for managing the complexities of organizational structures, processes, and technologies. It offers a holistic way to align business requirements and operations with strategic objectives (Josey & Hornford, 2022). A key element of EA is the Capability Map, which serves as a high-level tool connecting strategic decisions to operational processes, structures, and IT systems. This alignment ensures that an organization's physical and digital assets effectively support its objectives (The Open Group 1995-2004). However, traditional capability mapping methods often struggle to keep pace with the rapid evolution of business and technology environments.

As organizations face increasingly complex environments, they are seeking more dynamic approaches to manage their operations and align with strategic goals. To address this challenge, there is a growing interest in leveraging advanced technologies such as AI (World Intellectual Property Organization 2019, 13). AI has the potential to enhance the capability mapping process by improving its accuracy, efficiency, and adaptability, making it more responsive and agile (Goyal, Hussein, Thukral, and Trimble 21 May 2024). Integrating AI into capability mapping provides promising results for organizations, enabling them to remain competitive and proactive in an ever-evolving market (Baragry 12 September 2024).

This thesis examines the role of the Capability Map within EA, focusing on its utilization, strategic alignment, and the integration of advanced technologies like AI. It employs a high-level architectural approach to leverage AI for capability design and usage, excluding non-capability-related elements. The research is based on a case study exploring specific use cases where AI contributes to the creation and utilization of the Capability Map with AI-driven elements. The outcome will be a detailed process model and an AI-based development framework, providing practical methodologies for organizations. By investigating AI's benefits, challenges, and potential in advancing capability-driven processes, this research aims to offer insights into how organizations can align their business strategies with evolving technological landscapes.

1.1 The Client Organization and the Research Context

The client organization for this thesis research is a globally recognized and prominent provider of ICT consultancy services based in Finland, serving a diverse range of industries. The company operates multiple offices across the country and is committed to delivering tailored solutions that meet the evolving needs of its clients. Its business model emphasizes leveraging the latest technologies and innovative business solutions to enhance and accelerate business outcomes. In the research, the client company has appointed a specialist with a strong background in EA. This specialist plays a key role in providing expert insights and guidance on EA solutions, aligning with the company's strategic goals and supporting the research objectives.

The thesis is based on a case study methodology that supports the development of the EA Capability Map (The Open Group 2022b, Business Scenarios). According to Gillham (2010, 10), this methodology allows for deeper understanding of what is going on, illuminating issues and revealing explanations. To complement the case study methodology and facilitate the AI-based technical solution design, DSR methodology is well-suited. It is particularly relevant for research involving the development of AI-driven IT artifacts, where iterative design, validation, and theoretical grounding are essential, such as high-level architecture plans for AI-based systems and new process models. According to Vaishnavi and Kuechler (2015, 9), DSR focuses on the creation and evaluation of IT artifacts, contributing new knowledge, typically in the form of theory. By combining these methodologies, the research integrates case study insights with DSR's principles to form a cohesive research approach, leading to a compatible whole solution.

In line with the customer company's goals and constraints, the research limits its scope to the development and maintenance of EA capabilities, particularly focusing on the utilization of the capability map through advanced technology, especially AI. These constraints establish guidelines for technological selection, which represents one of the core aspects of research. Based on these guidelines, the study develops a conceptual model for an AI-driven design that supports the development of EA. Throughout the research process, the scope is further refined and aligned with the objectives by applying research methodology practices to fine-tune the thesis and achieve the desired outcomes.

1.2 Research Questions

- *What is the role of capabilities and the Capability Map in an organization's EA and strategic processes?*
- *How is the Capability Map created, maintained, and utilized?*

- *How can AI be utilized in the creation, maintenance, and utilization of the Capability Map?*
- *What are the benefits and challenges of using AI to create, maintain, and utilize the Capability Map in EA?*

The defined research questions guide the study and provide a structured framework that supports the needs of the customer company while also advancing research. As Bondel, Faber, and Matthes (15 November 2018) emphasize, the business capability map (BCM) serves as a valuable tool for addressing challenges arising from digitization and aligning business and IT within EA. Understanding the scope of the capabilities and the capability map is crucial. Bondel et al. (15 November 2018) discuss the implementation of the capability map, outlining the tools and techniques used in its construction, including various approaches for gathering the required information. Additionally, Bondel et al. (15 November 2018) present a BCM in Figure 1 (page 12), which illustrates how business capabilities are structured into layers or logical categories with hierarchical dependencies. Finally, the fourth research question facilitates an evaluation of the possibilities and challenges associated with both the practical and theoretical aspects of the thesis, as discussed in chapter 5.1 Research Question Analysis (page 67).

1.3 Key Concepts

Concept	Definition
Enterprise Architecture	EA refers to the overall structure of an enterprise, encompassing its functional groups, multiple systems, and all its information systems. It is embodied the organization’s components, their relationships with each other and the environment, and the principles guiding its design and evolution (The Open Group 1999-2006).
Enterprise Architecture framework	EA framework is a model for guiding EA work. It is often referred to as a structuring framework and a method that guides development of EA in a structured manner (Hosiaisluoma 2015, 26).
Architecture Development Method	Architecture Development Method (ADM) is a method for developing and managing the lifecycle of an EA and forms the core of the TOGAF standard. It combines elements of the TOGAF standard along with other architectural resources to address and organization’s business needs (The Open Group 1999-2022a).

Capabilities	In the context of EA, capabilities are formed by the combination of three key elements: operational models and processes, employees and skills, and knowledge and systems. These capabilities are essential for implementing business strategies and business models (Digital and Population Data Services Agency 2017a, 3).
Business Capability Map	The BCM provides a visual depiction of all business capabilities, decomposed at the appropriate level to form a stable set of the capabilities. According to The Open Group (2022a), it offers a self-contained view of the business, breaking down the organizational structure, business processes, IT systems and applications, and the product and service portfolio in a logical manner. This provides greater insight into aligning and optimizing each of these domains, once mapped back to the top-level business objectives.
Artificial Intelligence	The definition of AI is defined as a technology that enables machines and computers to mimic and simulate human abilities such as learning, problem-solving, comprehensions, decision-making, creativity, and autonomy (Stryker & Kavlakoglu 16 August 2024).
Natural Language Processing	To enable machines and algorithms to understand text or characters, it is essential to convert the data into machine-understandable format. Natural language processing (NLP) allows machines to understand and interpret human language (Kulkarni & Shivananda 2021, Introduction).
Large Language Model	Large Language Model (LLM) provides a technique that supports full open-ended support for human-like conversations. It can perform a wide range of NLP-based tasks, such as paragraph writing, text summarization and generation, language translation, and more (Wassan & Ghuriani 2024).
Retrieval-Augmented Generation	RAG involves gathering relevant contextual information from a data source and supplying it to an LLM along with the user's prompt. his contextual information enhances the model’s output, whether text

	or images, by supplementing its foundational knowledge (OpenAI 2024).

1.4 Overview of the Research Structure

This section provides an overview of the research report structure, outlining the content of each chapter. The report is divided into five chapters, each focusing on a specific aspect of the research process. It follows logical progression, building the necessary information phase by phase to ensure a well-managed, focused, and coherent output. Each chapter contributes to a deeper understanding of the research topic and guides through the essential components of the study. Below is a brief description of each chapter and its content:

- *Chapter 1: Introduction, Research Context, and Research Questions*

This chapter introduces the research topic, explains its relevance and the customer company context, and defines the research objectives and questions guiding the study.

- *Chapter 2: Literature Review and Previous Research*

A review of relevant literature provides a theoretical framework for the study, focusing on EA and AI. It identifies key EA solutions, emphasizing methods and the definition of capabilities that align with customer requirements and support AI outcomes, followed by the selection of essential AI solutions to achieve the research objectives. Additionally, the review examines the interconnections between EA and AI to ensure alignment and discusses previous research findings.

- *Chapter 3: Methods and Implementation*

The research methods used in conducting the research are explained, presented, and justified in detail, including the study design, data collection techniques, and analysis methods.

- *Chapter 4: Presentation of Results*

This chapter outlines the findings of the research and presents the necessary and defined solutions, justifying the choices made and providing data and analyses that address the research questions.

- *Chapter 5: Conclusions and Discussion*

The final chapter interprets the findings and answers the research questions. It also offers conclusions, suggestions for future research, practical applications, and an evaluation of the thesis. Furthermore, the chapter includes reflections on the responsibility of research, as well as considerations of reliability, ethics, and the learning process involved in conducting the study.

1.5 Utilizing Artificial Intelligence in the Thesis

AI has been ethically applied in this thesis, supporting various stages of the research process. For instance, OpenAI's ChatGPT-3.5 and Microsoft's Copilot language models have been utilized to assist in proofreading and refining the text. While AI has not been employed to generate content, it has been used to process the researcher’s text from a proofreading and grammatical perspective. These tools have been ensuring clarity, accuracy, and readability of the thesis, helping the researcher focus more on the substance of the research rather than on mechanical writing tasks.

Additionally, AI has been integrated into the development of the Grounded Theory (GT) research method, as described in Appendix 2 (page 85–88). This process has been involved in putting the thesis's topic areas into the language model, based on GT principles, to ensure alignment with the methodological requirements. The language model has then assisted in constructing the necessary frameworks for the GT model, which has allowed the researcher to implement the required content analysis effectively and facilitated the GT analysis process.

2 Literature Review

The Literature review is structured around three main areas: a review of previous research, capability management within EA, and AI solution identification. Together, these areas provide a solid foundation for understanding and addressing the core research objectives that guide development work. The research focuses on utilization of capability maps, particularly designing a process model that integrates AI into the EA tool, aiming to enhance organizational agility, drive automation, improve accuracy, and enable more informed decision-making, ultimately fostering operational excellence

The EA section examines EA from a strategic perspective, emphasizing its relationship to organizational capabilities and the role of capability maps. It explores the definitions of capability maps within EA and addresses key questions, such as the concept of a capability map, its connection to business processes and operations, and the tools and methods available to develop and manage it. This includes an in-depth look at a pre-selected open-source EA tool that supports the creation and maintenance of EA capabilities and the capability maps.

The AI section explores AI´s potential to enhance capability map utilization within the EA tool. The research focuses on identifying AI-driven solutions aligned with principles and guidelines that are defined by interviewing the customer company’s specialist. This section centers on identifying design solutions that advance the research development goals, particularly by producing AI-based design solution that contributes to a process model aligned with the project’s objectives.

2.1 Previous Research

This section focuses on identifying the most relevant previous research that relates to the thesis development tasks and determining the most effective ways to achieve results for research questions. Previous research guides the study in the right direction and helps align the scope to support the thesis objectives. In this study, previous research on EA capability maps is examined to identify the associated capabilities and their role in supporting organizational goals. Additionally, the reviewed studies explore the integration of AI into EA. Specifically, the study aims to identify and align prior studies related to EA development that incorporate AI capabilities.

The following subsections have been structured to address specific themes derived from prior research. These themes highlight the relevance of the development task, as previous research emphasizes the growing importance of AI in enhancing EA processes and the challenges of integrating AI technologies. They provide insights into how AI can optimize EA solutions, directly contributing to the study's objectives. The findings from these studies inform the thesis by exploring AI's role in EA tools, ensuring the development task is grounded in current research and practices.

- *Enterprise Architecture and Artificial Intelligence:*

This section aims to identify how AI is exploited in implementing EA-based solutions. AI is increasingly leveraged to enhance business development and decision making that are closely related to EA. Given the rapid advancements in AI technologies, understanding their integration into EA is crucial for developing an effective process model in this study. The section focuses on revealing the most relevant AI-driven methods and technologies to exploit in EA development work. This aligns with the development task of exploring AI's role in EA tools and processes.

Granger and Baragry (20 Feb 2024) are recognized specialists in EA, who seek to exploit the capabilities of AI to enhance EA modelling processes. By integrating the latest AI technologies and techniques, they aim to refine and advance EA practices. Similarly, Goyal et al. (21 May 2024) leverage AI-based NLP in the development of the EA Assistant (EAA), a tool designed to enhance time efficiency and strategic alignment within EA. This AI-driven approach optimizes workflows and enhances the alignment between business and IT strategies in EA.

- *Enterprise Architecture and Business Capability Maps:*

This subsection explores the role of BCMs in EA, particularly aligning business and IT to EA capability maps. It examines how previous studies contributed to the design and development of BCMs, with a focus on identifying tools, processes, and techniques for their effective implementation. The primary goal is to gain understanding of BCMs within an organizational context and to evaluate their value in addressing digital transformation challenges.

Bondel et al. (15 November 2018) highlights the importance of the BCM’s when aligning challenges arising from rapidly increasing digitization and aligning business and IT in EA. Bondel et al. (15 November 2018) study also identifies processes and structure to implement BCM. Additionally, Baragry (12 September 2024) highlights in his blog the use of AI-based tools in developing EA capability maps. His blog specifically discusses how AI can support the maintenance of a BCM within EA tool, offering an innovative perspective on automating and enhancing traditional capability mapping processes.

2.1.1 Enterprise Architecture and Artificial Intelligence

Granger and Baragry (20 Feb 2024) discuss the use of generative AI for generating EA in their blog post. Baragry is the Chief Enterprise Architect at Ardoq, a company founded in 2013, known for assisting its clients with digital transformation projects. In 2023, Ardoq is recognized as a leader in EA tools by the Gartner Magic Quadrant (Ardoq 2013-2024). For over the past three decades, modelling EA using frameworks has been the foundation of the field. The key question now is what form these models take in the era of generative AI. In EA, it is important to have a common language, and standards like ArchiMate and Business Process Modeling Notation (BPMN) enable this. However, these standards present a learning hurdle, requiring organizations to learn to use them. Generative AI offers potential concept to explore, interrogate, and reframe these processes (Granger & Baragry 20 Feb 2024).

Granger and Baragry (20 Feb 2024) present LLMs, including ChatGPT-4 Turbo, as dynamic processes and forms of knowledge representation utilizing neural networks. A key limitation of LLMs lies in their knowledge cutoff and the lack of enterprise-specific data. One promising approach for finetuning LLMs with an organization’s own data is RAG, which is becoming a leading method in this area. When generating EA models through the application of LLMs, two major issues arise. A fundamental limitation is input size, while the challenge of unstructured data formats has been addressed by leveraging LLMs in conjunction with vector databases, which improve the quality and resolution. Additionally, there are two key concerns: hallucinations and information silos. Hallucinations occur when results suggest that a business process and an application are related, though they may not be. Information silos, on the other hand, concerns how business and technological elements are related. When dealing with unstructured content, the data may be scattered across multiple documents, increasing the probability of missing relationships. The LLMs only connect based on the underlying information they are trained on.

Time constraint is often related to the roles and tasks of enterprise architects. Goyal et al. (21 May 2024) conduct a project aimed at leveraging NLP to enhance time efficiency and strategic alignment within EA. Their project aims to alleviate the time challenges by introducing NLP-based EA Assistant (EAA), which is used for tasks such as documentation, analysis, and decision-making processes. The EAA utilizes NLP on top of a chatbot to filter documents based on their relevance, subsequently summarizes them, and produces radial graph visualizations for further investigation by the enterprise architect.

The EAA system, as outlined by Goyal et al. (21 May 2024), consists of three distinct modules: a chatbot, a relevance checker, and a user interface for automatic relationship visualizations between stakeholders from the documents. The relevance checker's responsibility is to evaluate the documents' pertinence based on their similarity to the topic. Meanwhile, the chatbot aims to provide summaries and respond to inquiries related to the documents. The user interface enables users to examine interconnections between stakeholders visually formatted from the content mentioned in the documents. Additionally, Goyal et al. (21 May 2024) collect EA-related data using web scraping and simulation techniques from various reputable repositories, including scholarly articles and white papers. PDF-formatted content is also converted into machine-readable text, ensuring compatibility with the NLP-based analysis and processing pipeline.

According to Goyal et al. (21 May 2024) the chatbot is specifically designed to serve the needs of enterprise architects. The GPT-3.5-turbo-instruct model is chosen for its instructional capabilities and its ability to assist users with complex procedures, by its effectiveness in performing natural language tasks. It demonstrates the capability to handle both conversational and instructional interactions, supported by its "RAG" capabilities. RAG is a paradigm in NLP that integrates information retrieval processes into text generation. They selected RAG to enable the system to incorporate production components and access external data sources, addressing the needs of both technical and non-technical people in the EA teams. Additionally, RAG's analytical prowess allows it to effectively derive meaningful recommendations and conclusions.

From a technical and security perspective, the backend is programmed in Python programming language. Secure PDF management is achieved with the Tempfile module, which prevents long-term storage. Environment variables and API keys are managed through the .env file and the dotenv library, ensuring secure connections to services like OpenAI. Additionally, text is extracted from PDFs using the PyPDF2 library. The LangChain package is used to leverage LLMs for answering queries, including splitting text into more manageable chunks with CharachterTextSplitter and creating a Chroma vector database to process relevant queries (Goyal et al. 21 May 2024).

Goyal et al. (21 May 2024) point out that their test results reveal some instances of hallucinations, which can be referenced in Granger and Baragry's (20 Feb 2024) blog post about two major issues: hallucinations and information siloes. However, this type of behavior can be reduced by enhancing training data and its diversity, implementing regular updates, and fine-tuning the model, among other similar strategies (Goyal et al. 21 May 2024). In conclusion, Goyal et al. (21 May 2024) encapsulates that the project has significant real-world impact, aiming to revolutionize EA processes through advanced technologies. These technologies have the potential to drive cost savings, improve productivity, and enhance the effectiveness of EA processes across various industries.

2.1.2 Enterprise Architecture and Business Capability Maps

To address challenges arising from digitization and to align business and IT in EA, the BCM tool can be implemented (Bondel et al. 15 November 2018). Bondel et al. (15 November 2018) study answers this question by implementing a case study of the BCM in a medium-sized organization, following The Open Group's implementation methods. Since business capability aligns an organization's ability to perform certain activities from a functional perspective, the BCM visually represents a group of business capabilities. The BCM provides a view for both business and IT that is neither too technical nor too strategic and can also be understood by relevant stakeholders.

Bondel et al. (15 November 2018) defines a BCM in Figure 1 (page 12), where business capabilities are grouped into layers or logical categories, with hierarchical dependencies between the business capabilities. Business capabilities can break down to support more detailed analysis and planning. For example, in Figure 1 (page 12), the business capability *Production* breaks down into the business capabilities *Procurement* and the *Plant Management.* For better decision-making, BCM can be combined with a heat map that provides more detailed information about factors such as strategic contribution, cost contribution, or the criticality of business capabilities.

The map visually represents capability characteristics from strongest to weakest using a color scheme of green, yellow, and red.

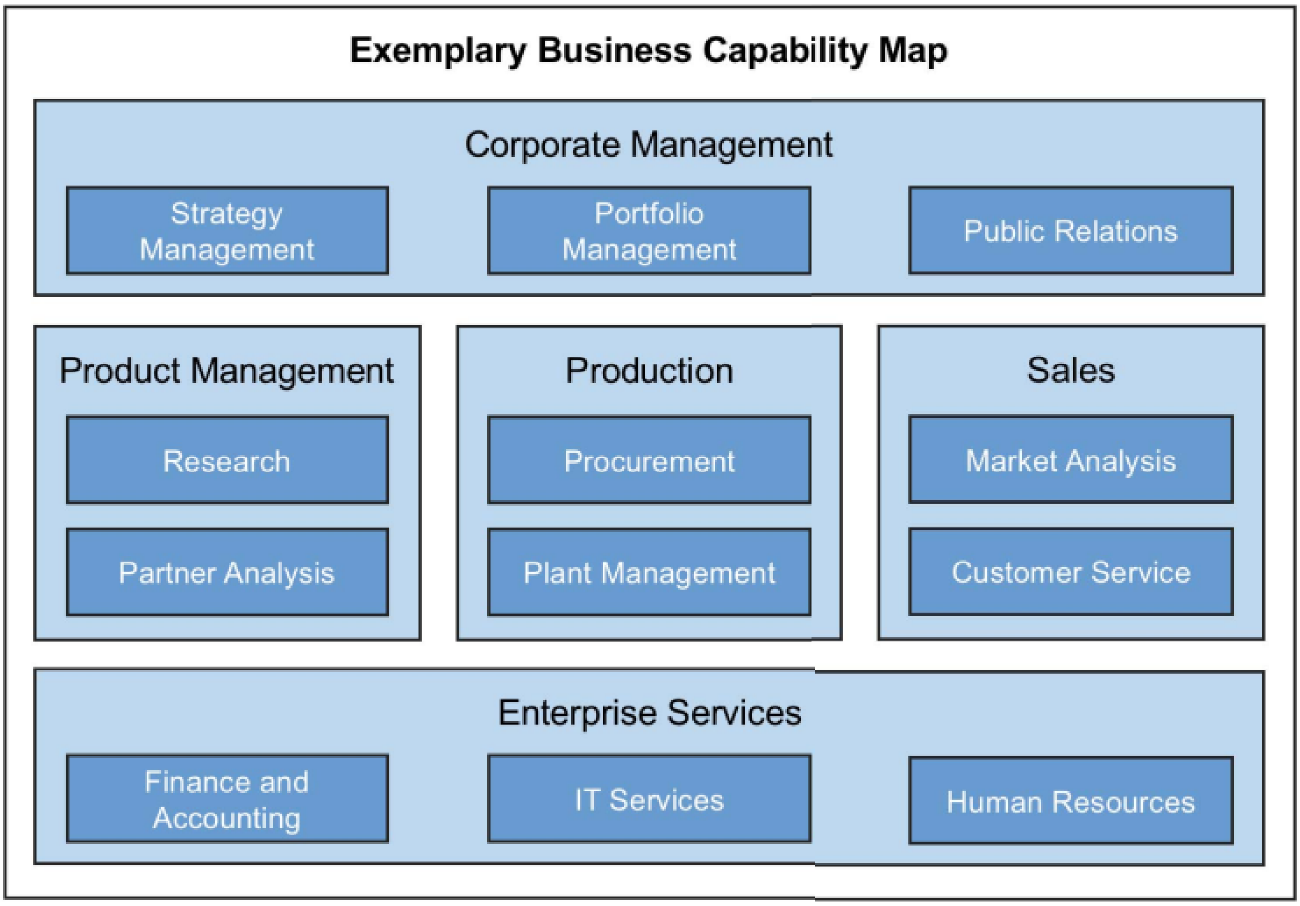

Figure 1. Example of a Business Capability Map with five business capabilities on level one and twelve business capabilities level two (Bondel et al. 15 November 2018)

Bondel et al. (15 November 2018) describe a BCM implementation that begins with the selection of a TOGAF-based top-down approach, identifying the five highest-level capabilities. To identify these capabilities, it is necessary to gather information such as business rules, business entities, business processes, and strategic objectives. They used several information-gathering techniques, including reading, writing, interviews, observations, and questionnaires to model business capabilities utilizing a framework. This framework provides a detailed description of business capabilities, with elements arranged as rows and external and internal environmental knowledge, ends, and means arranged as columns. Bondel et al. (15 November 2018) highlights that while several approaches are used by consultancy companies to provide BCMs, most of these approaches do not focus on the proper definition of the business capabilities.

The creation of the BCM is a comprehensive process aimed at representing the business capabilities of the entire organization. The initial version of the BCM, which serves as the foundation for further development, was collaboratively drafted by the managing director, the vice-managing director, the head of the strategy department, and a representative of the research group (Bondel et al. 15 November 2018). The study by Bondel et al. (15 November 2018) identifies four distinct levels of description related to BCMs, emphasizing the importance of structure and clarity in BCM development. A key lesson learned from the study is that involving business leadership in the project not only enhances a sense of shared ownership but also bridges the gap between business and IT. This equal involvement ensures that the BCM is not perceived as an isolated IT artifact but as a strategic tool that aligns business and IT objectives. Moreover, the active participation of business leaders fosters the development of a common organizational language. This shared language enables stakeholders to better identify and address pain points within the organization, promoting a more structured approach to strategy planning and communication across all units.

Baragry's (12 September 2024) blog article addresses how to use a LLM for RAG in combination with the Ardoq Tool. The LLM automatically suggests relationships between Application and business capabilities it realizes, adding those to Ardoq via the API. Business capabilities are not generated by LLM but are already implemented in Ardoq. Business capabilities represent general knowledge that LLM training models should encompass. A capability map also includes descriptive information provided by the organization that uses or offers it. One major concern is confidentiality, which must be ensured when configuring an LLM with an organization's data.

RAG's benefit compared to a traditional LLM is that it can incorporate specific data not included in the model's original training. By combining an external data source with the LLM request, it can generate more relevant and contextually accurate responses, including private or organization specific information based on the user's query. Traditionally, LLMs face limitations when embedding contextual information into queries. RAG systems must break external information into smaller chunks and identify the most relevant pieces to include in the user's query. This process is often implemented in traditional RAG applications using embeddings and vector databases (Baragry 12 September 2024).

Baragry's (12 September 2024) work highlights the use of an already implemented capability map implemented into Ardoq Tool into the request context for OpenAI's GPT-4 model using RAG. He provides various suggestions for combining and configuring LLMs, RAG, and the capability map to address similar challenges. A key element of this approach is that prompt design is critical to getting accurate and consistent results from an LLM; starting with a system message that instructs GPT-4 to output JSON data can significantly enhance the effectiveness of the model. However, Baragry's (12 September 2024) notes that there is a lot of work to do before this kind of solution can be standardized. A key concern is the unreliability of LLMs generated text. Nonetheless, several solutions exist to validate accuracy of the generated text and ensure that the information produced by the LLM meets user's needs.

2.2 Enterprise Architecture Framework: The TOGAF Standard and Business Capabilities

Josey and Hornford (2022) introduce TOGAF as a standardized EA framework in their E-book: *The TOGAF Standard, 10th Edition (2022) – A Pocket Guide*. It is an approach to develop, approve, use, and maintain EAs. It supports a reusable set of existing architectural assets and best practices based on an iterative process model. The TOGAF framework 10th edition is based on the earlier versions of the TOGAF Standards. The Open Group states in the TOGAF 9.2 Overview that version 9.2 is a proven EA methodology and framework used by the world's leading organizations (The Open Group 1995-2004). According to The Open Group, it is the most reliable and prominent EA standard. The standard highlights that it responds to today's organizational demands, including digital enterprise trends and changes, by being agile and transformative. Version 9.2 is redesigned and restructured into smaller guides and publications while retaining the core of the TOGAF framework.

According to Josey and Hornford (2022), the TOGAF standard version 1 is developed in 1995 based on the *US Department of Defense Technical Architecture Framework for Information Management*. The Open Group Architecture Forum develops the framework and has published multiple versions on The Open Group's public website. Josey and Hornford (2022) provides an overview of TOGAF standard in their book: *TOGAF Standards, 10th Edition Pocket Guide* introduction page, describing the content of the framework including the TOGAF Library that covers, pocket guides, white papers, reference architectures, guides, among others.

Due to TOGAF's structured methodology, its iterative development process, and widespread adoption by leading organizations worldwide, the framework serves as a solid and compelling foundation for this study. Its emphasis on best practices, flexibility, and continuous improvement makes it highly adaptable to the current organizational landscape, aligning with digital enterprise trends, technological advancements, and agile practices.

2.2.1 The TOGAF Framework and the Architecture Development Method

Figure 2 below describes how Business Vision and Drivers along with Business Capabilities are the driving forces around the central ADM. Capability and Governance act as operational drivers, supported by ADM Techniques and Applying the ADM as well as the TOGAF Series Guides and the TOGAF Library. The TOGAF Framework Overview in Figure 2 highlights the most important elements that guide the architecture process, ensuring a structured and methodical approach to EA (Josey & Hornford 2022). The ADM (Architecture Development Method) is a key component of the TOGAF Framework in this study, providing a structured approach that informs and guides the development of EA.

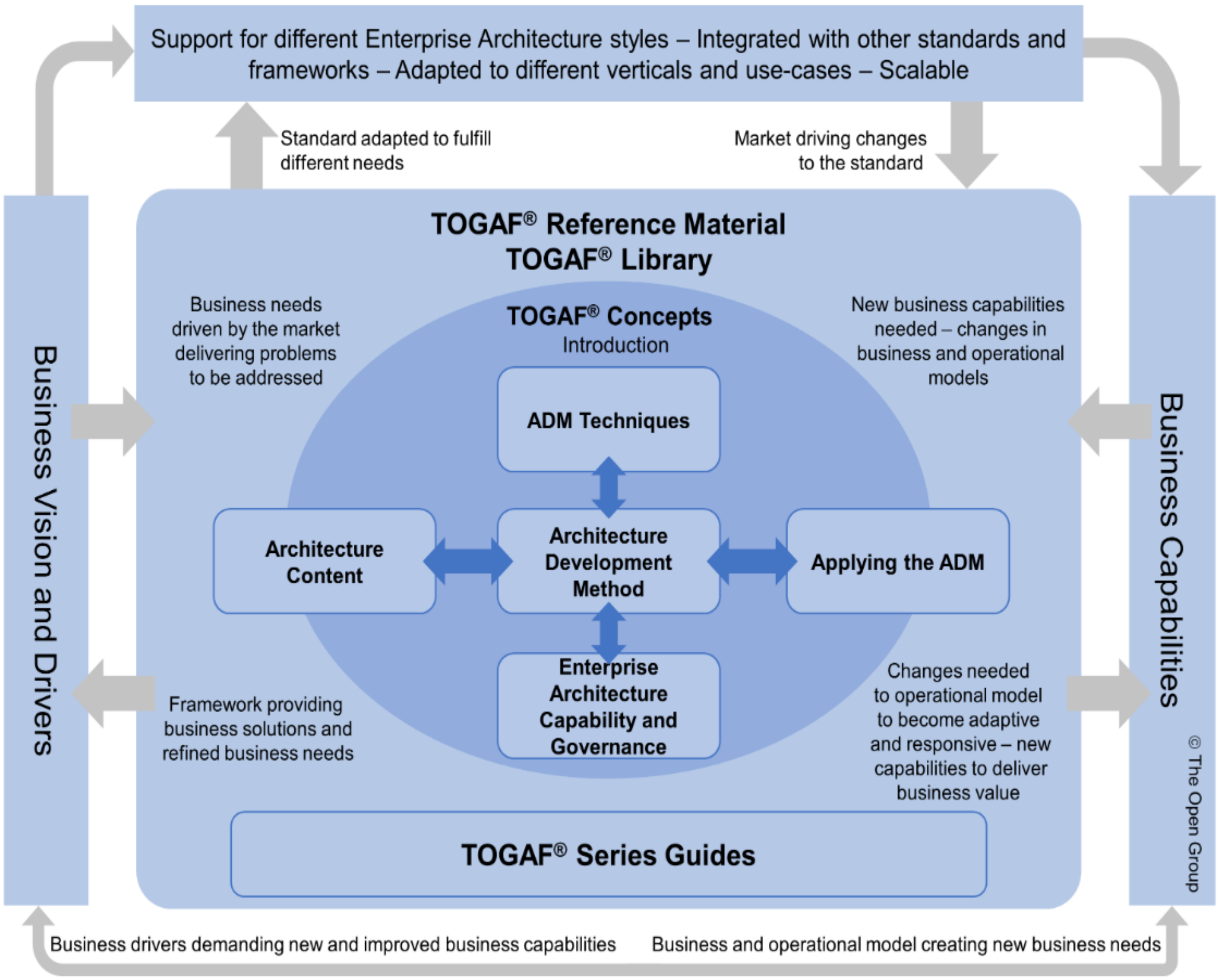

Figure 2. TOGAF Framework Overview (Josey & Hornford 2022)

Figure 3 below presents a detailed view of the ADM and its phases. As one of the key elements of TOGAF's EA, the ADM serves as the foundation for establishing an architecture framework, developing architecture content, and governing the implementation and transition of the architecture. The activities within the ADM are carried out in iterative cycles, enabling continuous architecture definition and realization. This approach allows organizations to manage controlled transformation in alignment with business goals and opportunities (The Open Group 1992-2022). The iterative phases of the ADM and their implementation are central to the development of this research. By aligning the capabilities and capability map with the corresponding phases of the ADM, a solid foundation for the study is established.

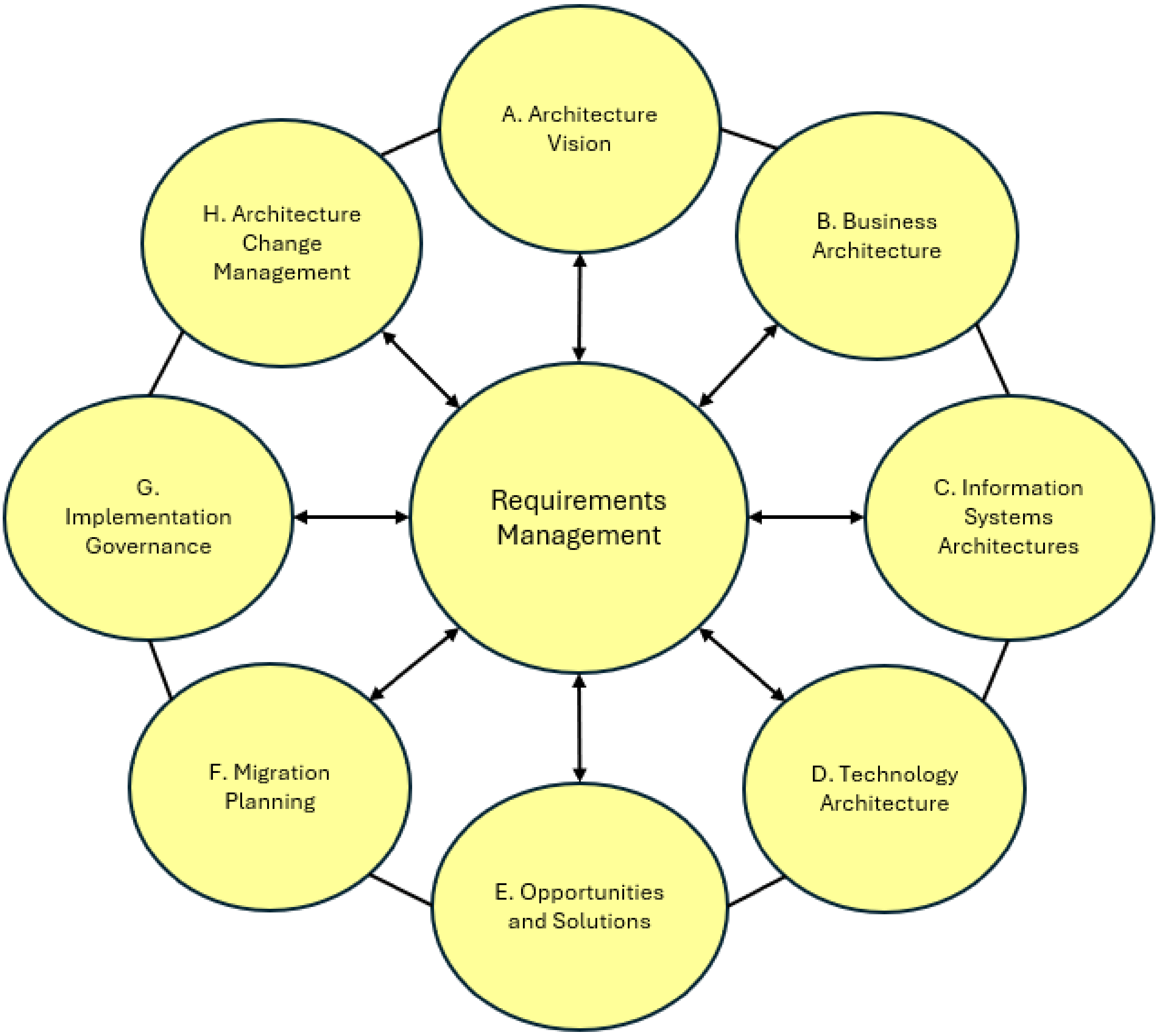

Figure 3. Architecture Development Cycle (adapted from The Open Group 1992-2022)

Business capabilities in the ADM are primarily associated with the Architecture Vision (Phase A) and Business Architecture (Phase B), as shown in Figure 3 (page 16). In Phase A, the required business capabilities are identified and evaluated based on strategic priorities and documented in a Capability Map, which serves as the basis for developing a high-level view of baseline and target architectures (The Open Group 2022a). Additionally, the Capability Assessment, which evaluates the maturity and level of capabilities, is an output of Phase A (The Open Group 1999-2022b).

Business Capability Mapping is implemented in Phase B. This process delivers value for stakeholders by identifying, categorizing, and decomposing the business capabilities (The Open Group 1999-2022b). A detailed analysis of business capability gaps is also performed in Phase B (The Open Group 2022a). Referencing the earlier case study by Bondel et al. (15 November 2018), The Open Group also recommends using a heat map in conjunction with the BCM in Phase B to provide more detailed information about strategic and financial contributions. In Phase B, business capabilities should be mapped back to value streams, organizational units, information systems, and strategic plans. These aspects are addressed in the ADM's phases A to D (The Open Group 1999-2022b).

2.2.2 Business Capabilities

The Open Group defines a capability as *an ability to do something*. Capabilities generally require a combination of tools, processes, and people that the organization possesses to perform that ability. A business capability aligns what business does without providing further details, such as why, where, how, and how the capability is used. A business capability can either exist as it is or be required to create new business value or strategy. When business capabilities are compiled into a catalog, they represent the abilities an enterprise must run its business. (The Open Group 2023)

Digital and Population Data Services Agency (2017a, 5), a key government body responsible for overseeing the Population Information System, a foundational element of Finland's societal infrastructure, suggests utilizing one or multiple tools to evaluate a capability's maturity at a strategic level. These metrics can be implemented using frameworks such as BSC, CAF, EFQM, and CCMI (Capability Maturity Model Integration). These tools are used to evaluate capability management as well as to assess and develop the current and target states of capabilities. Development targets for organizational capabilities can be identified through these metrics and directly incorporated into development programs, where they can be further refined into specific initiatives using EA.

In more detail, the Digital and Population Data Services Agency (2017b, 21) positions the capability to build a business model that implements strategy in the JUHTA – (Public Administration IT Management Advisory Board), a Ministry of Finance committee overseeing information management in Finland's public administration. This is demonstrated in the material JHS 179 Overall Architecture Planning and Development, as shown in the adapted Figure 4 below. The figure illustrates how People and Skills, Operating Model and Processes, and Data and Information Systems contribute to the development of the capability, which in turn further aligns the structure of the Business Model and Strategy. This visualization is essential as it provides a structured representation of the key components and their interrelations, illustrating how capabilities are positioned to support the implementation of strategy.

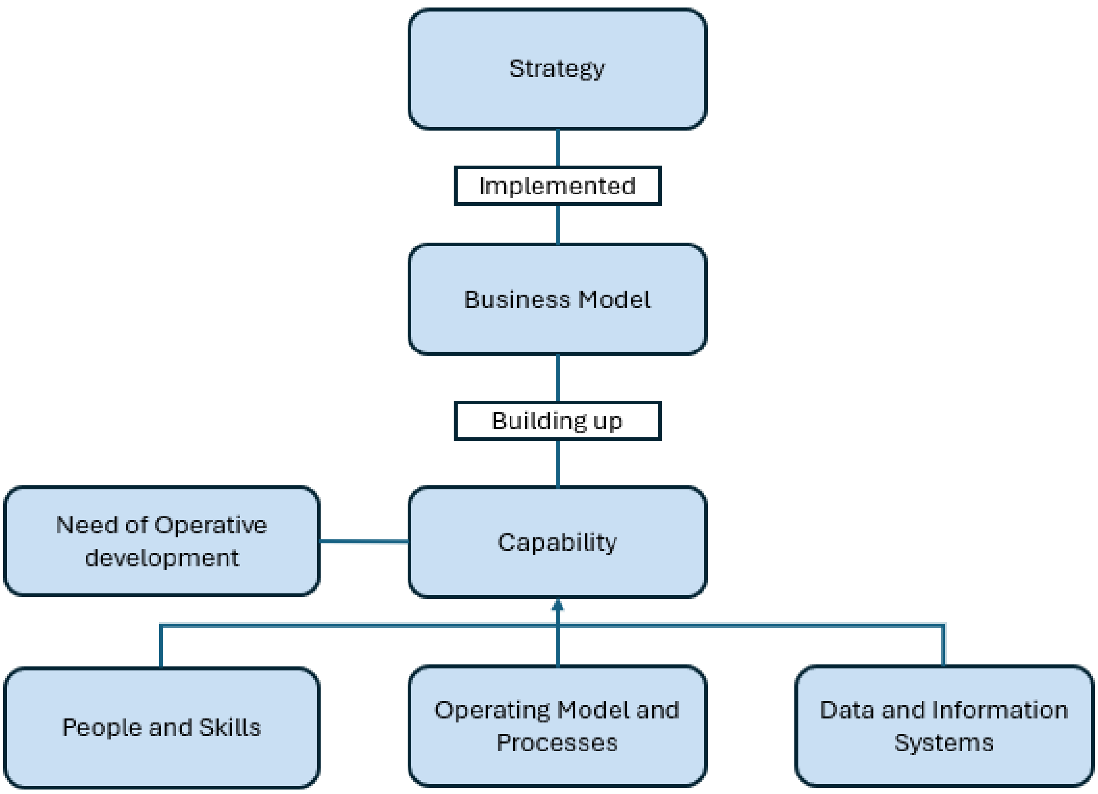

Figure 4. Capabilities Required by Business models (adapted from Digital and Population Data Services Agency 2017b, 21)

Value streams, which illustrate the flow of activities that create value to the customers, work together with business capabilities to offer visual presentation to communicate how an organization delivers value to stakeholders and customers. Mapping business capabilities to value streams helps clarify what is required for implementation and how value is delivered (The Open Group 2023). In this context, the alignment of business capabilities with value streams ensures that all necessary processes are in place to achieve strategic goals efficiently.

City of Espoo (2014-2020, 16) highlights the importance of providing value to a customer when prioritizing the focus of development and targets. One of the main targets is to reveal hidden capabilities and identify the topics that are relevant for responding to changing circumstances. In this context, the importance of thoroughly evaluating the current state and the target state emerges, along with the need to clarify the requirements for target state.

The Open Group recommends using the ArchiMate modelling language to provide a unified view for describing and managing diagrams, including those related to value streams and business capabilities. A business capability catalog serves as a resource that provides a definition for each business capability, along with a model to enable people to find the business capabilities they need, ensuring clarity and alignment across the organization (The Open Group 2023).

2.2.3 Business Capability Planning

Understanding business operations creates a foundation for initiating business capability planning activities. Business operation artifacts may not always exist in formal documents and might require EA practitioners to generate and analyze them to complete the analysis. The following elements are essential for the overall organization, and in larger organizations, they may be applied to each distinct business area individually (The Open Group 2023).

- *The organization structure*
- *The business model and associated objects*
- *Current strategic, business and financial plans*
- *Applications and information objects (The Open Group 2023)*

Organizations are generally structured to support business capabilities. Typically, people implement processes and allocate resources or tools, such as money and IT. The business capability catalog is frequently used to inform the organizational structure. Within an organization, multiple units are involved creating and delivering business capabilities, which tend to be more stable than the organizational structure itself. Additionally, capabilities should not be named too specifically to individual business units (The Open Group 2023).

To understand capabilities and business capabilities within an organization, as presented at a high level in Figure 4 (page 18), Figure 5 below provides a more detailed description of what business capabilities are and how they connect to business units. The example retail organization includes three main functional areas, each with smaller departments. This organization chart is not detailed enough to inform an overall capability catalog but can be used to indicate where the business capabilities are located. Within the Finance business, shown in Figure 5, there are three functional elements for Cash management, Accounts Receivable, and Accounts Payable, which presents business capabilities for most organizations. For the Warehouse management and Retail stores, the organizational hierarchy does not provide sufficient detail for capability catalog, either due to geographic or product-oriented division. (The Open Group 2023) To develop a capability map, it is necessary to understand the concept of capabilities and how they are positioned within the map, as described by Bondel et al. (15 November 2018) in Figure 1 (page 12) This concept provides the necessary information to define use cases for capability and capability mapping in this thesis.

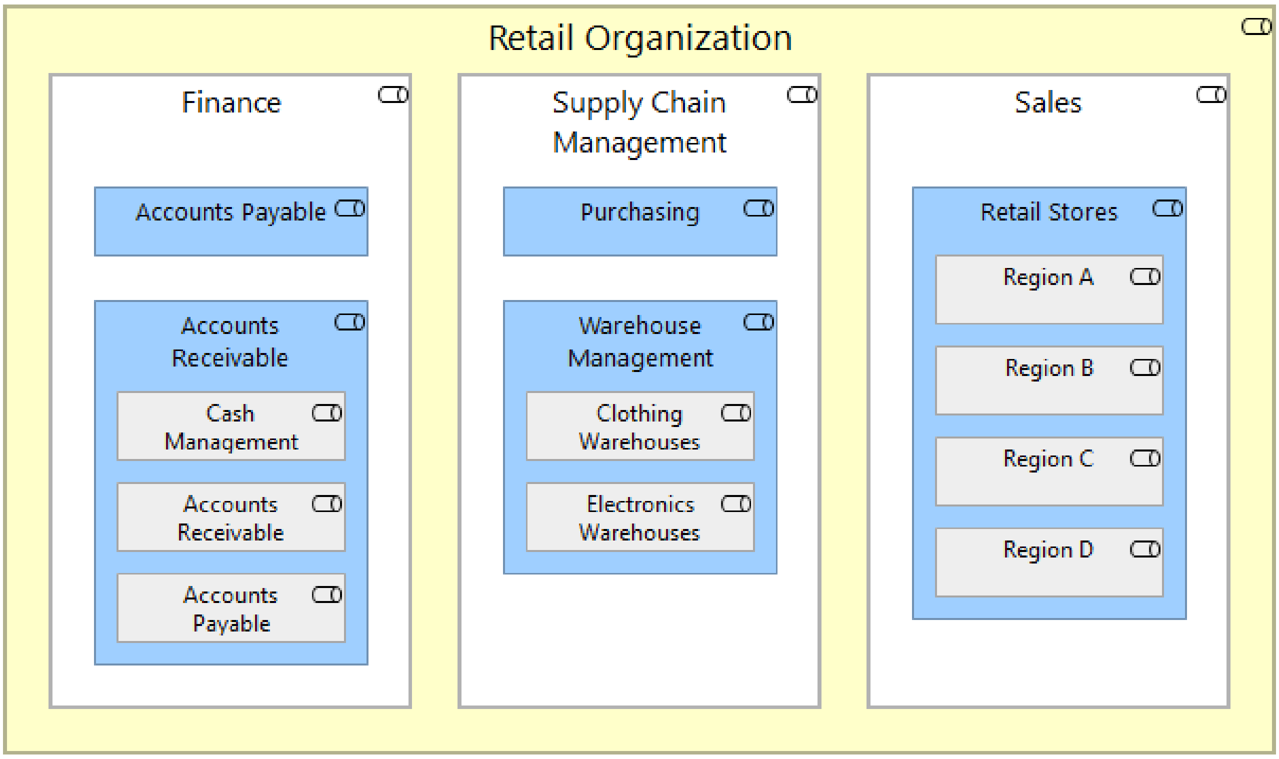

Figure 5. Retail Company Organization Chart (adapted from The Open Group 2023)

By using the Business Model Canvas, it is possible to identify and communicate the business model and associated objectives at a macro level (The Open Group 2023). This tool helps organizations visualize and understand how various components of their business interact and align with strategic goals. The Business Model Canvas also provides a structured approach to designing and analyzing business models in different contexts. Additionally, Financial, business, and strategic plans can be identified by using three different elements (The Open Group 2023):

- *A Financial statement provides insights into how various aspects of the business are prioritized and indicates areas of relative.*
- *A business plan is a narrative description of a company's intensions relating to the market, while a strategic plan outlines goals and objectives in response to external or internal stimuli.*
- *The Balanced Scorecard is the most widely used tool for measuring business performance and serves as a source of objectives and measurable goals.*

Digital and Population Data Services Agency (2017a, 3) suggests starting to identify capabilities by recognizing an organization's operating processes and services. A capability forms from several processes and resources (e.g., actors, information systems). Services are produced for customers through certain capabilities. The Digital and Population Data Services Agency (2017a, 3-4) categorizes capabilities into three different groups: operational capabilities, operational support capabilities, and strategic capabilities. Strategic capabilities are key capabilities needed to realize the organization's goals. They consist of both operational and operational support capabilities.

2.2.4 Business Capability Map and Business Capability Catalog

Business Capability Catalog consists of business capabilities that are captured and organized in a logical manner. There are three different approaches to identify and establish these capabilities: top-down, bottom-up, and incremental. These approaches guide the process of identifying and defining business capabilities within a company. The top-down approach begins by identifying capabilities from a strategic and management perspective, while the bottom-up approach starts from processes, tools, and resources. The incremental approach builds on already defined capabilities (The Open Group 2023).

The BCM visually depicts a complete and stable set of business capabilities, which are decomposed at the appropriate level of detail and logically grouped into categories enabling analysis and planning. Mapping business capabilities back into organizational units, value streams, IT architecture, and operational and strategic plans provides valuable insights for aligning and optimizing each of these domains (The Open Group 2022a).

Similar to business capability catalog, the Open Group offers two way to identify capabilities for the capability map, top-down and bottom-up. However, top-down approach represents enterprise-level perspective to identify 20 to 30 highest-level business capabilities. In contrast, the bottom-up approach involves building business capabilities from the bottom up, incorporating insights from various parts of the business. Combining both approaches is usually implemented, as it allows for more refinement of the map (The Open Group 2022a).

The Open Group (2022) recommends using the organizational structure, business model, and current strategic and financial plans when identifying capabilities for the Capability Map. These sources are comparable to those used for compiling the Business Capability Catalog, though the methods and emphasis may differ. While the BCM can be informed by organizational structures, it should not be tightly aligned with them. It is not uncommon to find silos or duplications of capabilities within organizations structures, as these do not always align neatly with business objects. Therefore, it is important to analyze an initial set of capabilities to clarify business objects and adjust the Capability Map to ensure a more accurate alignment with these objects. Business models can be used to derive individual capabilities for the Capability Map. Additionally, strategic and operational business plans as well as financial plans should be utilized during the implementation of the Capability Map.

Stratification is the step of organizing capabilities into a logical structure. The stratification essentially involves classifying, grouping, and aligning capabilities within tiers, categories, and layers. This approach provides a framework to group, analyze, and align the interest of different stakeholders. In contrast, leveling refers to approach of breaking-down top-level capabilities into more detailed, lower level capabilities for better communication with the audience and stakeholders (The Open Group 2022a).

2.3 Archi: ArchiMate Modelling

Archi is an open-source tool for creating ArchiMate models and sketches. Archi is particularly targeted for use with the ArchiMate modelling language, which is an EA standard supporting visualization, analysis, and description of architecture across different business domains. ArchiMate fully aligns with TOGAF and is hosted by The Open Group. Archi is a well-known open-source modelling toolkit that is used globally by banks, insurance companies, training organizations, universities, and EA consultants (Beauvoir & Sarrodie 2013-2024a).

Archi, together with the ArchiMate modelling language, supports creating EA views and viewpoints. If needed, it also offers a Hints View to facilitate the use of ArchiMate. In addition, Archi offers various elements for creating visualizations, sketching views, and canvas modeling toolkit, among other features. Archi includes the Visualizer feature, which presents selected model elements along with their connections in a radial-tree diagram. It is an interactive graphic that updates instantly based on selections made within the Model Tree, Navigator, or Diagram View, allowing users to explore elements in detail and gain a deeper understanding of their relationships (Beauvoir & Sarrodie 2013-2024a).

Referring to Beauvoir and Sarrodie (2013-2024b), Archi Tool offers different plug-ins to extend its additional functionality. These functionalities can provide features such as gallery previews, collaborative work through repository sharing and versioning of models, and exporting models to spreadsheet format. However, jArchi – Scripting for Archi, allows the use of JavaScript language to script components of an Archi model. With jArchi, users can construct custom tasks and queries to accomplish actions such as:

- *Generate reports*
- *Create heat maps*
- *Import: Various formats including CSV, and the ArchiMate Model Exchange File Format.*
- *Export: PDF, DOC, PPT, and HTML.*
- *Find and remove redundant elements*
- *Batch processing*
- *Query models. (Beauvoir & Sarrodie 2013-2024b)*
- *Import and export to and from different formats (Sarrodie 5 October 2018)*

For generating AI-based solutions using jArchi, Rohde (18 August 2024) shared scripts in the Archi Forum under the topic: *AI-Assisted Modelling Scripts.* These scripts are used to generate EA capability models, models form a schema, meta model, and strategy model by utilizing OpenAI's AI, to interpret user inputs and generate ArchiMate models. These scripts can be used to provide automated schemas and models, easing the workload of enterprise architects.

In a discussion on the Archi Forum under the topic; *Using OpenAI API*, Archi Admin Beauvoir (8 February 2024) explains how jArchi can be utilized to implement NLP-based automation with the OpenAI API. jArchi, which is built on JavaScript, allows for seamless integration with OpenAI, enabling the creation of automated NLP tasks. This integration provides opportunities for automating various EA processes and enhancing the functionality of ArchiMate models via NLP.

2.4 Artificial Intelligence and Advanced Technologies

This chapter provides an overview of AI, focusing on key innovations that support the utilization of capabilities and capability maps within an organization's EA. It delves into AI topics specifically designed to aid the thesis development work, emphasizing the role of AI in creating systems that support EA development. Additionally, the chapter examines pre-defined topics essential for implementing EA using AI, including ethical considerations, security issues, data privacy, and related challenges. The primary objective is to identify and categorize tools and technologies to ensure the system design incorporates appropriate recommendations for specific tools while excluding irrelevant components.

Huawei Technologies CO., Ltd. (2023, 1) presents the lifecycle of AI from its very early phases, as introduced by John McCarthy at the 1956 Dartmouth Conference, outlining the widely accepted definition of AI: *AI is about letting a machine stimulate the intelligent behavior of humans as precisely as it can be*. Schmidhuber (2022, 1-3) provides a long history of AI in his technical report: *Annotated History of Modern AI and Deep Learning*, covering mathematical breakthroughs that have influenced the development of AI from the early 17th century to the present day. However, Schmidhuber's report clarifies that modern AI, with its focus on deep learning and neural networks, is conceptually closer to cybernetics than to the traditional AI approaches defined by McCarthy in 1956, such as expert systems and logic programming.

The concept of neural networks has evolved significantly over time. Adrien-Marie Legendre introduced an early form of what is now called a linear neural network in 1805 (Schmidhuber 2022, 5). In 1965, Alexey Ivakhnenko and Valentin Lapa introduced deep learning by presenting the first general algorithms for deep multilayer perceptrons with many hidden layers. Their pioneering work laid the foundation for modern deep learning, which has become a major area of machine learning (ML), especially notable in eastern Europe where much of ML's development took place (Schmidhuber 2022, 8).

Stryker and Kavlakoglu (16 August 2024) depicts the timeline and relationship between modern AI in Figure 6 (page 25), which visually represents the different phases of AI development and its technological lifecycle. It illustrates the progression from AI to ML, to deep learning, and ultimately to generative AI across different periods. The timeline begins in the 1950s and continues through the ongoing advancements of the 2020s, providing a clearer understanding of the relationships and dependencies between different AI technologies. This aligns with Schmidhuber (2022, 1) account of neural networks and deep learning, which traces back to Alexey Ivakhnenko and Valentin Lapa's 1965 development of general algorithms for deep learning. This laid the foundation for the growth of deep learning as a key component of ML and AI. In a similar vein, Huawei Technologies CO., Ltd. (2023, 4) introduces ML as a technology that enables computers to perform human-like learning and develop new skills by acquiring new knowledge. Deep learning, a subset of ML that derives from artificial neural networks (ANN), focuses on mimicking the way the human brain interprets data, such as images, sound, and text.

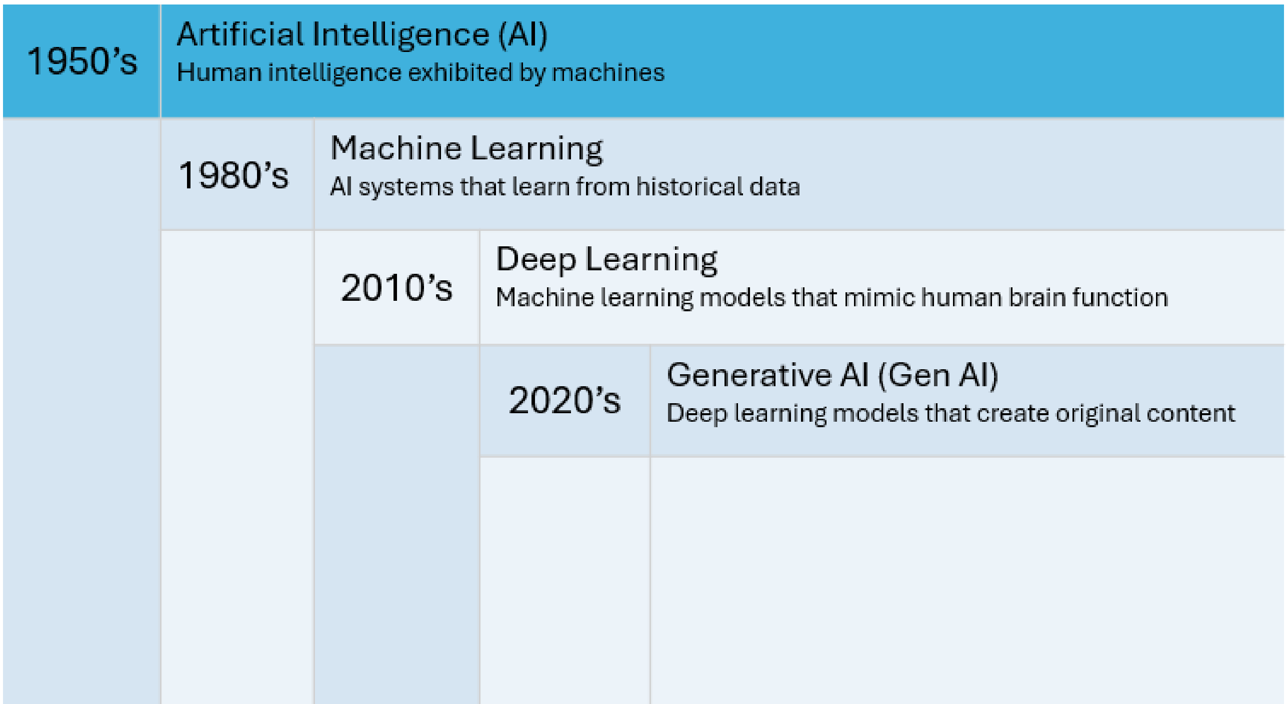

Figure 6. Artificial Intelligence as it has emerged over more than 70 years (adapted from Stryker & Kavlakoglu 16 August 2024)

In 2019, the World Intellectual Property Organization 2019, 13) summarizes how AI is increasingly driving important developments in business and technology. AI power can be harnessed for use in areas ranging from medical diagnosis to autonomous vehicles and advanced manufacturing. This rapid development is not limited to specific regions. According to Huawei Technologies CO., Ltd. (2023, v), the rapid development of information technology and AI in China significantly changes how people live, study, and work. AI is incorporated into China's national agenda, triggering interest across almost all industries. However, as AI technologies continue to expand, they also raise important ethical challenges. Stahl, Rodrigues, and Schroeder (2023, 1) highlight this issue, noting that the ethical challenges of AI are one of the biggest topics of the twenty-first century. While the benefits of AI are said to be numerous, ethical concerns are also associated with many AI solutions, ranging from algorithm bias to serious health and safety issues.

2.4.1 Deep Learning Frameworks

Deep Learning was the fastest-growing AI technique and sub-category of ML between 2013 and 2016 (World Intellectual Property Organization 2019, 32). Dominguez-Morales (2024), in his book *Deep Learning*, introduces the rapid development of deep learning techniques in recent years. In the preface, he briefly explains how deep learning techniques involve training artificial neural networks with large amounts of data. These networks make highly accurate predictions or decisions and continuously learn through these processes. He also presents one of the recent deep learning findings, generative adversarial networks (GANs), which are used in the field of image recognition. GANs generate realistic images by training two neural networks. However, Ping (2024, chapter 3) discusses the challenges GANs face with stabilization and convergence and describes newer technologies that are much more capable than GANs. According to Dominguez-Morales (2024), the second breakthrough in deep learning is transformer networks, which enhance NLP tasks such as language translation and sentiment analysis.

Instead of coding complex neural networks with backpropagation algorithms, it is possible to build a deep learning framework by configuring hyperparameters according to specific needs and adding custom layers to the existing model if required. These frameworks are considered a set of building blocks. Each Individual block or component is a model or algorithm that can be assembled with other components to meet the required architecture. The most widely used deep learning frameworks currently include TensorFlow, PyTorch, Caffe, and so on (Huawei Technologies CO., Ltd. 2023, 15).

TensorFlow has a long history, with more than 100 released versions. The product has developed a comprehensive ecosystem, evolving from the early prototyping stage to the productization of the model (Ganegerada 2022, chapter 1.1). Ping (2024, chapter 5) presents TensorFlow in his book *The Machine Learning Solutions Architect Handbook*. TensorFlow is an open-source ML library primarily backed up by Google. It can be used to build systems that support various use cases, including answer questioning, speech recognition, text summarization, forecasting, and computer vision.

TensorFlow is integrated with Keras, which serves as its high-level neural network API. Keras is used to create and train deep learning models within the TensorFlow (Kapoor, Gulli & Sujit 2022, chapter 1). TensorFlow can also be used to implement probabilistic ML models, perform computations related to computer graphics, reuse pre-trained models, and visualize/debug TensorFlow models. These models can be predefined or custom-layered implementations

(Ganegerada 2022, chapter 1.1). As a complete ML framework, TensorFlow supports various capabilities and stages of projects. Popular ML datasets can be downloaded into TensorFlow with a single line of code. Keras data generators provide access to different types of data from various sources, such as disks (Ganegerada 2022, chapter 1.1.1).

PyTorch, a popular open-source ML library for deep learning, is initially released in 2016. Many technology giants use PyTorch to train their own deep learning models, such as those for NLP and computer vision. PyTorch implements C++ programming language for the backend, offering ease of use while supporting interoperability and dynamic computational graphs, with the rest of the system relying on Python code. The PyTorch library consists of several key modules, including tensors, autograd, optimizers, and neural networks. Tensors store and manipulate multidimensional arrays of numbers. Various operations can be performed on tensors, such as matrix multiplication, transposition, finding the maximum value, and manipulating dimensions. One of the most essential aspects of the deep learning library is the training dataset, which needs to be prepared and identified to start using PyTorch (Ping 2024, chapter 5).

Nvidia (2024) lists various examples of PyTorch-related use cases on their webpage, emphasizing its application across various fields of research and industry. For example, Salesforce utilizes Pytorch for multi-task learning and NLP-based tasks. PyTorch's interface is easy to use, and its support for rapid prototyping and flexibility makes it well-suited for research. Stanford University develops and tests new algorithms using PyTorch, emphasizing its adaptability and efficiency.

2.4.2 Natural Language Processing and Large Language Models

NLP enables machines to interpret and understand human language in text form. Kulkarni and Shivananda (2021, Introduction) highlight how NLP, when coupled with ML and deep learning, enables the extraction of actionable and significant insights from text data. By leveraging these technologies, NLP offers numerous opportunities for addressing complex AI-related challenges and has emerged as a pivotal area in the development of intelligent, deep learning-based applications.

To implement NLP effectively in conjunction with ML and deep learning models typically requires installing Python-based libraries such as TextBlob, NLTK, Spacy, and others. Python provides a wide range of libraries for performing various NLP tasks and is the most widely used language for building NLP applications (Kulkarni & Shivananda 2021, Introduction). Gollnick (2024) discusses several NLP-based applications in his audiobook, including chatbots such as OpenAI’s ChatGPT, Google’s Bard, and Meta’s LLaMa. These applications are used for sentiment analysis, text translation, speech recognition, and much more. They utilize neural network models to process input text, generate predictions, and provide responses (Gollnick 2024, chapter 2).

Wassan and Ghuriani (2024) list various NLP modeling techniques that produce the required results. These techniques include logistic regression, a statistical model; Naïve Bayes algorithm, a probabilistic classifier; decision trees, used for classification and regression; the Hidden Markov Model, a stochastic model that describes a system transitioning between states at discrete time intervals; and convolutional neural networks (CNNs), which automatically detect and classify defined properties.

As NLP continues to advance, Generative Pre-trained Transformers (GPT), developed by OpenAI, represents a class of NLP models for text generation, translation, text completion, question answering, and more (Wassan & Ghuriani 2024). Looking to the future, Wassan and Ghuriani (2024) highlight LLMs such as OpenAI's GPT-3, which provide full open-ended support for human-like conversations. LLMs offer more versatile and nuanced responses and are better at understanding user intent compared to previous versions. Referencing the deep learning techniques discussed earlier, Gheorghiu (2024, chapter 1) presents the best-performing LLMs, such as Llama2, Claude 2.1, and GPT-4, which contain trillions of parameters and are trained on extensive internet-scaled datasets with advanced deep learning methods.

Although LLMs achieve top-level results when applied to downstream NLP tasks, they still face limitations in manipulating knowledge and generating accurate information, which affects their performance in task-specific architecture. Additionally, challenges in maintaining updated data further complicate their application in knowledge-intensive tasks. To address these challenges, RAG models are proposed as a potential solution. RAG models combine pre-trained parametric and non-parametric memory to provide more fact-based, accurate, and diverse answers. (Lewis et al. 2020, 1).

2.4.3 Retrieval-Augmented Generation

RAG offers several potential benefits for enhancing LLMs, as discussed by Taulli (2024). These benefits include improved timeliness, better handling of proprietary information, reducing hallucinations with transformational models, and increased cost-effectiveness. However, as Gheorghiu (2024, chapter 1) points out, while RAG enhances the reliability of LLMs and improves their performance, it does not make them entirely immune to generating incorrect or misleading answers. According to OpenAI (2024), RAG works by retrieving relevant contextual information from a specific data source. Figure 7 (page 29) illustrates the structure of RAG by showing how a specific data source connects to the LLM model, allowing it to return content in response to user prompts. Instead of relying merely on the LLM's general knowledge base, RAG accesses the connected data source, enabling it to analyze and answer questions related to the specific data. This figure provides a high-level overview of RAG's data flow, illustrating how RAG supplements LLM-generated responses with real-time or organization-specific information, thereby enhancing contextual accuracy and relevance.

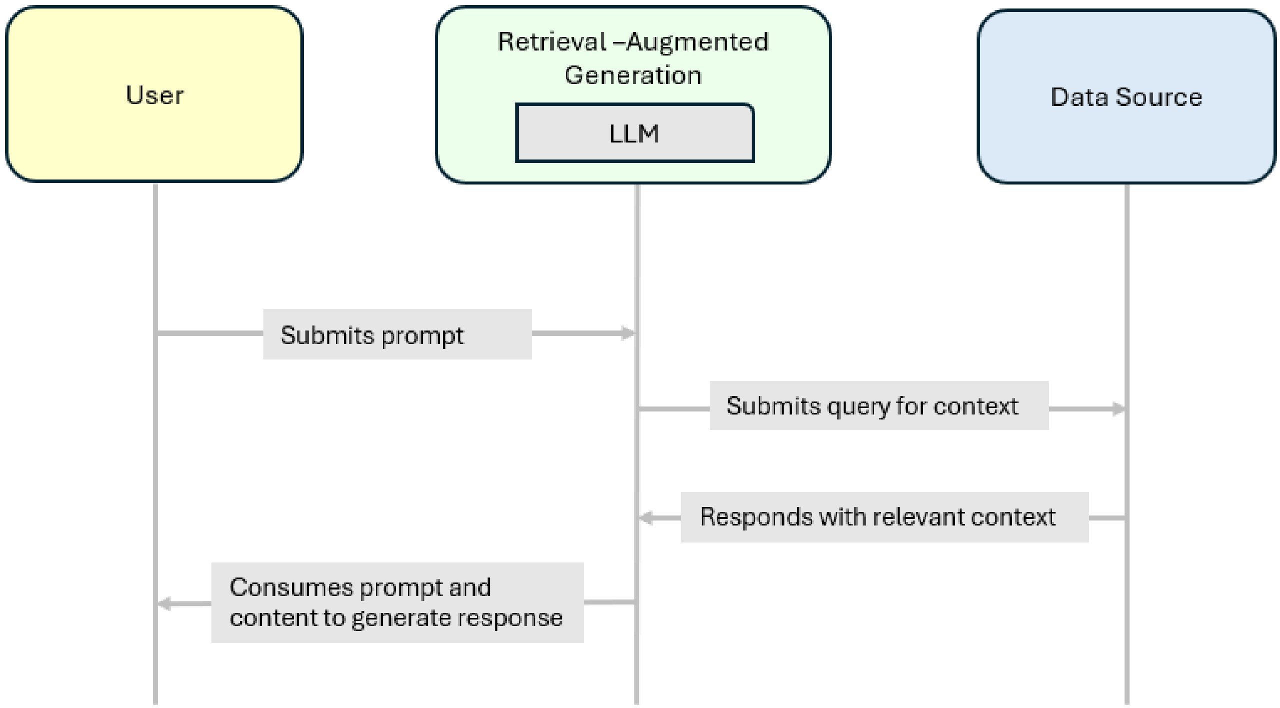

Figure 7. Basic RAG Workflow (adapted from Open AI 2024)

More specifically, Zeng et al. (2024, 4507) describe the typical process flow of the RAG Pipeline in mathematical terms and explain how the RAG system retrieves information. This process is expressed as $R(q, D) = \{d1, d2, ..., dk\} \subseteq D$, where the retriever R is responsible for selecting the top-k relevant documents from dataset D, referred to as the Data Source in Figure 7 above, that are associated with the query q. Extraction is based on calculating the similarity or distance between the query embeddings (eq) and the stored document embeddings (edi) using a ML technique (K-Nearest Neighbors), The selection process can be defined as: $R(q, D) = \{di \in D \mid dist(eq, edi) \text{ is in the top } k\}$. Once the relevant documents are retrieved, the RAG system concatenates these documents with the original query *q* to form a comprehensive input for the language model M, with the final input-output relationship expressed *as:* $a = M (R (q, D) \| q)$.

2.4.4 Retrieval-Augmented Generation System: Technical Workflow and Integration

To connect the data source to LLMs, as described in Figure 7 (page 29), Gheorghiu (2024, chapter 2) introduces LlamaIndex as an easy way solution for establishing the connection. LLMs such as Llama, Claude, and GPT-4 integrate custom knowledge through LlamaIndex, creating a bridge between specific data and the LLMs capabilities. Before data can be integrated, it must undergo preprocessing and validation steps to ensure it is in a usable format for LLMs. This approach mirrors the method described by Goyal et al. (21 May 2024) who extract text from PDFs using the PyPDF2 library. The LangChain package is then utilized to integrate LLMs with the custom knowledge base, including splitting text into more manageable chunks with CharacterTextSplitter and creating a Chroma vector database to process relevant queries.

Preprocessing and validation are critical steps in this pipeline. Gheorghiu (2024, chapter 4) introduces several tools for data preprocessing and validation in conjunction with LlamaIndex. These tools are designed to preprocess and clean data from various formats. Hugging Face is particularly effective for scrubbing personal and other sensitive information. In addition, simple text splitters can be used to break down documents into smaller pieces. The SentenceSplitter splits text and automatically extracts nodes from documents. The TokenTextSplitter operates at the token level and is capable of splitting text into boundaries suitable for NLP. The CodeSplitter interprets source code and splits text-based on programming languages, making it ideal for managing technical documentation. This preprocessing phase ensures that documents are parsed into smaller units, or nodes, allowing for better handling of internal structure and maintaining accuracy (Gheorghiu 2024, chapter 4).

Once the data has been preprocessed and validated, the next step involves embedding the data into a vector database. Vector databases store, manage, and query data represented in geometric formats. They are a special type of database that is popular in NLP and can be used to store high-dimensional data (embeddings) for similarity analysis and fast querying. They process images, texts, videos, and sounds (Gollnick 2024, chapter 5). In addition to Goyal et al. (21 May 2024) using the Chroma vector database in their project to implement an NLP system, Gollnick (2024, chapter 5) lists Chroma, Pinecone, and Redis as alternative options for vector databases.

In differentiating between vector databases, Dasgupta (2024, chapter 2) introduces Pinecone, Milvus, and Weaviate as vector databases with robust capabilities, in contrast to FAISS and ChromaDB, which primarily provide vectorized indexes. Gollnick (2024, chapter 5) describes the full picture of using vector database. The first step is to get a corpus, the text data to be added into the database. After that, the text data needs to be tokenized, and the database is set up. Finally, the tokenized data can be added to the database.

These tokenized data points, known as embeddings, represent the core data structure stored in vector databases. Dasgupta (2024, chapter 2) presents language embeddings or vector embeddings as a numerical representation of non-numerical data objects. These embeddings can include natural language, video, or sound. The semantic meaning of words and their relationships within language are captured by text embeddings. In this way, semantic similarities can be encoded between words, for example, *king* is closer to *queen* than to *car*. Embedding models are essential for converting text into a machine-readable format. Common models for implementing text embedding into a vector database are TF-IDF, Word2vec, and BERT. At a general level, TF-IDF creates sparse embeddings, whereas Word2Vec creates dense vector representations, and BERT creates context-rich embeddings. These text embeddings can be used in various ways, including similarity search, information retrieval, and visualization, among others.

After the data has been embedded and stored in a vector database, it is ready for retrieval and processing by LLMs. To interact with LLMs, Dasgupta (2024, chapter 1) introduces prompts that input instructions, external information, output indicator, and user input. This approach aims to streamline interactions with LLMs while enhancing data management and compliance issues related to the RAG system. The prompts are not hardcoded; rather, they are generic solutions that allow for specific values to be changed as needed.

Dasgupta (2024, chapter 1) presents Chain of Thought (CoT) prompting as a helpful technique for guiding language models to generate step-by-step reasoning through the prompt. Dasgupta (2024, chapter 1) explains how CoT prompting works within the LangChain framework to guide LLMs through a series of logical steps to complete complex tasks. Also, Goyal et al. (21 May 2024) used LangChain to leverage LLMs for answering queries in their project aimed at enhancing EA through NLP. Furthermore, Dasgupta (2024, chapter 1) introduces both LangChain and LlamaIndex as tools for creating and working with prompt templates, which are dynamic and model-agnostic.

In addition to prompting, several frameworks support the interaction between LLMs and custom data sources. In a manner comparable to LlamaIndex, Haystack provides an open-source framework for building RAG pipelines and production-ready LLM applications. It enables the development of advanced RAG systems using the latest retrieval and generation techniques, including generative multi-modal question answering on diverse knowledge bases containing images, text, audio, and tables. Haystack also supports the creation of chatbots and agents, as well as information extraction to input data into databases or knowledge graphs from documents. Haystack (2024)

Finally, the orchestration of the advanced RAG system seamlessly integrates all components. Dasgupta (2024, chapter 2) provides a detailed architecture diagram in Figure 8 below, illustrating the query flow. The diagram presents the query flow within the RAG system, showing a query is either passed directly to the LLM model to return an answer, or how it is routed through various processing steps including query transformation and query routing through either the Vector Store Index, Summary Index, or directly to the DB storage. The content retrieved from the DB storage is then re-ranked, and the final context is passed to the LLM to produce an answer. The orchestration illustrates how a query is forwarded to a vector database, from which it retrieves organization specific content for the LLM, allowing it to return an answer either by combining the query with the directly sent query or by returning a response directly from its own processing. The system processes the query to infer the correct content from the database vectors.

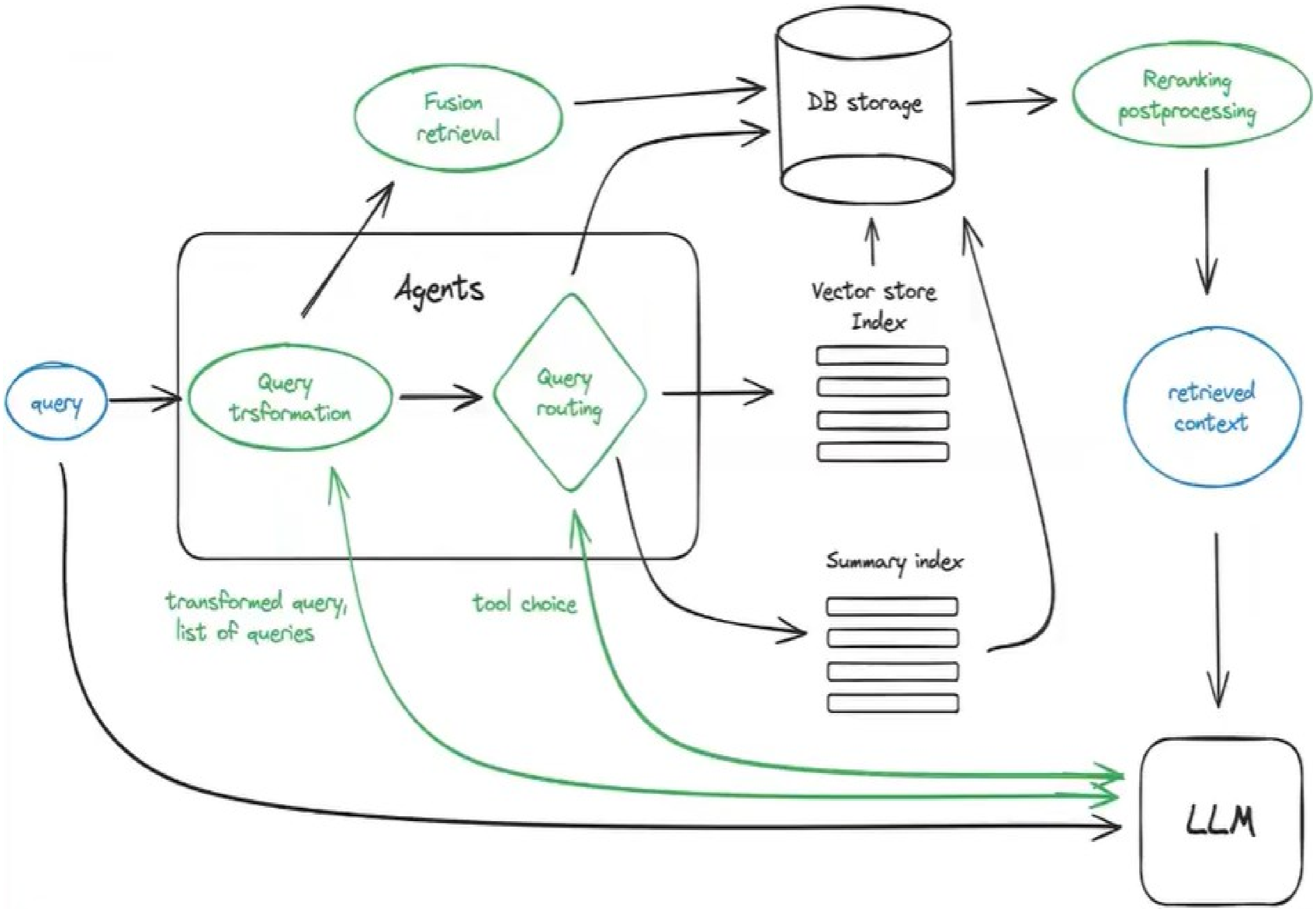

Figure 8. Advanced RAG Orchestration (adapted from Dasgupta 2024, chapter 2)

2.4.5 Compliance and Privacy in Retrieval-Augmented Generation System

Zeng et al. (2024, 4505-4506) have conducted an extensive empirical study on novel attack methods against RAG systems. Their research demonstrates how LLMs have been shown to pose a privacy risk and how RAG may potentially increase these privacy issues. Their research shows that LLMs often inadvertently reveal data from pre-trained corpora, a tendency attributed to the way LLMs memorize, recall, and reproduce their training data.

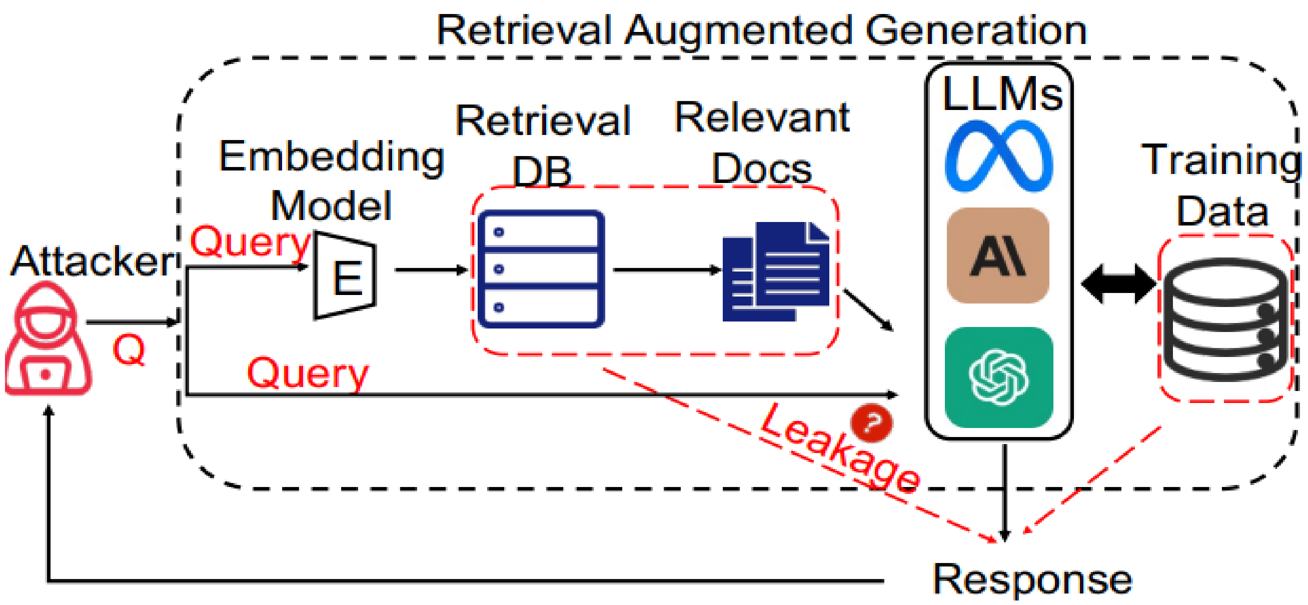

Figure 9. The RAG System and Potential Risks (Zeng et al. 2024, 4505)

Figure 9 above illustrates a potential vulnerability in the RAG process, where data leakage occurs through the LLM when data from the retrieval database is combined with the LLM's response and training data. The figure highlights how the interaction between the retrieval data, the LLM's response, and the trained data gives rise to privacy risks, emphasizing the potential for sensitive information to be inadvertently exposed during the generation of the LLM's response.

Zeng et al. (2024, 4506) note that numerous studies have highlighted the vulnerability of LLMs in inadvertently revealing data from pre-trained corpora. This tendency is attributed to the way LLMs memorize, recall, and reproduce their training data. Subsequent research has shown that factors such as model size, data duplication, and prompt length can significantly increase the risk of memorization. Fine-tuning processes further amplify the risk of memorization, especially when modifying model heads, which leads to greater memorization compared to smaller adjustments. Zeng et al. (2024, 4505-4506) study consists of two main research questions:

- *(RQ1) Can we extract private data from the external retrieval database in RAG?*

- *(RQ2) Can retrieval data affect the memorization of LLMs in RAG?*

To answer these research questions, they used various attack techniques against the API requests, prompts, and chatbots operating on top of the RAG framework's different components to reveal leakage risks and privacy vulnerabilities (Zeng et al. 2024, 4506-4507). Their research results for *RQ1* revealed the RAG system's high vulnerability when subjected to targeted and untargeted attack methods. The LLMs used in the RAG system were commonly safety-aligned models, including Llama-7b-chat, Llama-13b-chat, and GPT-3.5-turbo (Zeng et al. 2024, 4508).

For the *RQ2* Zeng et al. (2024, 4511) harnessed GPT-Neo-1.3B generative model in use. For the dataset, they integrate Enron_Mail dataset, a subset of pre-trained data for the GPT-Neo-1.3B, as well as several other datasets, including wikitext-103, HealthcareMagic, and w3c-email, to ensure coverage of different scenarios and improve the generalization of the study. To isolate the impact of retrieval data, the authors added baselines with 50 tokens of random noise and protective prompts, distinguishing retrieval augmentation from extra content.

The findings of the research revealed that when the retrieval data consists of entirely disparate data types, the LLM demonstrates a decreased ability to identify personally identifiable information (PII). However, when retrieval data includes a dataset with PII, such as W3C-Email, the LLM tends to produce more retrieval data rather than training data. Integrating retrieval data into the RAG system decreases risk of privacy leaks from LLMs' training data. By utilizing non-sensitive public data or carefully desensitized data as retrieval content, RAG can effectively protect private information and significantly reduce the risk of information leakage from LLMs. Implementing RAG alongside LLMs, as described, can help minimize potential data leakage (Zeng et al. 2024, 4512-4513).

2.5 Literature Review Summary

The literature review summary encapsulates the key observations from the three main topics addressed in the review. These topics integrate the relevant themes into a comprehensive context that supports the research objective of developing a process model and an AI-based solution for generating EA capabilities and capability maps. This includes previous research on both EA capability maps and AI usage to implement EA-based solutions, with a particular focus on capability maps within EA.

There have been a few previous studies available related to the two main research topics of this literature review: capability maps in EA and implementing EA using AI-based solutions. While EA has been widely researched, the concept of capability maps is relatively new. Findings emphasized the significance of the capability map in strategic decision-making, as it has served as a bridge between organizational management and operational development.

AI has advanced rapidly in recent years, introducing new concepts such as LLMs and deep learning. Granger and Baragry's (20 Feb 2024) have discussed in their project the relevance of these technologies, including ethical questions related to LLMs and RAG. Goyal et al. (21 May 2024) have implemented an EA-based solution with AI and RAG technology, highlighting how hallucinations are common issues with LLMs and suggesting mitigation through continuous data training, ensuring diversity in the RAG system's data, and regular data updates.

Although advanced AI has been a relatively novel technique, it has probably led to one of the most interesting research topics in the early 2020s. This has been evident in the research and literature related that has emerged over the past few years. Additionally, projects related to EA and capability map development using AI-based solutions have increased. Research by Goyal et al. (21 May 2024) suggests the substantial real-world impact of AI in improving cost savings, productivity, and overall process efficiency across industries.

The literature review has emphasized that capabilities and capability maps are essential for supporting an organization's strategic management. These concepts have been closely connected to business and strategy and have typically been developed early in the EA lifecycle. The capability map has proven to be flexible and can be constructed using various existing elements within the organization. Established norms and phases have guided the implementation of capability maps, and several tools have been made available to help organizations evaluate and utilize both capabilities and the capability map. The Open Group's TOGAF ADM framework has been selected for this study. According to Josey and Hornford (2022), it is widely recognized and used in the field of EA management. It provides clear guidelines and follows dynamic phases that can be applied in an agile manner.

AI has a long history and has made significant advancements in recent years. It is a complex and broad discipline with various forms and subfields. In this research, AI's potential for supporting the utilization of EA and capability maps has been explored, narrowing the focus to NLP and LLMs. As Lewis et al. (2020, 1) have concluded, while LLMs excel in many NLP tasks, they have faced limitations in accurately manipulating knowledge and maintaining updated data, which affects their performance in task-specific applications. To address these challenges, RAG models have been proposed as a potential solution. RAG models combine pre-trained parametric and non-parametric memory to provide more fact-based, accurate, and diverse answers.

This study has identified RAG as a key technology that supports the development of EA while emphasizing secure data handling and reducing hallucinations. RAG has shown significant advantages, especially when generating organization-specific content. RAG systems include various tools and techniques, for example, Gheorghiu (2024, chapter 2) who has introduced LlamaIndex as an easy way solution to connect a data source to LLMs. Dasgupta (2024, chapter 2) has highlighted various vector databases that are vital when implementing RAG system.

To implement the process model that integrates EA and AI-based RAG solution, it was essential to research a high-level architectural overview of the system, including the main technical phases required for the implementation of the RAG system and defining capabilities and a capability map in EA, as established in previous research and new formal literature. This composition of the literature review on existing EA theories and AI technologies has formed the foundation for the proposed process model and has guided the practical implementation of the RAG solution. It focused on key components, such as the integration of the Archi Tool into the RAG system design and the precise definition of the required components for the system's functionality. This approach has provided answers to all research questions.

3 Methods and Implementation

The Methods and Implementation chapter begins by outlining the aim of the thesis work, emphasizing the key topics addressed in the study. The chapter describes the research approaches and explains the rationale behind the selection of specific methods for the development work, presenting them in a logical order to demonstrate their alignment with the thesis objectives. Each method is detailed and justified within the context of the research, with the progression stages described chronologically to give the reader an understanding of the research implementation. Additionally, the data collection methods and analysis techniques are explained and justified in the context of their contribution to the research.

The aim of the thesis was to develop a comprehensive process model for capability mapping and its utilization within EA, supported by advanced technologies and AI. This addressed the inefficiencies and manual redundancies in creating and maintaining EA in fast-paced, agile business environments and supported the effective utilization of capability mapping. The research focused on optimizing EA content production through selected tools and methodologies, aiming to create a more dynamic and adaptive process.

The objective of this development task was to identify AI-based solutions to enhance EA development in an enterprise context. By integrating these AI solutions with EA frameworks, the thesis led to the creation of a semi-automated or automated process to streamline the creation of EA descriptions and models. The outcome was a more efficient and agile approach to EA management that ensured continuous alignment of organizational capabilities with strategic goals and operational needs.

3.1 Research Approach

The research approach for this study was based on the requirements outlined in the aims and development tasks, where the development and implementation of EA and AI-based solutions were grounded in information technology and technology development. This means that the data collected to be continuously renewed and processed iteratively. The result of the development task evolved iteratively throughout the course of the project. Lincoln (2021, 17) presented qualitative research methods as highly suitable for gathering data consistently during project implementation by employing techniques such as interviews, questionnaires, and observations. To design and develop the EA process model with AI technology, a comprehensive literature review along with ongoing evaluation and planning were essential. Lincoln (2021, 4-5) aligns qualitative research methodology with the production of descriptive statistical methods, distinguishing it from quantitative methods, which rely more on inferential statistics.

The research was grounded in real-world case scenarios, as outlined in Appendix 1 (pages 81–84), with the objective of developing an effective and robust model for practical applications. The Open Group (2022b, Business Scenarios) recommended that architects strengthen business scenarios with case studies. They outline several key steps to consider for case study activities, such as data gathering, identifying key questions to be answered, data preprocessing, and conducting one or more workshops, among others. This evidence was supported by Bondel et al. (15 November 2018), who described the case study as an appropriate research methodology for studying phenomena in their natural setting within the field of information system science. They implemented a capability map for a medium-sized organization using an exploratory approach together with a single use case. Farquhar (2012) defined case study methods as a strategy that supports the collection of data for qualitative, quantitative, or mixed-method research.

In this study, the choice of qualitative research methods was motivated by the need to understand complex processes and organizational context, which could not be easily quantified. The use of case studies was considered particularly relevant, as it allowed the examination of EA implementation in its natural setting. By employing DSR alongside case studies, the research was able to leverage iterative feedback and refinement cycles, ensuring that the developed EA model remained practical and grounded in real-world scenarios.

To ensure that all necessary information and data of the research were gathered and analyzed as effectively as possible, and to support the use cases in the study, the DSR method was selected as an excellent choice alongside case studies. The DSR method was particularly well-suited for this study because of its focus on solving real-world problems and creating innovative solutions through the development of artifacts. According to Vaishnavi and Kuechler (2015), the DSR method has been widely recognized for its effectiveness in guiding the development and innovation of information and communication technology artifacts. The choice of DSR was driven by its ability to contribute new knowledge, typically in the form of theory, which aligns with the objectives of this research to advance the understanding of EA and AI integration.

The DSR supported various artifact development techniques for problem-solving, such as case studies or experimentation. Additionally, DSR supported both qualitative and quantitative object solutions (vom Brocke, Hevner & Maedche 24 September 2020). Vaishnavi and Kuechler (2015, 18) emphasized how DSR was particularly effective in complex technical projects, where the DSR process model supported iterative approaches across its phases to develop, evaluate, and suggest changes throughout the circumscription of knowledge contribution.

3.1.1 Case Study

Research suggests that the case study method is a versatile and multidisciplinary research approach that can be applied for business-related scenarios in EA as well as developing complex technical solutions in information technology. The primary challenge of this study involved identifying specific use cases, their phases, and their interrelationships. This challenge was closely tied to process development and the need to connect AI to the process model. To support this and the need to collect qualitative data, Gillham (2010, 10) argues that the case study method enables understanding the meaning of what is going on. It illuminates issues and reveals explanations, especially when searching for meaning. However, as noted by Gillham (2010, 1), the case study approach has its limitations in that it can be difficult to establish clear and precise boundaries around the case.

Gillham (2010, 1) lists four points to understand and define a case:

- *a unit of human activity embedded in the real world*
- *which can only be studied or understood in context*
- *which exist here and now*
- *that merges in with its context so that precise boundaries are difficult to draw*

In this research, the case study approach is vital for understanding work phases involved in capability and capability map definition, their utilization in context, and especially how these contexts are redefined and changed. Yin's (2009, in Farquhar 2012, chapter 1) aptly defines case study research as an empirical method used to explore a contemporary phenomenon within its real-life context, especially in situations where it is difficult to distinguish the phenomenon from its surrounding context.

When selecting research methods, it was essential to find mutually supportive approaches whose practical implementations differ, but whose data collection and analysis support each other. Farquhar (2012, chapter 2) emphasizes how mixed methods are an increasingly recognized approach offering significant benefits for research. The case study method supports other research methods very well in its approach to collecting and analyzing data.

Case studies are often categorized into deductive and inductive research logics, with inductive being more common. It produces theory from data by exploiting patterns within the data and is concerned with explorations and understanding, whereas the deductive approach involves developing a conceptual framework that is then tested (Farquhar 2012, chapter 2). Inductive research logic was chosen, as its phases serve from the early stages to the end of the research process. Eisenhardt (1989 in Farquhar 2012, chapter 2) outlines phases for defining and selecting cases, collecting and analyzing qualitative data, and iteratively shaping findings, concluding with comparing the results to similar and conflicting literature.

To support the case study, a model was selected from Yin's (2009, in Farquhar 2012, chapter 3) five components of case study research design:

1. *Component: Case study*
2. *Research Question: Refining research questions by using literature and exploring a phenomenon in a new context*
3. *Propositions/Objectives: Develop a set of propositions or objectives that will guide the study of a specific aspect of the research question and determine how it may be investigated*
4. *Unit of Analysis: Fundamental question of what constitutes the case. What are the boundaries of the case, and what is the focus of the investigation?*
5. *Linking Data to Propositions: This pertains to data analysis. If theory is driving the data collection and analysis (deductive approach), then the propositions structure the data and discussion. In contrast, if the research is inductive, emerging patterns will inform the structure*
6. *Interpretation of Findings: The literature will have examined competing theories and studies, and it is crucial to return to these discussions when interpreting your findings*

This model was selected to ensure the structure and consistency of the case study approach. Yin's five components provided a clear framework for guiding the entire study. The Research Question component helped refine the focus, particularly on understanding how capability mapping and AI integration could be applied in EA. The Propositions/Objectives component allowed the study to focus on specific aspects of the research question, ensuring that the investigation remained targeted and organized. The Unit of Analysis defined the scope, identifying relevant elements of the EA process. The Linking Data to Propositions component ensured that patterns from qualitative data were captured and analyzed in a way that aligned with the theoretical framework. Finally, the Interpretation of Findings allowed the study to place the results within the context of existing literature, ensuring that the findings were compared and validated against similar studies.

For the Propositions and Objectives phase in the five components model, three use cases, as outlined in Appendix 1 (pages 81–84), were defined based on Yin's (2009, in Farquhar 2012, chapter 3) framework. These use cases facilitated a detailed analysis of capability mapping and its implementation for further development, aligned with the guidelines from the city of Helsinki.fi (5 November 2021). The use cases included specific scenarios describing interactions between users and systems, illustrating how different capabilities are applied in real-world contexts. These use cases provided a structured framework for analyzing the integration of AI into the EA process and its impact on capabilities and Capability Map development. The use cases included:

1. *User*
2. *Target*
3. *Trigger*
4. *Use case Flow*
5. *Exceptions*
6. *Notes*

To support Propositions and Objectives phase, it was necessary for the data collection and analysis to be in harmony and maintain consistency. For these use cases, the most suitable data gathering methods were interviews and documents in both unstructured and semi-structured (thematic) forms. Lincoln (2021, 183) presents case studies that utilize multiple techniques of data collection, emphasizing the importance of planning interview questions to effectively gather data (Lincoln 2021, 109). Additionally, Farquhar (2012, chapter 5) lists various data gathering techniques for case studies, placing interviews, observations, archives, and documents at the center of these techniques. A more detailed outline of the thematic interview framework, including the questions and responses, is presented in Appendix 3 (pages 90–94).

Research on data collection and analysis in qualitative case study research indicates that data collection and analysis are often combined, with data frequently being written up and analyzed concurrently. Farquhar (2012, chapter 6) recommends implementing short notes right after in semi-structured interviews. However, the data may not always be in the right format for the main analysis, where researchers analyze what informants have or have not said. This kind of preliminary analysis creates a foundation for subsequent data collection. Analysis options are driven by two approaches: deductive and inductive. Inductive analysis provides insights from the collected data by identifying emerging theoretical constructions that are supported across the interviews. Another limitation of the case study method, as noted by Yin (2009, in Farquhar 2012, chapter 3), is that the data collection and analysis process can be influenced by the researcher's subjective interpretations. This can introduce bias, especially if the researcher has preconceived notions or assumptions about the phenomena being studied. This subjectivity can sometimes challenge the validity and generalizability of the findings, especially in the "Interpretation of Findings" phase, where literature and competing theories should be revisited to frame the analysis.

3.1.2 Design Science Research

To support the development of AI and EA in this study, the DSR framework offers a suitable approach that is useful for the innovation and development of information and communication technology artifacts. Vaishnavi and Kuechler (2015, 9) describe DSR as a methodology suited for development tasks that contribute new knowledge, typically in the form of theory. Building on this, DSR also addresses the creation of new artifacts in areas where existing knowledge is inadequate, especially in complex technical projects. Its structured approach to innovation is crucial for integrating AI into EA. The iterative nature of DSR allows for ongoing feedback and refinement throughout the research process, enabling adaptation to new findings and insights. Vaishnavi and Kuechler (2015, 14) highlight DSR's ability to address the challenges associated with the research case study by producing new knowledge, particularly in areas where existing knowledge is lacking and where the associated risks are significant.

Furthermore, DSR is particularly effective in complex technical projects, where the DSR process model supports iterative approaches across its phases, allowing for continuous development, evaluation, and the proposal of changes at all stages as the research progresses (Vaishnavi & Kuechler 2015, 18). To build upon this, one of the key reasons for choosing DSR alongside the case study research method was its ability to support various artifact techniques for problem-solving, such as case studies or experimentation (vom Brocke et al. 24 September 2020). However, throughout the study, the DSR methodology proved to be time-consuming and required a high degree of precision at each stage. The iterative nature of DSR, while beneficial for adapting to new insights, made the process complex and at times demanding, especially during the evaluation and development phases, where careful attention to detail was necessary to ensure the artifacts addressed the identified problems effectively. Additionally, the repeated analysis and implementation of interviews added further complexity, requiring careful synthesis of qualitative data over multiple rounds to ensure that the evolving artifacts remained aligned with requirements and feedback.

The computable design process model presented by Takeda, Veerkamp, Tomiyama, and Yoshikawa (1990, in Vaishnavi & Kuechler 2015, 15) is shown in Figure 10 (page 43). It highlights the importance of *Knowledge contribution* in the DSR process. Knowledge is created and accumulated through action, following a cycle where actions are taken, outcomes are evaluated, and the process is repeated. This cycle illustrates how knowledge is creatively applied to produce outcomes, and how these outcomes are evaluated to further enrich and enhance knowledge.

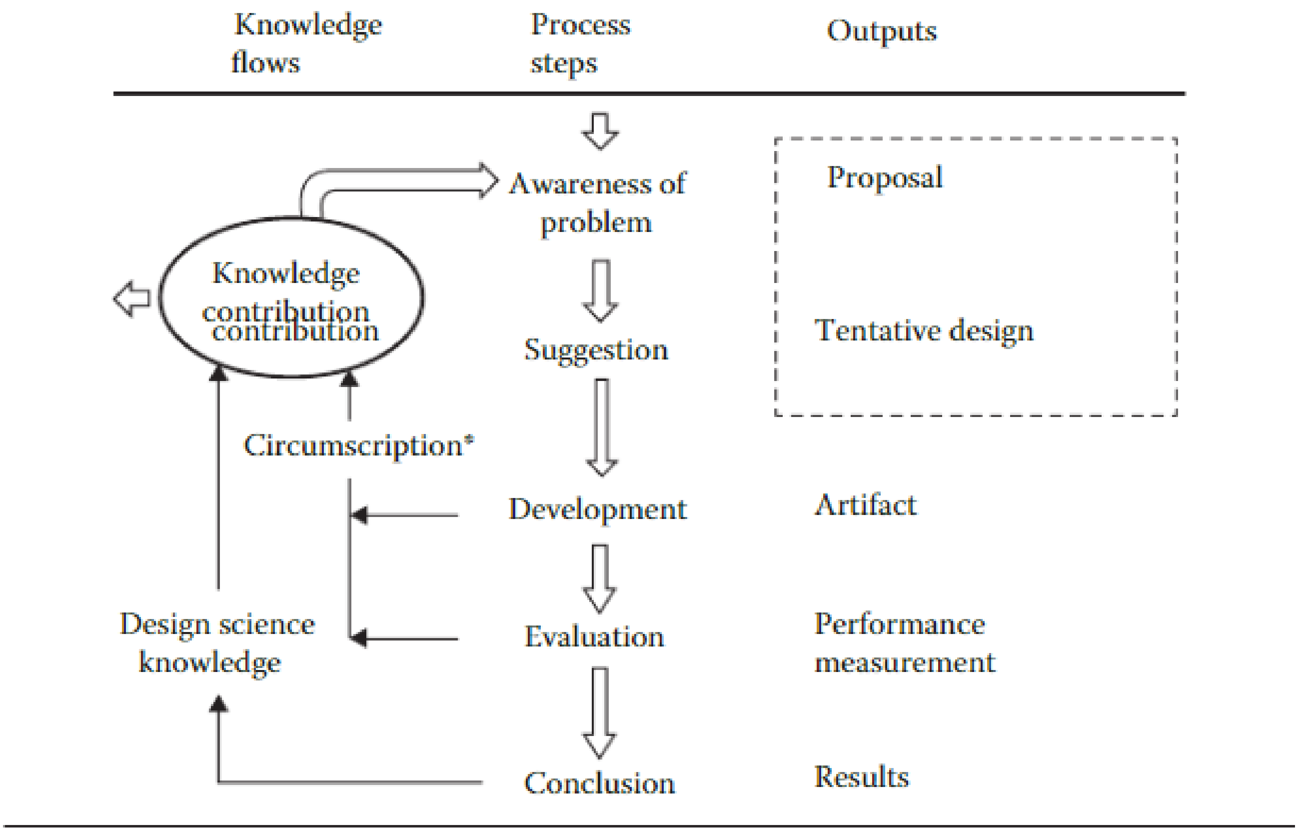

Figure 10. Design Science Research Process Model (DSR Cycle) (Vaishnavi & Kuechler 2015, 15)

Figure 10 above illustrates how knowledge flows through different process steps, producing specific outputs at each step of the process. According to Vaishnavi and Kuechler (2015, 15-17) the tentative design and performance of a prototype are an essential objective of the development task, contributing significantly to the proposal output. The suggestion phase is a creative step where new or existing elements, combined with new configurations, are used to envision new functionality. The output of suggestion phase, which is the tentative design, is further developed and implemented during the development phase, aiming at producing an artifact. This artifact is then evaluated based on criteria established in the awareness of the problem phase. The evaluation can be implemented using either qualitative or quantitative methods. The conclusion phase often produces results that are satisfactory, leading to the need for a new research cycle, or the results may be adjudged "good enough". However, the results can also be the final product of the specific research effort.

Sonnenberg and vom Brocke (2012, in Gimbel 2024a, min. 5:37–9:05) present design principles of DSR that follows the steps of identifying a problem, designing, constructing, using, and iteratively returning to the problem identification if necessary. Their presentation emphasizes the iterative nature of the process, showcasing the ongoing interaction between each phase. Referring to Prof. Dr. Gimbel (2024) from the University of Hohenheim, DSR must produce a viable artifact, which can be a model, a method, a construct, or an instantiation. The goal is to develop a technology-based solution to the business problem (Prof. Dr. Gimbel 2024a, min. 3:00-4:00).

Prof. Dr. Gimbel (2024a, min. 5:37–9:05) recommends using methods such as surveys, expert interviews, and structured literature research to *identify the problem* and assess its relevance. The design phase facilitates the evaluation of the identified problem, which has been validated, and the generation of new knowledge based on design principles such as feasibility, completeness, simulation, and demonstration. The construction phase involves building a prototype based on the abstract design. Once the prototype is developed, it can be evaluated through demonstrations to a focus group, comparing its feasibility and efficiency against the interests of the focus group. The use phase relates to the instantiating of the prototype. The fourth and the last evaluation phase in the design principles focuses on assessing the prototype's applicability and efficiency through field experiments conducted with the focus group.

Numerous techniques can be used to collect both qualitative and quantitative data in DSR. Prof. Dr. Gimbel (2024b, min. 0:28–1:17) identifies behavior-oriented qualitative data collection methods, including qualitative interviews, case studies, GT, and literature reviews, among others. Additionally, he highlights design-oriented methods suitable for formalized and structured qualitative approaches, such as prototyping and commenting. Supporting Dr. Gimbel (2024b, min. 0:28–1:17) definition of DSR data collection methods, vom Brocke et al. (24 September 2020) reference DSR-related studies that apply various data collection methods, including interviews, literature reviews, surveys, among others. When explored in more depth, Galletta and Cross (2013, 75) introduce interviews as particularly effective in prompting participants, allowing for the rephrasing of questions and adjustments during the interview. Reciprocity between participants facilitates more in-depth analysis and reflexivity regarding the outcomes produced.

Vaishnavi and Kuechler (2015, 139-140) delve deeper into specific data gathering techniques for suggestion and development phases, as illustrated in Figure 10 (page 43) Data collection techniques in the suggestion phase include industry/practice awareness, brainstorming, complex system analysis, and various organizational analysis matrices. Among these, research conversation is one of the most utilized methods. In the development phase, techniques such as sketching and design solutions are recommended.

The evaluation phase in the DSR process, as shown in Figure 10 (page 43) plays a crucial role in analyzing the artifact in an iterative manner. The primary goal is to measure how well the artifact addresses the identified or supports the proposed solution. This iterative approach allows researchers to return to earlier stages of the DSR process to enhance or modify the artifact based on evaluation outcomes. Furthermore, the iterative process involves effective communication with relevant stakeholders regarding any changes or improvements related to the artifact. The evaluation can be demonstrated through several different techniques, such as experimentation, case studies, or simulation technique (vom Brocke et al. 24 September 2020).

Tie, Birks, and Francis (2 January 2019) introduce GT research in their research article: *Grounded Theory research: A design framework for novice researchers.* In the Results section, they describe GT research as an iterative and recursive methodology. GT methods and processes involve gathering data using various procedures or tools, along with systematic modes of collecting and analyzing data. Continuous data collection and analysis, following multiple stages of coding and comparative analysis, leads to a framework that produces the final GT results.

Phases of Grounded Theory Analysis

According to Koskennurmi-Sivonen (2007), GT analysis involves three main phases:

1. *Open Coding:*
- *This phase involves conceptualizing the data. In this phase, the researcher analyzes the collected data that needs to identify and recognize relevant concepts. These related conceptual perceptions are then categorized together. Dimensions arise from the relationships and behaviors of concepts and categories. They represent various characteristics and properties that illustrate how these concepts are interrelated. However, dimensions may not always emerge if the concepts are too similar. In that case, the context may take a different.*
2. *Axial Coding:*
- *This phase establishes connections between concepts/categories. The researcher identifies connections withing the same phenomenon, such as causal relationships, intervening factors, and other linking elements. The key is to identify which higher-level concepts define the ones created during open coding. Although the phases are presented*

sequentially, they often occur simultaneously or overlap in the researcher's mind in practice.

3. *Selective Coding:*

- *Selective coding involves identifying the core category and relating it to other categories.*
- *Validating these relationships*
- *Further developing categories that appear incomplete*
- *Core Category**:***
  The core category represents the central aspect of the phenomenon to which all other categories are integrated. In GT research, it is emphasized that there should be only one core category that unifies all other findings within the research process.

3.2 Research Progression and Stages

The aim of the thesis evolved from the development of the author's skills and expertise in EA and the management of the organization's strategy. The customer organization needed to adopt approaches for developing AI-based solutions in alignment with EA practices. This need formed the foundation upon which the concept of the case study and the overall research process were built. The organization needed to build a robust and agile process and system to automate the creation and utilization of capabilities and capability maps that aligned with the scope and the case study. The case study was based on Yin's (2009, in Farquhar 2012, chapter 3) five components of case study research design:

- *Research question*
- *Propositions/objectives*
- *Unit of analysis*
- *Linking data to propositions*
- *Interpretation of findings*

The first phase involved defining the research questions, which evolved as the study progressed. To support the Proposition/Objectives phase of the five-component model, three use cases were defined in parallel with the literature review, as depicted in Appendix 1 (pages 81–84). Together, these elements formed the structure for identifying interview questions and propositions related to the unit of analysis, which ultimately informed the proposed solution. When linking data to propositions, the inductive approach generated theory as the study proceeded, revealing the emerging patterns. The case study connected the research to a real-life scenario and formed a structure around the case, guiding the exploration in the right direction by filtering out irrelevant topics. Throughout the research, the theory evolved from literature and previous studies, while the use cases, developed in parallel with them, played a key role in shaping the interview questions and directing the research process.

The DSR method complemented the case study methodology, implementing similar phases within the research. DSR was implemented to support the development of an agile EA process model. Specifically, the DSR phases of suggestion, development, and evaluation were employed iteratively. During these phases, the emerging prototypes were presented to the company's lead enterprise architect, who provided feedback and development suggestions, ensuring an iterative and collaborative development process. Additionally, the feedback and development suggestions were compared with the emerging literature and applied the GT method suggested by Prof. Dr. Gimbel (2024b, min. 0:28–1:17) to support the contents of the case study, literature review, and other constructed materials. GT was exceptionally helpful in constructing a thematic interview that addressed the most relevant and in-depth topics, facilitating the identification and understanding of key themes that guided the interview structure and detailed questions, as detailed in Appendix 2 (pages 85–89).

The mind map presented below in Figure 11 is constructed based on the axial coding results of GT analysis, as described in Table 2 of Appendix 2 (page 87). This table, based on the method described by Koskennurmi-Sivonen (2007), facilitates the identification of all relevant topics and their subtopics. In particular, the mind map plays a pivotal role in developing the theme-interview, as it provides alignment between interview questions and the main topics and subthemes uncovered during the coding process. This ensures that data collection is directly informed by the themes that emerged from the analysis.

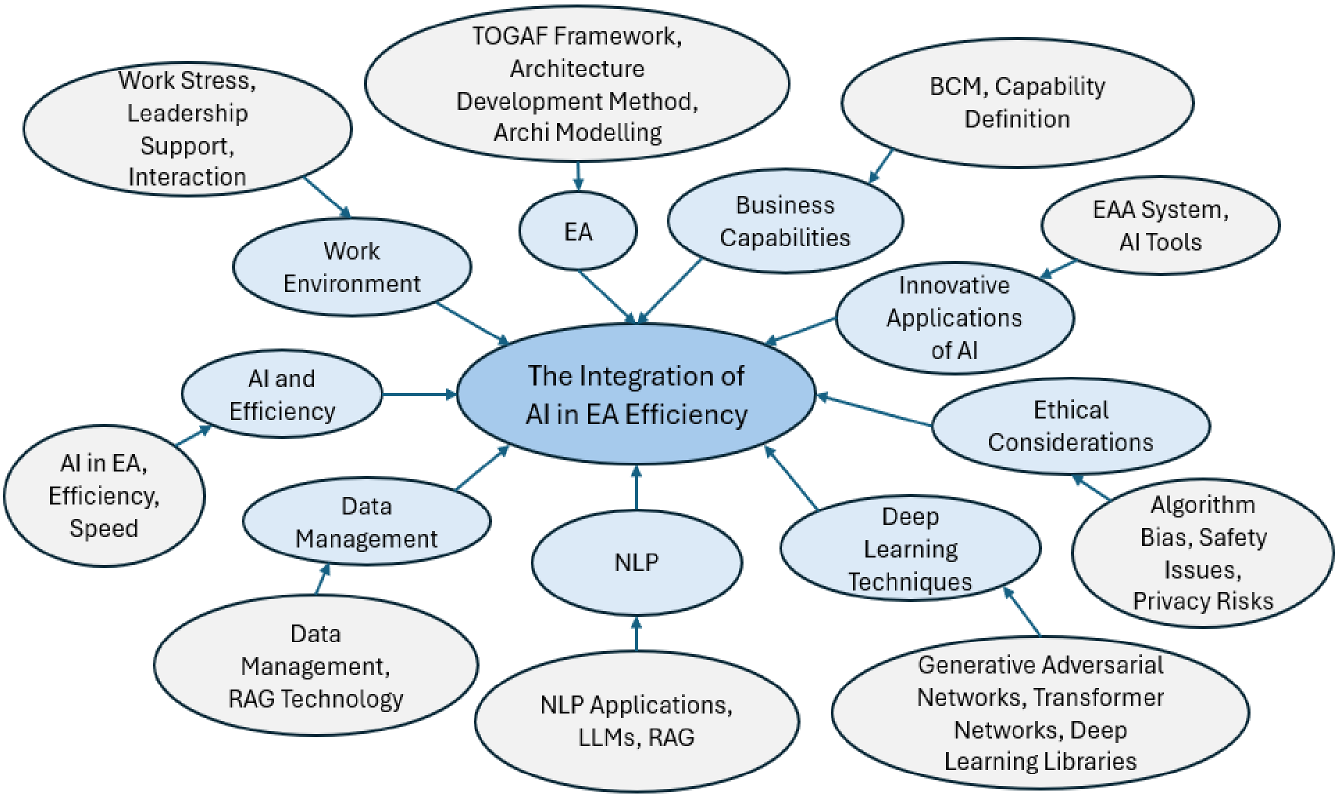

Figure 11. Mind Map of Grounded Theory Axial Coding Results

The following figure (Figure 12, page 49), titled The Logical Flow of the Research Progression, illustrates the main methodologies of the study and their sequential progression. The diagram provides a structured overview, facilitating a comprehensive understanding of the study's methodological framework. It shows how the research questions, literature review, and content analysis logically progressed, forming the foundation upon which content analysis and interview research methods reinforced one another. The figure also demonstrates the sequential structure of these phases, progressing systematically from the initial research design to the final implementation of the AI-based solution. Additionally, it highlights the iterative role of interviews and content analysis within the DSR methodology, showing how continuous refinement contributed to the AI-based solution's development. This iterative process was key in shaping the EA-based process model, ensuring that the research outcomes were systematically derived and aligned with the study's objectives.

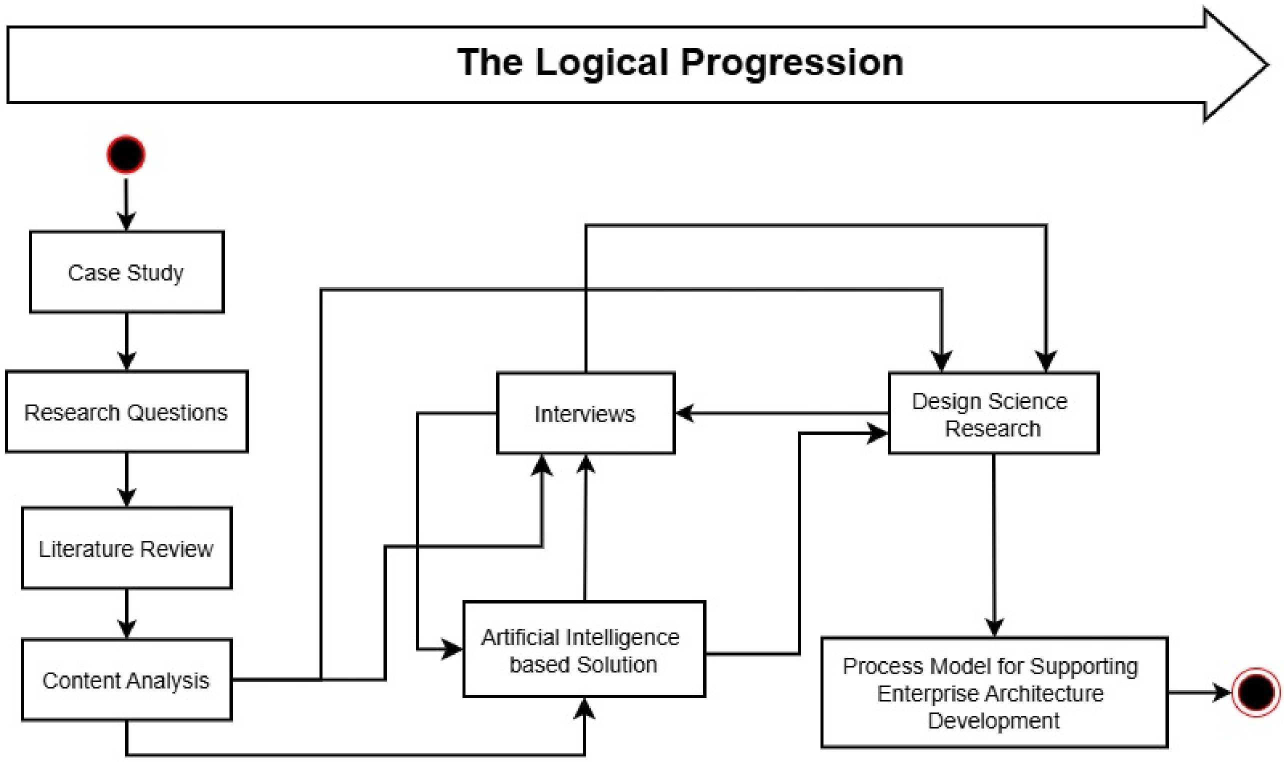

Figure 12. The Logical Flow of the Research Progression

3.3 Methods of Data Collection and Analysis

In the research and development project, the focus was on collecting and analyzing data to create a viable AI system that supports a TOGAF-based EA process model in an agile and effective manner. These collection and analysis methods were specifically tailored for EA and strategic capability management in collaboration with the customer company specialist. The project progressed systematically, using structured data collection and analysis methods to guide critical decisions and refinements throughout the development process. The empirical method was based on Yin's (2009, in Farquhar 2012, chapter 1) case study research approach, which explores contemporary phenomena within real-life contexts through qualitative data collection and analysis. To support this approach, Zhang and Wildemuth's (2009, 2-3) qualitative content analysis was implemented during the literature review, while Farquhar's (2012, chapter 6) approach was applied for interview content analysis, using immediate short notes and timely analysis after interviews and theme interviews.

The primary data collection methods, including Teams meetings, emails, and in-depth interviews, were conducted during the early phases of the project to discuss various aspects of the AI system's design and requirements. Furthermore, these data collection methods were applied iteratively throughout the development of both the AI system design and the EA process model, depicted in Figure 12 (page 49). Two in-depth Teams interviews were conducted with the customer's specialist. The sampling method of this study was determined by the client company, which selected the most capable experts for participation. The first interview lasted approximately half an hour and followed a discussion-based approach to explore the primary objectives for the AI system's functionality and refine EA-related definitions. This iterative conversation provided valuable insights and feedback, which informed subsequent project stages to ensure alignment with evolving project goals.

Between these main interviews, additional email and Teams discussions were held around use cases, aiming to refine and modify stages in the implementation process for capabilities, capability maps, and their utilization. During these discussions, the company's aim was to exploit the developed system for automated content, including content exploitation, which required the system to generate elements directly into the selected Archi Tool. The interview iterative process continued through the production of the DSR artifact, gathering additional information by interviewing the customer's specialist relating to the required features as well as incorporating questions from the literature review's content analysis. The DRS artifact evolved to its final form through these stages.

The second main interview was structured around GT principles, utilizing a thematic approach to review and analyze the developed DSR artifact, discussing recent modifications, and exploring potential future applications. Key themes covered included data security, strategies for integrating with other EA tools, and models for EA management. In addition, critical project-related inquiries such as the implications of data security and the architectural alignment with other organizational tools and management models were explored.

All collected data were handled in strict compliance with the customer's information systems protocols and data encryption standards. This included the use of the company's preconfigured secure networks, designated user credentials, and password-protected directories restricted to project members only, ensuring robust data security. Interview transcriptions, raw data, and notes were securely stored and managed in access-controlled folders to prevent unauthorized access.

For the analysis, qualitative inductive research methodology guidelines were followed, including systematic review, transcription, and structured documentation of interview responses and insights. The first interview was transcribed and analyzed to shape the foundational structure of the AI system architecture. In contrast, the thematic interview underwent a detailed transcription and categorization process. This approach yielded in-depth insights into each topic, which were refined by comparing findings with relevant literature and the project's guiding research questions. Throughout the analysis, insights were cross-referenced with key use cases and existing literature to support theoretical grounding and practical relevance in line with Yin's (2009, in Farquhar 2012, chapter 3) five components for case study research design.

After completing the analyses, the findings were organized into a coherent structure to support the thesis. All materials, including interview transcripts, supplementary notes, and synthesized findings, were compiled into a comprehensive dataset for future reference. The data used in this research was not sensitive to the company, thereby minimizing confidentiality concerns. This dataset was securely stored according to the company's and the Haaga-Helia University of Sciences data retention policy and will be disposed of upon the project's completion. Interview outlines, survey templates, consent forms, and use case templates utilized during the study have been included as appendices in the final thesis report. This detailed documentation ensures that future researchers or developers can build upon the work with a well-structured foundation and transparency.

4 Results

This section presents the results of the development research for the case study on the utilization of AI-driven capability maps in EA. The case study was conducted from September 2024 to January 2025 as part of the master's thesis project. The research was carried out in collaboration with an expert from the client organization, a seasoned enterprise architect with several decades of experience in EA. The author conducted the research, specifically focusing on aspects defined by the research questions. The study was scoped to address these key areas within the broader context of the thesis project, ensuring a targeted investigation into the use of AI-driven capability maps in EA. After the data collection and analysis, as described in chapter *3.3 Methods of Data Collection and Analysis*, the results were compiled and are presented in the following sections.

Initially, the research explored four use cases. As the study progressed and the solution evolved, these were refined and consolidated into three distinct use cases, as outlined in Appendix 1 (pages 81–84). This adjustment was made to ensure clarity and focus, and to maintain a manageable research timeline. To further ensure coherence, the fourth use case, *Identify Processes, Systems, and Resources for Capability Development*, was omitted. Although AI-based capability definition was not originally within the study's scope, insights from the literature review and AI system specification highlighted its role as an integral component of capability map utilization, leading to its incorporation into the finalized use cases. These final use cases provided practical examples that guided the study's primary objective: implementation of a process model artifact using the DSR methodology. This artifact integrates an AI-driven model to enhance capability mapping and utilization within EA practices.

The results are structured into two main parts. First, insights from the use cases are introduced, forming the foundation for developing both the AI-driven solution and the DSR artifact, which represents the process model. These use cases demonstrate how AI can support enterprise-level decision-making and capability planning. Second, the developed process model is presented, showcasing how the iterative application of the DSR methodology enabled the creation of a scalable and practical solution. By aligning the case study findings with the research objectives, the process model addresses both practical challenges and theoretical gaps. Overall, these findings emphasize the importance of the AI-driven approach in enhancing EA practices and contributing to more dynamic, agile decision-making and capability planning within organizations.

4.1 Case Study: AI-Driven Capability Mapping and High-Level Architectural Design

The case study aimed to explore the potential of utilizing AI for the creation and utilization of capabilities and a capability map. One of the main objectives was to identify AI-based solutions that support the development of capabilities and capability map utilization, such as capability assessment and prioritization for strategic purposes. Data collected from a comprehensive literature review, interviews, and relevant sources, including books, official websites, and prior studies, contributed to the development of an overview of the proposed solution. The use cases, outlined in Appendix 1 (pages 81–84), were developed during collaborative planning sessions with the customer organization's specialists, where the focus was on identifying the scope to guide the development of the process model. Three primary use cases were selected, each related to the development of an operating model for utilizing capabilities: defining the current state, assessing the current state, and prioritization based on strategic goals:

- *Use Case 1: Define Capabilities and Create a Capability Map Using AI Based Tools and Features*
- *Use Case 2: Current State Definition Using Advanced Technology and AI*
- *Use Case 3: Prioritize the Capability Map Capabilities Based on Strategic Goals*

These use cases were defined based on the theme of creating a capability map and identifying ways to utilize the capabilities identified in the map. There was also a need to delineate the scope addressed by the developed process model. To achieve this, The Open Group's TOGAF EA management framework was selected, specifically the ADM. This method supports reusable architectural assets and an iterative process model approach, aligned with distinct phases that outline the inputs and outputs for each stage of the ADM. According to Josey and Hornford (2022), the TOGAF ADM is widely known and recognized being one of the most widely used EA methodologies and frameworks by the world's leading organizations. The development of TOGAF reflects an understanding of contemporary enterprise demands, including trends and transformations associated with the digital enterprise.

The AI-based solution was conceptually designed as a high-level architecture for a RAG system, outlining the components and their interactions to achieve the intended objectives. Appendix 4 (page 95) visually presents the high-level architecture of the defined Retrieval-Augmented Generation system, highlighting its key components and their interrelationships. These form the foundation for the system's application in capability mapping and decision-making processes. Although a practical application was not implemented during the research, the design framework serves as a blueprint for utilizing the system in capability mapping and decision-making. The system architecture consists of the following components, which are numbered sequentially in Appendix 4, corresponding to the sequential phases outlined in the following process descriptions.

1. *Archi: Enterprise Architecture Modelling tool for ArchiMate modelling*

The Archi Tool was selected for research based on the requirements of the customer company. It is a widely known open-source tool for managing EA. The selection was made during the research development in collaboration with the customer's specialists when the scope of the AI system was defined. Beauvoir and Sarrodie (2013-2024a) highlight that Archi is specifically designed for creating ArchiMate models, an EA standard that supports the visualization, analysis, and description of architecture across various business domains, and fully aligns with TOGAF.

2. *Automated data processing from Archi using jArchi*

The aim of the system design was to automate certain critical steps. These steps included text-based data export from Archi to the next phase: *3. Preprocess and Validate data*. This idea was supported by Sarrodie (5 October 2018) who stated that Archi can Import and export data to and from different formats. Additionally, Beauvoir and Sarrodie (2013-2024b) discussed the implementation of batch processing and query models in Archi, facilitating the automation of code and data processing.

3. *Preprocess and validate the data*

Preprocessing and data validation are critical to ensure the integrity and relevance of the data, confirming that it is accurate, complete, and suitable for further processing. This validation process was conducted immediately following the data export from Archi using jArchi. During this phase, the data underwent rigorous checks to verify its accuracy, consistency, and alignment with the specified requirements before proceeding to the subsequent stage of the workflow.

In this system, the automation can be enhanced using libraries that support automated document processing. For example, commonly used Word and PDF files can be processed using various libraries. In this case, the Python-docx (2013) library allows for processing documents from specified file paths or directories. Similar tools are available in many other formats as well. As discussed by Zeng et al. (2024, 4512-4513), data desensitization and the use of non-sensitive public data are critical when implementing a RAG system. Preprocessing can be implemented using various tools, depending on the context of input. Gheorghiu (2024, chapter 4) introduces

Hugging Face for processing sensitive data, along with different tools for preprocessing various text formats. Kulkarni & Shivananda (2021, chapter 2.3) note that the NLTK library offers a simple way to preprocess different texts. To validate XML format data exported from Archi, JSON Schema (2020) serves a powerful tool for validating data before further processing. JSON Schema, which stands for JavaScript Object Notation provides the foundation for generating data in Archi, especially in conjunction with the jArchi plugin.

4. *Text translation into vectors*

This vital phase transforms the validated and preprocessed data into machine readable format and ready for importing it into suitable vector database. The text translation phase includes multiple steps, such as Gollnick's (2024, chapter 5) definition of first step is to get a corpus, which is output from the "preprocess and validate data" phase. After that, the text data needs to be tokenized. The final step in vectorizing the tokenized text data involves converting these tokens into embeddings.

Models that could be used include Dasgupta's (2024, chapter 2) examples such as TF-IDF, Word2vec, and BERT for generating embeddings. Additionally, Kulkarni and Shivananda (2021, Introduction) discuss the importance of Python-based libraries such as TextBlob, NLTK, and SpaCy in NLP and deep learning to ensure the system effectiveness, noting that all three libraries are suitable for tokenization (Kulkarni & Shivananda 2021, chapter 2.3).

5. *Vector Database*

Many vector databases are available for use in AI-based solutions. Baragry (12 September 2024) aligns how traditional RAG applications integrate with vector databases. Vector databases offer different features for different needs, and there is no single suitable option. However, these databases have certain basic similarities. According to Gollnick (2024, chapter 5), these specialized databases store and query high-dimensional data (embeddings) for similarity analysis in NLP. They can also process images, text, video, and audio. Goyal et al. (21 May 2024) utilize the Chroma vector database in their NLP project, while Gollnick (2024, chapter 5) lists Chroma, Pinecone, and Redis as alternative options. Additionally, Dasgupta (2024, chapter 2) highlights Pinecone, Milvus, and Weaviate as robust vector database choices, whereas FAISS and ChromaDB focus primarily on vectorized indexing.

6. *The Large Language Model Query Pipeline Library*

Connecting a vector database to an LLM is essential for choosing the right tools and techniques that meet the requirements of the RAG system. Dasgupta (2024, chapter 1) presents Chain of Thought (CoT) prompting, which is used to guide language models to generate step-by-step reasoning for advanced systems. CoT can dynamically implement project related tasks, such as capability definition from the data input into a database, as well as creating capability maps.

These prompts can also be coded to implement capability assessment using various models available for LLMs, such as TOGAF-based capability assessment and others. Similarly, capability assessment and capability prioritization can be implemented by leveraging these methods and tools. Tools like Dasgupta’s (2024, chapter 1) presented LangChain framework guide LLMs through a series of logical steps to complete complex tasks. Similarly, to leverage queries to LLM LangChain and LlamaIndex as tools for creating and working with prompt templates, which are dynamic and model-agnostic.

7. *Retrieval- Augmented Generation and Large Language Model*

The driving force in the RAG system to generate AI-based responses to user queries is LLM. In this design the user queries are presented as dynamic prompts that provide EA-based content such as request to define capabilities, form the capability map, and assess and prioritize them. Baragry’s (12 September 2024) implementation highlighted the updating and utilization of the EA capability map in Ardoq Tool. The implementation exploited RAG alongside LLMs, utilizing prompt design and OpenAI’s GPT-4 model. The output generated by the GPT-4 model was structured in JSON format to facilitate integration with the Ardoq Tool.

8. *Open-Source Pre-trained Model for Large Language Model Queries*

In this thesis work, which includes use cases that can be implemented with non-sensitive data, the LLMs presented by Lewis et al. (2020, 1) were shown to be sufficient for producing relevant results in such a system. These models effectively leverage pre-trained LLMs integrated with external datasets, such as Wikipedia, through a retrieval mechanism. When selecting an LLM model for this research design, Gheorghiu’s (2024, chapter 1) identifies the best-performing LLMs, such as Llama2, Claude 2.1, and GPT-4. These models, which contain trillions of parameters and are trained on extensive internet-scale datasets, represent excellent choices for this system design.

9. *Custom-Trained Open-Source Model for the Large Language Model Queries (Optional)*

Custom-trained models can be developed to further optimize system performance for specific queries or tasks. Using frameworks like TensorFlow and PyTorch, it is possible to fine-tune large LLMs on specific datasets, making them better suited to the unique requirements of the thesis use cases and potential future scenarios involving more secure data. Kapoor et al. 2022, chapter 1) introduce Keras, a high-level neural network API that can be used to create and train deep learning models within TensorFlow. Ganegerada (2022, chapter 1.1) suggests that these kinds of predefined or custom-layered implementations can perform computations related to computer graphics, reuse pre-trained models, and visualize/debug TensorFlow models.

PyTorch is another popular open-source ML library for train custom models. Nvidia (2024) highlights various use cases where PyTorch excels due to its flexibility, rapid prototyping, and efficiency. According to Granger and Baragry (20 Feb 2024), a custom-trained model can offer numerous benefits. It enhances the security of personal data by restricting data communications to personal use only. Additionally, it addresses known issues associated with LLMs, such as "hallucinations and information silos".

10. Automated Data processing for The Generated Results

The system designed in this case, there were two endpoints to save the LLM generated results. The primary endpoint is EA tool Archi, which utilizes ArchiMate modeling notations. The second endpoint is an architecture repository that generally saves EA-based documentation within *EA Repository* refers to Appendix 4. To save results automatically, prompting and other relevant techniques can be employed. Compatible formats and models, such as graphs, can be processed using ArchiMate and similar languages.

Sarrodie's (5 October 2018) definition of Archi's capability to import and export data to and from different formats supports this evidence. In addition, the integration of jArchi scripts and NLP capabilities, as discussed by Rohde (18 August 2024) and Beauvoir (8 February 2024), significantly enhances the system's functionality. These scripts facilitate the automatic generation of EA models, such as schemas and strategy models, by processing LLM outputs and converting them into ArchiMate-compatible formats, as well as enabling the production of content in formats supported by the architecture repository.

4.2 Design Science Research: AI-Integrated Process Model in Enterprise Architecture

The AI-based RAG system design, as introduced in the case study chapter, was intended to serve as a foundational component of the overall process model and was built alongside it. The implementation of the DSR methodology began after the necessary information was gathered through literature reviews and interviews. The EA methodology TOGAF provided the foundation, aligning specific phases to implement the selected use cases – in this case, the developed process model. The focus was on TOGAF's inputs and outputs phases, as defined by The Open Group (2022a). In Phase A, the required business capabilities must be identified and evaluated based on strategic priorities, and then documented in a Capability Map. The outputs of *Phase A*, which include the identified capabilities and the Capability Map, are then used as inputs for *Phase B*. This can be seen in the *Identify Scope and Objectives* and *Approval and Publishing* phases in Figure 13 below, where arrows labeled *Phase A*. and *Phase B*. indicate the data flow in the process model. The figure highlights the key elements of the process model, particularly illustrating the input and output outcomes at each phase within the TOGAF framework.

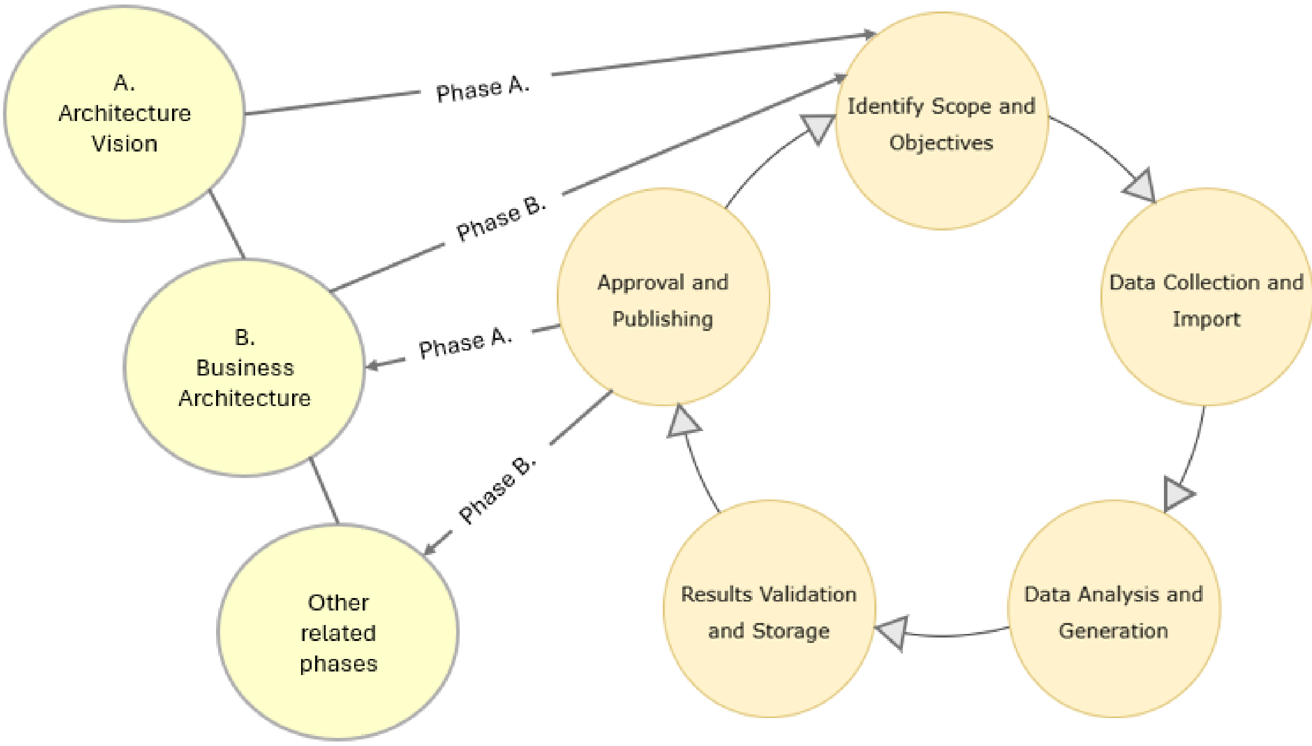

Figure 13. The Process Development Model for Capability Utilization in TOGAF's ADM

The *Data collection, preprocessing and import* phase, as shown in Figure 14 (page 59), involves gathering the necessary information related to the identified scope and objectives. The Open Group (2023) defines capabilities as a combination of tools, processes, and people. Additionally, The Open Group (2022) recommends considering the organizational structure, business model, and current strategic and financial plans when identifying capabilities for the Capability Map. Figure 14 (page 59) illustrates the importance of data collection by depicting how different roles and responsibilities, highlighted within yellow boxes, are defined around the process model, represented by gray boxes. The figure also illustrates how these roles and tools are interconnected within the process model, reinforcing the systematic flow of data and its impact on capability mapping. In the process, the enterprise architect is responsible for gathering relevant information from the strategy team, relevant stakeholders, and the Archi Tool, as well as managing the data flow.

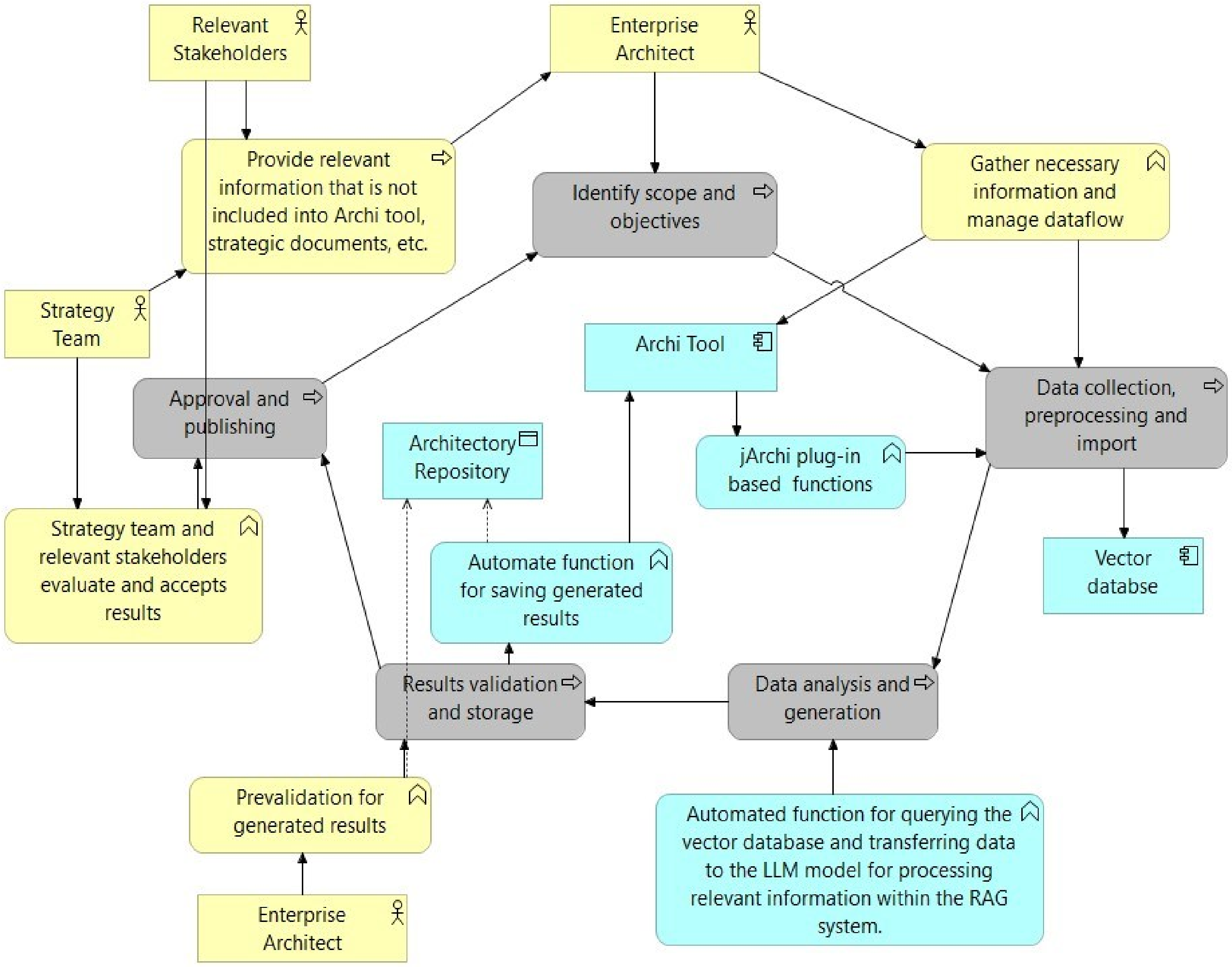

Figure 14. The Process Model Including Description of Tools and Business Phases

After the relevant data is collected, preprocessing and importing can be implemented. Figure 15 (page 60) illustrates the phases of the AI-based solution, which are programmatically automated within the overall process model. This figure provides a higher-level overview by breaking down the technical phases of the AI solution into three distinct steps, depicted within gray boxes, which emphasize its role in the process model. These phases are designed to follow automation through various techniques presented in the *process description* of the previous section (see Chapter 4.1). Specifically, the *Preprocess and Validate Data* phase (steps 2 and 3) explains how data from different directories is processed using appropriate libraries. The *Data Analysis and Generation* phase corresponds to steps six, seven, and eight (optionally step nine), while the *Results Validation and Storage* phase is automated through coding practices, as described in step 10 of the *process description* in Chapter 4.1.

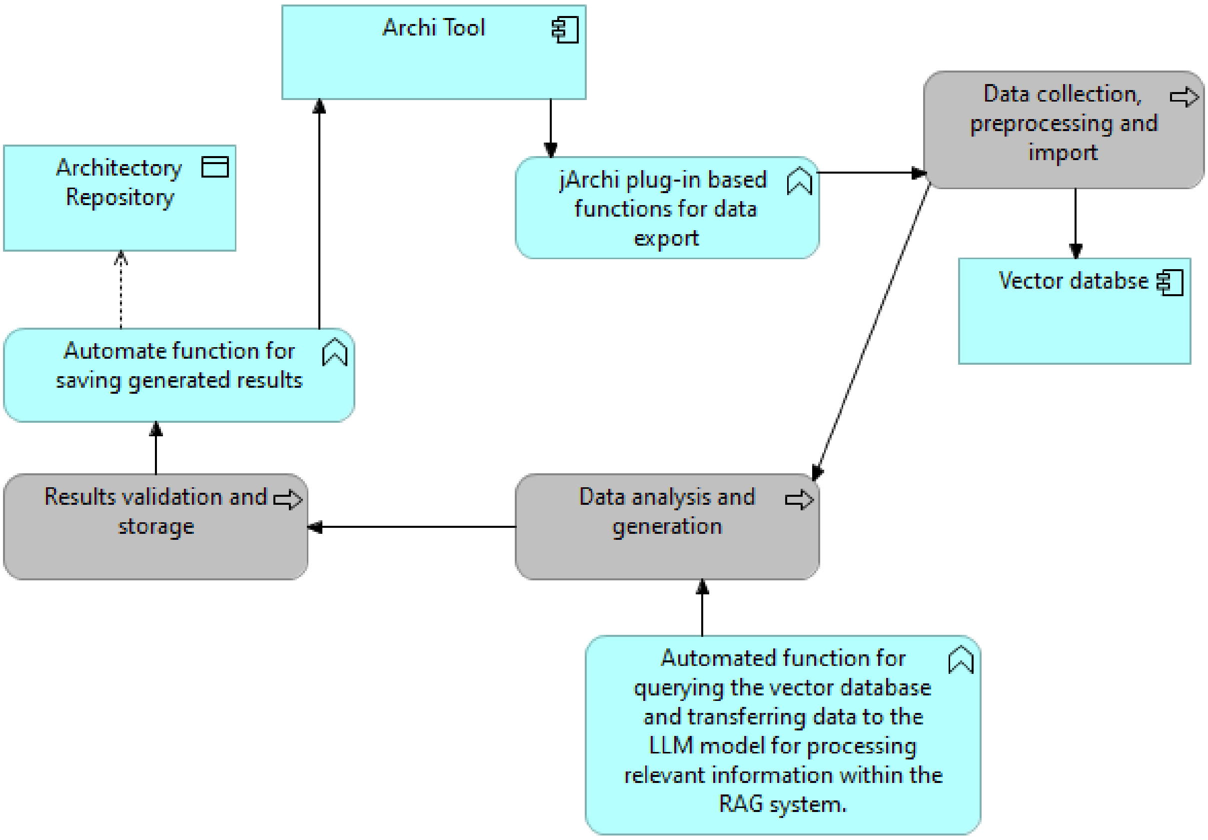

Figure 15. The Retrieval-Augmented Generation System: High-Level Description from a Technical Process Perspective

Manual processing is described in more detail in Figure 16 (page 61). This figure illustrates how the *Approval and Publishing* phase follows manual intervention when the strategy team or relevant stakeholders request changes to the pre-validated architecture content, which was initially implemented by the EA architect. This process ensures alignment with the organizational strategy and stakeholder requirements. Additionally, it serves as a quality control mechanism, ensuring that any modifications are both well-informed and contextually relevant within the scope of the EA framework. It is essential to understand the manual steps required to pre-validate or revalidate AI-processed content, highlighting the iterative nature of EA development. In this process, stakeholder feedback directly informs content updates. In its simplest form, this involves a meeting or a brief exchange of information to address irrelevant content, clarify any ambiguities, and subsequently update the architecture in the Archi Tool or EA repository.

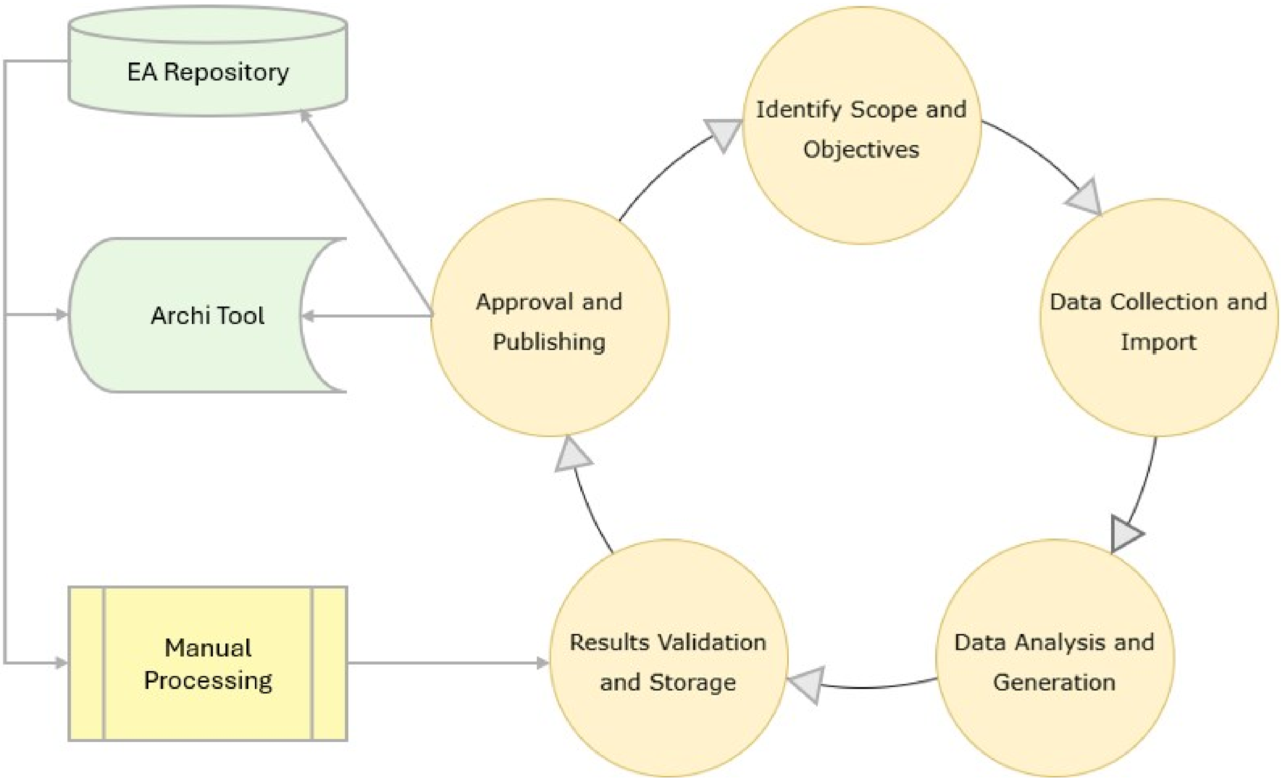

Figure 16. Manual Implementation within the Process Model

The *Approval and Publishing* phase, illustrated in the figure (Figure 17, page 62), which represents the high-level process model, is designed to fulfill approval and publishing requirements. In general, the implemented EA artifacts and models must be approved by the relevant stakeholders, including the strategy team, both before and after input and output phases, as described in Figure 13 (page 58). However, this does not prevent initiating a new iterative process for non-dependent deliverables to identify a new scope and objectives.

To support this, Bondel et al. (15 November 2018) described a BCM implementation based on TOGAF's principles in their case study. BCM implementation involves gathering information related to business rules, business entities, business processes, and strategic objectives, utilizing techniques such as reading, writing, interviews, observations, and questionnaires. These phases are more closely related to the identifying scope and objectives, followed by further data collection. Additionally, findings from the literature and interviews with the customer company's specialists highlighted the importance of involving key stakeholders and business leadership in the process. As Bondel et al. (15 November 2018) emphasized, the development of the initial version of the BCM was developed in collaboration with the managing director, vice-managing director, and head of department strategy to ensure further development.

The developed process model in the Figure 17 below emphasizes not only the potential for automation through AI but also the critical role of human involvement in the implementation process. The research findings show that while certain phases can be automated to generate machine-based solutions, the need for human input remains indispensable. The process model highlights that successful implementation requires a balanced approach, where technology and human decision-making work in tandem to achieve optimal results.

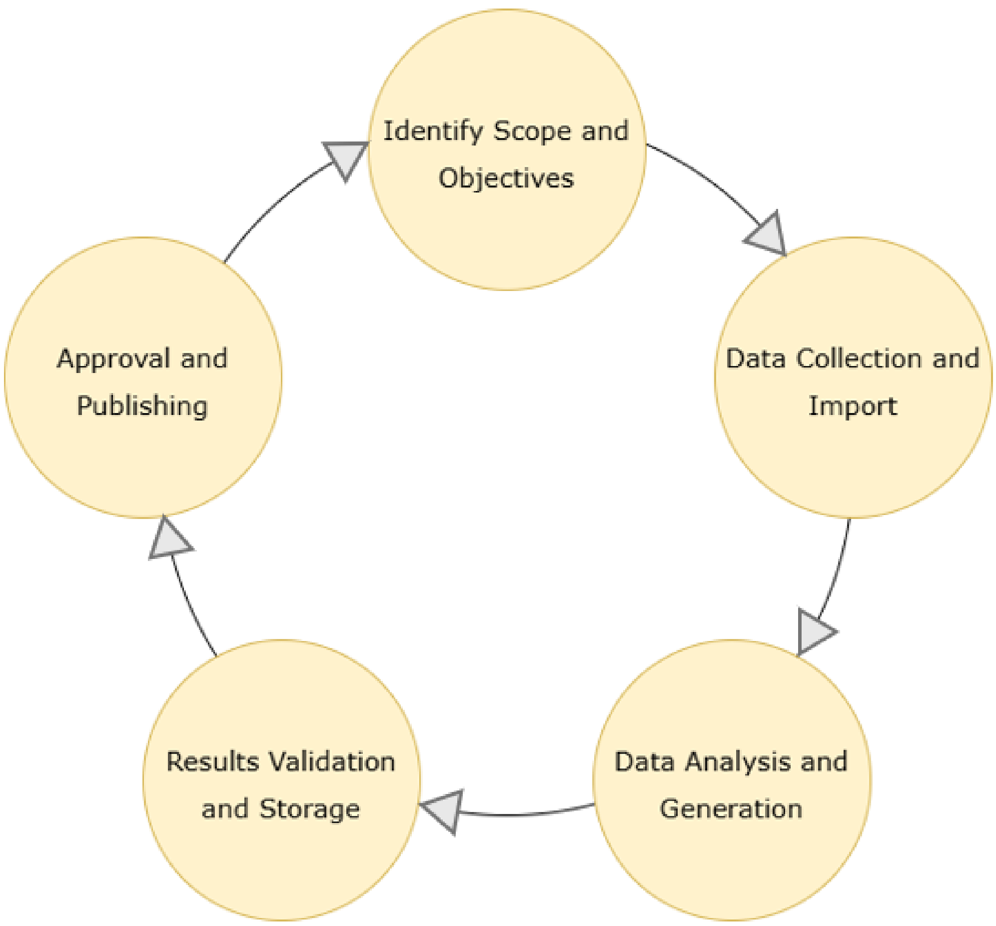

Figure 17. The Process Model for Implementing Capability Map Utilization Using AI

5 Conclusions and Discussion

This chapter summarizes the research reflections through the following topics. The *Conclusions and Discussion* chapter outlines the key findings of the thesis. In the first subsection, the research questions are addressed in detail, with the answers being processed based on the results and the final outcomes analyzed. The *Further Research and Development* section concludes the findings and suggests directions for future research, drawing on insights from the literature review, the defined results, and the analysis of the research questions. The *Summary of Key Findings* subsection summarizes the key outcomes of the research, focusing on the development of an AI-based solution and a process model for EA. It highlights the model's practical applications, addresses challenges, and discusses its broader implications and prerequisites for implementation.

The *Reliability and Ethics* subsection discusses the ethical principles and research practices followed throughout the thesis to ensure reliability, validity, and confidentiality in data handling and reporting. It highlights the measures taken to protect participants and maintain the integrity of the research process. The *Reflection on the Responsibility of Research* subsection examines how the study aligns sustainable development principles, emphasizing the responsible use of AI and EA. It explores the broader societal, economic, and environmental implications of the research and its contributions to long-term sustainability and innovation. The *Evaluation of the Thesis and Learning* chapter evaluates the success of the thesis and reflects on the author's learning throughout the research process. This chapter discusses how well the study addresses the research questions and achieves its objectives, including identifying an AI-based system and integrating it with a process model for EA capability map development. Additionally, it reflects on the author's enhanced understanding of EA and AI technologies, as well as the insights gained into research methodologies for future studies.

5.1 Research Question Analysis

This section presents the analysis of the research questions and justifies the choice of analytical methods. The study employed a combination of qualitative content analysis, a case study approach (Yin 2009, in Farquhar 2012, chapter 1), and GT methodology (Koskennurmi-Sivonen 2007), as detailed in Appendix 2 (page 85), to address research questions related to organizational business capabilities and their modeling. The case study approach was particularly suitable for exploring the practical aspects of business capability modeling, while the literature review provided theoretical grounding. A semi-structured thematic interview, employing GT methodology, further enriched the study by facilitating the theoretical development of the phenomenon directly from the data, as outlined in Appendix 3 (pages 90–94).

Research Questions 1 and 2 were analyzed applying the case study approach in conjunction with the literature review and the supported use cases, which provided the necessary theoretical context. Research Question 3 was supported by qualitative interviews, the literature review, and the results section of the study. For Research Question 4, focusing on the benefits and challenges of utilizing AI in EA Capability development, the analysis incorporated all previously mentioned methods. The thematic interview based on GT provided detailed insights, particularly regarding AI's role and its associated opportunities and challenges.

The empirical method was grounded in Yin’s (2009, in Farquhar 2012, chapter 1) case study research approach, which explores contemporary phenomena within real-life contexts through qualitative data collection and analysis. Zhang and Wildemuth’s (2009, 2-3) qualitative content analysis was implemented during the literature review, while Farquhar’s (2012, chapter 6) approach was applied for interview content analysis, using immediate short notes and timely analysis after interviews and semi-structured thematic interview based on GT methodology.

Research Question 1: What is the role of capabilities and the Capability Map in an organization’s EA and strategic processes?

The research revealed that capabilities and a capability map are strongly connected to an organization’s strategic perspective. They link the organization’s development, from high-level strategic thinking and planning to its tools, processes, people, and culture. Capabilities serve as the driving force behind the organization’s operational management and development. Bondel et al. (15 November 2018) highlighted in their project research results how a Capability Map improves understanding between business and IT by creating a common language that helps to better identify and address pain points.

Research Question 2: How is the Capability Map created, maintained, and utilized?

The creation and utilization of a Capability Map follows specific guidelines and principles that guide the process. The Open Group suggests implementing a Capability Map through defined phases in alignment with TOGAF’s ADM (The Open Group 2022a). Various data gathering techniques are essential. For example, Bondel et al. (15 November 2018) employed methods such as reading, writing, interviews, observations, and questionnaires in their project to define capabilities and create a Capability Map.

The Open Group (2023) identifies financial, business, and strategic plans as key resources for identifying capabilities. These plans include financial statements, a business plan, and the Balanced Scorecard (BSC). Similar tools can be applied in different phases. For instance, the Digital and Population Data Services Agency (2017a, 5) suggests using tools such as the BSC, CAF, EFQM, and CCMI (Capability Maturity Model Integration) to evaluate a capability's maturity at a strategic level. These tools can be utilized to identify and assess capabilities, and they serve as foundational tools for defining a Capability Map.

The creation of Capability Maps has become an emerging trend in EA development in recent years. Previous studies, such as Baragry (2024), which include AI-based implementations, highlight the evolving role of Capability Maps within EA and its tools, supporting the progression of this approach. Additionally, it is important to note that the latest version of TOGAF (The Open Group 2023) has further developed the use of capabilities, particularly focusing on integrating and utilizing Capability Maps, including Business Capability Catalogs, within TOGAF's principles. The creation of a Capability Map represents a cornerstone of EA, and its accurate and thorough execution ensures guidance for the overall architecture, influencing subsequent phases of implementation and management. By doing so, it supports the successful deployment and governance of the EA.

Research Question 3: How can AI be utilized in the creation, maintenance, and utilization of the Capability Map?

AI offers various solutions to support the creation, maintenance, and utilization of EA capability maps. Several white papers and related studies on EA tools emphasize the role of AI in advancing EA development. Many of these tools incorporate add-ins or integrated features that provide AI capabilities, including Archi, which is incorporated in this study. These capabilities often include LLM solutions, which aim to enhance the efficient development and understanding of EA. These features are relatively new at the time of writing this thesis and have emerged in the past few years.

AI has a long history, but it has taken significant steps in its development, especially with the latest innovations in AI, including deep learning and NLP models, and their fine-grained subdivisions. These subdivisions include language and speech processing, as well as image and video analysis and generation. In addition to analyzing existing data, AI can generate both graphical and textual content to EA tools and repositories. For example, AI can create visual diagrams for capability maps or draft narrative explanations of EA concepts. This is supported by Gollnick (2024, chapter

5), who defines vector databases as systems capable of supporting various formats, including images, text, videos, and sound. Vector databases are a primary data store that can handle data processed by LLMs.

In relation to EA capability maps, AI can read text- and image-based data and, correspondingly, generate new content into EA tools and architecture repositories. According to the client company's expert, capability determination itself can be implemented quite easily with the assistance of AI. For instance, this can involve collecting information from the company's website and generating results through an AI-based chatbot. This demonstrates the public availability of information and its straightforward processing. Goyal et al. (21 May 2024) exploited the GPT-3.5-turbo-instruct model to generate user guided queries to implement EA content in their project. Similarly, Gheorghiu (2024, chapter 1) lists the best-performing LLMs, such as Llama2, Claude 2.1, and GPT-4 chatbots that can be used as licensed or open-source options.

When analyzing the study by Goyal et al. (21 May 2024), it can be concluded that the implementation of capabilities and the capability map does not necessarily require advanced AI solutions. However, the implementation is more manual, with the AI and AI chatbot functioning as assistants rather than as an automated AI process system that generates results. Accordingly, AI contributes to the improvement of task efficiency, but human expertise remains a crucial component in interpreting and refining the outcomes. When generating EA content and capability-related information using AI, it is necessary to understand the requirements related to various organizational needs, such as business strategy, technological infrastructure, and operational goals. AI can guide the process, but the final outputs need careful validation to ensure they align with the organization's strategic direction and objectives. This refinement can be identified in the Figure 17 (page 62), where the manual validation and expert review process is illustrated to ensure the generated content fits organizational needs and strategic goals. Without proper validation, AI-generated results may lead to several risks and challenges. For example, unvalidated data or insufficiently reviewed outputs could result in misaligned business strategies, incorrect capability mapping, or operational inefficiencies. These issues could further disrupt EA processes and their subsequent phases, cascading from top-level strategic decisions down to operational and organizational levels.

Capabilities and a capability map are strategic tools for organizations. These can be constructed using various levels of confidential materials, as Bondel et al. (15 November 2018) describe. Capabilities are identified by gathering information such as business rules, business entities, business processes, and strategic objectives, which can then be used to further refine the capability map. Many of these elements can be identified through publicly available sources, such as an organization's web site. However, as the development of EA progresses into further and more advanced scenarios, the importance of security and robust system requirements increases. Currently, RAG offers the most advanced features to implement secure and efficient AI generation for an organization. Granger and Baragry (20 February 2024) present RAG as an innovative approach to fine-tuning LLMs with an organization's own data, and it is quickly becoming a leading method in this domain.

The question of creating, maintaining, and utilizing a capability map in EA through the implementation of AI lies at the core of the effective development process, as described in the Results paragraph and illustrated in Figure 16 (page 61). This approach provides a continuous and systematic method for implementing the EA capability mapping process. Based on the literature review, two different options for the RAG system were identified: pre-programmed system automation and a query-based chatbot. According to interviews and discussions with customer specialists, the chatbot implementation would be suitable for content analysis and for leveraging organization-wide analysis using queries and analytics. However, the need for this system is defined as an automated system that generates the required pre-programmed results using either the Archi Tool or the EA repository.

Research Question 4: What are the benefits and challenges of using AI to create, maintain, and utilize the Capability Map in EA?

The answers to this question are based on the literature review, the analysis of the designed technical RAG system solution, and the content from the thematic interviews. There are numerous benefits that can be achieved by using AI integrated with EA implementation. AI provides significant advantages in implementing EA capabilities and developing a capability map. RAG systems represent one of the most advanced AI applications. As Taulli (2024) highlights, RAG offers several benefits, including improved timeliness, enhanced handling of proprietary information, reduced hallucinations in transformational models, and increased cost-effectiveness. AI reduces manual work and delivers more accurate, less error-prone content, enabling EA efforts to focus on higher-value activities. These advantages can also be effectively utilized by integrating AI-powered RAG systems into EA implementation.

Based on the evaluation provided by the customer company specialist and the insights from the thematic interview, as outlined in Appendix 3 (pages 90–94), the developed RAG system does not sufficiently justify the resources invested or the time required for its full implementation. This conclusion considers the benefits derived from implementing the use cases defined in the study, such as capability implementation, capability map creation, and prioritization of capabilities for utilization. As a result, the system does not offer a satisfactory benefit-to-cost ratio. However, when integrated with the developed process model, the RAG system demonstrates potential for generating multiple similar EA development use cases for practical implementation.

However, it must be noted at this point that implementing TOGAF in practice is complex and a multi-level process. TOGAF 10^{th} edition defines two key concepts for managing the EA lifecycle; iteration and levels, which are closely interconnected. Iteration occurs inside a comprehensive Architecture Landscape (The Open Group 1999-2022c) that includes three different levels, Strategic Architecture, Segment Architecture, and Capability Architecture, all of which manage EA at different levels of an organization (The Open Group 1999-2022d).

Managing EA within the comprehensive Architecture Landscape in an iterative manner involves executing projects through the entire ADM, starting from Phase A. Each ADM cycle generates output that contributes to the Architecture Landscape, extending or modifying it as needed. Separate projects may operate through ADM cycles simultaneously and may potentially trigger other related projects. Another approach focuses on managing EA iteration within an ADM cycle (Segment Architecture). In this case multiple projects may operate concurrently, maintaining coherent relationships between different phases, such as B, C, and D phases. In this way, the projects may cycle between phases and may return work packages to the previous phases to update work products with new information. Lastly, Capability Architecture iteration involves revisiting the Preliminary Phase to refine elements of the Architecture Capability established in Phase A. Adjustments may also be required to address new or revised requirements identified in Phase H, such as those emerging from Change Requests (The Open Group 1999–2022c). These iterative methods ensure that EA practices remain adaptive and responsive to organizational needs.

Although the RAG system provides enhanced benefits and some level of security compared to traditional AI systems, it still suffers from several limitations related to data privacy, security, and hallucinations. One of the main concerns is the reliability of generated results. For example, Granger and Baragry (20 Feb 2024) highlight how information silos in AI systems can contribute to unreliable outcomes. Goyal et al. (21 May 2024) suggest implementing improvements such as enhanced training data diversity, regular data updates, and fine-tuning of the model to address these issues.

Moreover, Zeng et al. (2024, 4506) state that LLMs tend to inadvertently reveal data from pre-trained corpora. This behavior arises due to the way LLMs memorize, recall, and reproduce their training data. When implementing an RAG system, whether using a chatbot or a preprogrammed system, it is vital to address organizational requirements related to data privacy and security. Zeng et al. (2024, 4512-4513) suggest building the system carefully by utilizing non-sensitive public data or thoroughly desensitized data as retrieval content before incorporating it into the RAG.

AI and RAG systems offer a lot of possibilities; however, different data formats in conjunction with an EA tool seem to present a challenge. Archi provides versatile tools to produce elements with or without AI using JavaScript plugins. Still, there is a challenge in producing formats that are not natively supported by the tool. For example, Archi's (2024) "Frequently Asked Questions" section reveals that Archi does not support UML or business process model notation (BPMN). When considering the thesis research and the designed system at a practical level, this presents one of the main challenges: designing and producing the correct formats in the right place.

5.2 Further Research and Development

Thesis research regarding both BCMs and AI offers a lot of potential to define new research and development cases. Drawing from the existing literature, including key studies such as that of Bondel et al. (15 November 2018), BCMs have been discussed in literature and practice and are gaining recognition. However, the development of systematic approaches to their implementation remains in its early stages. When considering EA, capabilities, and capability maps as a part of the concept, positioning these maps at a high strategic level within EA enhances their value and importance in EA development. According to Bondel et al. (15 November 2018) study, capability maps foster a common language that enables stakeholders to better identify and address pain points within the organization. The study emphasizes that this shared language is crucial for bridging the gap between business and IT, promoting collaboration, and ensuring that BCMs are perceived as strategic tools rather than isolated IT artifacts. The integration of BCMs at this strategic level enables organizations to use them as frameworks for guiding transformation efforts and addressing inefficiencies.

Combining AI with EA provides an opportunity to develop high-level strategic cases that enhance both EA and business. Together with the customer company's specialists, ideas were generated to expand the system's capabilities. The ideal concept would be to integrate most of the organization's material into the RAG system, generating EA-based visual, graphical, and textual content directly into the EA repository or tool. This content could support business-related strategies, highlight development opportunities, and identify gaps in EA and business operations. By exploiting organizational data through analytics, businesses could evaluate and refine their operations. To support ongoing development, additional use cases could be outlined, aligning with TOGAF's continuous input-output paradigm. This approach ensures the process model evolves dynamically, providing valuable insights across EA development stages. Automating manual phases and refining the process would further increase efficiency, requiring deeper understanding of EA implementation.

Integrating digital twins into the EA framework takes this further by enabling real-time data utilization together with the RAG system to create virtual representations of physical systems or processes. Digital twins leverage real-time data to create a virtual representation of a physical system or process, offering enhanced understanding and predictive insights. They integrate geometric, functional, and behavioral characteristics, utilizing IoT, AI, and data analytics. This allows organizations to model and analyze systems, identify inefficiencies, and improve performance (Korhan 2023). Digital twin refers to a virtual representation of a real object, such as a system, process, or product. This concept has gained significant attention across various industries due to its potential to fundamentally change how we design, monitor, and optimize real-world systems. To achieve the goals of Industry 4.0, which emphasizes smart factories, intelligent infrastructure, and efficient supply chains, the digital twin concept has been recognized as a critical enabler (Korhan 2023).

However, the integration of digital twins in this scenario remains a future-oriented vision. While technological potential is evident, the process and system design are still in their early stages. More practical development, testing, and refinement are necessary before digital twins can be fully implemented and leveraged within the EA framework. This includes addressing challenges related to data accuracy, system integration, and the development of standardized practices for digital twin integration.

The further development of the designed process model focuses on defining new use cases and refining the phases within the model. A key objective is to automate manual phases to enhance the model's flexibility and dynamism. Achieving this requires a deeper understanding of how EA is practically implemented and how its phases can be optimized for automation. This will involve integrating advanced technologies such as AI and analytics to ensure the process evolves continuously, providing actionable insights at each stage of development. Further exploration of these advancements will be essential to fully realize the potential of the process model.

In this context, tools like the QPR ProcessAnalyzer (2024) can play a pivotal role in optimizing and analyzing business processes. This tool enables the evaluation and enhancement of organizational processes using AI-based business optimization analytics. With features such as KPI-driven forecasting and automated corrective measures, QPR ProcessAnalyzer can support real-time refinement of processes, further driving efficiency and ensuring the dynamic evolution of the process model.

5.3 Summary of Key Findings

The research focused on designing an AI-based solution, along with a process model that supports dynamic and agile principles, to produce and exploit an EA capability map. The research successfully developed a continuous process model that enables the creation of the EA with the assistance of AI, addressing key issues related to the company's EA, such as data usage, information security, the implementation of various levels and phases of the architecture, and the integration of these into the most efficient and suitable AI-based RAG system solution. According to the customer's specialist evaluation in the theme interview, the designed process model and its phases will align with other EA models. For example, it will be used in conjunction with the standardized JHS 179 Overall Architecture Planning and Development principles. Furthermore, the process model is industry-independent, following principles that enable its application across various industries that implement appropriate EA practices.

The restrictions and challenges of system design integrated with the process model arise from the complexity of organizational environments, involving multiple software systems and rapid digitalization. To address these, the system must be designed to adapt to fast-changing business and technological landscapes, ensuring it remains renewable, agile, and capable of supporting ongoing innovation. Incorporating seasonality into the design is also essential to maintain long-term relevance and effectiveness in diverse operational contexts.

This research offers a new perspective on the integration of AI and EA, providing organizations with the opportunity to develop more dynamic and agile processes (see Chapter 4.2). It also enhances the understanding of how AI can improve the management and utilization of organizational capabilities while supporting the overall development of EA. The presented model offers a practical tool for organizations looking to develop and manage EA in a more dynamic and flexible way. The implementation of the model requires organizations to have a deep understanding of both EA and AI utilization, as well as the necessary technical resources, including personnel with decades of experience in EA and deep expertise in AI, along with robust infrastructure.

5.4 Reliability and Ethics

This thesis followed the practices and requirements of research ethics and reliability. Lincoln (2021, chapter 3) lists eight different topics for research ethics in a qualitative study. This research included the following principles from Lincoln's eight phases, including informing participants of the study scope and their role in it. The study was approved by an institutional and organizational review board and contact information for the boards was provided for any inquiries. Data was securely protected, with all identifying information anonymized in reports and publications, including data storage and processing during the research implementation. Steps were taken to ensure participant identity was protected, including the separation of personal data from research data and the use of aggregated data in publications to maintain confidentiality. Although the study did not involve sensitive or critical information, all documents and data were securely disposed of upon completion.

The data collection and analysis methods were selected to align with the study’s objectives, ensuring the reliability and validity of the findings. These methods were implemented consistently throughout the research to minimize bias and ensure accurate results, following established best practices in qualitative research. The literature review concentrated on the research questions and methodology, ensuring it supported the study's objectives and provided relevant context.

5.5 Reflection on the Responsibility of Research

This research follows the principles of the Foreign Ministry’s (2025) Agenda 2030 – Sustainable Development Goals, specifically addressing aspects related to sustainable economic growth, full and productive employment, sustainable infrastructure, and strengthening global partnerships for sustainable development. By following these principles, the thesis contributes to the broader objectives of social, economic, and environmental responsibility. The designed process model supports dynamic and agile practices, enabling adaptation to complex and rapidly changing technological and business environments. Combined with an AI-based RAG solution, the model emphasizes the efficient use of data, robust information security, and ethical handling of organizational data, promoting innovation, the development of technological innovations, and long-term economic growth. Additionally, the iterative process model's industry-independent nature, integrated with the RAG solution, ensures equitable access, supporting various organizations in adopting inclusive EA practices.

The integration of AI into EA highlights the responsible use of technology to streamline operations, reduce manual workloads, and enable meaningful, technology-enhanced roles for the workforce. This approach minimizes risks of exclusion or job displacement while encouraging workforce transformation and full, productive employment. Finally, the development of EA through collaboration underscores the importance of global partnerships. Pooling resources, sharing knowledge, and adopting best practices ensure a unified effort to address shared challenges, advance innovation, and support sustainable development on a broader scale.

5.6 Evaluation of the Thesis and Learning

The thesis successfully answered all research questions and provided solutions to all defined objectives. The scope of the research was extensive and focused on two main targets: identifying an AI-based system design and developing a process model for EA capability map development that optimally supports one another. Both targets can be evaluated as ongoing developmental objectives, as AI is a rapidly evolving discipline, and business environments, along with their associated processes, continue to change dynamically. The thesis addressed timely topics related to agile and dynamic business development, while also leaving many opportunities for future research and development. As AI continues to evolve rapidly and business environments undergo dynamic changes, the need for constant updates and improvements becomes clear.

The author deepened their understanding of EA, particularly in relation to TOGAF, its implementation, and phases, while also enhancing their knowledge of AI, LLMs, and RAG system technologies. Furthermore, the author successfully completed the TOGAF Enterprise Architecture Foundation certification, with support from the client company. Moreover, the author's perspective on organizational security and data management expanded and became more refined through the research process, drawing on their prior professional experience. These insights were enhanced by the author’s prior professional experience, allowing for the integration of theory and practice. The application of diverse research methodologies further broadened the research perspective, providing valuable insights for conducting similar studies in the future.

References

Archi. 2024. Frequently Asked Questions. URL: https://www.archimatetool.com/faq/ Accessed: 4 January 2025.

Ardoq, 2013-2024. ABOUT ARDOQ. URL: About Us | Ardoq Accessed: 10 September 2024.

Baragry, J. 12 September 2024. AI-based reference creation in Ardoq with simple Retrieval Augmented Generation. Blog. URL: AI-based reference creation in Ardoq with simple Retrieval Augmented Generation | Ardoq Help Accessed: 12 September 2024.

Beauvoir, P. 8 February 2024. Using OpenAI API. URL: https://forum.archimatetool.com/index.php?topic=1495.0 Accessed: 7 October 2024.

Beauvoir, P. & Sarrodie, J-P. 2013-2024a. Archi ArchiMate Modelling. URL: https://www.archimatetool.com/ Accessed: 5 October 2024.

Beauvoir, P. & Sarrodie, J-P. 2013-2024b. Archi ArchiMate Modelling. URL: https://www.archimatetool.com/plugins/ Accessed: 7 October 2024.

Bondel, G., Faber, A. & Matthes, F. 15 November 2018. Reporting from the Implementation of a Business Capability Map as Business-IT Alignment Tool. 2018 IEEE 22nd International Enterprise Distributed Object Computing Workshop. Electronic dataset. DOI: 10.1109/EDOCW.2018.00027 Accessed: 10 September 2024.

Dasgupta, M. 2024. Gen AI - RAG Application Development using LlamaIndex. Packt Publishing, 2024. Audiobook. Accessed: 19 September 2024.

Digital and Population Data Services Agency. 2017a. JHS 179 Planning and Development of Enterprise Architecture. Appendix 2. Business Models and Capabilities in EA Planning. URL: https://dvv.fi/documents/16079645/202143806/JHS179_liite2.doc/ebd8bc04-8440-1049-7746-9b2dcf1da896?t=1709296356687 Accessed: 9 October 2023.

Digital and Population Data Services Agency. 2017b. JHS 179 Overall Architecture Planning and Development. URL: https://dvv.fi/documents/16079645/202143806/JHS179.doc/62b572fa-0a4f-f247-8beb-73c7501be68c?t=1709296308300 Accessed: 8 October 2024.

Dominguez-Morales, M. 2024. Deep Learning, IntechOpen, 2024. E-book. Accessed: 9 September 2024.

Espoo. 2014-2020. Capability Management Handbook. URL: Handbook_Capability_Management_Handbook_Not_Accessible.pdf (espoo.fi) Accessed: 10 October 2024.

Farquhar, J.D. 2012. Case Study Research for Business. SAGE Publications Ltd, 2012 E-book. Accessed: 24 September 2024.

Foreign Ministry. 2025. Agenda 2030 – Sustainable Development Goals. URL: Agenda 2030 – Sustainable Development Goals (um.fi) Accessed: 18 January 2025.

Galletta, A. & Cross, W.E. 2013. Mastering the Semi-Structured Interview and Beyond: From Research Design to Analysis and Publication. NYU Press 2013. E-book. Accessed: 24 September 2024.

Ganegerada, T. 2022. TensorFlow in Action. Simon & Schuster 2022. E-book. Accessed: 9 September 2024.

Gheorghiu, A. 2024. Building Data-Driven Applications with LlamaIndex. Packt Publishing, 2024. Birmingham. E-book. Accessed: 14 September 2024.

Gillham, B. 2010. Case Study Research Methods. Bloomsbury Publishing 2010. E-book. Accessed: 30 September 2024.

Gimbel, H. 2024a. Design Science Research. University of Hohenheim. Online video. URL: https://www.youtube.com/watch?v=P6BfVaxqbOw Accessed: 2 October 2024.

Gimbel, H. 2024b. Design Science Research. University of Hohenheim. Online video. URL: https://www.youtube.com/watch?v=Kybe0KfC75Y Accessed: 2 October 2024.

Gollnick, B. 2024. Applied Generative AI and Natural Language Processing with Python. Packt Publishing, 2024. Audiobook. Accessed: 13 September 2024.

Goyal, S., Hussein, F., Thukral, I. & Trimble, T. 21 May 2024. Empowering Enterprise Architecture: Leveraging NLP for Time Efficiency and Strategic Alignment. 2024 Systems and Information Engineering Design Symposium (SIEDS). Electronic dataset. DOI: 10.1109/SIEDS61124.2024.10534749 Accessed: 10 September 2024.

Granger, E. & Baragry, J. 20 Feb 2024. Generative AI and Enterprise Architecture: Modeling Enterprise. URL: Generative AI and EA: Modeling the Enterprise | Ardoq Accessed: 10 September 2024.

Haystack. 2024. What is Haystack? URL: https://haystack.deepset.ai/overview/intro Accessed: 19 September 2024.

Helsinki.fi. 5 November 2021. Use Case Model – Software Modeling. URL: https://www.cs.helsinki.fi/u/mluukkai/ohmas10/luentokalvot/luento2 Accessed: 30 September 2024.

Hosiaisluoma, E. 2015. Model-Based Development of Organizational Operations and Structure. Master's thesis. University of Helsinki. Faculty of Mathematics and Natural Sciences. URL: Master's Thesis: Model-Based Development of Organizational Operations and Structure Accessed: 7 October 2023.

Huawei Technologies CO., Ltd. 2023. AI Technology. Springer Nature. E-book. Accessed: 5 September 2024.

Josey, A. & Hornford, D. 2022. The TOGAF Standard, 10th Edition - A Pocket Guide. First edition, first impression, April 2022. Van Haren Publishing, s-hertogenbosch. E-book. Accessed: 31 August 2024.

JSON Schema. 2020. What is a Schema? URL: https://json-schema.org/understanding-json-schema/about Accessed. 30 October 2024.

Kapoor, A., Gulli, A. & Sujit, P. 2022. Deep Learning with TensorFlow and Keras - Third Edition. Packt Publishing. E-book. Accessed: 13 September 2024.

Korhan, O. 2023. Digital Twin Technology. IntechOpen, 2023. Electronic dataset. DOI: 10.5772/intechopen.104048 Accessed: 5 January 2025.

Koskennurmi-Sivonen, R. 2007. Grounded Theory. URL: https://rkosken.kapsi.fi/gt.html Accessed: 24 September 2024.

Kulkarni, A. & Shivananda, A. 2021. Natural Language Processing Recipes Unlocking Text Data with Machine Learning and Deep Learning using Python. Apress, 2019. E-book. Accessed: 10 October 2023.

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S. & Kiela, D. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. 34th Conference on Neural Information Processing Systems (NeurIPS 2020). Vancouver, Canada. URL: https://proceedings.neurips.cc/paper_files/paper/2020/file/6b493230205f780e1bc26945df7481e5-Paper.pdf Accessed: 16 September 2024.

Lincoln, T. 2021. Qualitative Research. Atla Open Press, 2021. E-book. Accessed: 24 September 2024.

Nvidia. 2024. PyTorch. URL: https://www.nvidia.com/en-us/glossary/pytorch/ Accessed: 15 October 2024.

OpenAI. 2024. Retrieval Augmented Generation (RAG) and Semantic Search for GPTs. URL: https://help.openai.com/en/articles/8868588-retrieval-augmented-generation-rag-and-semantic-search-for-gpts Accessed: 3 September 2024.

Ping, D. 2024. Machine Learning Solutions Architect Handbook – Second Edition. Birmingham Packt Publishing, Limited, 2024. Limited 2024. E-book. Accessed: 9 September 2024.

Python-docx. 2013. Working with Documents. URL: https://python-docx.readthedocs.io/en/latest/user/documents.html Accessed: 22 October 2024.

QPR. 2024. QPR ProcessAnalyzer. URL: https://www.qpr.com/fi/prosessilouhinta/qpr-processanalyzer Accessed: 11 January 2025.

Rohde, T. 18 August 2024. AI-Assisted Modeling Scripts. URL: https://forum.archimatetool.com/index.php?topic=1554.msg7974#msg7974 Accessed: 7 October 2024.

Sarrodie, J-P. 5 October 2018. Why Archi. URL: Why Archi? – Archi (archimatetool.com) Accessed: 14 October 2024.

Schmidhuber, J. 2022. Annotated History of Modern AI and Deep Learning. Technical Report IDSIA-22-22 (v2), IDSIA, 12/29/2022. Swiss AI Lab IDSIA, USI. URL:https://arxiv.org/pdf/2212.11279 Accessed: 5 September 2024.

Stahl, B.C., Rodrigues, R. & Schroeder, D. 2023. Ethics of AI. Springer International Publishing, 2023. E-book. Accessed: 5 September 2024.

Stryker, C. & Kavlakoglu, E. 16 August 2024. What is AI? URL: https://www.ibm.com/topics/artificial-intelligence Accessed: 8 October 2024.

Taulli, T. 2024. What is RAG (Retrieval Augmented Generation) for an LLM? O'Reilly Media, Inc, 2024. Audiobook. Accessed: 16 September 2024.

The Open Group. 1992-2022. 3. Core Concepts. URL: https://pubs.opengroup.org/togaf-standard/introduction/chap03.html Accessed: 7 September 2024.

The Open Group. 1995-2024. TOGAF 9.2 Overview. URL: TOGAF 9.2 Overview | www.opengroup.org Accessed: 1 September 2024.

The Open Group. 1999-2006. Welcome to TOGAF™ - The Open Group Architecture Framework. URL: https://pubs.opengroup.org/architecture/togaf8-doc/arch/ Accessed: 7 October 2023.

The Open Group. 1999-2022a. 1. Introduction. URL: https://pubs.opengroup.org/togaf-standard/adm/chap01.html Accessed: 9 October 2023.

The Open Group. 1999-2022b. 3. Phase A: Architecture Vision. URL: https://pubs.opengroup.org/togaf-standard/adm/chap03.html Accessed: 18 September 2024.

The Open Group. 1999-2022c. 2. Applying Iteration to the ADM. URL: https://pubs.opengroup.org/togaf-standard/applying-the-adm/chap02.html#tag_02_02 Accessed: 29 December 2024.

The Open Group. 1999-2022d. 3. Applying the ADM Across the Architecture Landscape. URL: https://pubs.opengroup.org/togaf-standard/applying-the-adm/chap03.html#tag_03 Accessed: 29 December 2024.

The Open Group. 2022a. Business Capabilities, Version 2. URL: https://pubs.opengroup.org/togaf-standard/business-architecture/business-capabilities.html Accessed: 9 October 2023.

The Open Group. 2022b. Business Scenarios. URL: https://pubs.opengroup.org/togaf-standard/business-architecture/business-scenarios.html Accessed: 24 September 2024.

The Open Group. 2023. Business Capability Planning. URL: https://pubs.opengroup.org/togaf-standard/business-architecture/business-capability-planning.html#_Toc130906758 Accessed: 7 September 2024.

Tie, Y.C., Birks, M. & Francis, K. 2 January 2019. Grounded theory research: A design framework for novice researchers. SageJournals. Electronic dataset. DOI: https://doi.org/10.1177/2050312118822927 Accessed: 22 January 2025.

Vaishnavi, V.K. & Kuechler, W. 2015. Design science research methods and patterns: innovating information and communication technology. CRC Press 2015. E-book. Accessed: 25 September 2024.

Vom Brocke, J., Hevner, A. & Maedche, A. 24 September 2020. Introduction to Design Science Research. Design Science Research. Cases (1-13). Progress in IS. Springer, Cham. Electronic dataset. DOI: 10.1007/978-3-030-46781-4_1 Accessed: 25 September 2024.

Wassan, J.T. & Ghuriani, V. 2024. Deep Learning – Recent Findings and Research. IntechOpen. Electronic dataset. DOI: 10.5772/intechopen.107725 Accessed: 14 September 2024.

World Intellectual Property Organization 2019. 2019. WIPO Technology Trends 2019: AI. World Intellectual Property Organization. E-book. Accessed: 5 September 2024.

Zeng, S., Zhang, J., He, P., Xing, Y., Liu, Y., Xu, H., Ren, J., Wang, S., Yin, D., Chang, Y. & Tang, J. August 2024. The Good and The Bad: Exploring Privacy Issues in Retrieval-Augmented Generation (RAG). In Findings of Association for Computational Linguistics, pp. 4505-4524. URL: https://aclanthology.org/2024.findings-acl.267 Accessed: 4 September 2024.

Zhang, Y. & Wildemuth, B. M. 2009. Qualitative analysis of content. In B. Wildemuth (Ed.), Applications of Social Research Methods to Questions in Information and Library Science, pp. 308-319. Westport, CT: Libraries Unlimited. URL: https://pages.ischool.utexas.edu/yanz/Content_analysis.pdf Accessed: 8 February 2025.

Appendices

Appendix 1. Use Cases

The following three use cases presented in the study's appendices contribute to enhancing the EA process through the integration of AI-driven tools, aiming to streamline capability mapping and improve strategic alignment. These use cases were created in accordance with the City of Helsinki.fi (5 November 2021) reference as detailed in chapter 3.1.1 Case Study. Use Case 1 focuses on defining organizational capabilities and creating a dynamic capability map using AI, helping to refine and accelerate the mapping process. Use Case 2 aims to assess the current state of the organization’s capabilities through AI, providing insights that inform future planning and development. Finally, Use Case 3 prioritizes capabilities based on strategic goals using AI to guide decision-making, ensuring that the most important capabilities are aligned with the organization's evolving objectives. Together, these use cases illustrate how AI can shift EA from manual processes to more automated and efficient methods, enhancing both agility and strategic decision-making within the organization.

Use Case 1 – Define Capabilities and Create a Capability Map Using AI-Based Tools and Features

Actors:

- Enterprise Architect
- Strategy Team (Architecture Board)
- Key Stakeholders

Objective:
Define capabilities and create a Capability Map utilizing advanced technologies and AI based on the defined current state.

Trigger:
Need to develop a dynamic and agile capability definition and capability mapping process.

Process Flow:

1. Identify Scope and Objectives:
    - Align with the strategic team and key stakeholders to determine the scope and objectives for defining organizational capabilities and the scope of the capability map.
2. Data Collection and Import:
    - Collect relevant information and insights from stakeholders regarding the definitions of capabilities and the capability map.

    - Validate collected information and add the information into the Retrieval-Augmented Generation (RAG) system.
3. Data Analysis and Generation:
    - Trigger predefined RAG system code to analyze input data and generate the required outputs.
4. Results Validation and Storage:
    - Review and refine outputs if necessary.
    - Accept capabilities and the capability map, then document results.
5. Approval and Publishing:
    - Obtain approval from relevant stakeholders.
    - Make changes if needed.
    - Finalize and publish the Capability Map.

Exceptions:

- No major exceptions are needed for this case.

Notes:

- Ensure that the defined capabilities align with organizational strategy.
- Strategic data can be leveraged to guide the creation of the capability map.

Use Case 2 – Current State Definition Using Advanced Technology and AI

Actors:

- Enterprise Architect
- Strategy Team (Architecture Board)
- Key Stakeholders

Objective:
Define the current state of the organization’s capabilities using advanced technology and AI.

Trigger:
The need to assess and understand the current state of the organization’s capabilities through the application of artificial intelligence and other advanced technologies.

Process Flow:

1. Identify Scope and Objectives:
    - Define what aspects of the current state need evaluation.
    - Align objectives with key stakeholders and clarify decision-making needs.
2. Data Collection and Import:

    - Additional qualitative data may be required to gather. (Reflects the need for extra context in the assessment.)
3. Data Analysis and Generation:
    - Trigger predefined RAG system code to analyze input data and generate the required outputs.
4. Results Validation and Storage:
    - Validate the results and adjust as necessary.
    - Document the final capability assessment into Archi Tool.
5. Approval and Publishing:
    - Obtain formal approval from relevant stakeholders.
    - Make changes if needed.
    - Publish the finalized current state assessment for further strategic use.

Exceptions:

- No major exceptions are needed for this case.

Notes:

- Customization of tools and methods may be necessary to meet the specific needs of the organization and to ensure alignment with strategic objectives.
- Strategic data gathered in Use Case 1 plays an important role in defining the current state.

Use Case 3 – Prioritize the Capability Map Capabilities Based on Strategic Goals

Actors:

- Enterprise Architect
- Strategy Team (Architecture Board)
- Key Stakeholders

Objective:
Prioritize capabilities based on the highest priority strategic goals, utilizing AI and advanced technologies to align with and support the organization's strategic objectives.

Trigger:
The need to select and prioritize capabilities according to the organization's most important strategic goals.

Process Flow:

1. Identify Scope and Objectives:
    - Align with the strategic team and relevant stakeholders to determine prioritization criteria.

   - Define how AI-driven insights will be used to support the decision-making process.
2. Data Collection and Import:
   - Not required, as data has already been collected in the previous stages.
3. Data Analysis and Generation:
   - Trigger predefined RAG system code to analyze input data and generate the required outputs.
4. Results Validation and Storage:
   - Review AI-generated prioritization results and adjust if needed.
   - Document the final prioritized capability list into Archi Tool.
5. Approval and Publishing:
   - Obtain approval of the prioritized capabilities from the strategy team and relevant stakeholders.
   - Make changes if needed.
   - Document the prioritization process and the results, including the rationale behind the decisions, for future reference and further development.

Exceptions:

- No major exceptions are needed for this case.

Notes:

- Strategic information needs to be integrated into the prioritization process.
- The quality of the prioritization results is the key factor guiding the evaluation of the prioritization process.
- The prioritization process may lead into the *Opportunities and Solutions* phase, as highlighted in the interview.

Appendix 2. Grounded Theory

Overview of GT Analysis Phases

According to Koskennurmi-Sivonen (2007), GT analysis is based on three key phases: Open Coding, Axial Coding, and Selective Coding. Each phase builds upon the previous one to refine the research and develop meaningful insights. In the context of this research, these phases were applied to analyze the relationships between various concepts and categories that emerged from the data.

The phases in this are outlined below:

- *Open Coding: The first phase involves identifying and categorizing relevant concepts from the data. Concepts are grouped based on similarities, and dimensions representing their relationships emerge. If concepts are too similar, the focus may shift to contextual factors.*
- *Axial Coding: In this phase, connections between categories are established, identifying causal relationships and other links within the phenomenon. This phase refines the structure and defines higher-level concepts explaining the relationships between categories.*
- *Selective Coding: The final phase focuses on identifying the core category and relating it to other categories. Relationships are validated, and incomplete categories are further developed. The core category integrates the findings, unifying the research around the central theme.*

Finally, the central narrative and identified core category are constructed based on the findings from these phases, providing a comprehensive synthesis of the analysis.

Table 1. Open Coding Table

Interview Quote	**Code**	**Category**
There is constant pressure, and no time to recover.	Work Stress	Work Load
The manager supports decisions and provides freedom.	Leadership Support	Meaningfulness of Work
Generative AI improves process efficiency.	Efficiency	AI in EA
Data management is a major challenge.	Data Management	Challenges in EA
EAA aids in documentation and analysis.	EAA System	Innovative Applications
LLMs enable quick information retrieval.	Speed	AI in EA

RAG integrates organization-specific data.	RAG Technology	Data Management
TOGAF is a standardized framework for EA development.	TOGAF Framework	Enterprise Architecture
It supports reusable architectural assets.	Reusable Assets	Enterprise Architecture
The ADM lays the foundation for an architecture framework.	Architecture Development Method	TOGAF Framework
Business capabilities are linked to the Architecture Vision.	Business Capabilities	Business Capability Planning
Value streams represent activities that create value.	Value Streams	Business Capabilities
Business Capability Mapping benefits stakeholders.	Business Capability Mapping	Business Capability Planning
Business capabilities define core business functions at a high level.	Capability Definition	Business Capability Planning
The Capability Map offers a high-level view of baseline and target architectures.	Capability Map	Business Capability Map
Mapping business capabilities clarifies implementation requirements.	Capability Mapping	Business Capability Map
AI emulates human behavior and capabilities.	AI Definition	AI Fundamentals
Deep learning and neural networks are central to modern AI.	Deep Learning	AI Fundamentals
Machine learning allows computers to improve from data.	Machine Learning	AI Fundamentals
GANs are used in image recognition and pose unique challenges.	GANs	AI Techniques
Transformers significantly enhance NLP tasks.	Transformer Networks	AI Techniques
NLP combined with ML and deep learning provides insights from text.	Natural Language Processing	AI Applications
RAG boosts LLM reliability but does not eliminate errors.	Retrieval-Augmented Generation	AI Techniques
LlamaIndex integrates custom knowledge with LLMs.	LlamaIndex	AI Tools
LangChain enables effective query processing from custom knowledge bases.	LangChain	AI Tools
Prompts improve LLM interactions and data management.	Prompting	AI Tools
Chain of Thought prompting guides LLMs through logical steps for complex tasks.	Chain of Thought	AI Techniques

Vector databases are used for similarity analysis and quick queries of high-dimensional data.	Vector Databases	AI Tools
Chroma, Pinecone, and Redis are examples of vector databases for NLP.	Vector Database Options	AI Tools
Archi is a modeling tool for enterprise architecture.	Archi Modeling Tool	EA Tools
Tokenization is the first step in preparing text for vector databases.	Tokenization	AI Data Preparation

Table 2. Axial Coding Table

Core Category	**Subcategories**	**Relationships/Connections**
Work Environment	Work Stress, Leadership Support, Interaction	Work stress reduces productivity, while leadership support and interaction improve team dynamics.
AI and Efficiency	AI in EA, Efficiency, Speed	Generative AI enhances process efficiency, and LLMs improve information retrieval speed.
Data Management	Data Management, RAG Technology	Data management challenges require tools like RAG to integrate organization-specific data.
Enterprise Architecture	TOGAF Framework, ADM, Archi Modelling	TOGAF standardizes EA, ADM is critical for architecture frameworks, and Archi is a tool for implementation.
Business Capabilities	BCM, Capability Definition	BCM delivers stakeholder value and aligns with architecture vision; capabilities define core business functions at a high level.
Innovative Applications of AI	EAA System, AI Tools	EAA supports documentation and analysis, while AI tools like LlamaIndex and LangChain improve data integration and retrieval. RAG enhances information retrieval by integrating organization-specific data.
Ethical Considerations	Algorithm Bias, Safety Issues, Privacy Risks	AI introduces ethical challenges, underscoring the need for responsible AI practices.
Deep Learning Techniques	Generative Adversarial Networks, Transformer Networks, PyTorch, TensorFlow	GANs drive image recognition and frameworks like PyTorch and TensorFlow enable the development and deployment of deep learning models, reflecting AI's expanding role.
NLP	NLP Applications, LLMs	NLP enables understanding of human language, and LLMs enhance conversational abilities.

Table 3. Selective Coding Table

Core Category	**Description**	**Connections to Other Categories**

Work Environment	Workplace dynamics affect productivity and innovation.	Tied to AI and Efficiency (work stress impacts performance), Ethical Considerations (leadership supports ethical culture).
AI and Efficiency	Generative AI and LLMs improve process efficiency and decision-making.	Influences Data Management (enhancing data handling), Business Capabilities (boosts operational efficiency).
Data Management	Effective data management is key for AI-driven business improvements.	Supports Enterprise Architecture (enables data-driven frameworks), ties into Innovative Applications of AI.
Enterprise Architecture	TOGAF and other frameworks align business goals with AI.	Defines Business Capabilities and ensures AI integration with strategic goals.
Business Capabilities	Mapping capabilities aligns AI tools with organizational goals.	Influences Work Environment (improves morale) and AI and Efficiency (enhancing operations).
Innovative Applications of AI	Advanced AI tools lead to innovative solutions in processes and documentation.	Tied to Data Management (enhances integration), Ethical Considerations (ensures responsible AI use), and RAG Technology (improves data retrieval and integration).
Ethical Considerations	Responsible AI practices mitigate risks like bias and safety concerns.	Affects Work Environment (shapes ethical culture), AI and Efficiency (can slow innovation).
Deep Learning Techniques	GANs, PyTorch, and TensorFlow are advanced AI technologies driving new applications.	Supports Innovative Applications of AI (powers AI tools), impacts NLP (enhances LLM capabilities).
Natural Language Processing (NLP)	NLP facilitates communication between machines and humans, improving retrieval.	Linked to AI and Efficiency (improves retrieval) and Innovative Applications of AI (powers advanced applications).

Central Narrative

This research identifies the **Work Environment** as a critical factor influencing both **AI and Efficiency**, which are essential for enterprise productivity. Effective **Data Management**, particularly in data collection and processing, is vital for leveraging Generative AI and LLMs, driving business process improvements. Frameworks like **TOGAF** provide a structured approach to align enterprise goals with AI capabilities, ultimately enhancing **Business Capabilities**.

The **Innovative Applications of AI**, supported by advanced **NLP** techniques such as LLMs and **RAG systems**, along with refined **Deep Learning Techniques** like GANs, **PyTorch**, and **TensorFlow**, illustrate how enterprises can optimize efficiency and decision-making through AI. However, these technological advancements must be carefully balanced with **Ethical Considerations** to mitigate risks such as algorithmic bias and safety concerns.

Finally, **Natural Language Processing (NLP)** plays a crucial role in facilitating seamless human-machine interactions, which enhances AI's overall effectiveness in achieving enterprise objectives.

Identifying the Core Category:

Core Category: **The Integration of AI in EA Efficiency**

Appendix 3. Theme-Interview

The thematic interview covers various topics aligned with the research aim, drawn from pre-selected and relevant perspectives. These topics are structured into themes, with subtopics further dividing the themes into more detailed levels. Some of the content serves as introductory material and setting the stage for the topic and theme discussion within the interview. Questions that are not in the form of direct questions are intended as discussion topics designed to prompt the interviewee and uncover important insights during the interview. The interview was conducted with a senior expert in EA from the client company. The expert has decades of experience in developing and managing EA and possesses an in-depth understanding of its practical implementation within the organization.

The thematic interview is based on the results of GT analysis, which was conducted based on the literature review for this study, in accordance with Koskennurmi-Sivonen's (2007) guidelines in chapter 3.1.2 Design Science Research. This theme interview applies Farquhar's (2012, chapter 5) principles, which place interviews, observations, archives, and documents at the core of data collection. Additionally, it follows Farquhar's (2012, chapter 6) definition of the concurrent process of data collection and analysis, where data is often written up and analyzed as it is gathered, as discussed in chapter 3.1.1 Case Study.

1. Theme: Enterprise Architecture and Efficiency:

Comment: At the beginning of the project, the artificial intelligence (Retrieval-Augmented Generation) environment must be implemented, after which the data of the overall architecture can be entered into the system and processed.
Comment: Relevant models and tools for capability assessment and prioritization need to be redefined and coded.

2. Theme: Business Capabilities and Capability Mapping:

The following use cases are defined to identify and design an EA-based process model supported by AI. The use cases are listed sequentially below:

1. *Define capabilities and create a capability map using AI-based tools and features*
2. *Define the current state using advanced technology and AI*

- *Interviewee's response:*
- *Define the scope and objectives.*
- *Additional qualitative data may be required.*

3. *Prioritize the capabilities in the capability map based on strategic goals*

- *Interviewee's response:*
- *Strategic information is needed to support this.*
- *This process leads to the "Opportunities and Solutions" phase, with potential extensions to other phases.*

Question: How is strategic data collected during the process? Is it gathered all at once, or in multiple phases?
Interviewee's response: It is collected one phase at a time.

Question: How is information approved by the strategy team and stakeholders? Are capabilities, the capability map, and the maturity assessment approved individually or as a whole?
Interviewee's response: They can be sent for approval as needed or immediately upon completion.

3. Theme: Enterprise Architecture Frameworks

Question: How is TOGAF's ADM used as a framework to implement the process model?
Interviewee's response: Capability definition, capability map creation, and utilization are mainly implemented in phases A and B of the ADM cycle.
Comment: TOGAF is beneficial as it provides clearer boundaries compared to many other similar models.

Question: What are the alternative options for TOGAF's ADM and capability mapping?
Interviewee's response: Alternatives include **Julkisen Hallinnon Suositus (JHS)** or other frameworks.
Comment: The model can also be adapted to work within different governance models and standardized frameworks.

4. Theme: Innovative Applications of Artificial Intelligence

Question: Is a fully automated process the target?

Interviewee's response: Yes

Comment: There are two different options for large language model (LLM) prompting:

- *Instructions (Automated: the system is coded to produce results).*

- *External information is required in the RAG system to process organizational data through the LLM.*
- *User Input or Query (Chatbot): Output Indicator, the output is modeled or structured as needed.*
- *Interviewee's response:*
    - Suitable for content analysis.
    - Could be integrated into the EA tool to implement individual solutions.
    - In future applications, the entire material could be used for queries and analytics.

Question: What are the main goals for LLM implementation?

Interviewee's response: The tool should be marketable from a business perspective.

Question: What is the scope of the final AI-based product?

Interviewee's response:

- A high-level description of the main tools and components, including libraries and programming languages used to develop the system.
- Justification of the system’s structure based on literature.
- A functional description is sufficient.

Theme 5: Ethical Considerations

Question: How can potential ethical risks be identified and managed?

Interviewee's response and the author’s reflection:

- *Data privacy risks:*
    - No personal data or interview data is used.
    - Automated processes ensure data modification and deletion.
- *Algorithm bias:*
- *Safety issues:*

Comment: Using a training set prevents user queries and LLM data from flowing into LLM providers’ databases.

Theme 6: Open Questions

Question: Are there any open questions or ideas for further development?

Interviewee's response and the author’s reflection:

- Other use cases and development planning:
  - Creation and evaluation of other content.
  - Comprehensive analysis.
  - Chatbot development.
  - Further development ideas for the chatbot.
    - The entire architecture content.
    - Strategy development and new ideas.
    - Miscellaneous items.
- Operational data (Digital Twins):
  - Used to guide decision-making.
  - Description of implementing capabilities with AI assistance (to model, analyze, or design using AI).
    - Definition of the contents of capabilities.
    - Identification of organizational units at a deeper level.
- Development suggestions:
  - ArchiMate modeling: Is it sufficient? What about other notations?
    - If capabilities or processes need further detail, BPMN may be required, which ArchiMate may not support.
    - Specific process descriptions or data models (such as UML) may be necessary, which ArchiMate does not support.
  - Archi-content development and interpretation:

- AI can generate content, but humans must be able to edit it. All content must be editable.
- The ability to modify data is essential.

Reflection of the Benefits:

- The system's overall benefits were considered limited in relation to the specific use cases and system implementation. However, the potential future outcomes became clearer when considering the design and implementation of additional use cases.
- Considering other tools in addition to Archi and the possibility of implementing the system without Archi, updating the information to the repository is essential. However, it was noted that Archi is beneficial for tasks like managing input/output operations and handling file format challenges between different tools.

Regarding the future, it was noted that a similar or even more advanced system will likely be adopted in the future, given the evolution of enterprise architecture tools, planning, and development.

Appendix 4. High-Level Architecture of the Retrieval-Augmented Generation System