BISE Student

https://bise-student.io

MASTER'S THESIS

Turning Disruption into Potential: Key Capabilities for Leveraging Generative Artificial Intelligence Successfully

Publication Date: 2024-11-27

Author
Julian FABER
Karlsruhe Institute of Technology
Karlsruhe, Germany
julian.faber@student.kit.edu
0x85e46b14015fB8a4F7965bAca85034F68315d968

Abstract

The thesis investigates what capabilities are necessary to use the transformative potential of Generative AI (GAI) in companies, highlighting how early usage can be a source of competitive advantage. Given the rapid advancements in GAI since 2022 and its impact on business, this research identifies key capabilities to use GAI successfully. Addressing a gap in current literature, which is often limited by industry-specific perspectives and pre-2022 developments, this study adopts a cross-industry explanatory approach. It builds upon insights from eleven GAI experts across corporates, SMEs, startups, consultancies, and venture capital funds, while capturing diverse perspectives. The research initially establishes a fundamental understanding of GAI, exploring leading providers, market trends, and limitations. GAI's unique sequence-to-sequence modeling automates repetitive tasks and enhances productivity, but also presents challenges like data inconsistencies, often requiring a human-in-the-loop. The study underscores the...

Keywords: Genertaive AI, AI, Dynamic Capabilities
Methods: Literature Review, Expert Interviews

Submission Date: 2024-10-31
Submission Contract: 0x518840b329308cc0d36Ea74b4C9037B2B654C789

Turning Disruption into Potential: Key Capabilities for Leveraging Generative Artificial Intelligence Successfully

Master Thesis

by Julian Faber

Degree Course: Industrial Engineering and Management M.Sc.

Matriculation Number: 2151144

Institute for Applied Informatics and Formal Description Methods (AIFB)

KIT Department of Economics and Management

Advisor:	Prof. Dr. Ali Sunyaev
Second Advisor:	Prof. Dr. Andreas Oberweis
Supervisor:	Dr. Sebastian Lins
Submitted:	18. October 2024

Table of Contents

List of Abbreviations

AI	Artificial Intelligence
API	Application programming interface
CAGR	Compound annual growth rate
GAI	Generative AI
GPT	Generative pre-trained transformer
IP	Intellectual property
LLM	Large language model
POC	Proof-of-concept
ROI	Return-on-investment
SBOM	Software-bill-of-material
SMEs	Small and medium-sized enterprises

List of Figures

List of Tables

1. Introduction

Generative Artificial Intelligence (GAI) is an emerging technology that has achieved remarkable advances since late 2022 (Douglas, 2023, pp. 1–2) and profoundly influenced global dynamics (Mondal et al., 2023, p. 3). Notably, ChatGPT, which reached over a million users in just five days (Mondal et al., 2023, p. 3) made society realize the far-reaching impact that Artificial Intelligence (AI) will generate (Morris, 2023, p. 23). In the upcoming years, companies must adapt and leverage GAI successfully. Rapid technological changes as the digital disruption already demonstrated how companies that could not adapt failed while their competitors gained an advantage by using new technologies (Ho & Chen, 2018a, p. 1). For instance, the use of the internet had a remarkable impact on the competitive advantage of companies in the first years (Teo & Pian, 2003, p. 90). The impact of GAI is also far-reaching and will bring long-term implications, from upskilling efforts to workforce reductions (McKinsey, n.d.). In the next years, AI software is anticipated to cause the most disruption in the realm of technology (Bloomberg, 2017). Therefore, companies that implement AI technologies successfully can achieve a considerable competitive advantage (Climent et al., 2024, p. 1), while those who overlook it may risk encountering difficulties in the long term.

There exists certain research on how to use AI in companies successfully (Brenner et al., 2021, p. 15; Wagner, 2020, p. 19) and how AI can improve organizations (Hercheui & Ranjith, Rishikesh, 2020, p. 87; Wamba-Taguimdje et al., 2020, p. 3). Some authors identified maturity levels for AI implementation (Lichtenthaler, 2020, p. 39) or a scale to assess the AI usage in the core services of a company (Drydakis, 2022, p. 1223). However, even if the results provide valuable insights the biggest caveat remains, that the research was published before the hype of GAI which started between 2022 and 2023 (Leaver & Srdarov, 2023, p. 1) and does not incorporate latest developments and circumstances. Further research was focused on specific industries like the Norwegian marketing sector (Mikalef et al., 2021, p. 80), the banking industry (Gallego-Gomez & De-Pablos-Heredero, 2020, p. 20), the manufacturing industry (Abou-Foul et al., 2023, p. 1; Sjödin et al., 2023, p. 1), and GAI in the Chinese construction and the French-American supply chain industries (Fosso Wamba et al., 2024, p. 1; Liu et al., 2024, p. 64).

Thus, most results do not consider the latest developments in GAI or focus on industry-specific perspectives. This creates a gap for a cross-industry explanatory approach, as it has not yet been sufficiently investigated how companies can use GAI successfully and which capabilities are needed. Therefore, the objective of this thesis is to distill cross-industry GAI capabilities to use the technology successfully. In doing so, I answer the following research question:

Research Question: *What capabilities are needed for the successful usage of GAI?*

To answer the research question, the first step consisted of understanding current developments within the GAI landscape. Subsequently, I explored how companies can integrate GAI into their business use cases, optimize organizational processes, and leverage the necessary resources for effective implementation. These findings are used to identify cross-industry capabilities to successfully use GAI.

Such a cross-industry explanatory approach should include the perspectives of corporates, small and medium-sized enterprises (SMEs), startups, consultants and investors. For this reason, a qualitative research method is used by conducting interviews with various experts from relevant industries to nurture multi-faceted perspectives.

To derive an initial guideline for the expert interviews, the theory of dynamic capabilities was used. The theory of dynamic capabilities offers a useful framework to identify the necessary capabilities for leveraging new technologies (Teece et al., 1997, p. 509). The theory explains how companies can sense, seize, and reconfigure in times of rapid technological change (Hercheui & Ranjith, Rishikesh, 2020, p. 87; Teece et al., 1997, p. 509) and was also used by, e.g., Drydakis (2022), Hercheui & Ranjith, Rishikesh (2020), Wamba-Taguimdje et al. (2020) in the context of AI. As expressed in the theory of dynamic capabilities it is not possible to provide an overall handbook on wealth creation for all companies, but rather “suggest overall direction[s]” (Teece et al., 1997, p. 528) by working with an appropriate framework (Teece et al., 1997, p. 526).

This cross-industry explanatory approach for GAI capabilities can serve as a foundation for the decision-making process of C-level executives and managers in formulating GAI strategies, as it becomes a focus topic for leaders (McKinsey, n.d.). Furthermore, it supports future research on how to identify capabilities to use specific technologies.

The thesis proceeds as follows: The theoretical background of GAI and dynamic capabilities are elaborated, succeeded by the methodology, which primarily consists of the qualitative research method of expert interviews. After that, the current market developments are analyzed. Subsequently, the results of the expert interviews are examined to identify GAI capabilities. The thesis concludes with a discussion section that addresses the principal findings, implications for research and practice, suggestions for future research, and limitations.

2. Theoretical Background

2.1. Generative Artificial Intelligence

GAI systems are intended to produce text resembling human language, respond to queries, and perform various tasks in natural language. (Floridi & Chiriatti, 2020, p. 684; Kasneci et al., 2023, p. 1; Sejnowski, 2023, p. 1; Thirunavukarasu et al., 2023, p. 1). GAI systems are based on foundation models, which are trained with extensive datasets, to tackle new tasks derived by instructions in natural language (Scao et al., 2023, p. 3). Besides language-related operations, advanced foundation models are multimodal, which means that they possess the ability to comprehend, generalize, and function with various input forms, including images, audio, or video, alongside text (OpenAI, n.d.-a; Pichai & Hassabis, 2024).

However, in this thesis the focus remains on text-based language processing as provided particularly by Large Language Models (LLMs) with less focus on image, audio, and video generation, as those use cases are currently less common for the B2B segment (AWS, n.d.). Technically, LLMs represent a subspace of GAI, illustrating a statistical probability for a certain sequence of words occurring in a language (Luitse & Denkena, 2021, p. 1).

In 2017 Google researchers published a paper about transformer model architecture, which inspired many LLM providers, among them also OpenAI, to develop generative pre-trained transformers (GPTs) of the first generation (Radford et al., 2018, p. 1; Scao et al., 2023, p. 3; Vaswani et al., 2023, p. 1). Further LLMs followed and have been improved continuously leading to altered versions of the initial model architecture, including OpenAI which published GPT-4 (OpenAI, n.d.-a), Google with its Gemini 1.5 (Pichai & Hassabis, 2024), and Meta with Llama 2 (Meta, n.d.), among many others. Those foundation LLMs, are also called foundation models in the further course of the thesis, as they are the base to power GAI applications (Geirhofer & McKinney, 2023, p. 3).

Currently, foundation models are divided into two main categories consisting of open-source and closed-source models (Geirhofer & McKinney, 2023, p. 6). Examples of open-source foundation models are developed by Mistral (MistralAI, n.d.), Meta (Meta, n.d.), Hugging Face (HuggingFace, 2024), and Google (Banks & Warkentin, 2024; Schmid et al., 2024). Even if the foundation models of these providers are open-source, some of them have additional streams of income through charging customers for platform access or infrastructure services (HuggingFace, n.d.; MistralAI, n.d.).

In contrast, there are also closed-source foundation models offered by Adept or Cohere (Adept, 2022; Cohere, 2024c). Nevertheless, both providers also launched open-source foundation models besides their closed-source offerings (Cohere, 2024b, 2024a; Ranjan, 2024).

Furthermore, there are also provider in the realm of LLMs, that do not develop foundation models themselves but build on top of them to power various GAI application (Geirhofer & McKinney, 2023, pp. 2–4). In my thesis, these providers are referred to as layer providers of GAI applications, as they build a layer of additional utility for customers on top of the foundation models and manage the layer of interaction between the foundation model and end-users by providing a user interface. Those layer providers can use open-source foundation models, as well as pay for access to closed-source models. The added utility of layer providers may lie on combining several foundation models and fine-tuning them with proprietary data to achieve increased performance for specific use cases or industries. Listing of layer providers is not possible within the scope of this thesis, as there are countless start-ups and corporate inhouse solutions that represent such layers.

Well-known examples can be found at Microsoft and SAP, which started to build GAI applications within their product portfolio based on various foundation models (Marr, 2023; SAP, 2024).

However, there are also providers who develop closed-source foundation models and directly offer layer applications to end-users. These include Aleph Alpha, Anthropic, and OpenAI (AlephAlpha, 2023; Anthropic, 2023; OpenAI, 2022). Besides offering direct layers for end-users, OpenAI and Anthropic also offer access to their closed-source model via application programming interfaces (APIs), while Aleph Alpha also published an open-source model recently (AlephAlpha, 2024). Stability AI and xAI on the other hand also offer direct layers for end-users based on their own open-source models (Reuters, 2023; Stability AI, n.d.; x.ai, n.d.).

In addition, there are cloud platforms that also provide useful AI platforms for training, managing and running models (Oracle, 2024). These cloud and AI platforms are for example provided by AWS, Microsoft Azure, Google Cloud and IBM (Law, 2023).

These relationships between foundation models and layer providers alongside cloud and AI platforms are illustrated in Figure 1. Additionally, examples of prominent providers from each segment are also illustrated, enabling a better understanding of relevant market players.

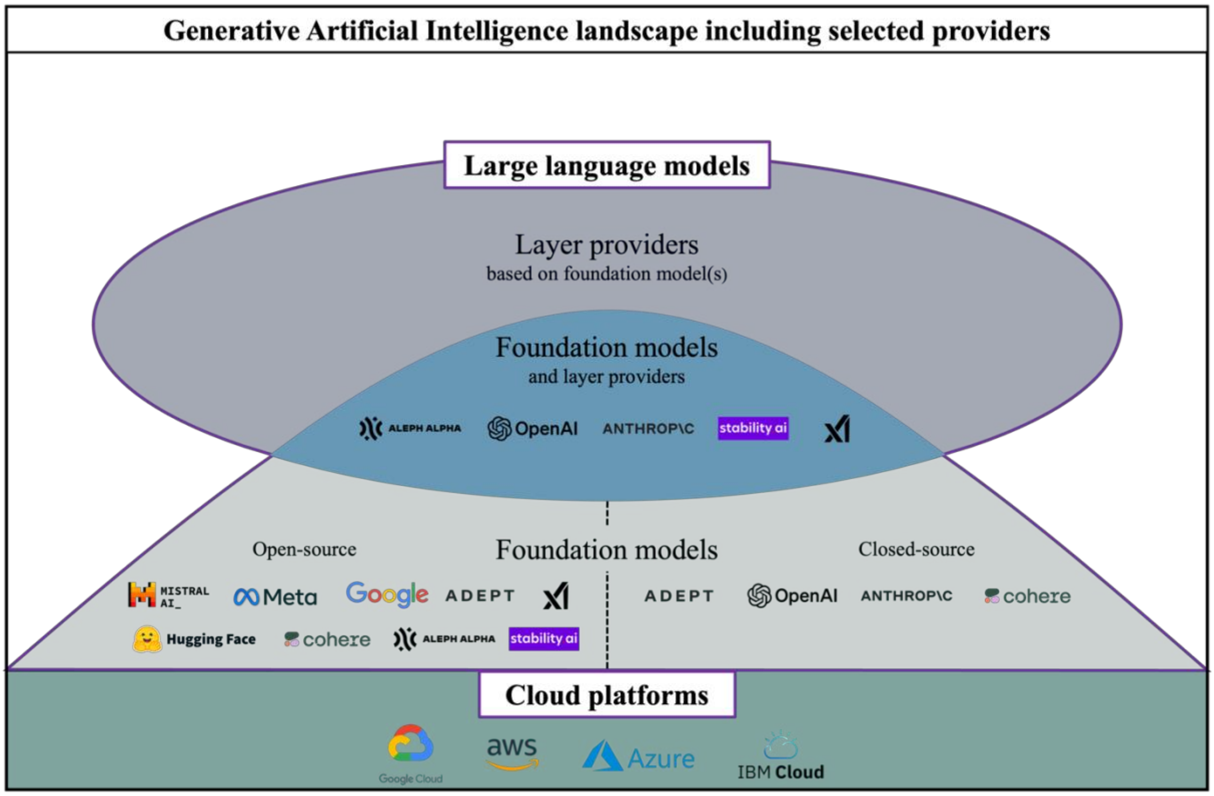

Figure 1: Visualization of the GAI subset of foundation models and layer providers, including cloud platforms

2.2. Dynamic Capabilities

A crucial topic in strategic management is how a company can achieve and obtain competitive advantages (Teece et al., 1997, p. 509). Many theories focus on strategies for maintaining the competitive advantage that was already achieved, hence featuring a privileged perspective (Teece et al., 1997, p. 510). One of these theories is the resource-based view, this theory assumes that competitive advantage is achieved through difficult-to-imitate superiority in structures and systems, which are forming the company specific resources (Teece et al., 1997, p. 513). The theory of dynamic capabilities extends the resource-based view and focuses on how specific companies can build competitive advantages in times of dynamic technological and market-related changes (Teece et al., 1997, p. 512). The theory proves to have particularly relevant insights into the competitive dynamics of high-tech industries such as software and semiconductors (Teece et al., 1997, p. 515), making it also suitable for new emerging technologies as GAI. Moreover, the theory of dynamic capabilities discusses business models that integrate enabling technologies via licensing agreements and identifies profitability challenges for licensors that charge royalty fees (Teece, 2018, p. 47). In complete contrast to these findings, leading foundation model providers (as OpenAI) chose to monetize the usage of their models through third parties by charging them fees for accessing their foundation model via an API (OpenAI, n.d.-b), which provides new momentum to innovative business models.

Companies need to build management capabilities to adapt, integrate, and reconfigure skills regarding technology, organization, and functionality (Teece et al., 1997, p. 511). Even if big companies (e.g., IBM) seem to accumulate technological resources through aggressive intellectual property (IP) claims, long-term winners at a global scale will be companies that can handle rapid technological innovations and poses the managerial capability to execute strategies (Teece et al., 1997, p. 515). Therefore, dynamic capabilities enable the development of new types of competitive advantages during rapid technological changes.

To identify the dynamic capabilities of a company there are various factors which can be divided into three main categories, which are “processes, positions, and paths” (Teece et al., 1997, p. 518): The competitive advantage of a specific company arises from its organizational processes, consisting of static coordination, dynamic learning, and transformational reconfiguration (Teece et al., 1997, p. 518).

Managerial coordination involves the integration of external and internal activities into the firm (Teece et al., 1997, p. 518), which makes efficiency and effectivity key components for success. This also includes the integration of technologies to achieve strategic advantages. Hence, performance can be achieved through superior organizational routines (Teece et al., 1997, p. 519). Learning on the other hand is an iterative concept of repeating and experimenting on different tasks to enable performance improvements and identification of opportunities (Teece et al., 1997, p. 520). Key characteristics regarding the learning capability are individual and collective skills, and routines to solve problems within the organization. The ability of reconfiguration involves sensing the necessity to reconfigure the firm's structures and accomplish such transformations (Teece et al., 1997, p. 520). It is a skill that can be trained through frequent repetition and subsequently represents a competitive advantage for companies that transformed in the past. Decentralized companies can often contribute to these transformational processes and sometimes are called high flex.

All these processes are shaped by the firm's asset positions consisting of technological, complementary, financial, reputational, structural, institutional, and market assets (Teece et al., 1997, pp. 521–522). The processes and asset positions described above influence certain path dependencies, as historical decisions on established routines and acquired assets determine the pool of available strategic options for future behavior (Teece et al., 1997, pp. 522–523). One example of this is an early investment in a certain technology, which can also serve as the basis for many other business areas later (e.g. early entry into the cloud computing). Such examples show that early investments can secure a first mover advantage but are by no means self-fulfilling prophecies for long-term success (Teece et al., 1997, p. 523). At the same time, early decisions can lead to a lock-in effect on non-competitive technologies and cause high switching costs (Arthur, 1989, pp. 116–117). However, switching costs should always be compared to the potential switching benefits (Teece et al., 1997, p. 523). Furthermore, technological opportunities are not always exogenous, as they often require previous research and development activities (Teece et al., 1997, p. 523). Hence, firms' dynamic capabilities stem from the organizational processes that are shaped by its asset positions and past decisions (Teece et al., 1997, p. 524), as illustrated in Figure 2.

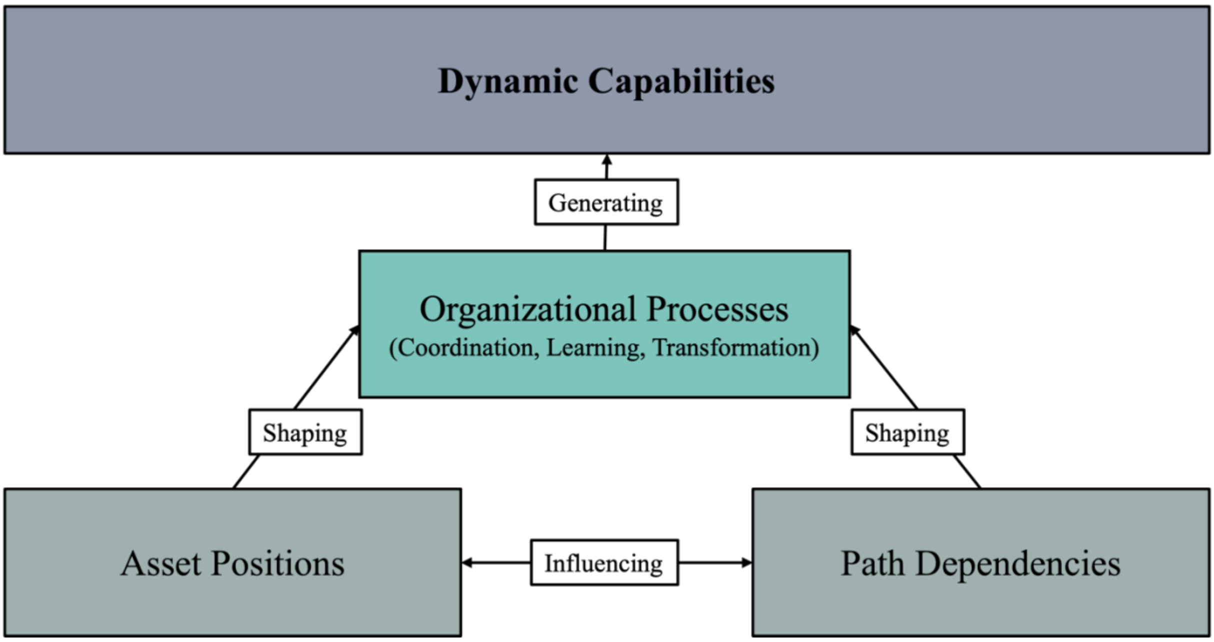

Figure 2: Visualization of the theory of dynamic capabilities by Teece et al. (1997)

Over time, further insights were incorporated into the theory of dynamic capabilities, with an emphasis on its effects on business model design and strategy (Teece, 2018, p. 40). A business model streamlines technological innovation, knowhow, and assets to generate profit.

As the business model design of a firm adjusts and transforms over time to redirect resources and ensure long term profitability, it is highly dependent on the firm's underlying dynamic capabilities (Teece, 2018, pp. 40–41). Dynamic capabilities enable sensing of technological opportunities and seizing new business models and resources within the competitor landscape, leading to successful transformations of the organizational structures (Teece, 2018, p. 41). However, the capability of an organization to sense new opportunities is shaped by its managerial competences. Strategy, on the other hand, is a pool of analyses, actions, concepts, and arguments to react to challenges when stakes are high (Rumelt, 2012, p. 16). Therefore, strategy leads to selecting a specific business model over others, but also choosing new ones over time (Casadesus-Masanell & Ricart, 2011; Teece, 2018, p. 44). However, in times of rapid technological changes new business models are established, which influence the corporate strategy as well (Teece, 2018, p. 44). For that reason, the effects between dynamic capabilities, business model, and strategy shape a firm's competitive advantage. The combination of dynamic capabilities with strategy creates a business model that enables organizational transformation (Teece, 2018, p. 44). These interconnections are visualized in Figure 3, which is based on the illustration of Teece (2018, p. 44).

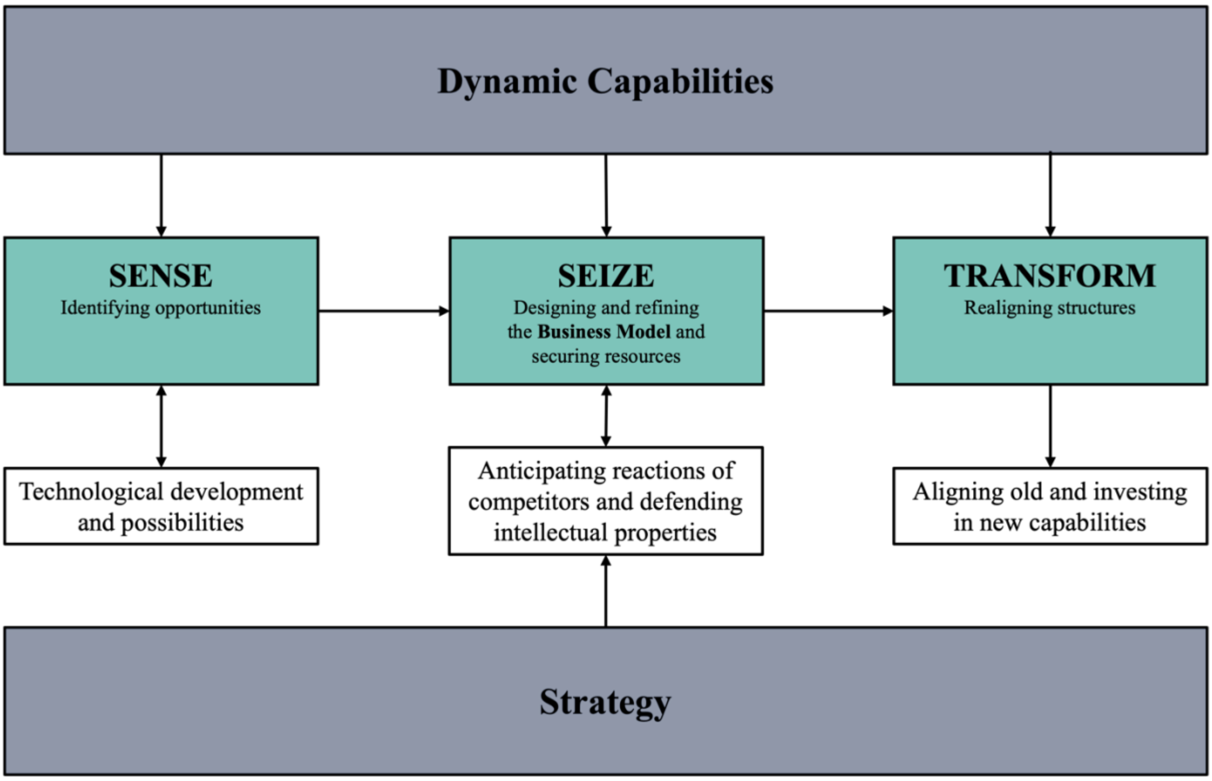

Figure 3: Visualization of dynamic capabilities, business model, and strategy (Teece, 2018, p. 44)

3. Methodology

The methodology is based on a combination of literature review and expert interviews. While the focus of this thesis lies on the expert interviews primarily, a literature review was carried out beforehand to establish a knowledge base and familiarize with previous results.

3.1. Literature Review

The literature review used a multifaceted approach to gather insights from a diverse range of sources. However, this review extends beyond traditional academic literature and incorporates gray literature, which includes news articles, market analyses, industry reports, and other non-peer-reviewed materials. The purpose of this inclusive approach is to obtain a comprehensive and nuanced understanding of the existing capabilities on how to use GAI successfully.

To initiate the process, a search strategy is devised, combining both academic databases and specialized repositories for gray literature. This strategy involved querying databases such as EBCSOHost, Proquest, and ScienceDirect with combinations of various terms as generative or gen combined with AI or Artificial Intelligence and capabilities or capability and use or usage in the title and abstract. Most of the literature focuses on specific industry use cases. Besides that, Google Scholar was used for journal articles, conference papers, and white papers related to technological disruption through GAI. Simultaneously, research is conducted on news websites, industry publications, and market analysis platforms to capture real-time developments, trends, and expert opinions.

This comprehensive review formed the knowledge base for the study, allowing the identification of current market dynamics, and further insights regarding GAI. This literature review established the theoretical knowledge for the following analysis, focusing on the capabilities of companies to successfully leverage GAI.

3.2. Expert Interviews

3.2.1. Research Design

To obtain insights directly from industry experts, a qualitative research method is used by conducting interviews based on predefined guidelines. These guidelines are developed in alignment with the research questions and the theoretical framework, drawing upon the methods outlined by DeMarrais and Lapan (2004, pp. 61–63) to ensure that interviews produce information pertinent to the study objectives. By engaging with industry professionals, the objective is to capture nuanced perspectives, strategic insights, and industry-specific knowledge that may not be readily available through desk research alone. Interviews allow for the collection of qualitative data, which is essential for understanding the context, trends, and nuances within the GAI technology (Rosenthal, 2016, p. 510). This type of information is often difficult to capture using quantitative methods alone, especially in emerging markets.

3.2.2. Interview Partner Acquisition and Interview Conduction

To gather comprehensive information for this study, a multifaceted approach was employed to identify suitable interview partners. Initially, personal connections played a crucial role, taking advantage of existing networks to establish contacts with individuals who possess relevant expertise in the field. In addition, expert workshops and events on GAI were visited to network with leading companies and experts with the objective to secure further expert interviews. Another important aspect were introductions to new interviewees by the experts already interviewed. Furthermore, outreach efforts were extended to key personnel within prominent companies. This included a proactive engagement with potential interviewees through professional platforms, such as LinkedIn. Cold outreach through this platform allowed the identification and connection with individuals who exhibited expertise in the subject.

The focus was specifically on individuals working in corporations undergoing technological changes due to GAI, who have already used GAI applications or been involved in making strategic decisions related to GAI. In addition, founders or consultants who are developing and implementing GAI applications, and investors who invested in GAI companies, were interviewed. This method ensured a diverse range of perspectives from industry professionals.

The final sample size consisted of a selected group of eleven interviewees consisting of AI experts, C-level executives, investors, consultants and founders from small (n = 5), medium-sized (n = 2) and large companies (n = 4). These companies also encompass industry leaders from sectors such as strategy consulting, automotive, and software technology.

All interviewees have in-depth industry knowledge regarding GAI and a focus on technical implementation, strategy development, investing or compliance. An overview of all the interview partners with further information, business segment, job title, and respective perspective is presented in Table 1.

The semi-structured interview format was chosen to provide an architecture that allowed flexibility in exploring diverse perspectives while ensuring that key themes were consistently addressed across all interviews (Myers & Newman, 2007, p. 4). Within the scope of this, an interview guideline based on the theory of dynamic capabilities, introduced in section 2.2, was prepared (see Appendix A). Besides general questions on GAI, theoretical aspects of coordination, learning, transformation, asset positions, path dependencies and business model were used to derive overall categories with various questions adapted to the context of GAI. For example, two questions were derived for the theoretical aspect of learning. One of the interview questions was stated as follows: “How can organizations iterate and learn fast during the integration process of GAI?”.

Each interview was designed to last around 45 minutes. To preserve the accuracy and integrity of the data, the audio of all interviews was recorded during video calls. This method not only facilitated transcription and subsequent analysis but also eliminated a major source of interviewer bias, enriching the depth of the qualitative data collected (Bucher et al., 1956, p. 360).

Table 1: Information on interview partners

No.	Industry	Employees	Job Title	Perspective	Experience with AI	Professional Experience	Interview Length
I1	Globally leading technology corporate	>100,000	AI Product Expert	Corporate, User, Technical	4 years	17 years	30:53 min
I2	Globally leading technology corporate	10,000 – 100,000	AI Compliance Professional	Corporate, User, Compliance	2 years	2.5 years	47:24 min
I3	Venture Capital unit of a globally leading technology corporate	>100,000	Investment Director	Corporate, Investor, Technical	-	13 years	46:13 min
I4	Venture Capital Fund	1 – 50	Investment Manager	Investor	5 years	5 years	30:45 min
I5	Venture Capital Fund	1 – 50	Co-Founder, Partner	Investor	8 years	14 years	42:22 min
I6	Robotics and GAI Startup	1 – 50	Chief Technology Officer, Co-Founder	Startup, Technical, Provider	-	8 years	42:27 min
I7	GAI Startup	1 – 50	Chief Technology Officer, Co-Founder	Startup, Technical Provider	8 years	15 years	53:01 min
I8	Consultancy for GAI startups and investors	1 – 50	Technology Advisor	Consulting, Startup, Technical	5 years	24 years	71:04 min
I9	Globally leading consultancy	10,000 – 100,000	Practice Manager	Consulting, Provider	-	6 .5 years	40:28 min
I10	Computer Vision and GAI Provider	50 – 250	Chairman of the Board	Consulting, Technical Provider	6 years	14 years	69:44 min
I11	Leading German (G)AI Consultancy	50 – 250	Chief Executive Officer	Consulting, Provider	7 years	14 years	31:35 min

3.2.3. Data Analysis Procedures

To enable a robust analysis, the post-interview phase was initiated through the transcription of the collected audio data. This step involved the transformation of qualitative spoken content into a written format, preserving the insights of participant responses (Bailey, 2008, p. 3). Recognizing the importance of precision and consistency in capturing participant responses, the transcription procedure followed established guidelines for a comprehensive representation of the interview content.

In this study, data analysis revolved around a robust coding framework (Hsieh & Shannon, 2005, p. 1282), primarily employing a mixture of *deductive coding* and *open coding* as the cornerstone methodology. In the process of deductive coding, several codes were predetermined before the interviews were analyzed. The reuse and application of established concepts was used to validate or expend upon a theoretical framework (Abraham et al., 2013, p. 6). The theory of dynamic capabilities, introduced in section 2.2, served as initial starting point for the deductive codes of the analysis. Henceforth, codes as business model, processes and organization, strategy, coordination, learning, path dependencies, resources, switching costs, flexibility, and IP were derived. Additionally, open coding introduced new codes if certain text segments could not be assigned to the deductive codes and surfaced during the expert interviews. The open codes grouped similar text passages that were frequently mentioned by the experts. Thus, the deductive codes (n = 10) provided an initial starting point, which is further expanded by open codes (n = 22) from the expert interviews. A total of 845 text passages from the transcripts were assigned to the 32 codes, of which 280 text passages belong to the deductive codes and 565 to the open codes. For example, the text passage "So I don't think it's likely in the near future, let's say, that the real expert will be replaced by this [GAI]. I don't think so, at least not soon" of Interviewee 10 was assigned to the open code "General Developments_Limitations".

After assigning all relevant text passages to the codes, all codes (deductive and open) were again grouped into categories. These categories are representing codes (as background, general developments, use cases, process and organization, path dependencies, compliance, security, and international comparison) to which further sub-codes were assigned. Furthermore, these categories determine the structure of the result section, as first general perspectives on GAI and the identification of compatible use cases are examined. Following that, an effective organizational design is analyzed. Furthermore, advantageous resources and finally compliance and data security are discussed. The category of international comparison is featured in the discussion section, while the category of background provided information on the interview partners, which is presented in Table 1.

Figure 4 provides a detailed overview of all the codes and the number of assigned text passages. Codes that were created during deductive coding are illustrated in blue, while codes that were created by open coding are marked in green. The categories are presented on the far left and let the sub-codes branch off from themselves.

The goal was to extract meaningful insights from the conducted interviews with a diverse array of stakeholders, including consultants, investors, founders, and corporate executives. Simultaneously, memoing was used to capture important thoughts regarding the interviews and to build on them later in the results section (Wiesche et al., 2017, p. 688).

To enrich the analysis, a *constant comparison method* was used to examine and contrast data across various participant groups (Wiesche et al., 2017, p. 688). Through the comparison of data and concepts in all interviews, distinctions, similarities, and characteristics were examined to ensure the reliability of the findings. Given the different professional backgrounds of the interviewees, this comparative analysis revealed potential divergences and convergences in their perceptions of the GAI capabilities. Above all, the different perspectives of investors compared to technical experts, from small companies to corporates, from customers to providers, complement a multi-faceted picture. This facilitated the identification of patterns and trends, which enabled a comprehensive understanding of how different stakeholders conceptualize needed capabilities to use GAI successfully (Strauss & Corbin, 2003, p. 67).

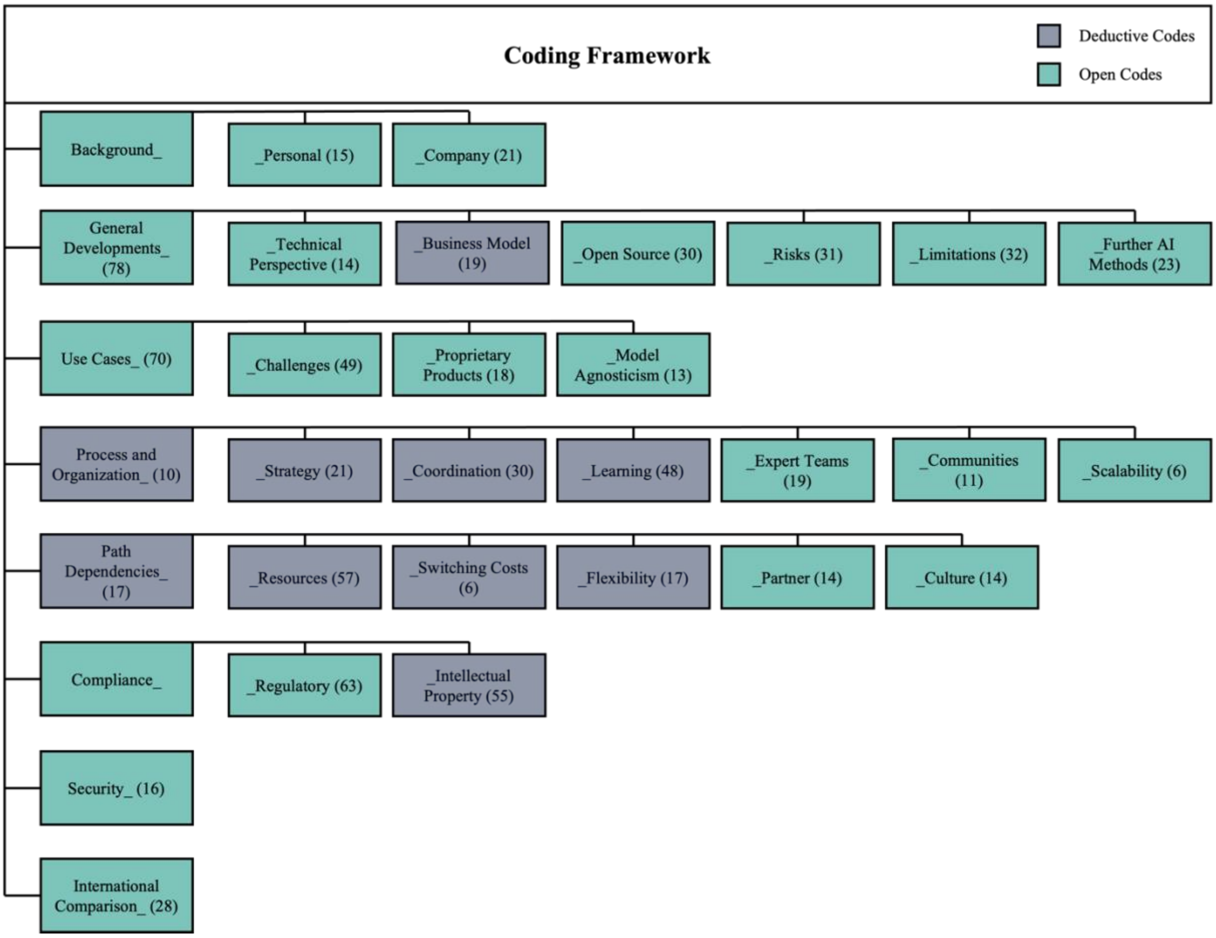

Figure 4: Tree diagram of the coding framework with number of mapped text passages

4. Market Developments

All started when Google presented the underlying architecture of transformers which paved the way for foundation models in 2017 (Radford et al., 2018, p. 1; Scao et al., 2023, p. 3; Vaswani et al., 2023, p. 1).[1] In simple terms, the transformer architecture performs sequence-to-sequence modeling and returns the most probabilistic output based on a certain input, including multimodal tasks (I6). The probabilistic retrieval based on the input uses all the data with which the foundation model was trained and can operate on very large matrices, unstructured data, and broad contexts (I5, I6, I7).

In the following years the foundation model architecture improved remarkably so that even short inputs led to even more accurate completion of sentences and output predictions (I6). This was possible through large amounts of training data and GPUs that made these technologies scalable and enabled the development of better foundation model to power GAI applications (I2, I6, I7, I10). Those foundation models could be used for a variety of tasks, with text-based tasks being particularly suitable (I6).

As noted by Interviewee 7, during this time, various companies noticed the enormous potential and started researching in this area. Between 2017 and 2020, the focus was presumably on foundation model performance (I7). Therefore, better foundation models were constantly developed and trained, while rarely incorporated into products. However, to impact industries, foundation models had to be integrated into value-creating products for customers. Interviewee 7 raises the suspicion, that these developments caused a split between academic researchers and the industrial world, as researchers continued to improve foundation models while commercial providers uploaded large amounts of data into the existing foundation models and sometimes achieved inaccurate but mostly useful results. One example for that is ChatGPT, which does not provide flawless accuracy but is already useful.

The release of ChatGPT from the lab into society enabled far-reaching diffusion among people, so they were able to brainstorm about beneficial use cases (I7). Since society became aware of the new technology, prejudices against AI also diminished at managerial levels, providing a major marketing campaign for AI in general (I10). Nevertheless, ChatGPT is used by the younger generation more than by older employees in many companies (I11). On the other hand, ChatGPT was also able to influence the opinions of skeptical employees and thus contribute to a change in AI usage (I11).

According to Interviewee 7 the developments showed that it was not only about the best foundation model which are becoming more and more similar in terms of performance among leading providers. It became more important to build an industrial machine with processes, data pipelines and GPU clusters for the required computing power (I7, I10). A leading example is OpenAI, that doesn't necessarily have the best foundation model and less proprietary data than big technology providers such as Amazon, Google and Microsoft, but early on established efficient processes and scaled faster (I7). In addition to

[1] Recapitulation: Foundation models are LLMs which power GAI applications.

establishing GPU clusters for the efficient training of foundation models, data pipelines are also crucial. However, many data source of OpenAI remains uncertain. Presumably, synthetic data and scraping huge amounts of data from the internet (e.g., YouTube) played a role in building a unique data set to train their foundation models (I7). This resulted in a time advantage of several months, which could not be completely minimized despite the large resources of big technology providers (I7). Nevertheless, OpenAI entered a partnership with Microsoft to receive various advantages as additional GPUs for training (I5).

Interviewee 5 remarks that Elon Musk's acquisition of Twitter blocked OpenAI's access to its data, probably preventing further model training. Thereby, Twitter's data was especially valuable for language processing, as the algorithm generates relevant retweets for individual feeds. Whereupon, xAI now uses this data to train its own foundation models (I5). Experts have also highlighted concerns about internal conflicting interests at Google regarding the development of GAI, as it changes the search market and their advertising business model (I5, I7). Consequently, their progress in developing foundation models has been delayed.

Strong competitors, besides OpenAI and Google, include Anthropic followed by open-source providers such as Cohere, Meta and Mistral (I7). Some of these foundation model providers purchased GPUs early on, or presumably also cross-trained their models with GPT-4. However, as it is a competitive market where high amounts of funding are required, with most GPUs already being stored in the USA (I7). This competitive market is also characterized by customer requirements, as they mostly want to use the best performing foundation model and will rather not fall back on inferior ones (I7, I8, I9). Thus, almost all layer providers and inhouse solutions build on one of the leading foundation models (I7). Besides the constantly improving model architecture market growth in the GAI sector was strongly influenced by the advancing developments in the required GPUs (Grand View Research, 2024), with NVIDIA having a market share of 92% in the GPU segment of data centers (IOT Analytics, 2023, p. 2).

The market research revealed following valuations over 1 billion USD for private companies that are developing leading foundation models: OpenAI 80 billion USD, xAI 18 billion USD, Anthropic 15 billion USD, Mistral 6.4 billion USD, Cohere 5.5 billion USD, Hugging Face 4.5 billion USD, Stability AI 1 billion USD, and Adept 1 billion USD (Dealroom, 2024). The key players and their respective funding are displayed in Figure 5. Other leading foundation models are also developed by various publicly traded corporates, which have allocated a part of their resources towards the development of foundation models. The enterprise valuation of Apple is the highest with 3.5 trillion USD, followed by Amazon with 2 trillion USD, Google with 1.9 trillion USD and Meta with 1.4 trillion USD (Dealroom, 2024). All these corporates are incorporated in the USA, while fifth and sixth place are taken by two Chinese companies, namely Alibaba (240.6 billion USD) and Baidu (25.1 billion USD). Figure 6 presents the enterprise valuation of the respective corporates. The respective data is based on the September 28, 2024.

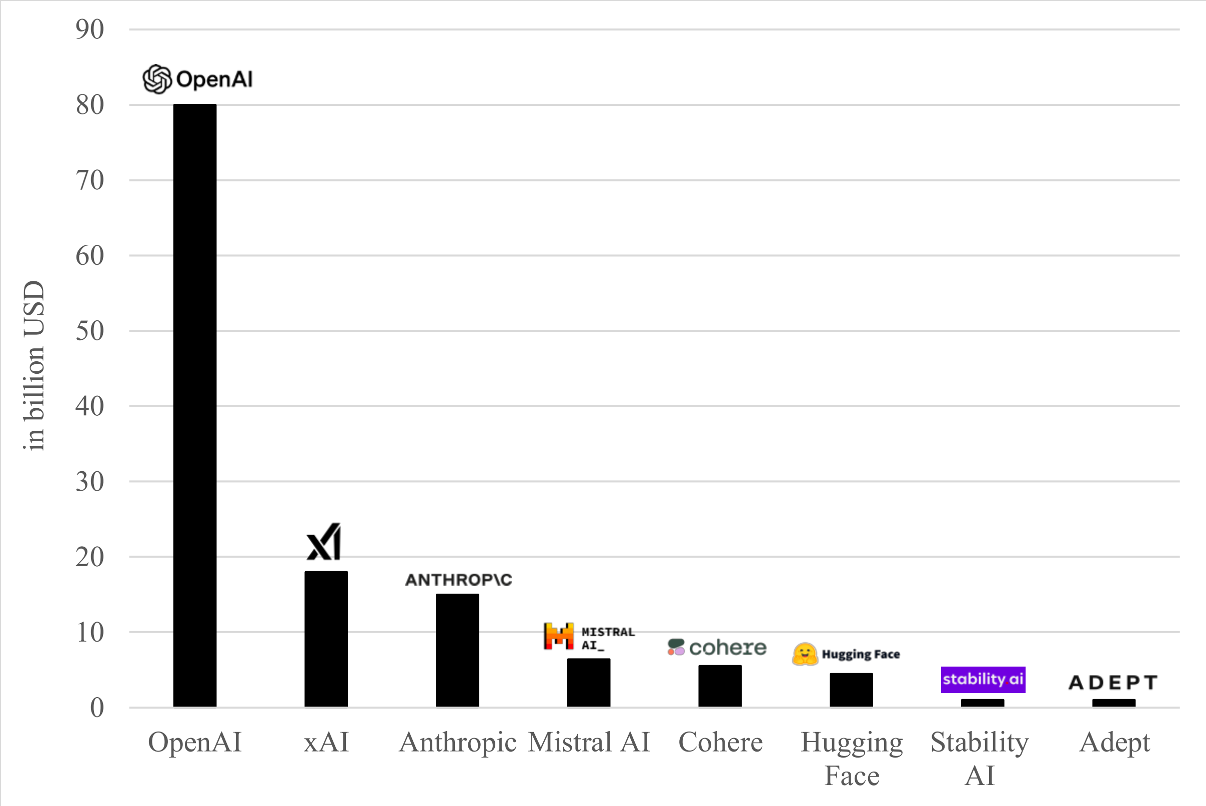

Figure 5: Valuation of private companies that develop foundation models based on Dealroom (2024)

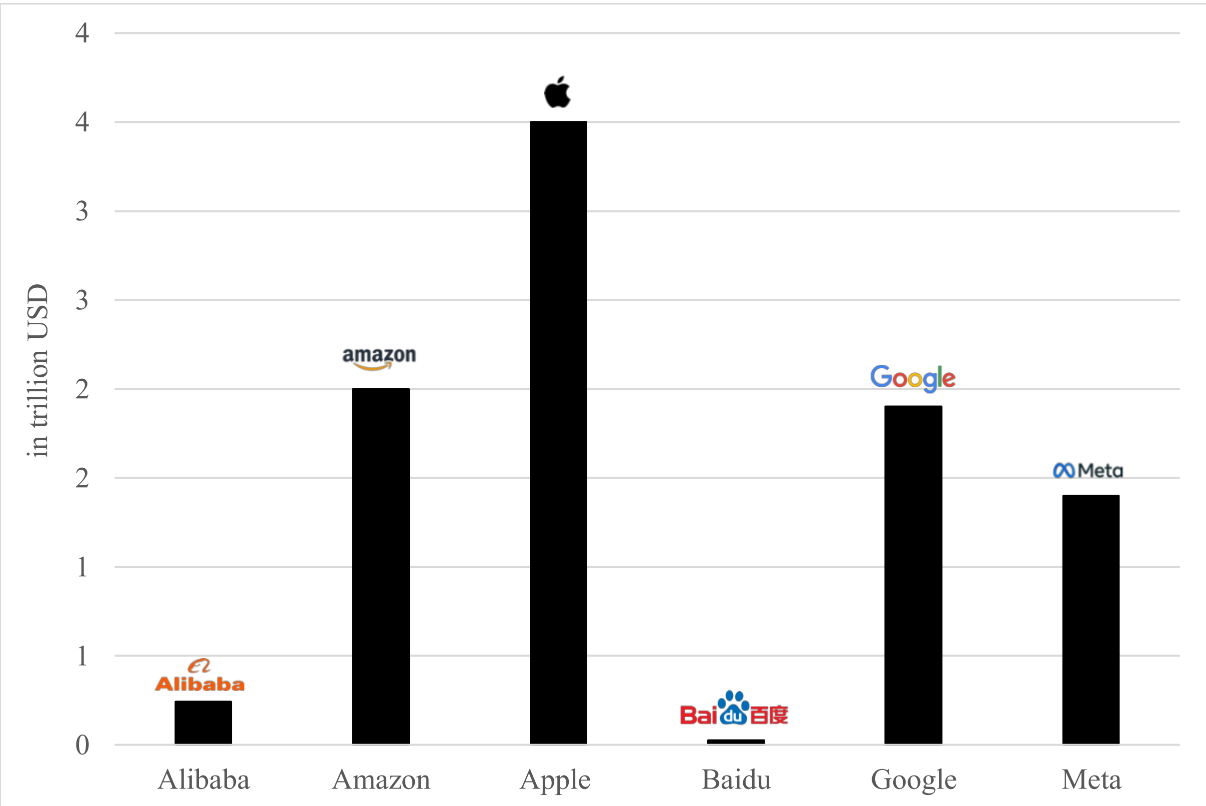

Figure 6: Valuation of public corporates that develop foundation models based on Dealroom (2024)

OpenAI and Microsoft have a far-reaching partnership and together hold a market share of 69% for platforms and foundation models (IOT Analytics, 2023, p. 2). Besides that, the market size of GAI worldwide amounted to 5.51 billion USD in 2020 and grew to 36.06 billion USD in 2024, while a market size of 365.1 billion USD is expected to be reached by 2030 (Statista, n.d.). In addition, the market size in the Europe amounted to 1.17 billion USD in 2020 grew to 11.22 billion USD in 2024 and will prospectively reach 110.8 billion USD in 2030. Over the entire period, Europe will account for 31% of the global GAI market. The compound annual growth rate (CAGR) amounts to 46.47% from 2024 to 2020. The corresponding market size developments over time are illustrated in Figure 7.

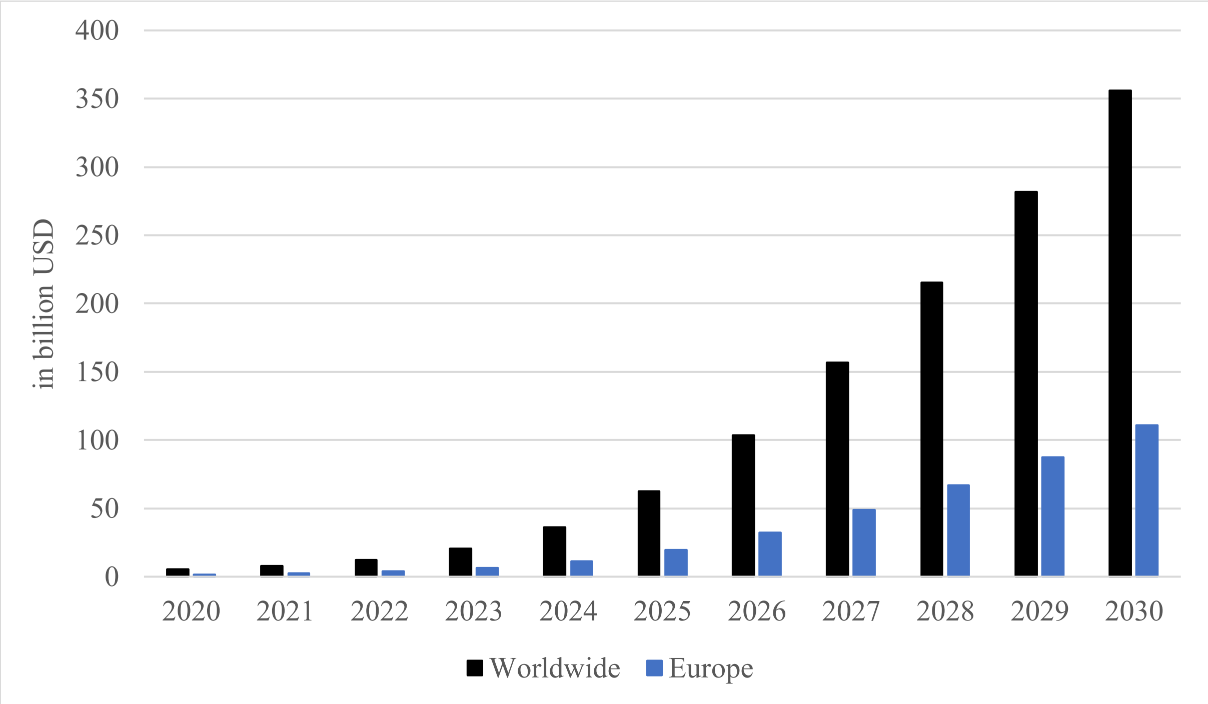

Figure 7: Development of the GAI market size based on Statista (n.d.)

As estimated market sizes of GAI fluctuate strongly, they result in substantially different forecasts. Appendix B highlights these fluctuations while also underscoring the consensus on remarkable market growth over the next years.

Another interesting insight is the change in market size over the next years. The annual change in market size in 2021 amounted to 40% and grew to 76% by 2024 (Statista, n.d.). From that point onward, the annual change in market size will gradually decline, reaching approximately 26% by 2030. It is important to emphasize that this still represents strong growth in terms of market size for each year, however the market is expected to expand at a slower pace from 2024 onwards compared to the earlier acceleration phase. The graphic suggests that a period of hype significantly boosted the GAI market between 2021 and 2024, providing a strong initial surge.

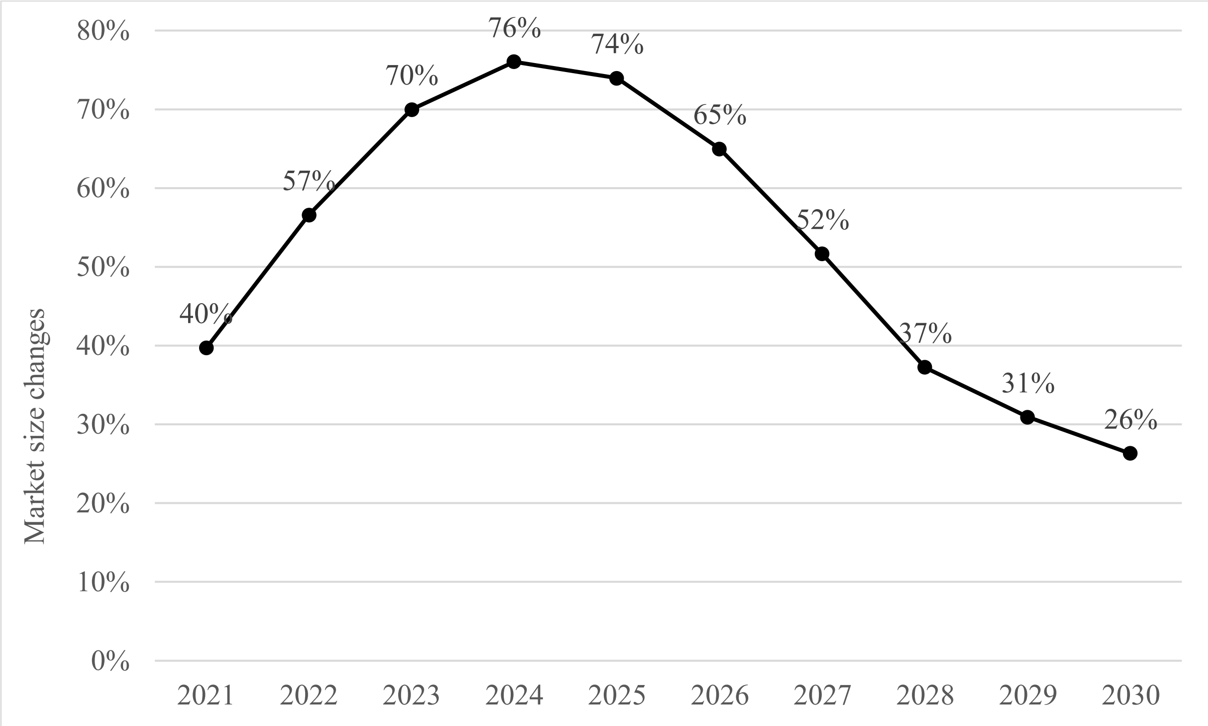

Figure 8: Market size changes for GAI based on Statista (n.d.)

Future developments are difficult to predict, although experts assume that in the next decade robotics will become increasingly important alongside GAI (I6, I7). According to OpenAI's five development stages (Forbes, 2024), the Chatbot phase has already been reached, with the next stages focusing on advancing reasoning capabilities. Following this, the progression includes innovators, agents, and eventually organizations. At the agent stage, systems should operate autonomously, introducing new processes to optimize or, where necessary, replace existing workflows through improved logic and efficiency.

This will be the next step toward General AI (I7). It is also possible that the current transformer architecture can be the foundation for this General AI or used to develop better architectures by itself in the future (I7). On the other hand, we could also find ourselves in a local minimum with the current models and be further away from General AI than many assume, as text-based foundation models alone cannot map the complexity of the environment for robotics, as they were not trained with complex real-world modalities (I5, I6, I7). Therefore, foundation models can be a useful as an additional building block for some use cases as robotics, but are not able to represent such a complex system completely.

While some individuals overestimate the potential of GAI, others regard it as less useful (I6). Following that, Interviewee 4 suggested that the abilities of AI are often overhyped in the short term, yet long-term impacts tend to be underestimated (I4).

5. Identification of GAI Capabilities

5.1. General Perspectives on GAI

A relevant part of the expert interviews consisted of gathering industry perceptions on the status of GAI. These perspectives reveal expert sentiments towards current and future developments. In addition, they feature technical aspects, viewpoints on monetization and open-source foundation models. Risks, challenges and limitations of GAI are also addressed, followed by differences to other AI technologies to understand where and how GAI can be used. GAI is not a completely new phenomenon, as programmers and data scientist already dealt with the basic concepts from machine learning and statistics for many years (I2).

However, it is not trivial to build a GAI application based on a foundation model (I8). The development from proof-of-concept (POC) to a scalable product is often underestimated and requires strong expertise in data management, privacy and security (I8). The integration of GAI into software products increases the complexity of the entire software architecture, as infrastructure, data security and disaster recovery must be covered as well (I8). It is important to monitor and measure what goes in and out of these systems (I8). Anything regarding larger customers that exceeds individual use cases must be scalable and requires not only a foundation model but also a platform that offers general services and infrastructures (I3, I4).

Furthermore, GAI is often used to promote software products and integrated where it doesn't necessarily make sense (I3). This does not solve an actual pain point, but rather uses the hype to benefit from it. However, the phase of "happy hacking" in which employees played around with GAI seems to be over (I3). The hype cycle was relatively short-lived and is also slowly approaching the end, especially in the B2C segment (I3, I10). The current focus lies more on evaluating in which segments it can provide real value (I10). Jeff Bezos suggested that LLMs, in their current form, should not be viewed as inventions but as discoveries, highlighting the need to reflect on how we harness and apply this newly uncovered technology (I10). At the same time, GAI proved that it is not vaporware, but a useful technology (I3). GAI can be used in various industries, whereby not every company will build an inhouse solution for its specific use case but use layer providers for certain industries (I4). This also includes very large customer groups such as governments (I4).

So far, the interaction between humans and computers changed to text-based communication (I1). This leads to the elimination of software tutorials through self-explanatory products with better user experience (I1). Many repetitive and time-consuming tasks can be automated (I8), which will increase productivity and counteract the shortage of skilled labor (I4). Tasks can be solved more efficiently than before, or further training opportunities can be created for employees (I4, I8). However, this development can lead to social distortions that cannot yet be accurately predicted, while past disruptions demonstrate that output increased, and new professions arise (I4). At the same time, AI cannot be panacea for all business challenges, considering the digitalization delays in many companies.

When analyzing the business models of foundation models on the other hand, it becomes evident that many closed-source providers charge users to access foundation models based on usage volume via APIs (I1). In usage-based pricing, typically packages that bill per thousand requests at a set rate are offered (I1). While this is an operationally appealing business model, its long-term sustainability remains to be validated (I1). How foundation model providers will generate profit in the long-term remains uncertain, as a high level of expenditure is required to develop state-of-the-art foundation models (I4). The extent to which these costs can be recovered from the use of the foundation model to break even before a new one has to be trained in a few months is very controversial (I4). Additionally, due to the relatively high API costs at this moment, it is likely that there will be a demand to decrease these costs in the near future (I4). Closing this gap between expenditures and revenue will be a challenge for foundation model providers within their business models (I4). Henceforth, there exists no consensus on which business model will establish itself as financially viable in the long-term (I3).

To generate revenue companies must solve a problem for their customer, so the business model for layer providers remains the same as for software-as-a-service companies, even if usage-based-billing may be an additional component (I3, I4). Other established software providers will integrate GAI features into their existing software solution and continue with a software-as-a-service business model (I3). However, even if GAI solves business problems, customers must quantify the value it provides (I11). Usage-based pricing can become costly, especially on a large scale when expensive models are used for a high number of simple queries (I11). At this point, it is necessary to compare the costs and the needed performance with the value created by GAI (I11).

Simultaneously, there is an opposing trend as many foundation models are available for free, while monetization is realized through resulting services surrounding it, such as fine-tuning for special use cases (I1). This idea of open-source foundation models plays into the trend of open-source projects, which are generally a very important building block in business software (I1). The users of open-source benefit from the ideas of peers who work on the same challenges, usually resulting in comprehensive learning and a better solution for all (I1). Open-source has the advantage of transparency, as users can trace back what code was used (I2, I8). However, participants with certain interests can shift open-source projects into their desired directions or contaminate the architecture (I2).

Nevertheless, the use of open-source has other advantages, as such models are often smaller and can be used locally without a data loss to the large providers of closed-sourced models in the USA (I8). Furthermore, this advantage must be counterbalanced with less cross-application options for these open-source alternatives compared to more generalist closed-source models (I8). Thus, every company must compare benchmarks and decide which LLM is best suited for their requirements and use cases (I9, I11). Open-source models are useful for customers with very specific use cases looking for an optimized model with higher speed, accuracy, and transparency while using confidential data (I8, I11). While open-source models seem cheaper as no usage-based pricing is applied, they lead to increased costs of implementation, infrastructure set-up, and maintenance on the other hand (I11). Additionally, users of open-source models cannot be sure whether optimized models will be released in the future (I11). This is particularly important as companies need to be future-oriented, because new foundation models with improved performance are available every few months (I11). That's why it can be an advantage to choose a foundation model provider that will keep up with these developments (I11). This also includes assessing whether the company has the human resources to implement open-source models (e.g., skilled software engineers) or whether it should fall back on services from commercial providers (I11).

If small open-source models are chosen, they can be fine-tuned for specific tasks (I8). Some layer providers of GAI applications already use several foundation models in the background (I8). Sometimes even simultaneously, which has the advantage that the GAI application still runs despite the crash of a foundation model and increases independence from one provider (I8). Conversely, if a company relies on a single foundation model, it can save time and costs due to its familiarity with the system (I9). When companies update or introduce new foundation models, users also must ensure that the new model provides equal or better performance (I11). That cannot be deduced from benchmarking alone, as it always must be tested for each specific use case setting (I11).

On the other hand, the use of GAI also harbors various risks and challenges (I3). Major risks arise in the context of social media, as misinformation and deep fakes can be created effortlessly with GAI tools (I2). Furthermore, unrestricted open-source foundation models that respond to queries without filtering can potentially provide dangerous information, such as instructions to create bioweapons (I5). Besides that, criminals can enhance their phishing and spam attacks through GAI (I5). Another risk is mass surveillance, as comprehensive monitoring could become viable with the ability to harness large volumes of unstructured data (I5). Although it seems currently too expensive, it may become feasible in the future (I5).

In the B2B context, many companies can fall for consulting boutiques with strong sales departments that demonstrate spectacular POC and showcases (I3). However, if the company does not have employees with the necessary technical skills to maintain inhouse solution appropriately, problems may arise afterward. That's why companies that want to use GAI should already have a certain level of digital maturity, otherwise frustration during the use of GAI may arise (I3). Therefore, the digital maturity level of companies is essential, as if it is too low, GAI initiatives will most likely fail (I3). At the same time, C-level executives of SMEs report that in recent years, leading consultancies repeatedly recommended outsourcing IT services (I3). Hence, many SMEs lack digital and technical know-how, representing a major risk for future technological usage (I3, I4). In contrast to that, there are start-ups, which may have software engineers skilled in GAI, but operate negligently, because they underestimate underlying implications regarding data protection or compliance (I8).

The resulting challenges of GAI are diverse, as it is unclear how decisions of companies and authorities will be influenced by non-explainable AI recommendations, which may be based on biased, opaque, and limited training data (I5, I10). For example, if a certain appearance is associated with negative behavioral patterns by the AI, or certain appearances are not recognized as they were not observed in the training data at all (I10). Thus, companies and authorities must reflect about the underlying data and models to interpret the recommendations of such systems appropriately (I10). This is one reason for the EU AI Act, which will regulate certain use cases and could lead to special market characteristics in the EU with far fewer models available (I11).

Another big challenge is the energy consumption caused by all the foundation models, leading to big technology providers as Google to adjust their targets for net-zero (I3). Furthermore, the CEO of OpenAI expressed the idea of using nuclear power plants to provide additional energy for model training. Considering the total number of GPUs already sold by Nvidia, the magnitude of the future energy consumption becomes evident. This energy consumption will be reflected in a company's overall CO2 footprint for information technology and will probably require new regulations.

In addition to certain risks and challenges regarding GAI, the abilities of this technology are often overestimated by many, who do not understand the underlying statistical principles (I10). Hereby, the models are not explicitly useful for previously unobserved action predictions, because they are inherently limited by the provided training data (I6). So, they can interpolate, but they are not able to extrapolate beyond its data. As a result, LLMs would struggle as soon as there are inputs which were not observed in the past, but since they have been trained with the entire internet, their knowledge base is extensive (I6). Despite that, current foundation models are not capable of lateral thinking, which consists of independently producing a creative combination of two contexts (I5). Thus, they are efficient in certain tasks, cannot create truly new solutions (I5). Another aspect which is often overestimated is the ability to write programming code. Even if recurring tasks such as programming websites or apps can be simplified. However, as soon as programmers need to tackle complex new problems in software engineering, incorporating replicability, extensibility, understandability, and scalability, LLMs cannot provide new algorithms (I6).

Generally, it is useful for less qualitative repetitive simple tasks to save time (I10), whereupon it is crucial to make users aware that the output generated may not always be accurate (I11). This can be caused by data inconsistencies (I11) or hallucinations, a major problem of GAI (Maleki et al., 2024, p. 127). Hallucinations occur if output appears comprehensible but is incorrect, and not connected to the input or any information at all (Wiseman et al., 2017, p. 9). So, for complex technical tasks users need a baseline of knowledge to validate the output, suggesting that specialist will not be completely replaced by GAI in the upcoming years (I10). Understanding how LLMs work and where their limits lie makes it clear in which domains it can or cannot be used in a performance-enhancing way (I3, I6). In addition to that, it is important to remain aware of other AI technologies and their abilities. The current hype surrounding GAI often leads to the assumption that it can solve almost any pain point, when in fact, more suitable methods exist for certain use cases (I3, I9). There is some confusion around GAI and where other AI technologies tie in, which is why education on GAI should always cover further aspects of AI to create holistic knowledge (I6, I9, I10). For example, predictive maintenance methods are more efficient in the evaluation of machine sensor data and have been used by many companies for years to reduce costs in manufacturing (I2). Also, computer vision models for video and image detection are methods that can save costs in manufacturing (I2). Further examples that can solve various use cases more effectively than GAI include robotics, AI recommendation systems for pricing, and established mathematical models (I2, I3). Many of these methods have been used successfully for years and were developed inhouse by advanced industrial companies or with the help of external partners (I2, I8, I10). There is already more consolidation for such technologies than in the dynamic GAI market, where new models are released every few months (I11). Thus, GAI can be seen as a valuable technology for specific use cases, but beside that, it can serve also as an additional layer that enhances the interaction between existing AI technologies and users (I9). As a result of the unifying transformer architecture, we can expect to see several future use cases that are more integrated (I10).

In summary, companies must cultivate and acquire *fundamental understanding and continuous learning* regarding AI, particularly GAI, to keep pace with the dynamic advancements in the field (I7, I11). It is also essential to evaluate the distinctions between closed-source and open-source foundation models, while considering the integration of established and additional software components (I1, I3, I4, I8). Thus, organizations need to develop a clear understanding of the technological principles and limitations of GAI. While some companies already possess this capability, it is important for others to develop it over time, which is possible through upskilling efforts, hiring new employees with the respective expertise, or building a network of potential technology partners (I1, I7, I9, I11). These factors form the fundamental understanding and continuous learning that precede any concrete use case implementation (I3, I8, I10, I11).

5.2. Identification of Compatible Use Cases

Once current developments of GAI, limitations and differences to other AI technologies are investigated, it is particularly relevant to identify which compatible use cases exist for GAI. Based on fundamental understanding and continuous learning, companies need the capability to assess *use case compatibility*. This enables companies to accurately identify which use cases (problems) can be addressed by which GAI application (solutions). This requires a clear understanding of the technological limitations and familiarity with other AI systems that may be better suited for certain tasks (I3, I8, I10, I11). Additionally, knowledge of existing use cases, whether used by competitors or available from off-the-shelf providers, is valuable (I9). It’s also essential to assess the implementation costs and determine the return-on-investment (ROI) regarding a specific use case (I3, I4). Following that, the use cases can be evaluated to create a roadmap for implementation (I2, I9, I11). Defining the specific requirements and pain points is crucial, along with deep knowledge of the industry, legacy systems, and internal data management processes (I4, I10, I11). Subsequently, a decision must be made for each use case on whether to buy or build the GAI application (I2, I9). The resulting GAI capability is referred to as use case compatibility in the further course of the thesis, while the building blocks for this capability are summarized in Table 2.

Table 2: The GAI capability of use case compatibility with the respective building blocks

Description of the building blocks for use case compatibility	**Citations**
Fundamental understanding and continuous learning: Acquire and cultivate general expertise in AI (particularly GAI) through upskilling or hiring new employees. Understand the technological principles, challenges, and limitations of GAI compared to other AI technologies to decrease false expectations. Follow continuously new developments and build a network to exchange with industry experts, initiatives, legal advisors, decision-makers, researchers, competitors and technology providers.	I1, I2, I3, I4, I5, I7, I8, I9, I10, I11
Identify portfolio: Identify which use cases exist and if they can be solved with GAI. Develop a use case portfolio (long list of use cases). Examine what others are implementing and, if necessary, get support from internal or external experts to identify more use cases. Describe the use cases in detail and translate broad concepts into specified requirements, pain points, and milestones to avoid moving targets.	I9, I10, I11
Prioritized roadmap: Evaluate use cases by quantifying their ROI (calculate the total costs of infrastructure, foundation model, licenses, usage-based pricing, maintenance and compare them to the efficiency gains). Rank all use cases according to their quantified value (ROI) or other criteria (e.g., ease of implementation) and create a prioritized roadmap with waves of implementation over the next two to three years.	I4, I9, I11
Make-or-buy decisions: Understand the differences between off-the-shelf GAI applications and inhouse solutions and consider the underlying foundation models. Decide whether you build an inhouse solution or buy an off-the-shelf GAI application for a use case, and to which extent you want to customize it. Consider proprietary resources, challenges, risks, total costs, requirements and objectives. If building an inhouse solution is the preferred choice, evaluate advantages and disadvantaged of closed-source and open-source foundation models.	I2, I3, I4, I7, I8, I9, I11

Identify the use case portfolio and roadmap

Companies already use a wide range of GAI applications, through employees who access foundation models via application layers as ChatGPT or established software products which were extended by GAI features (I2, I3, I6). For example, Microsoft has strategically integrated its GAI copilot into its product ecosystem so that many companies already using the basic products will operate the copilot as well, making it difficult for competitors to reach the same number of users (I2). Besides established software solutions that integrate GAI features into their existing offers, there are also layer providers that address industry-specific use cases with off-the-shelf GAI applications and are used by certain divisions (I3). As a result, customers may not always be aware of which foundation models are powering the application and features behind the software (I9).

GAI is applicable across departments, professions and industries, as it makes many repetitive and time-consuming day-to-day tasks easier (I1, I8). GAI is primarily used for qualitative data, but since it is not always accurate, a human-in-the-loop is often required to review and improve the initial suggestions (I9). Some industry experts compare using ChatGPT or Microsoft Copilot to hiring interns who prepare first drafts for a presentation, code documentations or texts (I2, I6).

Typical use cases for GAI are magic buttons to summarize texts or gain important key insights (I1). Besides that, familiarization with new software is considerably simplified (I1). It is strongly used in marketing for text generation, advertising, creating presentations, submitting initial proposals and generating images or videos (I2, I3, I9, I11). Moreover, warehouse operators can manage new product requests through GAI powered chat interfaces, which, with appropriate access rights to other systems, streamline customer communication and enable direct product bookings (I3). Other examples include AI agents for copy and paste tasks from a database to an ERP-system (I5). Furthermore, human resource departments can benefit through preparing more efficiently for interviews or directly extract skills from CVs, while software developers will use code suggestions or documentation tools (I1, I8, I11).

But there are even more areas of application, such as developing a program plan, analyzing and evaluating overlaps in meetings, or basic calculations (I2). One industry expert reported that a company had to change the shirt color, a person and a vehicle in a company photo, which was done conveniently within seconds through GAI (I2). Further use cases include knowledge management systems for information retrieval, which can be fed with qualitative and quantitative company data and then report key performance indicators (I9, I11).

GAI can then be used as a chatbot for internal or external frequently asked questions, as it serves as a knowledge management system (I10). For external use cases, it can also be used as a substitute for a support hotline or call center (I10). Customer services are generally common use case that are often realized through personalization and automation in communication (I4, I9). One example of such a use case is a customer service agent for doctors that schedules appointments automatically via the hotline, freeing up employees for other tasks (I4). This GAI application must integrate with doctors' calendars and needs the ability to read or update customer information in the system (I4). This is especially relevant for senior citizens, who are more likely to call the doctor than to use software interfaces (I4).

GAI applications for sales departments prepare quotations, enabling SMEs to create and propose quotations more quickly (I4). In the USA, for instance, GAI applications for tradespeople exist to create a range of offers with varying levels of quality and service for customers (I4). Another use case arises in the real estate market, where asset managers must analyze huge numbers of sustainability reports so that GAI can process them faster and provide recommendations (I4, I8). Moreover, the real estate industry involves substantial regulatory documentation, which small companies often struggle to manage, leading them to rely on GAI applications (I4). Likewise, there are an increasing amount of GAI applications to support due diligence during mergers and acquisitions (I8). The management and analysis of contracts is also a huge pain point for many companies and brings potential for automation through GAI (I4). Further use cases include procurement, finance, and market analysis bots (I9).

GAI layer providers should develop applications tailored to industry-specific use cases (I4). GAI applications can be tailored to text-based processes of specific industries, even if market sizes are limited (I4). Alternatively, they can address broader, cross-industry tasks, such as requirement management for a variety of regulatory affairs across industries (I4). Most of these use cases apply to white-collar professions, although there are some applications for text-based documentation in the blue-collar segment (I2, I4). In the medium term, best practices and certain foundation models will become established for specific use cases (I1).

Familiarizing with existing use cases for GAI helps to create a long list of potential use cases for the own company (I9, I11). It is also recommended to examine what competitors are implementing and seek external or internal support if needed (I9). A comprehensive long list, also called use case portfolio should be developed, including all company applications that can be automated through GAI (I9, I11).

One of the key factors for developing GAI solutions is the companies' ability to translate broad concepts into specified requirements (I10). Without this, objectives represent a moving target, that can lead to substantial challenges. Providers must also build trust, as there may occur deviations between the management ideas and technical possibilities. Furthermore, it is useful if the company already identified specific pain points, that it wants to solve with GAI, as it can be difficult for providers to develop a solution if the company simply wants to try something with GAI. Companies often take a long time with their decision-making processes yet expect quick results afterward. This presents a challenge for providers implementing inhouse solutions, as they must manage both timelines effectively.

Therefore, the use case portfolio should be extended by the specific requirements, including pain points, feedback loops and milestones (I9). The more detailed a use case is described, the better requirements, pain points, and milestones, can be distilled to prevent moving targets (I0). Henceforth, *identifying the portfolio* of use cases in detail is an essential building block of use case compatibility.

Using GAI can leverage efficiencies in corporates considerably, but it can be expensive as well, while the impact is often difficult to quantify (I3). A high usage volume for GAI often results in a positive ROI (I4). A positive ROI indicates that the value generated through cost savings exceeds the expenses incurred in acquiring the technology. Such use cases often exist in customer service, marketing and sales

(I4). For example, support hotlines or call centers for customer services are strongly budget-driven, so any automation can be directly translated into fewer call center agents and thus has a quantifiable impact on ROI (I4).

Most companies expect to achieve a positive ROI in the short to medium term, while sometimes the focus is not on financial expectations, but rather on creating momentum through a flagship marketing project (I9). Thus, it is important for providers to identify the objectives of every customer before implementing GAI (I9). Many SMEs discover the limitations of GAI during implementation and usage, often realizing that achieving a positive ROI is not feasible (I4). Unlike blockchain, however, there are already practical use cases for GAI (I4).

Besides that, the implementation costs should be compared to the efficiency gains to calculate a potential ROI for each use case (I3, I4). Following that, a roadmap for the realization prioritizes use cases in a funnel-shaped pyramid according to certain criteria (e.g., ease of implementation or ROI) and defines three waves for the implementation over the next two to three years (I2, I9, I11). Therefore, this process represents the building block of the *prioritized roadmap*, which is also part of the use case compatibility.

Moreover, providers must be familiar with processes and historically grown legacy software, which makes it very challenging to connect new GAI applications (I4). Such GAI applications become powerful when they have access rights to other systems and databases so that they can read and change data (I4). These API connections to other systems are particularly complex, but there are workarounds that reduce dependence on legacy systems (I4). For instance, instead of developing a complete API for an outdated warehouse management system, GAI software can be linked to a single storage bin, representing an entirely new virtual warehouse management system. This allows businesses to avoid rebuilding their logic within old systems and instead implement small workarounds that simplify operations (I4). For all the use cases users cannot expect seamless accuracy, as a human-in-the-loop is required, or safety systems need to be installed (e.g. if the caller argues with the call center bot it is connected to a real human) (I4, I9). The lack of accuracy in GAI needs to be considered for each use case (e.g., in the legal field) as it makes it more challenging or nearly unusable (I5).

GAI providers must have industry-specific knowledge for certain use cases, as a deep understanding of the requirements is often necessary (I4). For example, in the USA, there are GAI tools for lawyers that search for precedents relevant to current cases (I4). However, some precedents are stored in non-digitized databases, meaning the software would lack access to all the necessary data and be unable to process precedents accurately (I4). To develop effective GAI solutions, it is crucial to understand the industry and incorporate older databases as well (I4).

At the same time, companies need to control their data appropriately, as otherwise different versions of documents (e.g., operating manuals) can lead to confusion and misinformation (I11). For example, a potential customer asks a chatbot about a product feature and the GAI confuses it with an outdated offer

(I11). This is why companies need to control and monitor data sources, inputs and outputs (I11). In addition, there are the employees who do not want to use such GAI applications, so it is a crucial point to introduce employees' step by step to gain acceptance (I9). Simultaneously, there often exists a conflict of interest, as management naturally wants to increase efficiencies through new technologies and reduce costs, which can lead to skepticism and fears among employees (I10). However, the greatest difficulty with many use cases lies in moving from a single pilot to an operating model which leverages the value of GAI on a large scale (I3, I8, I9). This is where many users of GAI get stuck and fail to take this important step (I9).

Make-or-buy decision

Interviewee 2 highlights "...that in the end, the same decision applies here as always in business, make or buy with intermediate steps, because you buy many things anyway and then adapt them for your own use. Very few companies now build their own LLMs completely". This *make-or-buy decision* involves choosing whether to use an off-the-shelf GAI application or develop a customized inhouse solution based on foundation models and depends on several factors such as ease of implementation or internal company resources (I9).

It does not make sense for every company to develop an inhouse solution, as off-the-shelf GAI applications can also solve certain pain points successfully (I2). Especially, smaller companies often turn to GAI applications of external providers and customize them to a certain extent (I3, I8). However, underlying foundation models should be considered as well (I8, I11).

For inhouse solutions, companies must decide between different foundation models (open- or closed-source) based on their requirements (I9, I11). Large corporates develop inhouse solutions especially for core business use cases to avoid data loss, increase control, and provide higher accuracy (I2, I3). It depends on the company's resources whether these inhouse solutions are built independently or external consulting and implementation partners are involved (I2, I8, I10). If companies decide to build an inhouse solution with external partners, it is important to involve partners who have a track record and can execute the respective project successfully (I4). While such proprietary inhouse solutions are usually more expensive, they can ensure better data security (I2). Finally, such inhouse solutions can not only be used to increase internal productivity but can also be integrated into the software and hardware products that the company sells (I3, I10).

5.3. Effective Organizational Design

The competitive advantage of a company relies heavily on their organizational processes (Teece et al., 1997, p. 518). This is even more relevant for large companies, where effective processes scale (I6, I10, I11). The management of a company must sense opportunities and integrate them through routines into the firm's activities to increase efficiency and build strategic advantages (Teece et al., 1997, p. 519).

The success of a company depends not only on a superior product or service, but also on the underlying industrial machine, i.e. the processes and resources that drive the organization (I7). Also, in terms of GAI, companies can create new entities to achieve value through the technology (positive ROI), while positioning them ideally for future applications. Thus, the analysis focuses on how this design for GAI usage is shaped, exploring strategies, coordination efforts, and insights for leveraging learnings through new entities. According to Interviewee 11 using GAI successfully "… is not just technological, but there's also the whole organizational aspect, including further training […] as the most important factors". Thus, the following chapter distills the capability of *organizational design*, which consists of various building blocks reaching from C-level support, additional teams, to processes for implementing, scaling and managing GAI applications. All these building blocks are summarized in Table 3.

Table 3: The GAI capability of organizational design with the respective building blocks

Description of the building blocks for organizational design	**Citations**
C-level commitment: The combination of technology and effective organization is a key factor for successful usage of GAI and can create far-reaching efficiency gains for the respective company. C-level executives and the board should have courage and prioritize AI while providing the resources to enable it. Trying different solutions through a trial-and-error approach should be encouraged. C-level executives and the board should remove initial barriers that require top-down decisions. This includes formulating a strategy, that considers ethical, compliance and technological feasibility. Defining such a strategy can incorporate internal or external domain experts. The strategy needs to be aligned with the use cases (see use case portfolio) and subsequent implementation of these use cases (according to the roadmap) must fulfill the milestones to achieve the overall strategy.	I1, I2, I3, I5, I9, I11
Core team: Establish a core team as the central point of contact for all topics regarding (G)AI, which directly reports to the board and receives instructions from the C-level. It is superordinate to all departments and cross-functional, as it pools domain experts from various departments who dedicate a certain amount of their capacity to the core team. Completely organized along the use cases, it serves as a catalyst for top-down and bottom-up innovation regarding GAI use cases. The core team evaluates use cases, leveraging various perspectives (strategy, compliance, software engineering). After a decision is made, tasks are cascaded throughout the organization to fulfill concrete steps of the roadmap.	I1, I2, I3, I7, I9,
Implementation: Enable departments to implement use cases independently with internal resources. Establish operational support through a technical inhouse consulting team of data scientists, if needed. As these additional services affect the budget of the respective department, they should also have the alternative of choosing external	I1, I2, I3, I9

consulting partners. Continuous upskilling increases the abilities for future implementations.	
Digital Hub: Form a digital hub, initially managed by the core team, to provide an education and learning platform where experts and users meet. This can ensure scalability, through answering frequently asked questions via digital content (videos, intranet posts, articles, blogs, podcasts, tutorials). Simultaneously, it guides employees how to use GAI, offers more acceptance and upskilling. The digital hub can also serve as a community or network, where like-minded users can interact and exchange best practices to promote grassroot movements and bottom-up innovation.	I1, I2, I3, I9, I11
Pilot trap avoidance: To ensure successful realization of the use cases a scalable structure must be defined from the beginning promoting company-wide rollout capabilities. Transitioning from an initial pilot to a company-wide rollout is a challenge. Therefore, building a scalable framework and identifying high impact use cases is crucial. The development of a specific use case requires strategic and technical skills and test-and-learn loops for iterative improvement. The development follows the rapid-prototyping approach with multiple feedback loops to build user trust through early results. Furthermore, it ensures continuous evaluation of new features, requirements, and pain points, legacy systems, and internal data management. With that approach, the prototype develops through extensive testing prior to launch into a final product. Simultaneously, additional use cases can be implemented. During this, developers must manage timelines successfully, as often decision-making processes take time and results are expected quickly.	I3, I8, I9, I10, I11
Product Management: Assign dedicated product managers to launch, monitor and maintain GAI applications. Effective product management is a key factor to ensure the ongoing success of the respective applications. Product managers control the interface of underlying technology and the user facing product. They must minimize deviations between performance and user expectations. As GAI requires additional knowledge that not all existing product managers possess, upskilling existing and hiring new product managers is essential.	I7, I9

C-level commitment

AI in general, and especially GAI are integral part of corporate digitalization strategies and currently represent the biggest strategic goals of many companies around the world, which is reflected in various conferences and roadmaps developed specifically for this purpose (I1, I3). GAI should always be conceived in the context of an overall AI strategy (I9, I11). It is a relevant topic that C-level executives cannot ignore, requiring fundamental knowledge and the constant development of additional expertise in this area (I1, I2, I3, I9). However, it is not purposeful to tackle this topic half-heartedly, as there must be a strong commitment to handle such technologies successfully (I1, I11). This includes a culture of openness towards innovation, so that C-level executives and board members are interested in using new technologies (I9). C-level executives need to be confident with introducing GAI, as they are often incentivized not to take such risks, as a mistake could affect their career negatively (I5). Openness towards new technologies should be anchored in the corporate culture, there should be an understanding that companies need to experiment and develop appropriate solutions through a process of trial and error (I2, I9). The same applied in the past for other major technological shifts, as cloud computing or the internet (I1, I7). If such disruptive technologies are handled effectively, they can bring great value to the company (I1).

However, companies must reflect about what value they want to achieve and how to measure it (I9). The strategy must be derived from the vision and incorporate desired targets (I9). The strategy making process includes initial top-down decisions from the C-level to remove all general barriers and lay the foundation for all further GAI developments (e.g., general agreements with technology partners) (I3). For that, various domain experts can be involved, such as the operational IT, data science, strategic communication, compliance and finance departments (I2). Above the desired outcome, it is essential to determine what is legally and ethically appropriate and what is technologically feasible (I2). These recommendations can support C-level executives to make their decisions considering important enablers, while removing initial barriers (I2). Fundamental decisions for a general setting can also be brainstormed with external partners, including market and competitor analyses (I9). The strategy must then be executed through use cases and their respective milestones (I9). Therefore, the use case portfolio and the roadmap should be aligned with the strategy, to achieve the desired strategic outcome in various waves of implementation over the next two to three years (I2, I9, I11). The C-level and board members should be transparent with their plans to provide effective change management (I9). Hence, *C-level commitment* is a fundamental building block to enable an effective organizational design.

New organizational entities

Organizations need to be extended by new entities to create an improved organizational design for using GAI at scale. First an interdisciplinary and cross-functional *core team* with domain experts must be established to streamline bottom-up and top-down innovations regarding GAI alongside the use cases, decide which use cases are feasible, and cascade the roadmap to the respective departments (I2). Those departments must demonstrate a certain independence for *implementation* of the assigned use cases (I2). If needed, internal or external consulting teams can support the implementation (I3, I9). Additionally, a *digital hub* should provide general educational material and communities to exchange with peers (I1, I11). Following that, the scalability of use cases must be ensured from the beginning to ensure *pilot trap avoidance*, while professionals for *product management* must be trained or hired to manage the GAI applications after the product launch successfully (I3, I8, I9). These new entities and their integration with existing entities are illustrated in the organizational chart of Figure 9. Furthermore, the main tasks of each entity are also specified.[2]

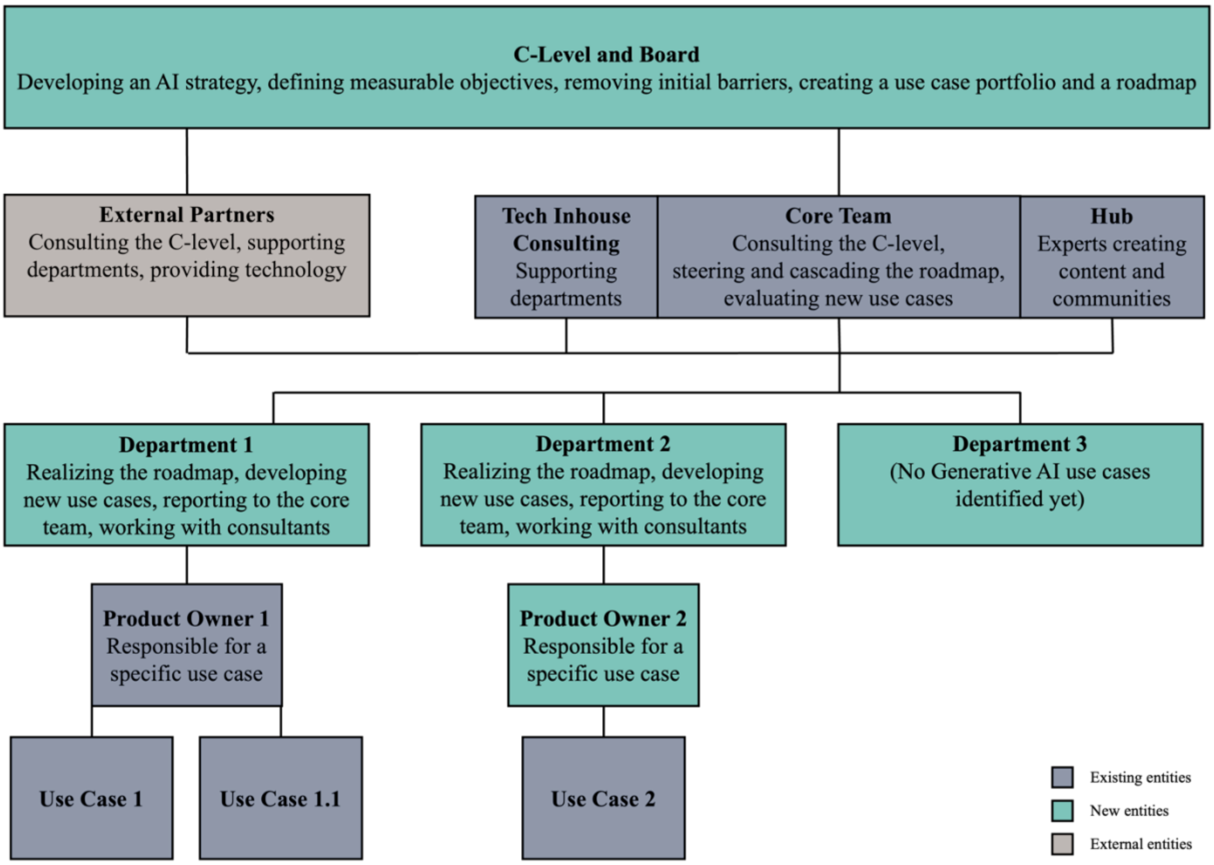

Figure 9: Simplified organizational chart for the integration of GAI structures

First, companies must appoint a core team where the responsibilities with the corresponding decision-making rights must be mapped to manage the further implementation (I7, I9, I11). This core team, sometimes referred to as the center of excellence, is the central point of contact for all topics related to GAI use cases (I2, I3). The core team reports directly to the board, takes instructions from the C-level, and is superordinate to all departments (I3). It is interdisciplinary and organized along the use cases, as it is the interface that streamlines top-down and bottom-up innovation regarding GAI (I1, I3, I9). It serves as the central entity that collects, reflects and discusses use cases to assess them regarding value, technical feasibility and ethical standards (I1). Besides that, it allocates resources and incorporates important stakeholder perspectives (I1, I2). It consists of interdisciplinary experts with deep domain expertise in strategy, product, compliance, and software engineering (I1). All experts belong to respective departments, having direct access to those networks and resources (I2). As an example, a core team expert for cyber security can validate certain pain points in his department and present the findings

[2] As companies are very individual, the organizational design serves as an inspiration for new entities to manage GAI successfully.

to the core team (I2). Having experts, directly from departments, serves as additional connection between the departments and the core team (I7).

After a decision is made, tasks are cascaded throughout the organization to fulfill concrete steps of the roadmap (I2). Departments should implement their tasks and use cases independently with their internal resources (I2). If departments are not able to implement use cases by their own, an inhouse consulting team with technical expertise serves as an enabler (I3, I9). The services are billed internally, affecting the budget of the respective department (I2, I3, I9). Alternatively, departments can work with external consulting teams (I3, I9, I11).

Over time, a digital hub is formed by the core team, which gradually incorporates more experts and forms an expert ecosystem (I1, I11). These experts serve internal matters but can also help external partners (I1). Companies must consider scalability, as many requests can arise quickly, which is why they should also provide reference material for internal and external requests (I1). The digital hub should document best practices for use cases, foundation models and technical procedures (I1). Moreover, frequently asked questions can be answered for all employees (I1). The materials should fulfill educational and training purposes (I3). All of this can be achieved through various formats, including online videos, intranet posts, articles, blogs, podcasts, tutorials, experience-sharing sessions, webinars, and interviews with technology providers (I1, I2, I9). The aim is to reduce employee concerns, raise awareness of GAI applications and share technical experiences and learnings (I2, I9).

However, the overall aim is to empower departments and employees to implement their own use cases and develop AI expertise, which is important for scaling (I1, I2). On the other hand, it is often difficult to introduce GAI applications in teams, as potential users are critical (I9). The benefits should be demonstrated directly, and hands-on training can build acceptance (I9, I11). Moreover, transparency regarding GAI ambitions can decrease prejudices and concerns, thus increasing overall employee acceptance (I9). The focus should primarily be on reducing prejudices against GAI (I2). At the same time, users need to understand where the limits of GAI lie (e.g., in the accuracy of outputs) (I11). Employees must therefore not only be incentivized to use GAI at all, but also to use it securely (I11). The digital hub includes platforms with decentralized communities where like-minded employees can exchange their experiences and create networks (I3, I11). As a result, many synergies can be gained from shared experiences and best practices (I9). This combination of upskilling opportunities and communities to exchange is important (I3).

As development advances, the digital hub becomes more decentralized (I11). The decentralized development is intended to facilitate grassroot movements and bottom-up innovation within the organization so that individual teams can also incorporate new ideas (I3, I11). Many employees already showed the motivation to test GAI independently (I2). At the beginning of GAI, there was chaos and a period of trial-and-error (I2). Even if some work was made twice, it enabled open brainstorming (I2, I7). The management should encourage employees to reflect about potential use cases (I2). Many are trying out different things, but everyone is still learning what can be useful (I4). Risk management can be conducted by defining "dos and don'ts" to guide employees on what can be done with GAI and what should be avoided (I9).

Scaling and managing GAI

Finally, POCs must be initiated, ensuring scalability from the outset and capable of being implemented company-wide to ensure *pilot trap avoidance* (I9, I10, I11). This is more difficult than building individual POCs, as a scalable approach requires the combination of strategic and technical skills (I11). Besides that, companies that want to successfully introduce GAI need to be agile, which leads to fast test-and-learn loops with iterative improvement (I9). The development follows the rapid-prototyping approach with multiple feedback loops to ensure continuous evaluation of new features, requirements, and pain points (I9, I10). Besides that, trust can be built through early results (I10). Following the POC, a prototype is created, which evolves into the final product after extensive testing prior to launch. Simultaneously, additional use cases should be developed (I9).

To launch and monitor the final product, dedicated product owners are assigned, ensuring ongoing maintenance (I9). For this reason, the product management is a key factor for success as these professionals manage the further developments of GAI use cases (I7).

Product manager should be aware of the risk that specific departments have individual interests that may conflict with the general GAI strategy (I7). It is very important to develop an understanding for the end-users of these GAI applications and to test them constantly after launch (I7). AI can also help with testing products (I7). Furthermore, timelines must be managed effectively (I10). Regarding these developments, the product managers are becoming increasingly important, as they manage the interface between technology and product (I7). However, GAI requires new skills and knowledge that not all existing product managers possess (I7). That is why upskilling existing and hiring new product managers is relevant to use GAI successfully (I7)

5.4. Advantageous Resources

Independently of current developments, compatible use cases, and organizational design, a company's resources have a major impact on its future scope of action. While decisions on resources lead to path dependencies, the right decisions on *advantageous resources* can build *firs mover advantages* (Teece et al., 1997, pp. 522–523), an increased *digital maturity*, *human capital and partnerships*. While wrong decisions can lead to a lock-in effect with non-competitive technologies or partners and create high switching costs (Arthur, 1989, pp. 116–117). The building blocks for the GAI capability of advantageous resources are presented in Table 4:

Table 4: The GAI capability of advantageous resources with the respective building blocks

Description of the building blocks for advantageous resources	**Citations**
First mover advantage: Companies that already used AI technologies (as machine learning and robotics) gained a faster understanding of GAI right from the start, could transfer upskilling efforts of employees, and best practices in implementing use cases and rollouts. As they have already gained learnings and experiences in dealing with AI. Furthermore, an already established network can be used and expanded. It is an advantage to define the strategic direction and the key objective early on. The right timing is crucial and increases the fundamental understanding and continuous learning regarding GAI. Nevertheless, moving early on requires financial resources. However, companies that did not prioritize AI at an early stage, can still develop into this direction through training employees. Furthermore, the fast advancements enable such companies to start now with state-of-the-art foundation models, which perform better and cost less. There are also companies that are very flexible and can adapt their organizations fast or have a lot of resources (customer data) and thus quickly catch up with the first movers. Nevertheless, the first mover advantages could also be observed with the internet and the cloud, both technologies which represent the indispensable foundation for GAI.	I1, I2, I3, I7, I9
Digital maturity: The digital maturity of the company must be assessed, including the cloud platforms (the backbone for GAI inhouse solutions). Anything extending one individual use case must be scalable and requires a cloud platform that offers general services and infrastructure (e.g., AWS, Microsoft Azure, and Google Cloud). These cloud platforms provide various foundation models and development tools. With state-of-the-art cloud platform, data scientists can work effectively with the latest models and tools in an iterative test and learn approach. Digital maturity encompasses technological standards for the effective use of GAI, including efficient data management, and precise fine-tuning. Additionally, collecting proprietary business and customer data provides a competitive advantage for training and refining GAI applications. However, many companies have a low level of digital maturity and struggle to integrate GAI into existing legacy system. Therefore, the level of digital maturity combined with human capital is essential, as if it is too low, GAI initiatives will most likely fail.	I1, I2, I3, I4, I5, I7, I8, I9, I11
Human capital: Even if skilled data scientists are in shortage, it is important to hire and retain them in the long term. A balanced team of data scientists, including software engineers and AI experts, is crucial for GAI applications. Moreover, it is essential to have interdisciplinary human capital including data scientists, product managers, and legal or compliance experts. The combination of technical and legal expertise has become important as companies must interpret an increasing number of regulations. This interdisciplinary human capital is useful to build the core team or simplify the	I2, I3, I7, I9

implementation through a technical inhouse consulting team and the necessary knowledge within departments. Furthermore, key employees and C-level executives need further training to develop proprietary knowledge within the company.	
Partnerships: Partnerships are an important resource, as they expand fundamental understanding and continuous learning, can provide first mover advantages, and inspire to improve the technology stack. This includes partnering with providers of foundation models, cloud platform, external consulting services, and GAI applications. Chosen technology partners (depending on the contracts) often remain long-term commitments. As the proprietary knowledge of such technology partnerships (e.g., a trained foundation model) is hardly transferable if the respective partner is changed, organizational efforts, sunk costs and switching costs arise. Thus, it is important to partner with providers that remain competitive in the long-term and committed to update their technology constantly, while providing excellent service.	I1, I2, I4, I7, I8, I9, I11

Advantageous resources can already be established or should be developed in a certain company (I3). Especially, fundamental understanding and continuous learning can lead to advantageous timing and experiences. Moreover, helpful partnerships can be closed early on and help to develop the technology stack (I1, I2, I3). A state-of-the-art technology stack includes effective data management and a cloud platform to use GAI (I3, I7, I8, I11). The cloud platform can enable data scientists to work with leading tools efficiently (I7). Consequently, an interdisciplinary human capital is a key building block and should include data scientist with a balanced background, compliance and legal experts (I2, I3, I8). Enabled by fundamental understanding and financial resources, the advantageous resources are created. These relationships and influences of the individual building blocks are illustrated in Figure 10.

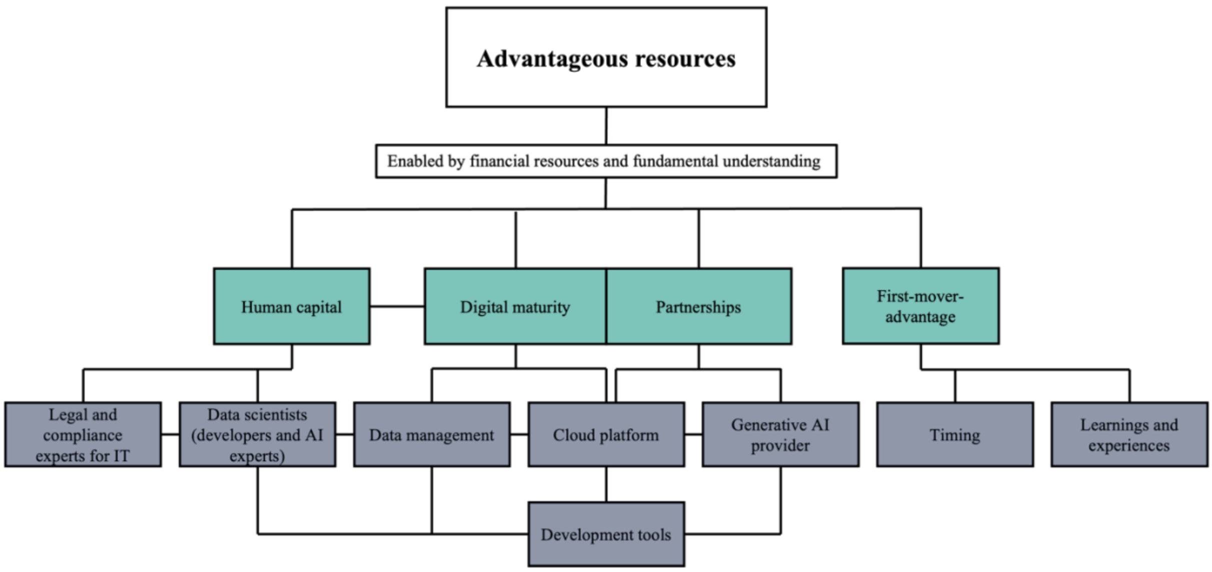

Figure 10: Advantageous resources and building blocks

First mover advantage

The impact of a first mover advantage was demonstrated by OpenAI, which operated fast and could establish itself as one if not the leading foundation model providers against big tech (I7). This is precisely why it is advantageous to define a strategy and the key objective early on (I2).

Early entry into AI especially for the core business, even before the hype, in areas like machine learning, automation, robotics and IoT, enabled a basic understanding of GAI right from the start (I1). Accordingly, some companies had already implemented use cases and rollouts with other AI technologies before the release of ChatGPT (I1). This includes training material and requirements that already existed and could also be transferred to GAI (I1). Companies that strategically approached challenges regarding AI early on could join initiatives and build a network (I1). These companies exchanged with peers about the challenge of building a competitive tech stack by discussing which methods are being utilized, conducting benchmarking and managing dependencies (I11). Additionally, connections were established with scientists and universities conducting AI research, allowing for the integration of more insights and innovations (I1). All these network activities increase the fundamental understanding and continuous learning regarding AI and complement use case compatibility. Such companies gained insight into this broad topic and recognized the importance of continuous education for all employees and managers (I3). Gaining and applying this knowledge early, while leveraging it for GAI, offers the opportunity to achieve a competitive edge through strategic timing (I2). This early development of AI competencies combined with a high degree of digital maturity can lead to successful usage of GAI from the beginning (I3). If a company was not a first mover, it can be developed into this direction through training employees. (I3).

On the other hand, Interviewee I2 argues that disadvantages of being late to GAI can be mitigated by the rapid advancements in technology, as better and more affordable models are continually being introduced. This allows companies to start directly with improved GAI solutions (I2). However, the extent to which GAI can provide value largely depends on the business model and the specific products and services involved (I2). Furthermore, the possession of customer or personnel data can be an advantage, as it can be used to train or fine-tune foundation models to achieve improved outputs (I8). Despite that, it does not make sense for every company to develop an inhouse solution, as off-the-shelf GAI applications can also solve certain pain points successfully or can be customized to a certain extent (I2, I9).

Another first mover advantage can be obtained through flexibility to integrate effective organizational designs quickly. Generally, startups tend to be more flexible, while corporates have more financial, technological, and personnel resources compared to small organizations (I11). At the same time, startups also need to raise funds for hiring skilled employees to remain competitive, while offering speed in the development of GAI applications, which is important (I2, I7). Startups have a strong focus on one target in contrast to corporates that must steer various business areas and are therefore usually slower (I1, I2).

On the other hand, corporates can leverage the usage of GAI even more, as they can profit from company-wide rollouts (I6, I10, I11). Furthermore, corporates have a large customer base, years of experience, domain experts and can transfer this directly to innovations and products (I1). However, making generalized statements is difficult, as both startups and large corporates excel or fall short in different areas, and much depends on the specific organization. However, there also industries with a more conservative culture, as the automotive industry or mechanical engineering (I2, I10). As emphasized by Interviewee 10, the openness towards GAI often depends on the industrial background of the respective addressee.

Digital maturity and human capital

The digital maturity of an organization plays a crucial role in shaping enablers for GAI (I9, I11). It starts with providing certain technological standards combined with effective technical implementation, data orchestration, management, and infrastructure (I7, I8, I9, I11).

In the past, many companies were late to use the internet or the cloud, which are both technologies that probably had at least a comparable impact to GAI (I7). However, these advancements are interconnected, making the cloud essential for the functionality of GAI as well (I9). Anything regarding lager customers that exceeds individual use cases must be scalable and requires not only a foundation model but also a platform that offers general services and infrastructure (I3, I4). Such cloud platforms for GAI are often operated by big tech, e.g., AWS, Microsoft Azure, and Google Cloud (I3). The Google platform, for example, offers a variety of models that can be deployed on it and covers various use cases (I3). Most companies deploying multiple GAI use cases rely on cloud platforms, which support various foundation models and enable scalable and efficient model operations (I11). Evaluating each platform is essential, as it allows companies to utilize certain models tailored to a variety of use cases, while other models may not be available at all (I11).

In some corporates, different models can be selected and then deployed via the cloud platform, depending on the specific use case requirements (I1, I3). Most companies will only use one platform but can use different models (I11). Moreover, cloud platforms give access data center capacities, ultimately offering significant cost advantages to end customers (I3). It is important for companies to choose performant state-of-the-art cloud providers, which provide all relevant development and support tools, so that software developers can work flawlessly (I7). Thus, companies should carefully reflect the most suitable cloud platforms, as these form the backbone for implementing GAI solutions (I3, I4, I11). If state-of-the-art cloud platform is not used, data scientists cannot work effectively with the latest models and tools in an iterative test and learn approach (I7). As Interviewee 7 explains, the EU and especially Germany, appear to remain behind the USA, as less companies have implemented a state-of-the-art cloud platform (I7).

Many companies currently have a low level of digital maturity, which is why attention should be paid to how GAI can be integrated into existing systems and whether the basis for it is provided by such a cloud platform (I3, I4, I7). Even though most clouds now provide highly secure environments, not all companies use them (I8). The strong focus on data protection, privacy and explainability in Germany and the EU are delaying developments in GAI (I7). Lacking a cloud platform and overemphasizing data protection leads to longer development cycles for German companies (I7). Leading German companies decided to launch on-premises projects in the past to increase data security (I5). However, the efforts to use this at scale are huge, while cloud services are part of the core business of many big technology providers and often result in a better performing and more cost-effective alternative.

Another important point is data collection, because the possession of customer or personnel data is an advantage, as it can be used to train or fine-tune foundation models (I8). Certain methods of data acquisition can also represent a competitive advantage (I8). Simultaneously, many corporates have a large customer base and thus can leverage the existing data (1). This leads directly to the topic of data management, as companies not only have to collect data, but also must handle it securely (I8). It's essential to understand how different types of data can be used and how they should be processed (I3). This knowledge is crucial for addressing subsequent tasks regarding GAI (I3). Companies that are data-centric have a remarkable advantage by knowing what data exists, where it is stored, and how it can be updated and accessed (I9). Also, unstructured data can be improved through tools for labeling and cleaning (I9). The more advanced a company becomes in its data management, the better positioned it is for GAI usage (I9). At the same time, a GAI providers can also increase their defensibility and minimize replicability by training their foundation models with each new customer (I5).

In addition to the fundamental work in terms of technology stack, the company must be able to afford GAI (I2). Licensing software, buying hardware (such as servers) or using cloud platforms is expensive (I2). In addition, skilled personnel as domain experts need to invest time to form part of various teams (e.g., the core team) (I2). Henceforth, in terms of financial and personnel resources, not all companies (e.g., small businesses) can leverage GAI to the same extent (I2).

While many manufacturing companies in the EU have the personnel and financial resources combined with effective operations, they seem not as competitive as US companies, as they often lack the implementation of leading software technologies and stuck to legacy systems and outdated tools (I4, I7). As a result, leading programming languages, development tools, and managed services are less common compared to the USA (I7, I8). According to experts, this results in more effort to deploy, iterate, secure, maintain, and replace software on the cloud, which is snowballing to the entire pipeline of skilled data scientists who are already familiar with these tools (I7, I8).

Interviewee 7 raised concerns that data scientists who are familiar with cloud platforms may be less frequently available in Germany, as these skills are less applied here. Thus, Interviewee 7 stated there exists a shortage of human capital which is used to leading tools (I7). This also affects other areas of software engineering as well, emphasizing the importance of digital maturity and skilled employees (I9). However, the difficulty is not only finding the right people, but also retaining them in the long term (I9). This leads to more time being invested in human resources, whereas companies in the US can operate faster (I7). In addition to hiring new employees, a focus should lie on upskilling key employees (I7). Especially, manufacturing companies with non-digital employees are naturally slower, while many SMEs do not have the required human capital at all (I2, I4).

A key factor to successfully use GAI lies in building up interdisciplinary teams with data scientists who understand technological fundamentals, programming and the IT infrastructure, which includes a balance between classic software developers and technical experts for AI (I8). In addition, companies need employees at the economic-technical interface and legal or compliance experts for risk classification management (I2). A technical background is helpful to develop expertise regarding managing emerging technologies (I2, I3). The combination of technical and legal expertise has become more important as companies must interpret increasing regulations (I3). Accordingly, large corporates have more employees with expertise in these areas and are more sensitive to the issue of compliance (I8). This makes human capital a key building block of the advantageous resources' capability.

Partnerships

Another important aspect that strongly influences resources are long-term partnerships. These can be, determined through using a provider of a GAI application, a certain foundation model, a cloud platform, or external consultancies (I7, I9). Many corporates already have partnerships with leading cloud providers, which can then be deepened regarding GAI, while the sharing of data and prompts can be regulated (I2). To address these needs, Microsoft Azure among others enable customers to host their models in the EU to provide additional data privacy (I4, I8).

Furthermore, it should be considered that specific cloud platforms do not allow the usage of all foundation models (I2). It's important to keep in mind that the chosen technology partner will be a long-term commitment (I2). Ending a partnership depends on the respective contracts and leads to organizational effort (I2). This is the cause why such partnerships lead to path dependencies that remain for several years (I2). These partners are checked similarly to other areas in terms of products and services through compliance departments (I2).

Particularly when companies collaborate with GAI providers to fine-tune foundation models for specific requirements, they are unable to transfer these customizations to other providers (I2). This is comparable to changing the ERP system or the cloud platform and involves substantial sunk and switching costs (I2). In addition, GAI is mostly a service-heavy business, and companies should choose partners who can also support them in the long term (I4). For that reason, it is important to ensure that the technology partner remains competitive over time and updates the model constantly to keep pace with rapid developments (I2, 11). Once a critical mass is reached and many customers are using a product, like

Microsoft Copilot, the level of convenience becomes high, making a switch to an alternative unlikely (I2). As mentioned, there is also a risk with open-source models as in some cases their development will no longer be continued (I11).

5.5. Compliance and Data Security

Regulatory awareness to ensure compliance is important for GAI as foundation models work with huge amounts of data and can give crucial recommendations to decision makers, often impacting humans (I10). As many GAI models lack explainability (I1) and hallucination may occur (Maleki et al., 2024, p. 127), certain use cases will be regulated. As a result, companies must ensure *compliance robustness*. This involves ensuring *regulatory proofness* for use cases and transparency through the AI software-bill-of-materials (*AI-SBOM*) regarding licensing rights for open-source components (I8). Furthermore, it is necessary to establish *data security,* so company data and IP is not leaking into foundation models (I8, I11). Since GAI applications can be influenced by certain data and prompts, they are vulnerable and require security mechanisms (I8). These fundamental building blocks to achieve the capability of compliance robustness are depicted in Table 5.

Table 5: The GAI capability of compliance robustness with the respective building blocks

Description of the building blocks for compliance robustness	**Citations**
Regulatory proofness: Prioritize established and pending regulatory compliance regarding (G)AI. Analyze the EU AI Act and understand relevant AI risk classes. Also, reflect if established or planned use cases (on the roadmap) may face future regulatory restrictions or prohibition. Examine the requirements for your use cases and adapt to become AI Act compliant. Especially use cases where personally identifiable information is involved are sensitive. Also, training foundation models with proprietary data will require a certain representativeness within the data set, meaning that a limited customer base may not meet regulatory compliance for representative data. Prepare for reporting obligations to increase transparency towards authorities. Implementing the EU AI Act will also have a transitional phase (like GDPR) until the requirements must be met. Also pay attention to GDPR compliance, especially around data privacy and storage. Consider further recommendations such as the NIST AI risk management framework, which lists helpful aspects for trustworthy AI and ISO/IEC 42001 which could also develop into a market-standard. Create a clear framework for your company, with internal requirements and conditions to guide developers. Establish use case management processes with AI checklists and internal reports to summarize all relevant information regarding the GAI applications, also for future reporting obligations.	I1, I2, I7, I8,
AI-SBOM: Prioritize licensing rights for open-source components in GAI applications, to avoid being sued if used commercially. Establish the concept of the AI-SBOM,	I8

describing all components and the respective licenses with which the software is built. This includes model weights and data with which the foundation model was trained as well. While public data is available to everyone, crawling and scraping data is a legal gray zone. This must be considered for inhouse solutions, but also for external GAI, as customers often do not know which software components and data sources are hidden underneath the application. Thus, customers are often unaware of which foundation models are used by a certain layer provider, where the training data originated, how it was collected, and if copyright-protected content was used. Although providers are often reluctant to disclose their AI-SBOM, companies must decide whether to accept this risk or seek alternative solutions.	
Data Security: Data security can be endangered when sensitive data is used for training and prompting, leading to leak outs into publicly available foundation models in the further process. Henceforth, sensitive and malicious data must be cleaned up beforehand, as deleting it afterward is not possible. Companies can implement additional layers, which filter the inputs of sensitive data before they flow into the foundation model. Nevertheless, employees should be properly trained and informed about guidelines to minimize the risk of data loss. To decrease the trade-off between working efficiently and protecting IP, compliant alternatives must be offered. Many larger companies with appropriate resources can develop inhouse solutions covering data protection concerns, while smaller companies often turn to GAI applications from external providers. However, data access regarding external GAI applications must be carefully reviewed and narrowed down through access to predefined shared folders with role-based permissions. Existing and future contracts, including their terms and conditions, should be reviewed to determine whether customer data can be used for training foundation models. Companies need to consider, where foundation models and the respective data are hosted and can increase data security additionally by hosting on EU or local servers only. In addition, it is important to have security whether GAI applications use customer data to train the entire foundation model, or only for the individual customer, so that proprietary data of one customer is not touching other data. To provide data security, a software orchestration layer can be implemented which includes monitoring prompt injection. This prevents hackers from bypassing security mechanisms by crafting clever prompts designed to “jailbreak” the model. Such activities aim to extract sensitive or restricted information that should not be disclosed. A software orchestration layer includes elaborated guardrails to monitor incoming prompts. It is trained with malicious prompts to recognize them and prohibit the use. This can also avoid bot attacks which inject high numbers of prompts to generate API call costs for the respective provider. Besides that, toxicity in the outputs can occur and should be blocked. Thus, an	I2, I3, I4, I5, I8, I9, I10, I11

observability layer must be developed to constantly track the number and quality of queries.	

Regulatory proofness

Some experts describe the current regulatory landscape of AI as “wild west” (I4, I8). However, additionally to the special features of the German and EU market due to a different technological set-up, this is changing due to far-reaching impacts of pending regulations (I7, I8).

A major legal framework is the EU AI Act, which aims to regulate AI within the European member states (PricewaterhouseCoopers, n.d.). On May 21, 2024, the member states approved the world’s first law regulating AI, which must now be implemented into national law (Bundesregierung, 2024). The EU AI Act will harmonize AI regulations and establish risk classes for AI, for which companies must then implement certain requirements within six to thirty-six months (PricewaterhouseCoopers, n.d.). Thus, companies will have an interim period to implement the regulations (comparable to GDPR) (I8). Furthermore, non-compliance leads to liability risks and considerable fines (PricewaterhouseCoopers, n.d.). While certain use cases are prohibited, others may be subject to specific requirements (PricewaterhouseCoopers, n.d.). Therefore, companies should carefully consider the use cases they implement and plan on their roadmap, assessing whether these might be restricted or even prohibited by the EU AI Act (I8, I9). This list of restricted use cases can be extended by the EU Commission at any time (PricewaterhouseCoopers, n.d.). Henceforth, it can be helpful to stay in contact with political decision-makers to assess new regulatory developments (I2).

In addition to specific requirements, reporting obligations will be introduced, while transparency and compliance with copyright laws must be guaranteed (I8). Further regulations such as the EU Data Act, GDPR, or regulations for certain industries must also be considered (PricewaterhouseCoopers, n.d.). Also, the GDPR will impact GAI applications regrading data privacy in Europe (I8). Furthermore, there are regulations for cloud usage, specifying the requirements for certain use cases and where the data must be stored (I8). This has an impact on the entire cloud set-up, while many companies in Europe tend to be cloud-agnostic and cannot rely on additional services from the major cloud providers (I8).

Nevertheless, the EU AI Act remains controversial, as some experts argue that categorizing AI according to risk classes makes sense, while many far-reaching provisions seem ineffective in achieving the desired goals (Veale & Zuiderveen Borgesius, 2021, p. 97). For that reason, companies need to evaluate carefully what they implement and establish guiding principles (I9). Particularly affected areas include human resources, health technologies, defense and financial technologies, where sensitive data as personally identifiable information is involved and little tolerance for margins of error exist, as a bias can severely harm individuals (I3, I8, I10). For example, a GAI application for CV preselection could contain a bias through unrepresentative data and discriminate against a certain group of people (I8, I9).

Subsequently, the EU takes a profoundly regulatory approach, where certain use cases will be prohibited, while others will be heavily regulated and closely monitored by authorities (I8). It will be complex to become AI Act compliant, but it will also be the major regulatory framework to follow (I11). For that reason, companies must ensure regulatory proofness, to avoid being penalized in the future (I4, I8).

In addition to the legally binding EU AI Act, there exists also the voluntary and useful NIST AI risk management framework (I7) which, together with the private and public sector, has published recommendations to manage risks and trustworthiness in AI (NIST, 2021). Moreover, there also exists the ISO/IEC 42001 certification, which is relevant for AI systems (I7). Even if not all governments internationally implement AI regulations, potential customers of GAI providers may still require specific certifications and frameworks, leading to the establishment of market-standards (I7).

Regulations may be even stricter for some use cases beyond GAI, such as computer vision (I10). This includes use cases which are trained through observing humans and most likely will be restricted, having a potential negative impact on robotics companies (I5). Restricting theoretical use cases too early can have negative consequences, as many applications may be categorically prohibited without practical insights into their potential (I5). As a result, American investors sometimes demand that startups are based in the USA and store their data there because they perceive EU regulations as a risk for business success (I5). Interviewee 4 remarks that one could argue whether the EU is adopting too many regulations (I4). Through the sheer number of regulations, often too much time is spent on compliance issues and less on productive developments (I7).

Interviewee 1 addresses that even if some people perceive regulations as too extensive, however, it is important to regulate AI, as people should know whether information is generated by AI and reliable (I1). Regulation naturally slows down developments, as it will be more difficult to work AI Act compliant (I10, I11). Despite that, Interviewee 10 emphasizes that "… regulation, on the other hand, also creates equal market conditions and, let's say, it's not the ruthless one who wins in the end".

Pending regulations imply that certain international foundation models will not be released in the EU, or only with delays, which is also an advantage for local providers (I11). Due to the non-deterministic nature of GAI and its limitations in explainability, sources and references are usually not provided and biased data may lead to discriminatory outcomes (I1, I8). This makes the assessments regarding compliance important (I1). Besides that, it is often also useful to elaborate conditions and requirements for AI, which are helpful for developers who know exactly what is expected (I1). Further processes to ensure compliance include use case management, AI checklists, and internal reports, which summarize information on GAI applications (I2). Besides creating internal transparency, such information can also be used if auditing will become mandatory in the future (I2). Risk management can be improved by outlining specific "dos and don'ts" to instruct employees on how to properly use GAI (I9). This is becoming increasingly crucial for international corporations, as they must comply with increased regulations while rolling out and scaling GAI applications globally (I3).

This also includes checking where GAI providers store their data and whether they work GDPR-compliant (I3). Another risk involves the information generated by GAI on behalf of the company, raising the question to which extent the company bears liability in such cases (I11). In addition, the high energy consumption by GAI will probably also require regulations in the future (I3). However, ignoring regulations can remain undetected for individual cases but represents a long-term risk (I8). While small companies tend to stay under the radar, particularly large corporates are at risk of being fined (I4).

AI-SBOM

Some companies tend to ignore certain regulations or even licensing rights and implement non-commercially usable open-source models in their products (I8). Such decisions often are based on cost-benefit analysis, as ignoring certain regulations or even licensing rights may remain unnoticed. However, this carries long-term risks, with companies being penalized for integrating non-commercially usable components in their software architecture.

As explained by Interviewee 8 "in classical software engineering, you have the concept of SBOM", describing the components on which the software is based. With respect to AI, this concept is extended to the AI-SBOM which transparently lists all components and licenses of a GAI application, as not all open-source components are available for commercial use (I8). All the implemented open-source components must be commercially usable, since otherwise legal challenges may arise. This includes model weights and data with which the foundation model was trained. This creates risks for the GAI providers, but also for their respective customers. Companies using GAI from different layer providers often do not know which software components and data sources are hidden underneath it. Interviewee 8 mentioned the example of the U.S. Army, where GAI providers declined to disclose their AI-SBOM to the organization. This example illustrates that some GAI providers refuse to disclose their AI-SBOM, leading to unawareness among customers regarding underlying foundation models and training data. Thus, customers do not know how the data was collected and whether copyright-protected content may arise during usage. If GAI providers do not transparently disclose their AI-SBOM, potential customers must decide whether to accept these risks or seek alternative solutions.

Data security

Furthermore, training with sensitive data can lead to leak outs during the usage of GAI, so that sensitive data must be removed before training the model to ensure data security (I8). In addition, sensitive data cannot be deleted afterward, which means that retraining the model from scratch would be necessary. That is why there should be a software layer in use that run on the customer site and filter sensitive data inputs before they flow into the foundation model on the cloud. However, many companies will be unable to train their GAI applications using customer data, as a limited customer base may not meet regulatory compliance for representative data in the future. Henceforth, companies can create defensibility and competitive advantage through data acquisition of large representative data sets. The challenge lies in obtaining these representative data sets and cleaning the malicious data to fine-tune foundation models effectively.

However, while public data is available to everyone, crawling and scraping data remains a legal gray zone (I8). For that reason, some investors look for startups that develop GAI applications which receive proprietary data with each new customer to improve the accuracy of the foundation model (I5). IP can be created through the ingested data, the fine-tuning process, the foundation model choice, or the orchestration of various small open-source models which work in a performance-enhancing way together (I8).

Many companies are unsure how their own data is processed and whether others can access it through foundation models (I11). Thus, many want to host their data in the EU (I9). Special use cases like banking emphasize data protection more and tend to implement on-premises solutions (I8). The preference is often driven by a sense of increased safety and perceived security associated with on-premises solutions (I5, I8). Many companies fear data loss, as there is uncertainty about whether foundation models like OpenAI, despite promising data security for specific circumstances, may still use the data for model training (I5). In addition, there is the risk that external GAI applications do not pay attention to data protection and feed data directly into underlying foundation models (e.g., GPT-4) (I8).

The worst-case scenario is that GAI applications could sift through internal company servers and extract as much information as possible to feed into a model (I2). If these GAI applications do not rely on internal data, the risk is minimized, or if they only access predefined shared folders with role-based permissions, where approval is required step-by-step for each document accessed by the tool.

A potential risk emerges if employees use confidential data in GAI applications (as ChatGPT), resulting in data loss (I10). Data loss is a risk if employees are not properly trained, and no internal guidelines are established (I2). As a worst-case scenario, Interviewee 2 sets up the theoretical example that confidential proprietary data, such as a blueprint, is uploaded by employees into a GAI application (which is hosted overseas) with a prompt to adapt the blueprint according to a specific ISO certification (I2). This blueprint could then reappear through the model somewhere else (I2). Thus, by using the GAI application, the employee could work more efficiently as a result, but risked IP loss (I10).

Therefore, larger companies with appropriate resources develop inhouse solutions due to data protection concerns, while smaller companies often turn to GAI applications from external providers (I3, I8). Companies need to collaborate with GAI providers or develop inhouse solutions to ensure employees use compliant tools and avoid turning to non-compliant alternatives (I3). Achieving data security is only possible if viable alternatives are provided (I3). This concern also extends to other tools, such as translation software, which carries the risk of data loss and therefore requires guidelines for handling confidential data (I1).

Furthermore, it should be noted that many existing contracts between companies do not specify whether emerging data can be used for training foundation models (I5). This requires reviewing existing contracts, while new agreements with GAI providers, including terms and conditions, should be analyzed regarding data security (I2, I5). With GAI, companies come across various legal and compliance issues, especially if data sharing and storage occurs (I3).

As GAI providers are aware of the increasing focus on data protection, they offer options to host foundation models on EU servers (e.g., OpenAI via Azure) (I4, I8, I9). Some companies even host their inhouse solutions locally (I10). GAI applications can also have isolated deployments for each customer data, so that the data is virtually isolated for each customer and does not end up in the same dataset (I8). Henceforth, proprietary data of one customer is not touching other data (I8). Whether GAI is allowed to perform transfer learning with customer data to deliver the best industry-specific results to all customers depends on the contracts and the respective industries (I4). In the pharmaceutical industry, for example, a lot of money is invested in research and development, which means that no data exchange will be permitted (I4). If, on the other hand, the use case is less sensitive as for example warehouse management, customers tend to allow data sharing more often (I4).

Further challenges will arise if models use inputs (e.g., brand logos) and generate new similar outputs from that, as this may infringe copyright and lead to legal implications (I4). However, courts will determine how the training of foundation models using publicly available, but copyright-protected or private information should be evaluated (I2).

Moreover, cyber security of GAI is a relevant topic. While existing IT security testing frameworks can also be adopted and used as an inspiration (I2). Besides that, GAI applications must be tested for toxicity in their outputs (I8). There are also a few software providers that cover security aspects of GAI. Such functionalities can include monitoring prompt injections, through which hackers want to bypass security mechanisms through clever prompting to “jailbreak” the model, aiming to gain insights that should not be shared. This is why guardrails must be established as a defense mechanism. Guardrails entail establishing measures to consistently monitor incoming prompts, utilizing various techniques to identify any potential malicious intent. Such guardrails can also be trained with malicious prompts to recognize them in the future and to prevent them from being entered. However, it is also a time-consuming endeavor to implement such guardrails.

Furthermore, training data can also be poisoned, so that a small part distorts the complete accuracy of the model (I8). Another possible attack would consist of bots that access GAI applications and enter high numbers of prompts, thus generating costs (API calls) for the provider (I8). For that, an observability layer must be installed to constantly track the number and quality of queries (I8). This observability filter should monitor all inputs and outputs, while incorporating the necessary guardrails (I8). Companies must technically ensure that the outputs of the AI are not dangerous in any way, as they may be liable for them (I11). There is a particular risk in direct end-user interfaces, as hackers can use clever prompting to create undesired behavior and thus attack the company interests (e.g., sell me a product cheaper) (I11). This is why a well-protected GAI system is important, especially when direct end-user interfaces are provided (I11).

6. Discussion

6.1. Principal Findings

AI software is anticipated to cause the most disruption in the realm of technology over the next years (Bloomberg, 2017). Especially GAI exhibits remarkable advances since 2022 (Douglas, 2023, pp. 1–2), impacting global dynamics, societies, and companies (McKinsey, n.d.; Mondal et al., 2023, p. 3; Morris, 2023, p. 23). While some companies adapted certain emerging technologies in the past and increased their competitive advantage, others failed (Ho & Chen, 2018b, p. 1; Teo & Pian, 2003, p. 90). Currently, such a competitive advantage can also be achieved through using AI technologies (Climent et al., 2024, p. 1).

Even if certain research exists on how to use AI in companies successfully (Brenner et al., 2021, p. 15; Hercheui & Ranjith, Rishikesh, 2020, p. 87; Wagner, 2020, p. 19; Wamba-Taguimdje et al., 2020, p. 3), most of the literature is focusing on AI before the strong advances of GAI starting in 2022 and features highly industry-specific perspectives. This creates a gap for a cross-industry explanatory approach investigating which GAI capabilities are needed to use the technology successfully.

Therefore, the research question “What capabilities are needed for the successful usage of GAI?” emerged. To answer this research question comprehensively, an understanding of current developments within the GAI landscape must be achieved. Subsequently, business use cases in which companies integrate GAI should be investigated. Besides that, optimization of existing organizational processes and resources to realize an effective implementation should be discussed. These findings lead to the identification of cross-industry capabilities to successfully use GAI. The thesis set out to answer this research question and, in addition to an initial literature review, primarily used the qualitative research method of expert interviews. The interviews with eleven GAI domain-experts ensured multi-faceted perspectives of corporates, SMEs, startups, consultants and investors from various industries.

A semi-structured interview format was used to provide flexibility across diverse perspectives, while addressing constantly key themes in all interviews. An initial interview guideline was prepared by using the theory of dynamic capabilities, which is a useful framework to identify the necessary capabilities for leveraging new technologies (Teece et al., 1997, p. 509). The theory also supported the coding framework, as theoretical aspects were used to derive initial open codes, which were extended by deductive codes for further aspects that surfaced during the interviews and subsequent data analysis.

The preliminary literature research and the analysis of all interviews established a profound understanding of current developments within the GAI landscape and featured technological limitations, current market dynamics and insights into the leading providers.

The GAI provided by foundation model providers performs sequence-to-sequence modeling and returns the most probabilistic output based on a certain input, including multimodal tasks (I6). The probabilistic retrieval based on the input uses all the data with which the foundation model was trained and can operate on very large matrices, unstructured data, and broad contexts (I5, I6, I7). Thus, repetitive and time-consuming tasks can be automated (I8), to increase productivity through solving tasks more effectively (I4). Meanwhile, leading provider of GAI develop either open-source (Meta, Hugging Face, Google) or closed-source foundation models (OpenAI) (HuggingFace, n.d.; Meta, n.d.; OpenAI, n.d.-a; Pichai & Hassabis, 2024). While many published models in both segments (Cohere), and sometimes even provide direct layer applications for end-users (xAI) or access to their closed-source models via APIs (Anthropic) (Anthropic, 2023; Cohere, 2024c; x.ai, n.d.). Additionally, cloud platforms provided by AWS, Microsoft Azure, Google Cloud and IBM (Law, 2023) are important for training, managing and running foundation models (Oracle, 2024). Various layer providers use these foundation models to power their own GAI applications (Geirhofer & McKinney, 2023, p. 3). Companies that develop GAI inhouse solutions or use off-the-shelf applications should consider the differences between underlying open-source and closed-source models (I1, I2, I8, I11). Besides that, most inhouse solutions and off-the-shelf GAI applications are powered by one of the leading foundation models (I7, I8, I9). Henceforth, there exists strong market growth in the GAI market, while the initial surge with high market growth seems to decrease in the upcoming years (Statista, n.d.).

However, users must be aware that the output generated by GAI may not always be accurate (I11), as data inconsistencies or hallucinations may occur (Maleki et al., 2024, p. 127). Thus, a human-in-the-loop is often required (I10). In addition to the limitations of GAI, it is also important to consider when other AI technologies can be used more effectively (I3, I9). For that reason, education regarding GAI should always cover further aspects of AI to create a holistic understanding (I6, I9, I10). While GAI can also be seen as a technological layer to connect existing AI technologies across use cases (I9). The profound understanding of developments within the GAI landscape, current market dynamics and differentiation from other AI technologies enable fundamental understanding and continuous learning (I7, I11).

Furthermore, common business use cases for GAI were distilled, and based on the fundamental understanding and continuous learning, the capability to assess use case compatibility was introduced. Common use cases include established software that introduced GAI features (e.g., Microsoft Copilot), and GAI applications to summarize key insights, propose texts, create presentations, and generate images (I1, I2, I3, I9, I11). It enables chat interfaces and streamline customer communication with direct access to internal information systems (I3, I5). While it is usable across various industries, the most prominent use cases include human resources, software development, customer service, sales and marketing, and financial industries with extensive due diligence, contract management and regulations (I4, I8, I9). Learning about potential use cases is valuable, as it gives companies the possibility to identify a portfolio of use cases (I9). This portfolio can be extended by engaging with internal or external consultants as well (I9). Furthermore, use cases should be described in detail to translate broad concepts into specified requirements while preventing moving targets (I10). Companies should reflect about use cases in detail and quantify their ROI (I2, I9, I11). The ROI supports determining a prioritized roadmap with waves of implementation over the next years (I2, I3, I4, I9, I11). Finally, a make-or-buy decision must be made to realize the respective use cases by building an inhouse solutions or using an off-the-shelf GAI application (I2, I8, I9, I10).

Improved organizational processes were an important topic across all interviews, leading to the development of the organizational design capability, which includes C-level commitment, a core team, implementation, a digital hub, pilot trap avoidance, and effective product management. Thereby, C-level commitment enables successful usage of AI, as it prioritizes GAI and provides the necessary resources from the beginning (I1, I7, I9, I11). This includes removing initial barriers, which require top-down decisions (I2). Following that, the strategy should align with the portfolio and roadmap to fulfill the strategic milestones (I2, I9, I11). The core team is the central point of contact for all topics regarding (G)AI, directly reporting to the board and the C-level (I2, I3). It is superordinate to all departments and pools cross-functional domain experts from strategy, compliance, and software engineering (I1, I3, I7). Organized along the use cases, it serves as a catalyst for top-down and bottom-up innovation (I1). Simultaneously, it cascades tasks to the organization to fulfill concrete steps of the roadmap (I2). The implementation of all tasks should be carried out independently within the departments, while continuous upskilling increases these abilities (I1, I2). If additional support is needed, technical inhouse consulting or external consulting teams can be included (I3, I9, I11). The digital hub, on the other site, is an educational platform to connect experts and users, while answering frequently asked questions through digital content (I1, I2, I3, I9, I11). This also serves as a community for like-minded people and increases bottom-up innovation (I2, I3, I11). Besides that, a scalable structure for each use case should be ensured from the beginning, as transitioning from a pilot to a company-wide rollout is a challenge (I9, I10, I11). Dedicated product managers should manage the GAI applications during launch and monitor the performance and user expectations (I7). Existing product managers should learn about GAI, while new product managers with GAI expertise can be added to the team (I7).

Following that, advantageous resources to use GAI were analyzed, allowing the discovery of various building blocks as first mover advantage, digital maturity, human capital and partnerships. While a first mover advantage can be achieved through experience with other AI technologies in the past, having an extensive network regarding AI, or offering flexibility to adapt fast can be useful as well (I1, I3, I11).

Moreover, digital maturity is important, as using a cloud platform where data scientist can work with latest development tools and models is often the backbone for GAI inhouse solutions (I3, I4, I11). Furthermore, efficient data management, fine-tuning, collecting proprietary business and customer data can increase the successful usage of GAI (I7, I8). Simultaneously, companies with a low level of digital maturity struggle to integrate GAI (I3).

Finally, the capability of compliance robustness was elaborated, including regulatory proofness, AI-SBOM and Data Security. The building block of regulatory proofness comprises the anticipating fulfillment of pending regulations as the EU AI Act regarding established or planned use cases that may face restrictions (I8, I9, I11).

Besides that, GDPR compliance must be implemented, while further frameworks may support the development of trustworthy GAI applications (e.g., NIST AI risk management framework, ISO/IEC 42001) (I7, I8). Following that, licensing rights for open-source components in GAI application must be commercially usable as otherwise legal implications may follow (I8). The concept of AI-SBOM can help to create transparency regarding all open-source components, model weights and data sources on which the GAI is built on (I8). Finally, data security is an important aspect of GAI, as sensitive data can be inserted into publicly available foundation models by employees and create data loss (I2, I8, I10). Therefore, companies must offer compliant alternatives through inhouse solutions or secure GAI applications and decide where they should be hosted (I4, I8, I9, I10). Furthermore, an orchestration layer can be developed to guardrail the number and quality of incoming prompts to ensure security for GAI (I8).

These findings lead to the identification of cross-industry capabilities to successfully use GAI, consisting of the use case compatibility, effective organizational design, advantageous resources, and compliance robustness with their respective building blocks.

6.1.1. International Comparison

An interesting aspect which was repeatedly pointed out across various interviews, but which is not directly related to GAI capabilities, is an international comparison between Europe, the USA and China.

While the international comparison can also be useful for identifying new use cases, attention was drawn to the slow implementation of the internet, cloud and AI in Germany compared to the USA and China (I4, I7). Extensive regulations as the EU AI Act will further change the European market, and presumably slow down the release of leading foundation models in the EU (I5, I11). The large industrial companies in Europe, which have been market leaders for many years were also described as being sometimes critical towards innovation (I5, I7). Interviewee 5 also assumed that such companies probably influence regulations in order to slow down developments so that their business models are not disrupted.

The EU AI Act defines different risk classes to restrict certain use cases and can be compared in some respects to GDPR (I7). Interviewee 9 adds, however, that other countries will also introduce regulations.

On the other hand, the US seems to employ a rather market-based than regulatory approach (I5, I7, I8). Interviewee 7 summarizes that „the issue is that it's laws that keep things in check [in the EU] whereas in America. It's the market that keeps things in check". This means that companies can operate more freely, and potential violations will later be clarified in court (I5). However, this approach also allows large companies such as Google to operate freely and later search for settlements in court (I5). Accordingly, the US tends to have guidelines that are known but are not always adopted (I8). Presumably, the AI-SBOM will also find limited application in the USA (I8). Furthermore, many US investors demand AI companies to be registered in the USA to avoid EU regulations (I5).

However, an important topic remains the production of powerful GPUs that enable the training of AI (I2, I6, I7, I10). Most of these GPUs are available in the USA, while access to the Chinese is also made more difficult by the CHIPS and Science Act, as this is a national security issue for the USA (I5, I7). Technologically speaking, the USA already has a first mover advantage and will promote these developments further, which is reflected in their additional investments in AI (I10).

The required GPUs to enable high computing capabilities are not available to the same extent in China (I10). China is a relatively closed market for GAI, while the model architecture is becoming increasingly standardized internationally with many Chinese researchers involved (I3, I10). However, besides architecture and data, the hardware seems to be the limiting factor for the success of foundation models (I10). In China, there exists the risk of ideological aspects being integrated into the foundation models (I2). Thus, Interviewee 10 raises the concern that, while data is crucial for training foundation models, censorship in China restricts the use of certain content, as it may resurface in undesirable ways. To avoid complications, specific portions of the training data must be excluded, ensuring compliance (I10). However, this careful selection process may result in the removal of excessive data, out of caution, which further limits the available datasets for training.

6.2. Implications for Research and Practice

Previous research endeavors focused on capabilities of AI (Brenner et al., 2021, p. 15; Hercheui & Ranjith, Rishikesh, 2020, p. 87; Wagner, 2020, p. 19; Wamba-Taguimdje et al., 2020, p. 3), while most of the literature did not focus on GAI nor considered the extensive advances in the past two years or adopted highly industry-specific perspectives. From this arises a gap for a cross-industry explanatory approach, investigating which capabilities are needed to use GAI successfully. The corresponding research question is: "What capabilities are needed for the successful usage of GAI?".

To answer this research question, an understanding of current developments within the GAI landscape was elaborated, business use cases for GAI were investigated, and cross-industry GAI capabilities consisting of use case compatibility, organizational design, advantageous resources, and compliance robustness were identified.

This was achieved with an initial literature review, while primarily the qualitative research method of expert interviews led to valuable insights. The interviews with eleven GAI domain-experts ensured multi-faceted cross-industry perspectives of corporates, SMEs, startups, consultants and investors.

The thesis also includes a procedure for the general identification of cross-industry capabilities for emerging technologies (e.g., also applicable for quantum computing or biotechnology) (Verfassungsschutz, n.d.). Since there is only limited traditional academic literature, it is often necessary to incorporates gray literature, which includes news articles, market analyses, industry reports, and other non-peer-reviewed materials to build up a knowledge base. Finally, interviews are used to accumulate first-hand insights from experts. An interview guide can be created with the theory of dynamic capabilities. Finally, the transcription of the collected audio data, to transform qualitative spoken content into a written format, preserving the richness of participant responses, was conducted (Bailey, 2008, p. 3). The subsequent data analysis based on deductive and open codes completed the methodology and led to the identification of cross-industry capabilities.

In addition to being applied to other emerging technologies, a similar approach could also be used to analyze other areas of AI. The resulting findings can also be applied in practice for many cases. In principle, they are addressing two groups: companies across all industries that want to use GAI and providers of GAI applications.

Companies that want to use GAI within their business processes are provided with a handbook that highlights important GAI capabilities and addresses various aspects that are relevant. As these were developed by interviewing industry experts, the work also provides a summary of previous practical learnings and experiences. The investigation of current developments and market dynamics provides a foundational understanding of GAI advancements. A comprehensive list of potential use cases helps to identify new areas for application, while processes for creating a use case portfolio and metrics for selection are outlined as part of the use case compatibility. Additionally, a dedicated chapter focuses on organizational design for successfully implementing GAI in companies, offering inspiration for new entities. The chapter of advantageous resources also enables companies to assess where strong progress has been made and which areas require more investments. Lastly, emphasis is placed on ensuring compliance robustness to avoid pitfalls like fines or data loss in the implementation and usage of GAI. Thus, the cross-industry explanatory approach for GAI capabilities can serve as a foundation for the decision-making process of C-level executives and managers in formulating and improving AI strategies.

The other side consists of GAI application providers who offer off-the-shelf solutions, which are sometimes customizable to a certain extent. Through various discussions with industry experts and the development of GAI capabilities, perspective of what GAI providers need to deliver to be of interest to potential customers were discovered. Simultaneously, insights on what to consider when developing GAI can be useful to GAI providers.

GAI providers often enter long-term partnerships with their respective customers and should identify who will manage the application within the company (e.g., a product manager) (I2, I7). Furthermore, they should reflect which digital maturity the company possesses to determine how their solution can be implemented in the company (I3). They should also consider how much they want to customize their GAI applications within companies, or whether a standardized off-the-shelf solution is offered (I9). GAI providers should know who their first point-of-contact in the company is (e.g., a certain department or the core team) and what quantifiable value they deliver with their solution (I3, I4, I10). Besides that, they must reflect what customers want to achieve, what requirements exist, what timelines are present, and how to avoid moving targets (I9, I10). Meanwhile, the pilot trap must also be avoided (I9, I10, I11). GAI applications should offer compliance robustness and data security (e.g., with an orchestration layer), if necessary (I8). At the same time, consideration should be given to whether the underlying data, models, weights and components should also be disclosed transparently to customers (AI-SBOM) (I8). At the same time, GAI providers should consider how they can create a certain defensibility, or a competitive advantage compared to other GAI solutions (I5). As public data is available to everyone, they should consider whether customer data may be used to improve their own solution and evaluate if data crawling and scraping should be used at all, as these represent a legal gray zone (I5, I8).

6.3. Limitations and Future Research

As traditional academic literature is often limited regarding emerging technologies, the chosen methodology of this thesis consisted of expert interviews, to obtain first-hand perspectives and insights. When the available literature lacks comprehensive insights, interviews serve as a valuable tool to obtain additional knowledge. However, it is crucial to recognize certain limitations within the research method, especially concerning the limited sample group of only eleven domain experts. This small sample size may restrict the generalizability of results (Alshenqeeti, 2014, p. 43). It should be noted that certain capabilities were mentioned across the interviews (e.g., data security), while other capabilities were only emphasized by individual experts of certain domains (e.g., AI-SBOM). For that reason, further discussions with experts should be sought to verify such capabilities in more depth. In addition, international interviewees criticized the regulations in Europe and lower cloud usage compared to the USA (I5, I7). It is worth noting that more neutral perspectives on the regulations were also gathered (I9, I0, I11). However, it would have been valuable to engage with additional experts who could have highlighted the disadvantages of cloud platforms, as well.

Besides that, interviews can inherent a certain bias as they are carried out through human interactions. Thus, biases can be unconscious, as they are shaped by personal experiences and cultural background (Alsaawi, 2014, p. 153). Moreover, the interviews were conducted predominantly with experts located in Germany, with the exceptions of three international interviewees from France (I6), Austria (I7), and the USA (I5).

Henceforth, the cross-industry explanatory approach cannot be completely transferred to other countries and markets. Besides that, the evolving field of AI always exhibits new developments, trends and dynamics. Furthermore, the cross-industry explanatory approach is generalizing in its structure and cannot cover all industry-specific requirements.

Lastly, certain limitations regarding the analysis of market developments should be considered. The exact calculations for market forecasts, sizes, and CAGRs were not always traceable. The gray literature consisted of reports from commercially driven consulting and market research firms, which often limit access to paying clients and do not consistently disclose their research or data collection methods.

Over the last decade, research activities related to AI have doubled, with notable progress in GAI and a rapidly expanding market, as more companies integrate GAI into their operations (AI Index Stanford, 2023, p. 11; McKinsey, n.d.; Statista, n.d.). This thesis provides a valuable contribution to the understanding of the necessary capabilities that companies need to use GAI.

However, the level of abstraction of this thesis is relatively high, as various experts from different sectors were interviewed. As a result, initial results for cross-industry capabilities were elaborated. To validate the results, a higher number of experts could be interviewed. Additionally, future research could incorporate quantitative methods to complement the qualitative expert interviews.

Moreover, it would be interesting to intensify the research on certain capabilities to derive more specific recommendations by interviewing a representative number of experts from this area alone. One example of this would be the capability of compliance robustness, which could be researched in more detail. As the EU AI Act has already been passed but has yet to be transposed into national legislation (Bundesregierung, 2024), it would be insightful to interview experts and elaborate on the individual risk classes to analyze the capability in more detail. Such deep dives into certain capabilities could further validate and make them more tangible for practical application.

Following past research, these cross-industry GAI capabilities could also be used to investigate industry-specific capabilities as well. This could include, for example, interviewing various experts such as data scientists, consultants or compliance professionals from the automotive industry.

Subsequently, the entire ethical aspect, which is becoming increasingly important (AI Index Stanford, 2023, p. 13), could also be examined and possible effects on society and the labor market could be studied in greater depth. In addition, it would also be exciting to carry out case studies with the GAI capabilities to examine how these can be validated through practical application. Finally, a comparable analysis would also be possible for international markets, e.g. in Asia or America. Nevertheless, the topic of energy consumption of GAI should also be addressed, as it has an increasing an impact on the environment (I3).

6.4. Conclusion

Research into AI has been growing rapidly in the last decade, while GAI has been gaining momentum in recent years and has seen widespread use (AI Index Stanford, 2023, p. 11; Mondal et al., 2023, p. 3). A growing number of companies are implementing GAI, while considering its far-reaching future implications (McKinsey, n.d.). Nevertheless, existing literature could not provide an answer about which capabilities are particularly important to use GAI successfully.

For that reason, eleven interviews were conducted with various GAI experts from globally leading companies, internationally renowned consultancies, venture capital funds, and technology startups with a focus on software development, strategic advisory, compliance and investing. The interviews were complemented by a robust data analysis procedure. After the extensive identification of current market players and GAI developments, the analysis revealed four key capabilities that are essential for the successful usage of GAI.

The first capability of use case compatibility comprises the fundamental understanding and continuous learning of (G)AI, which leads to the identification of an appropriate use case portfolio, a roadmap for implementation and a make-or-buy decision regarding each use case. Following that, the organizational design is essential, as the competitive advantage of a specific company arises from its organizational processes. Therefore, C-level commitment must be present, while the organization must be expanded by new entities such as a core team, a technical inhouse consulting team for implementation, a digital hub, and AI product managers. Furthermore, the organizational implementation must avoid the pilot trap to scale GAI use cases company-wide. Moreover, there are advantageous resources that already exist or can be built up in companies. These include first mover advantages if helpful decisions have been made in the past regarding AI, influencing learnings and experiences, upskilling efforts, flexibility and timing. Equally relevant is the capability to develop digital maturity and human capital. In addition, partnerships with technology providers can also leverage the advantageous resource capability. Finally, attention should be paid to compliance robustness so that regulatory proofness, transparency about the AI-SBOM and data security are ensured.

In conclusion, the underlying thesis examined which key capabilities are necessary for the successful usage of GAI within companies and “suggest[s] overall direction[s]” (Teece et al., 1997, p. 528). Thus, the established key capabilities for GAI can help C-level executives and managers in formulating GAI strategies and bridge the existing research gap with a cross-industry explanatory approach.

Appendix

A. Interview Guideline

1. Introduction (5 mins)

Introduction and further explanation of the purpose of the interview. Clarifying organizational details (e.g., recording, time restrictions) and structure of the interview.

1.1 Interview objective and structure

1. **Interviewer:** Brief introduction to each other
2. **Further explanation of the purpose of the interview:**
    a. Experts' perspective on capabilities to successfully use GAI.
 	
Clarifying organizational details:
 	
3. **Interview Duration**
    a. Planned time frame 45 mins (scheduled 60 mins, incl. buffer)
4. **Interview Structure**
    a. Introduction
    (Interviewee background & experience, conceptual basis for the research subject)
    b. Question phase on the interviewee's experience with GAI to identify certain capabilities that might be an indicator for successful usage.
    c. Deep dive into usage aspects mentioned (e.g. challenges, risks, implementation, examples) as needed.
    d. Closure/ Next Steps
    e. This agenda is only an orientation. Feel free to ask questions anytime or interrupt me to share your view and address further aspects that might be important.
5. **Interview (Audio) Recording**
    a. The interview will be recorded.
    b. All data will be **anonymized** and will **not allow any conclusions** regarding individual persons or companies.
6. **Study results will be provided if you wish.**
 	
Do you have any further questions or requests for the interview?

1.2 Bridging the knowledge gap on how to use GAI

Generative Artificial Intelligence (GAI) represents a groundbreaking technology that has significantly impacted global business dynamics since its breakthrough in late 2022. The swift integration of applications like ChatGPT highlights the pressing need for companies to adapt to these technological changes to stay competitive. This thesis investigates capabilities for successful GAI usage within organizations. By examining the current market landscape and exploring how companies can effectively respond to the technological disruptions caused by GAI, the research aims to develop a comprehensive explanatory approach to guide executives and managers in formulating AI strategies. Therefore, expert interviews will be conducted with industry leaders, including corporate managers, CEOs, founders, investors, and consultants who are actively usage GAI or deal with strategic decisions regarding it. These interviews will uncover best practices, challenges, and strategic approaches for leveraging GAI for business growth. The findings will contribute to an explanatory approach that identifies key capabilities to successfully use GAI, providing actionable guidance for navigating the GAI landscape.

1.3 Personal Introduction/Who is in the room?

Start Recording

The interviewee's background and experience are gathered to understand the interviewee's general attitude towards the GAI usage in their job role.

- **Interviewee:** Corporate executives, founders, investors, and consultants with at least 1-2 years of professional experience in related fields such as (Generative) Artificial Intelligence, Large Language Models, etc.
- **Demographics**
  - Job title
  - Industry and Company Size
  - Years of professional experience
  - Working experience/ experience with GAI, perception of GAI (e.g. how would you rate the importance of GAI?)
  - Location

2. Question Phase (30 mins)

a) **General**

- What are your general thoughts about GenAI in the business context?
- What are your business experiences with GenAI?
- What GenAI business use cases have you experienced?

- In your experience, who are usually the end users of GenAI applications?

b) Coordination

- What is the design of your organizational processes to use GenAI applications successfully and who is involved?
- What is your experience with certain organizational routines that increase performance in GenAI usage?

c) Learning

- How can organizations iterate and learn fast during the integration process of GenAI?
- What relevant skills and routines are needed to enable organizational learning?

d) **Transformation**

- Have you witnessed any efforts to align and invest in GenAI expertise and resources? What exactly do these efforts look like?
- Had your company (*your clients*) to transform due to disruption in the past? Is this exexperience providing any advantages in using GenAI?
- Has your company (*your clients*) a high degree of flexibility which brings advantages in using GenAI?

e) Asset Positions

- Are there assets that your company (*your clients*) already posesses in other areas that could also be useful for GenAI (complementary assets)?
- *How can your firms* (*your clients*) *ensure ownership protection of its assets (IP), or can it even be attractive for you* (*your clients*) *to reduce ownership protection of others to leverage GenAI?*
- *How is the regulatory environment influencing the usage of GenAI? What skills are needed to handle it?*

f) Path Dependencies

- Have you witnessed past decisions that provide a first mover advantage in terms of GenAI usage?
- Have you witnessed wrong historical decisions that represent a challenge for GenAI usage?

g) Business Model

- How are managers identifying opportunities for using GenAI applications early on to increase profitability (*sense*)?
- What business models are relevant to use GenAI (*seize*)?
- How are managers securing and committing resources for GenAI projects (*seize*)?

3. Closing (5 mins)

Conclusion and Next Steps
• Over the next month, we will conduct further stakeholder interviews and inform you about the evaluation and results if you wish. • Are there any colleagues or contacts in your network who might also be interested in contributing to our research by providing insights during an interview? • How do you feel about the interview? Are there any further questions or suggestions/ recommendations you would like to make regarding the interview itself or the overall study? *Thank you for your time and the valuable insights provided!*
Stop Recording

B. Global Market Sizes of GAI with Forecasts

Source	**Start**	**End**	**CAGR**
Statista (n.d.)	USD 36.06 billion (2024)	USD 356.1 billion (2030)	46.47%
Markets and Markets (2024)	USD 20.9 billion (2024)	USD 136.7 billion (2030)	36.7%
Dimension Market Research (2023)	USD 22.2 billion (2023)	USD 488.1 billion (2033)	41%
Bloomberg (2023)	USD 40 billion (2022)	USD 1.3 trillion (2032)	42%
Grand View Research (2024a)	USD 13 billion (2023)	USD 114.78 billion (2030)	36.5%
Fortune Business Insights (2024)	USD 67.18 billion (2024)	USD 967.65 billion (2032)	49.78%

C. Aggregated GAI Capabilities with the Respective Building Blocks

GAI capabilities	Description of the building blocks	Citations
Use case compatibility	**Fundamental understanding and continuous learning:** Acquire and cultivate general expertise in AI (particularly GAI) through upskilling or hiring new employees. Understand the technological principles, challenges, and limitations of GAI compared to other AI technologies to decrease false expectations. Follow continuously new developments and build a network to exchange with industry experts, initiatives, legal advisors, decision-makers, researchers, competitors and technology providers.	I1, I2, I3, I4, I5, I7, I8, I9, I10, I11
	Identify portfolio: Identify which use cases exist and if they can be solved with GAI. Develop a use case portfolio (long list of use cases). Examine what others are implementing and, if necessary, get support from internal or external experts to identify more use cases. Describe the use cases in detail and translate broad concepts into specified requirements, pain points, and milestones to avoid moving targets.	I9, I10, I11
	Prioritized roadmap: Evaluate use cases by quantifying their ROI (calculate the total costs of infrastructure, foundation model, licenses, usage-based pricing, maintenance and compare them to the efficiency gains). Rank all use cases according to their quantified value (ROI) or other criteria (e.g., ease of implementation) and create a prioritized roadmap with waves of implementation over the next two to three years.	I4, I9, I11
	Make-or-buy decisions: Understand the differences between off-the-shelf GAI applications and inhouse solutions and consider the underlying foundation models. Decide whether you build an inhouse solution or buy an off-the-shelf GAI application for a use case, and to which extent you want to customize it. Consider proprietary resources, challenges, risks, total costs, requirements and objectives. If building an inhouse solution is the preferred choice, evaluate advantages and disadvantaged of closed-source and open-source foundation models.	I2, I3, I4, I7, I8, I9, I11

Organizational design	**C-level commitment:** The combination of technology and effective organization is a key factor for successful usage of GAI and can create far-reaching efficiency gains for the respective company. C-level executives and the board should have courage and prioritize AI while providing the resources to enable it. Trying different solutions through a trial-and-error approach should be encouraged. C-level executives and the board should remove initial barriers that require top-down decisions. This includes formulating a strategy, that considers ethical, compliance and technological feasibility. Defining such a strategy can incorporate internal or external domain experts. The strategy needs to be aligned with the use cases (see use case portfolio) and subsequent implementation of these use cases (according to the roadmap) must fulfill the milestones to achieve the overall strategy.	I1, I2, I3, I5, I9, I11
	Core team: Establish a core team as the central point of contact for all topics regarding (G)AI, which directly reports to the board and receives instructions from the C-level. It is superordinate to all departments and cross-functional, as it pools domain experts from various departments who dedicate a certain amount of their capacity to the core team. Completely organized along the use cases, it serves as a catalyst for top-down and bottom-up innovation regarding GAI use cases. The core team evaluates use cases, leveraging various perspectives (strategy, compliance, software engineering). After a decision is made, tasks are cascaded throughout the organization to fulfill concrete steps of the roadmap.	I1, I2, I3, I7, I9,
	Implementation: Enable departments to implement use cases independently with internal resources. Establish operational support through a technical inhouse consulting team of data scientists, if needed. As these additional services affect the budget of the respective department, they should also have the alternative of choosing external consulting partners. Continuous upskilling increases the abilities for future implementations.	I1, I2, I3, I9
	Digital Hub: Form a digital hub, initially managed by the core team, to provide an education and learning platform where experts and users meet. This can ensure scalability, through answering frequently asked questions via digital content (videos, intranet posts, articles, blogs, podcasts, tutorials). Simultaneously, it guides employees how to use GAI, offers more acceptance and upskilling. The digital hub can also serve as a community or network, where like-minded users can interact and exchange best practices to promote grassroot movements and bottom-up innovation.	I1, I2, I3, I9, I11

	Pilot trap avoidance: To ensure successful realization of the use cases a scalable structure must be defined from the beginning promoting rollout capabilities. Transitioning from an initial pilot to a company-wide rollout is a challenge. Therefore, building a scalable framework and identifying high impact use cases is crucial. The development of a specific use case requires strategic and technical skills and test-and-learn loops for iterative improvement. The development follows the rapid-prototyping approach with multiple feedback loops to build user trust through early results. Furthermore, it ensures continuous evaluation of new features, requirements, and pain points, legacy systems, and internal data management. With that approach, the prototype develops through extensive testing prior to launch into a final product. Simultaneously, additional use cases can be implemented. During this, developers must manage timelines successfully, as often decision-making processes take time and results are expected quickly.	I3, I8, I9, I10, I11
	Product Management: Assign dedicated product managers to launch, monitor and maintain GAI applications. Effective product management is a key factor to ensure the ongoing success of the respective applications. Product managers control the interface of underlying technology and the user facing product. They must minimize deviations between performance and user expectations. As GAI requires additional knowledge that not all existing product managers possess, upskilling existing and hiring new product managers is essential.	I7, I9
Advantageous resources	**First mover advantage:** Companies that already used AI technologies (as machine learning and robotics) gained a faster understanding of GAI right from the start, could transfer upskilling efforts of employees, and best practices in implementing use cases and rollouts. As they have already gained learnings and experiences in dealing with AI. Furthermore, an already established network can be used and expanded. It is an advantage to define the strategic direction and the key objective early on. The right timing is crucial and increases the fundamental understanding and continuous learning regarding GAI. Nevertheless, moving early on requires financial resources. However, companies that did not prioritize AI at an early stage, can still develop into this direction through training employees. Furthermore, the fast advancements enable such companies to start now with state-of-the-art foundation models, which perform better and cost less. There are also companies that are very flexible and can adapt their organizations fast or have a lot of resources	I1, I2, I3, I7, I9

	(customer data) and thus quickly catch up with the first movers. Nevertheless, the first mover advantages could also be observed with the internet and the cloud, both technologies which represent the indispensable foundation for GAI.	
	Digital maturity: The digital maturity of the company must be assessed, including the cloud platforms (the backbone for GAI inhouse solutions). Anything extending one individual use case must be scalable and requires a cloud platform that offers general services and infrastructure (e.g., AWS, Microsoft Azure, and Google Cloud). These cloud platforms provide various foundation models and development tools. With state-of-the-art cloud platform, data scientists can work effectively with the latest models and tools in an iterative test and learn approach. Digital maturity encompasses technological standards for the effective use of GAI, including efficient data management, and precise fine-tuning. Additionally, collecting proprietary business and customer data provides a competitive advantage for training and refining GAI applications. However, many companies have a low level of digital maturity and struggle to integrate GAI into existing legacy system. Therefore, the level of digital maturity combined with human capital is essential, as if it is too low, GAI initiatives will most likely fail.	I1, I2, I3, I4, I5, I7, I8, I9, I11
	Human capital: Even if skilled data scientists are in shortage, it is important to hire and retain them in the long term. A balanced team of data scientists, including software engineers and AI experts, is crucial for GAI applications. Moreover, it is essential to have interdisciplinary human capital including data scientists, product managers, and legal or compliance experts. The combination of technical and legal expertise has become important as companies must interpret an increasing number of regulations. This interdisciplinary human capital is useful to build the *core team* or simplify the *implementation* through a technical inhouse consulting team and the necessary knowledge within departments. Furthermore, key employees and C-level executives need further training to develop proprietary knowledge within the company.	I2, I3, I7, I9
	Partnerships: Partnerships are an important resource, as they expand fundamental understanding and continuous learning, can provide first mover advantages, and inspire to improve the technology stack. This includes partnering with providers of foundation models, cloud platforms, external consulting services, and GAI applications. Chosen technology partners (depending on the contracts) often remain long-term commitments. As the proprietary knowledge of such	I1, I2, I4, I7, I8, I9, I11

	technology partnerships (e.g., trained foundation model) is hardly transferable if the respective partner is changed, organizational efforts, sunk costs and switching costs arise. Thus, it is important to partner with providers that remain competitive in the long-term and committed to update their technology constantly, while providing excellent service.	
Compliance robustness	**Regulatory proofness:** Prioritize established and pending regulatory compliance regarding (G)AI. Analyze the EU AI Act and understand relevant AI risk classes. Also, reflect if established or planned use cases (on the roadmap) may face future regulatory restrictions or prohibition. Examine the requirements for your use cases and adapt to become AI Act compliant. Especially use cases where personally identifiable information is involved are sensitive. Also, training foundation models with proprietary data will require a certain representativeness within the data set, meaning that a limited customer base may not meet regulatory compliance for representative data. Prepare for reporting obligations to increase transparency towards authorities. Implementing the EU AI Act will also have a transitional phase (like GDPR) until the requirements must be met. Also pay attention to GDPR compliance, especially around data privacy and storage. Consider further recommendations such as the NIST AI risk management framework, which lists helpful aspects for trustworthy AI and ISO/IEC 42001 which could also develop into a market-standard. Create a clear framework for your company, with internal requirements and conditions to guide developers. Establish use case management processes with AI checklists and internal reports to summarize all relevant information regarding the GAI applications, also for future reporting obligations.	I1, I2, I7, I8,
	AI-SBOM: Prioritize licensing rights for open-source components in GAI applications, to avoid being sued if used commercially. Establish the concept of the AI-SBOM, describing all components and the respective licenses with which the software is built. This includes model weights and data with which the foundation model was trained as well. While public data is available to everyone, crawling and scraping data is a legal gray zone. This must be considered for inhouse solutions, but also for external GAI, as customers often do not know which software components and data sources are hidden underneath the application. Thus, customers are often unaware of which foundation models are used by a certain layer provider, where the training data originated, how it was collected, and if copyright-protected content was used.	I8

	Although providers are often reluctant to disclose their AI-SBOM, companies must decide whether to accept this risk or seek alternative solutions.	
	Data Security: Data security can be endangered when sensitive data is used for training and prompting, leading to leak outs into publicly available foundation models in the further process. Henceforth, sensitive and malicious data must be cleaned up beforehand, as deleting it afterward is not possible. Companies can implement additional layers, which filter the inputs of sensitive data before they flow into the foundation model. Nevertheless, employees should be properly trained and informed about guidelines to minimize the risk of data loss. To decrease the trade-off between working efficiently and protecting IP, compliant alternatives must be offered. Many larger companies with appropriate resources can develop inhouse solutions covering data protection concerns, while smaller companies often turn to GAI applications from external providers. However, data access regarding external GAI applications must be carefully reviewed and narrowed down through access to predefined shared folders with role-based permissions. Existing and future contracts, including their terms and conditions, should be reviewed to determine whether customer data can be used for training foundation models. Companies need to consider, where foundation models and the respective data are hosted and can increase data security additionally by hosting on EU or local servers only. In addition, it is important to have security whether GAI applications use customer data to train the entire foundation model, or only for the individual customer, so that proprietary data of one customer is not touching other data. To provide data security, a software orchestration layer can be implemented which includes monitoring prompt injection. This prevents hackers from bypassing security mechanisms by crafting clever prompts designed to "jailbreak" the model. Such activities aim to extract sensitive or restricted information that should not be disclosed. A software orchestration layer includes elaborated guardrails to monitor incoming prompts. It is trained with malicious prompts to recognize them and prohibit the use. This can also avoid bot attacks which inject high numbers of prompts to generate API call costs for the respective provider. Besides that, toxicity in the outputs can occur and should be blocked. Thus, an observability layer must be developed to constantly track the number and quality of queries.	I2, I3, I4, I5, I8, I9, I10, I11

References

Abou-Foul, M., Ruiz-Alba, J. L., & López-Tenorio, P. J. (2023). The impact of artificial intelligence capabilities on servitization: The moderating role of absorptive capacity-A dynamic capabilities perspective. *Journal of Business Research*, *157*, 113609. https://doi.org/10.1016/j.jbusres.2022.113609

Abraham, C., Boudreau, M.-C., Junglas, I., & Watson, R. (2013). Enriching our theoretical repertoire: The role of evolutionary psychology in technology acceptance. *European Journal of Information Systems*, *22*(1), 56–75. https://doi.org/10.1057/ejis.2011.25

Adept. (2022). *Introducing Adept*. https://www.adept.ai/blog/introducing-adept/

AI Index Stanford. (2023). *AI Index Report 2023 – Artificial Intelligence Index* (pp. 1–386) [Market Report]. Standford University. https://aiindex.stanford.edu/wp-content/uploads/2023/04/HAI_AI-Index-Report_2023.pdf

AlephAlpha. (2023). *Aleph Alpha reaches first milestone on the way to content-correct, explainable and trustworthy AI - ALEPH ALPHA - AI for Enterprises and Governments*. ALEPH ALPHA. https://aleph-alpha.com/aleph-alpha-reaches-first-milestone-on-the-way-to-content-correct-explainable-and-trustworthy-ai/

AlephAlpha. (2024). *What is Pharia-1-LLM-7B? | Aleph Alpha API*. https://docs.aleph-alpha.com/docs/introduction/what-is-pharia-1-llm/

Alsaawi, A. (2014). A Critical Review of Qualitative Interviews. *European Journal of Business and Social Sciences*, *3*(4), 149–156. https://doi.org/10.2139/ssrn.2819536

Alshenqeeti, H. (2014). Interviewing as a Data Collection Method: A Critical Review. *English Linguistics Research*, *3*(1), 39–45. https://doi.org/10.5430/elr.v3n1p39

Anthropic. (2023). *Claude 2*. https://www.anthropic.com/news/claude-2

Arthur, W. B. (1989). Competing Technologies, Increasing Returns, and Lock-In by Historical Events. *The Economic Journal*, *99*(394), 116. https://doi.org/10.2307/2234208

AWS. (n.d.). *Generative KI: Anwendungsfälle und Ressourcen – AWS*. Amazon Web Services, Inc. Retrieved October 6, 2024, from https://aws.amazon.com/de/ai/generative-ai/use-cases/

Bailey, J. (2008). First steps in qualitative data analysis: Transcribing. *Family Practice*, *25*(2), 127–131. https://doi.org/10.1093/fampra/cmn003

Banks, J., & Warkentin, T. (2024, February 21). *Gemma: Introducing new state-of-the-art open models*. Google. https://blog.google/technology/developers/gemma-open-models/

Bloomberg. (2017). *A new era: Artificial intelligence is now the biggest tech disrupter*. https://www.bloomberg.com/professional/insights/trading/new-era-artificial-intelligence-now-biggest-tech-disrupter/

Bloomberg. (2023). Generative AI to Become a $1.3 Trillion Market by 2032, Research Finds | Press | Bloomberg LP. *Bloomberg L.P.* https://www.bloomberg.com/company/press/generative-ai-to-become-a-1-3-trillion-market-by-2032-research-finds/

Brenner, W., Van Giffen, B., & Koehler, J. (2021). Management of Artificial Intelligence: Feasibility, Desirability and Viability. In S. Aier, P. Rohner, & J. Schelp (Eds.), *Engineering the Transformation of the Enterprise* (pp. 15–36). Springer International Publishing. https://doi.org/10.1007/978-3-030-84655-8_2

Bucher, R., Fritz, C. E., & Quarantelli, E. L. (1956). Tape Recorded Interviews in Social Research. *American Sociological Review*, *21*(3), 359–364. https://doi.org/10.2307/2089294

Bundesregierung. (2024, May 22). *EU verabschiedet erstes KI-Gesetz weltweit | Bundesregierung*. Die Bundesregierung informiert | Startseite. https://www.bundesregierung.de/breg-de/themen/digitalisierung/kuenstliche-intelligenz/ai-act-2285944

Casadesus-Masanell, R., & Ricart, J. E. (2011). *How to Design a Winning Business Model*. Harvard Business Review. https://hbr.org/2011/01/how-to-design-a-winning-business-model

Climent, R. C., Haftor, D. M., & Staniewski, M. W. (2024). AI-enabled business models for competitive advantage. *Journal of Innovation & Knowledge*, *9*(3), 100532. https://doi.org/10.1016/j.jik.2024.100532

Cohere. (2024a). *Cohere For AI Launches Aya 23, 8 and 35 Billion Parameter Open Weights Release*. Cohere. https://cohere.com/blog/aya23

Cohere. (2024b). *Cohere For AI Launches Aya, an LLM Covering More Than 100 Languages*. Cohere. https://cohere.com/blog/aya

Cohere. (2024c). *Introducing Command R+: A Scalable LLM Built for Business*. Cohere. https://cohere.com/blog/command-r-plus-microsoft-azure

Dealroom. (2024). *Dealroom.co | Identify promising companies before everyone else*. Dealroom.Co. https://dealroom.co/

DeMarrais, K., & Lapan, S. D. (Eds.). (2004). *Foundations for research: Methods of inquiry in education and the social sciences* (1st ed.). Lawrence Erlbaum.

Dimension Market Research. (2023). *Generative AI Market Growth, Size, Share, Trends and Forecast 2032*. https://dimensionmarketresearch.com/report/generative-ai-market/

Douglas, M. R. (2023). *Large Language Models*. 1–47. https://doi.org/10.48550/arXiv.2307.05782

Drydakis, N. (2022). Artificial Intelligence and Reduced SMEs' Business Risks. A Dynamic Capabilities Analysis During the COVID-19 Pandemic. *Information Systems Frontiers*, *24*(4), 1223–1247. https://doi.org/10.1007/s10796-022-10249-6

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its Nature, Scope, Limits, and Consequences. *Minds and Machines*, *30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Forbes. (2024). *OpenAI's 5 Levels Of 'Super AI' (AGI To Outperform Human Capability)*. https://www.forbes.com/sites/jodiecook/2024/07/16/openais-5-levels-of-super-ai-agi-to-outperform-human-capability/

Fortune Business Insights. (2023). *Generative AI Market Size, Share & Trends Analysis* (pp. 1–120)

[Market Report]. Fortune Business Insights. https://www.fortunebusinessinsights.com/generative-ai-market-107837

Fortune Business Insights. (2024). *Generative AI Market Size, Share | Research Report [2032]*. https://www.fortunebusinessinsights.com/generative-ai-market-107837

Fosso Wamba, S., Queiroz, M. M., & Trinchera, L. (2024). The role of artificial intelligence-enabled dynamic capability on environmental performance: The mediation effect of a data-driven culture in France and the USA. *International Journal of Production Economics*, *268*, 109131. https://doi.org/10.1016/j.ijpe.2023.109131

Gallego-Gomez, C., & De-Pablos-Heredero, C. (2020). Artificial Intelligence as an Enabling Tool for the Development of Dynamic Capabilities in the Banking Industry: *International Journal of Enterprise Information Systems*, *16*(3), 20–33. https://doi.org/10.4018/IJEIS.2020070102

Geirhofer, S., & McKinney, S. (2023). *Foundational Models: Building Blocks for Generative AI Applications* (TechREG CHRONICLE, p. 7) [Technology Report]. Wilson Sonsini Goodrich & Rosati. https://www.pymnts.com/wp-content/uploads/2023/09/2-FOUNDATIONAL-MODELS-BUILDING-BLOCKS-FOR-GENERATIVE-AI-APPLICATIONS-Stefan-Geirhofer-Scott-McKinney.pdf

Grand View Research. (2024). *Large Language Model Market Size And Share Report, 2030*. https://www.grandviewresearch.com/industry-analysis/large-language-model-llm-market-report

Hercheui, M., & Ranjith, Rishikesh. (2020). Improving Organization Dynamic Capabilities Using Artificial Intelligence. *Global Journal of Business Research*, *14*, 104.

Ho, J. C., & Chen, H. (2018a). Managing the Disruptive and Sustaining the Disrupted: The Case of Kodak and Fujifilm in the Face of Digital Disruption. *Review of Policy Research*, *35*(3), 352–371. https://doi.org/10.1111/ropr.12278

Ho, J. C., & Chen, H. (2018b). Managing the Disruptive and Sustaining the Disrupted: The Case of Kodak and Fujifilm in the Face of Digital Disruption. *Review of Policy Research*, *35*(3), 352–371. https://doi.org/10.1111/ropr.12278

Hsieh, H.-F., & Shannon, S. E. (2005). Three Approaches to Qualitative Content Analysis. *Qualitative Health Research*, *15*(9), 1277–1288. https://doi.org/10.1177/1049732305276687

HuggingFace. (n.d.). *Hugging Face – Pricing*. Retrieved June 18, 2024, from https://huggingface.co/pricing

HuggingFace. (2024, March 13). *IBM NASA Geospatial*. https://huggingface.co/ibm-nasa-geospatial

IOT Analytics. (2023, December 14). *Generative AI Market Report 2023–2030*. IoT Analytics. https://iot-analytics.com/wp/wp-content/uploads/2023/12/INSIGHTS-RELEASE-The-leading-generative-AI-companies.pdf

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, 1–9. https://doi.org/10.1016/j.lindif.2023.102274

Law, M. (2023, April 27). *Top 10 companies offering AI solutions over the cloud*. https://aimagazine.com/top10/top-10-companies-offering-ai-solutions-over-the-cloud

Leaver, T., & Srdarov, S. (2023). ChatGPT Isn't Magic: The Hype and Hypocrisy of Generative Artificial Intelligence (AI) Rhetoric. *M/C Journal*, *26*(5). https://doi.org/10.5204/mcj.3004

Lichtenthaler, U. (2020). Five Maturity Levels of Managing AI: From Isolated Ignorance to Integrated Intelligence. *Journal of Innovation Management*, *8*(1). https://doi.org/10.24840/2183-0606_008.001_0005

Liu, Y. (David), Sun, J., Zhang, Z. (Justin), Wu, M., Sima, H., & Ooi, Y. M. (2024). How AI Impacts Companies' Dynamic Capabilities: Lessons from Six Chinese Construction Firms. *Research-Technology Management*, *67*(3), 64–76. https://doi.org/10.1080/08956308.2024.2324407

Luitse, D., & Denkena, W. (2021). The great Transformer: Examining the role of large language models in the political economy of AI. *Big Data & Society*, *8*(2), 1–14. https://doi.org/10.1177/20539517211047734

Maleki, N., Padmanabhan, B., & Dutta, K. (2024). AI Hallucinations: A Misnomer Worth Clarifying. *2024 IEEE Conference on Artificial Intelligence (CAI)*, 133–138. https://doi.org/10.1109/CAI59869.2024.00033

Markets and Markets. (2024). *Generative AI Market Size, Share, Trends, Growth Forecast—2030*. MarketsandMarkets. https://www.marketsandmarkets.com/Market-Reports/generative-ai-market-142870584.html

Marr, B. (2023, May 30). *10 Amazing Real-World Examples Of How Companies Are Using ChatGPT In 2023*. Forbes. https://www.forbes.com/sites/bernardmarr/2023/05/30/10-amazing-real-world-examples-of-how-companies-are-using-chatgpt-in-2023/

McKinsey. (n.d.). *The state of AI in 2023: Generative AI's breakout year*. Retrieved May 6, 2024, from https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year

Meta. (n.d.). *Llama 2*. Meta AI. Retrieved January 16, 2024, from https://ai.meta.com/llama-project

Mikalef, P., Conboy, K., & Krogstie, J. (2021). Artificial intelligence as an enabler of B2B marketing: A dynamic capabilities micro-foundations approach. *Industrial Marketing Management*, *98*, 80–92. https://doi.org/10.1016/j.indmarman.2021.08.003

MistralAI. (n.d.). *Technology*. Retrieved June 18, 2024, from https://mistral.ai/technology/

Mondal, S., Das, S., & Vrana, V. G. (2023). How to Bell the Cat? A Theoretical Review of Generative Artificial Intelligence towards Digital Disruption in All Walks of Life. *Technologies*, *11*(2), 44. https://doi.org/10.3390/technologies11020044

Morris, M. R. (2023). Scientists' Perspectives on the Potential for Generative AI in their Fields. *arXiv*, 1–26. https://doi.org/10.48550/arXiv.2304.01420

Myers, M. D., & Newman, M. (2007). The qualitative interview in IS research: Examining the craft. *Information and Organization*, *17*(1), 2–26. https://doi.org/10.1016/j.infoandorg.2006.11.001

NIST. (2021, July 12). *AI Risk Management Framework*. NIST. https://www.nist.gov/itl/ai-risk-management-framework

OpenAI. (n.d.-a). *GPT-4*. Retrieved January 19, 2024, from https://openai.com/gpt-4

OpenAI. (n.d.-b). *Pricing*. Retrieved June 8, 2024, from https://openai.com/api/pricing/

OpenAI. (2022). *Introducing ChatGPT*. https://openai.com/index/chatgpt/

Oracle. (2024). *Supercharged Cloud: How AI is Revolutionizing Cloud Computing*. https://www.oracle.com/artificial-intelligence/ai-cloud-computing/

Pichai, S., & Hassabis, D. (2024, February 15). *Our next-generation model: Gemini 1.5*. Google. https://blog.google/technology/ai/google-gemini-next-generation-model-february-2024/

PricewaterhouseCoopers. (n.d.). *EU AI Act: Europäische KI-Regulierung und ihre Umsetzung*. PwC. Retrieved September 25, 2024, from https://www.pwc.de/de/risk-regulatory/responsible-ai/europaeische-ki-regulierung-und-ihre-umsetzung.html

Radford, A., Narasimhan, K., Salimans, T., & Sutskever, I. (2018). *Improving Language Understanding by Generative Pre-Training*. 1–12.

Ranjan, R. (2024). *Adept AI Introduces Fuyu-Heavy: A New Multimodal Model Designed Specifically for Digital Agents*. MarkTechPost. https://www.marktechpost.com/2024/01/27/adept-ai-introduces-fuyu-heavy-a-new-multimodal-model-designed-specifically-for-digital-agents/

Reuters. (2023, November 7). Musk to integrate xAI with social media platform X. *Reuters*. https://www.reuters.com/technology/musk-integrate-xai-with-social-media-platform-x-2023-11-05/

Rosenthal, M. (2016). Qualitative research methods: Why, when, and how to conduct interviews and focus groups in pharmacy research. *Currents in Pharmacy Teaching and Learning*, *8*(4), 509–516. https://doi.org/10.1016/j.cptl.2016.03.021

Rumelt, R. P. (2012). Good Strategy/Bad Strategy: The Difference and Why It Matters. *Strategic Direction*, *28*(8). https://doi.org/10.1108/sd.2012.05628haa.002

SAP. (2024). *SAP Infuses Business AI Throughout Its Enterprise Cloud Portfolio and Partners with Cutting-Edge AI Leaders to Bring Out Customers' Best*. SAP Southeast Asia News Center. https://news.sap.com/sea/2024/06/sap-infuses-business-ai-throughout-its-enterprise-cloud-portfolio-and-partners-with-cutting-edge-ai-leaders-to-bring-out-customers-best/

Scao, T. L., Fan, A., Akiki, C., Pavlick, E., Ilić, S., Hesslow, D., Castagné, R., Luccioni, A. S., Yvon, F., Gallé, M., Tow, J., Rush, A. M., Biderman, S., Webson, A., Ammanamanchi, P. S., Wang, T., Sagot, B., Muennighoff, N., del Moral, A. V., … Wolf, T. (2023). *BLOOM: A 176B-Parameter Open-Access Multilingual Language Model*. 1–73. https://doi.org/10.48550/arXiv.2211.05100

Schmid, P., Sanseviero, O., & Cuenca, P. (2024, February 21). *Welcome Gemma—Google's new open LLM*. Hugging Face. https://huggingface.co/blog/gemma

Sejnowski, T. J. (2023). Large Language Models and the Reverse Turing Test. *Neural Computation*, *35*(3), 309–342. https://doi.org/10.1162/neco_a_01563

Sjödin, D., Parida, V., & Kohtamäki, M. (2023). Artificial intelligence enabling circular business model innovation in digital servitization: Conceptualizing dynamic capabilities, AI capacities, business models and effects. *Technological Forecasting and Social Change*, *197*, 122903. https://doi.org/10.1016/j.techfore.2023.122903

Stability AI. (n.d.). *Stability AI*. Stability AI. Retrieved October 2, 2024, from https://stability.ai

Statista. (n.d.). *Generative KI - Weltweit*. Statista. Retrieved September 28, 2024, from https://de.statista.com/outlook/tmo/kuenstliche-intelligenz/generative-ki/custom

Strauss, A. L., & Corbin, J. M. (2003). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2. ed.). Sage Publ.

Teece, D. J. (2018). Business models and dynamic capabilities. *Long Range Planning*, *51*(1), 40–49. https://doi.org/10.1016/j.lrp.2017.06.007

Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, *18*(7), 509–533. https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z

Teo, T. S. H., & Pian, Y. (2003). A contingency perspective on Internet adoption and competitive advantage. *European Journal of Information Systems*, *12*(2), 78–92. https://doi.org/10.1057/palgrave.ejis.3000448

Thirunavukarasu, A. J., Ting, D. S. J., Elangovan, K., Gutierrez, L., Tan, T. F., & Ting, D. S. W. (2023). Large language models in medicine. *Nature Medicine*, *29*(8), 1930–1940. https://doi.org/10.1038/s41591-023-02448-8

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2023). Attention Is All You Need. *arXiv*, 1–15. https://doi.org/10.48550/arXiv.1706.03762

Veale, M., & Zuiderveen Borgesius, F. (2021). Demystifying the Draft EU Artificial Intelligence Act—Analysing the good, the bad, and the unclear elements of the proposed approach. *Computer Law Review International*, *22*(4), 97–112. https://doi.org/10.9785/cri-2021-220402

Verfassungsschutz. (n.d.). *Glossar*. BundesamtfuerVerfassungsschutz. Retrieved October 11, 2024, from http://www.verfassungsschutz.de/DE/service/glossar/glossar_node.html

Wagner, D. N. (2020). Strategically managing the artificially intelligent firm. *Strategy & Leadership*, *48*(3), 19–25. https://doi.org/10.1108/SL-08-2019-0119

Wamba-Taguimdje, S.-L., Wamba, S. F., Kamdjoug, J. R. K., & Wanko, C. E. T. (2020). Impact of Artificial Intelligence on Firm Performance: Exploring the Mediating Effect of Process-Oriented Dynamic Capabilities. In R. Agrifoglio, R. Lamboglia, D. Mancini, & F. Ricciardi (Eds.), *Digital Business Transformation* (Vol. 38, pp. 3–18). Springer International Publishing.

https://doi.org/10.1007/978-3-030-47355-6_1

Wiesche, M., Jurisch, M. C., Yetton, P. W., & Krcmar, H. (2017). Grounded Theory Methodology in Information Systems Research. *MIS Quarterly*, *41*(3), 685-A9. https://doi.org/10.25300/MISQ/2017/41.3.02

Wiseman, S., Shieber, S. M., & Rush, A. M. (2017). *Challenges in Data-to-Document Generation* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1707.08052

x.ai. (n.d.). *Open Release of Grok-1*. Retrieved October 2, 2024, from https://x.ai/blog/grok-os

Declaration about the Thesis

Ich versichere wahrheitsgemäß, die Arbeit selbstständig verfasst, alle benutzten Hilfsmittel vollständig und genau angegeben und alles kenntlich gemacht zu haben, was aus Arbeiten anderer unverändert oder mit Abänderungen entnommen wurde sowie die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet zu haben.

Karlsruhe, den 18. Oktober 2024 Julian Faber