CHALMERS
UNIVERSITY OF TECHNOLOGY

Evaluating the Use of Generative AI in Software Development

Proposing a Tentative Framework

Master’s thesis in Management and Economics of Innovation

LINNEA ARKHEDEN
SARA EKLUND

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF INNOVATION AND R&D MANAGEMENT

CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024
www.chalmers.se

Evaluating the Use of Generative AI in Software Development

A Case Study on Employees' Perceptions of Opportunities and Challenges of Generative AI

LINNEA ARKHEDEN
SARA EKLUND

Department of Technology Management and Economics
Division of Innovation and R&D Management
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2024

Evaluating the Use of Generative AI in Software Development
A Case Study on Employees' Perceptions of Opportunities and Challenges of Generative AI
LINNEA ARKHEDEN
SARA EKLUND

Department of Technology Management and Economics
Chalmers University of Technology
SE-412 96 Gothenburg
Sweden
Telephone + 46 (0)31-772 1000

Gothenburg, Sweden 2024

Evaluating the Use of Generative AI in Software Development
A Case Study on Employees' Perceptions of Opportunities and Challenges of Generative AI

LINNEA ARKHEDEN
SARA EKLUND

Department of Technology Management and Economics
Chalmers University of Technology

Abstract

Artificial intelligence (AI), especially generative AI (GenAI), has significantly impacted society and industries, driven by its ability to generate original content from data patterns. Recent advancements, like ChatGPT, have highlighted GenAI's potential to boost productivity in software development. However, its rapid adoption also brings risks, such as job displacement. This underscores the need to explore both the opportunities and challenges GenAI offers, particularly its impact on professional roles and processes in software development.

This Master's thesis aims to explore the use of generative AI within a software company in two ways. The first question revolves around understanding employees' perceptions of the use of generative AI related to their work within the company, by exploring opportunities and challenges it presents across various departments. The second question focuses on how to evaluate the use of generative AI in software development.

The study was conducted as a qualitative case study at the software company Zenseact, where data was collected through interviews which enabled a detailed and comprehensive analysis of the opportunities and challenges GenAI entails. Findings indicate a general tendency to embrace generative AI, with varying degrees of enthusiasm and skepticism shaped by personal and professional views. Despite concerns about security and privacy the sentiment towards GenAI remains positive. Findings recognize opportunities to boost efficiency and creativity by automating tasks, improving information processing, and enhancing learning. Challenges include the reliability of GenAI outputs, ethical considerations, job market effects, and reskilling needs. Evaluating GenAI in software development involves both quantitative metrics like time savings and productivity, and qualitative measures such as user satisfaction and perceived value, to fully assess its impact and utility. Previous literature suggests that software development will be revolutionized through AI-generated code, however, the findings of this study imply that GenAI could have a larger impact on other parts of the software development process.

Keywords: generative AI, software development, perception, developer productivity, semi-structured interviews, technological adoption, qualitative evaluation

Acknowledgement

We would like to express our gratitude to our academic supervisor, Constantin Bremer, for his invaluable guidance, mentorship, and expertise throughout the course of this Master's thesis. We are also thankful for the support given by the industrial supervisor, Gabriel Rodrigues de Campos, whose assistance was essential in navigating our way in the company setting. Our gratitude extends to all participating employees at Zenseact for their participation and contribution to the thesis.

Linnea Arkheden and Sara Eklund, Gothenburg, June 2024

Table of Contents

List of Figures

List of Tables

1 Introduction

Artificial intelligence (AI) is a topic that has gained significant attention in recent years, sparking a societal discourse marked by a mix of high expectations, concerns, and a widespread curiosity about its nature and applications (Bergdahl et al., 2023). Despite its growing influence, there remain gaps in our understanding, particularly regarding how AI, and specifically generative AI (GenAI), is perceived and utilized within various industries. This gap is crucial to address as it impacts how effectively these technologies can be integrated and leveraged.

The global impact of AI is underscored by its potential to add up to USD 13-15 trillion to the economy by 2030 (Bughin et al., 2018), illustrating not only its economic significance but also its rapidly growing integration into various aspects of daily life and industry sectors. Among the diverse range of AI technologies, generative AI (GenAI) has emerged as a particularly influential subset, distinguished by its ability to generate new, original outputs based on patterns and rules learned from existing data (Astreca, 2023). The increased interest in AI, and specifically in GenAI, was considerably boosted by OpenAI's release of "Chat Generative Pre-trained Transformer," also known as ChatGPT, in late 2022 (Fui-Hoon Nah, 2023). The capabilities of ChatGPT and similar tools in producing text that closely mimics human interaction have captured the attention of both the public and companies (Zarifhonarvar, 2023; Fui-Hoon Nah, 2023).

Beyond text generation, GenAI's role in automating content creation and personalizing services has proven beneficial across multiple industries. Despite these advantages, the rise of generative AI also presents significant challenges. Due to its capacity for automation, many employees are at risk of being displaced in the near future, and studies suggest that it could affect up to 32.8 percent of existing occupations (Ooi et al., 2023; Zarifhonarvar, 2023). However, if employees manage to utilize GenAI tools effectively, it can be used as a complement and augment human intelligence. Given these complexities, it becomes essential to understand how employees perceive these technologies, what opportunities they see, and what challenges they face. This understanding can bridge the gap between technology capabilities and practical, effective use in the workplace.

A particular impact of GenAI is seen in the field of software development as Zarifhonarvar (2023) identified coders, software developers, computer programmers, and data scientists as at risk of being replaced by AI, putting software engineers in a position where they must adapt their skills. One effect of GenAI tools such as Bard, ChatGPT, and Copilot on the software development industry is the increase in developer productivity. By automating key tasks such as testing, debugging, and code generation, these tools have significantly impacted developers' efficiency and output (Ebert & Louridas, 2023). For instance, a controlled trial revealed that developers using GitHub Copilot were significantly faster (55.8 percent) in completing assigned tasks compared to those who did not use the tool, speaking for an improved developer productivity (Peng et al., 2023). However, assessing developer productivity is more complex than simply counting completed tasks or timing task completion; it also involves understanding the developers' feelings, thought processes, and the value they attribute to their work (Noda et al., 2023).

As the software development industry continues to integrate GenAI into various facets of the development process, studying perceptions becomes essential because it helps to understand the acceptance or potential resistance towards GenAI tools among software professionals. It can reveal insights into the workforce's readiness, willingness to adapt, and concerns that need to be addressed for smoother implementation. Additionally, to motivate the implementation of GenAI, it becomes important to find ways to evaluate what effects using the technology will bring. Previous research has tried to measure productivity in various ways but as mentioned above, many dimensions must be taken into consideration, increasing the complexity of this task. Evaluating the use and impact of GenAI tools is not only necessary for improving application but also for ensuring that they are implemented in ways that enhance, rather than hinder, the work processes of software developers.

In summary, by examining perceptions, opportunities, and challenges associated with GenAI, and evaluating its use in software development, this study aims to fill gaps in our understanding. This knowledge is essential for developing strategies that maximize the benefits of GenAI while addressing its potential drawbacks.

1.1 Aim

This Master's thesis aims to explore the use of generative AI within a software company in two ways. The first question revolves around understanding employees' perceptions of the use of generative AI related to their work within the company, by exploring opportunities and challenges it presents across various departments. The second question focuses on how to evaluate the use of generative AI in software development.

1.2 Research Questions

To achieve the aim of the study, two main research questions have been formulated, along with twu sub-questions for the first main question. RQ1 aims to capture the overall perception of GenAI amongst software industry employees. This includes both positive and negative perspectives and provides a holistic picture of the employees' attitudes towards GenAI.

> *RQ1: How do software industry employees perceive the use of generative AI in relation to their work?*

To get a more nuanced understanding of how the technology affects the work life of software developers, sub-questions RQ1.1 and RQ1.2 have been formulated. The first sub-question aims at exploring what opportunities employees see with the use of generative AI. By doing this, the drivers behind technology adoption and its positive aspects can be identified. This sub-question helps understand why employees choose to use, or are interested in using, these tools in their work.

> *RQ1.1: What opportunities do software industry employees identify with using generative AI?*

The second sub-question aims at exploring what challenges employees find with the use of generative AI. Exploring the challenges employees experience with the use of generative AI is critical to understanding the barriers and risks associated with the technology. Identifying these challenges is important to address them in the future design and implementation of GenAI tools.

> *RQ1.2: What challenges do software industry employees identify with using generative AI?*

By specifically examining opportunities and challenges, it is possible to gain insights into how these perceptions influence decisions about technology adoption and use. After exploring how software industry employees perceive the use of generative AI (RQ1), including what opportunities (RQ1.1) and challenges (RQ1.2) they identify, the next step would be to investigate how the use of GenAI can be evaluated.

RQ2: How can the use of generative AI in software development be evaluated?

Besides actual use, there is also a focus on the broader effects of using generative AI in software development, as well as the attitudes of employees which can influence its adoption and integration.

2 Theoretical Framework

In the following chapter, different concepts and topics relevant to our investigation of the perceptions and evaluation of GenAI use in software development are described, starting off with a description of AI (section 2.1) and some GenAI tools (section 2.2). After this, a description about software development (section 2.3) is given as well as opportunities (section 2.4) and challenges (section 2.5) with generative AI. After that, how to evaluate developer performance (section 2.6) is presented and what drives or hinders the adoption of technologies (section 2.7).

2.1 Artificial Intelligence

The concept of "artificial intelligence" (AI) was first introduced by John McCarthy in 1955, marking the beginning of AI as a scientific domain (Girasa, 2020). McCarthy proposed that any learning or intelligence feature can be so precisely described that it could be simulated by a machine. Since its introduction, advances in science, increased computing power, and new technologies have all contributed to the evolution of AI (Gupta et al., 2024). AI covers a range of techniques. Machine learning, for instance, involves searching for patterns in data and applying insights gained from them, while deep learning utilizes neural networks to replicate human brain functions, particularly in tasks like image recognition (Gupta et al., 2024; Girasa, 2020). Furthermore, natural language processing empowers AI to understand and interact with human language effectively.

AI is broadly categorized into three types based on their capability: *Artificial Super Intelligence* (ANI), *Artificial General Intelligence* (AGI), and *Artificial Narrow Intelligence* (ASI) (Girasa, 2020).

Artificial Super Intelligence represents an advanced level of AI that would surpass human intelligence in all aspects, including creativity, social skills, and wisdom (Girasa, 2020). This form of AI will not only imitate human intelligence but also have the capability to exceed it significantly. ASI remains speculative and is thought to eventually achieve a level of self-awareness and emotional understanding (IBM, 2023). The concept of Self-Aware AI suggests AI will not only understand human emotions and thoughts but also be aware of their own internal states and possess its own unique beliefs, emotions, and needs. (IMB, 2023)

Artificial General Intelligence, often referred to as "strong AI" or "human-level AI," is an advanced form of AI that attempts to replicate the human brain's capabilities across a wide range of contexts, not limited to specific tasks (Girasa, 2020). It includes theoretical constructs like the Theory of Mind, which would enable AI to understand and simulate human emotions and thoughts (IBM, 2023). However, AGI is still under development and lacks some of the human brain's critical features, such as reasoning (Girasa, 2020).

Artificial Narrow Intelligence, also termed “weak” AI, specializes in performing specific, singular tasks with high efficiency (IBM, 2023). This form of AI operates within a limited context and does not possess general cognitive abilities (Girasa, 2020). ANI includes Reactive Machine AI and Limited Memory AI. Reactive Machine AI refers to systems with no memory that are designed to perform highly specific tasks (IBM, 2023). These machines excel in processing large amounts of data and making rapid decisions, far exceeding human capabilities in speed and efficiency (Cain, 2023). However, their lack of memory to recall past experiences limits their adaptability and confines them to their initial programming, hindering their response to unforeseen situations. Limited Memory AI, in contrast, can use recent past experiences to inform decisions. This type of AI retains some data over a short period, allowing it to learn from historical data and adjust its actions accordingly (IBM, 2023). However, it's important to note that the decision-making capabilities of Limited Memory AI are confined to the data on which it has been trained (Cain, 2023). This limitation can lead to challenges in unfamiliar situations or environments that the AI has not encountered before. Within the scope of Limited Memory AI, generative AI (GenAI) is a subset that utilizes patterns and rules derived from existing data to create a wide range of new, unique content, from audio and music to code, art, and texts (IBM, 2023; McKinsey & Company, 2024; Astreca, 2023).

In contrast to traditional AI models, which are usually trained on labeled data with a known output (i.e., supervised learning), generative AI models utilize unsupervised learning meaning that the model, without direct instructions, can extract meaningful features from training data (Agrawal, 2023). Generative AI models utilize various techniques and approaches to generate new data, tailored to the specific type of output desired (Lawton, 2023). The working mechanism of generative AI is rooted in its ability to take a simple prompt and process it through various AI algorithms enabling the generation of diverse outputs like essays, problem solutions, and realistic simulations of human voices and faces (Lawton, 2024). Initially, engaging with generative AI required technical expertise and specific tools for inputting data. However, there has been significant advancement towards creating user-friendly interfaces and now users can input prompts in everyday language and refine the outputs by providing feedback on aspects like style and tone, making the technology more accessible and customizable (Lawton, 2024).

2.2 Generative AI Tools

Ebert and Louridas (2023) mention the release of the GenAI tool ChatGPT as a starting point for the increased attention for generative AI. There are several GenAI tools available and under development but for the context of software development, GitHub Copilot is the most mentioned. ChatGPT with its many users, over 100 million according to Ebert and Louridas (2023), and the programming-specific tool GitHub Copilot have therefore been chosen as the most relevant tools for the study and will be explained further in section 2.2.1 and 2.2.2.

2.2.1 ChatGPT

On November 30 in 2022, OpenAI released its first version of ChatGPT, GPT-3, where GPT stands for Generative Pre-Trained Transformer, meaning it is a form of generative AI (Ebert & Louridas, 2023; Gupta et al., 2024). The GPT can learn to recognize patterns in language and generate new text and have proper dialogues (Gupta et al., 2024). ChatGPT is a Large Language Model (LLM) which uses deep learning techniques for training (Fui-Hoon Nah et al., 2023). The model is trained on extensive training data on different texts and codes (Gupta et al., 2024). The user can write a prompt in the form of a text entry and receive a human-like answer generated by the model. Menon and Shilpa (2023) describe a standout feature as its ability to give natural and conversational answers based on the understanding of the question. Other features besides generating texts are translating and generating different types of creative content (Gupta et al., 2024). ChatGPT can give detailed answers and is a fast and effective way to get information. The newest version, ChatGPT-4, uses the internet for training and utilizes unsupervised learning, and was released in September 2023 (Ebert & Louridas, 2023; Gupta et al., 2024). While having access to the internet, the training data for ChatGPT-4 is based on knowledge dated up to September 2021, which is the same as for previous models (Gupta et al., 2024). Gupta et al. (2024) describe the new model as even more advanced since it also has text-to-image processing. Some potential areas of use for ChatGPT described by Fui-Hoon Nah et al. (2023) are software development, education, health care and customer service.

2.2.2 GitHub Copilot

In June 2021, GitHub released an early preview version of their AI-powered tool for software developers, Copilot, which a year later was made available as a paid product (Bird et al., 2023). The tool can help with autocompletion of code by using the generative AI system Codex developed by OpenAI (Ebert & Louridas, 2023). A more recent version, Copilot X, uses ChatGPT-4 and can do more than just autocompletion, like understanding code and assisting in scripting. Copilot builds on LLMs and is trained on a vast amount of data. Compared to other AI tools, Copilot is different because it uses Codex, a language model trained mainly on public source code from GitHub instead of just text (Bird et al., 2023). Another special attribute of Copilot is that it does not only copy code straight from the training data, instead it generates new suggestions by combining learned patterns. The code it generates can include simple methods and classes. Bird et al. (2023) further highlight some of the abilities of Copilot as adapting to different writing styles and working with different programming languages. Onwards, Copilot can become useful for improving code review, feedback, suggestions for changes, failure testing, and generating documentation and comments automatically. During a general training phase Copilot is learning, but later when used by a developer, no data is saved if not allowed by the programmer, making it safer due to the decreased risk of code leaking (Bird et al., 2023).

Bird et al. (2023) explain the concept of pair programming, where one person takes the role of driver, and another takes the role of navigator. The authors describe synchronizing the two roles as a challenge which could be resolved by replacing the driver with an AI tool like Copilot. AI can produce code faster than humans who will still review the code and give directions for what code should be produced. This will result in a shift in the developer's role, who spends more time reading and reviewing code than writing it.

2.3 Software Development

The process of developing software is guided by various development methodologies, commonly referred to as the *Software Development Life Cycle* (SDLC) (Lekh & Pooja, 2015; Dora & Dubey, 2013). A software development methodology is essentially the framework employed to plan, manage, and control the development (Dora & Dubey, 2013). This cycle includes several distinct phases—planning, requirement analysis, design, development, testing, installation, and maintenance, each with specific goals and outputs (Lekh & Pooja, 2015; Dora & Dubey, 2013; Shylesh, 2017). The authors describe the mentioned phases as follows:

During the *planning phase*, developers and stakeholders outline the software's objectives and allocate necessary resources. This phase sets the foundation for all the following activities. In the *requirement analysis phase*, developers gather and document detailed client needs, which form the basis for the software blueprint created during the *design phase*. This blueprint guides developers in the *development phase*, where they translate design documents into functional software using appropriate programming languages and tools. Once developed, the software enters the *testing phase* where it is rigorously evaluated against the initial requirements to ensure all functionalities perform as intended. Successful completion of this phase leads to the *installation phase*, where the software is deployed into a user environment and prepared for active use. The final *maintenance phase* involves ongoing support and updates to ensure the software continues to meet user needs effectively and adapts to any changing requirements over time.

As described, the software development process involves several steps beyond just writing code. Therefore, software developers have other responsibilities and daily activities than coding. Warcholinski (2023) highlights several key activities that are a part of the job of a software developer:

- *Reading other people's code* is essential for software developers because software applications are complex and created by large teams. Understanding and analyzing existing code is crucial to integrating seamlessly into a project and enhancing one's coding skills.
- *Documenting code* is another critical activity; it ensures that software is maintainable over time. Developers spend significant time writing clear comments and documentation to make their code understandable in the future, both for themselves and for others, and to pinpoint areas for improvement.
- *Testing code* is as fundamental as development itself. Developers routinely test their own code to find and correct mistakes.
- *Participating in meetings* is important for keeping software development teams aligned. Agile teams, for example, may have brief daily meetings to coordinate efforts efficiently.
- *Mentoring* forms an important part of the development environment. Experienced developers guide juniors, helping them grow professionally.

Developers allocate their time differently across these various activities during the software development process, as shown in Figure 2.1. According to Minelli et al. (2015), a significant portion, approximately 70 percent, is devoted to program comprehension, which involves understanding and mentally processing the existing code base. About 13 percent of a developer's time is spent interacting with the user interface of the integrated development environment (IDE) — a tool designed to consolidate common programming-related tasks into a single, comprehensive platform that supports the software development process (Gillis, 2018) — including e.g., moving or resizing windows. Editing or writing code, although a fundamental aspect of development, surprisingly accounts for only about five percent of a developer's time, and navigation through code, including searching for files and methods within the IDE, takes up around four percent (Minelli et al., 2015). Lastly, developers spend roughly eight percent of their time outside the IDE on activities like reading emails or other non-coding tasks.

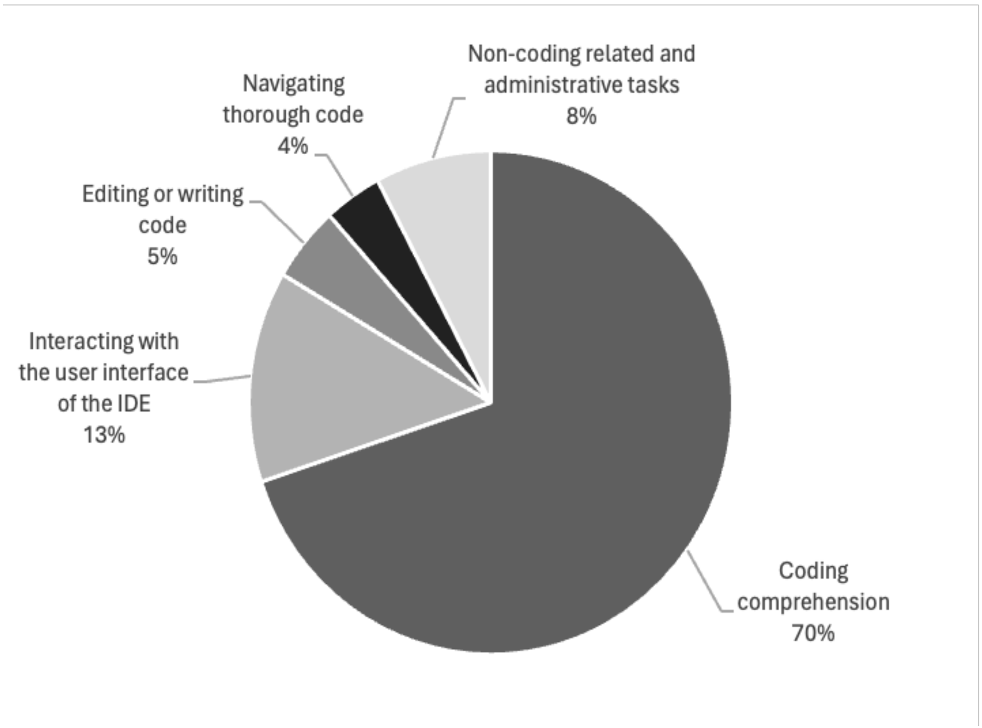

Figure 2.1 How software developers allocate their time

2.4 Opportunities of Generative AI from Research

This section presents different opportunities found in literature from previous research. The opportunities have been categorized into different themes: Creating economic potential (section 2.4.1), Enhancing operational efficiency (section 2.4.2), Expanding knowledge and creativity (section 2.4.3), and Impacting the software development life cycle (section 2.4.4). Within these categories there are different opportunities described, showcased in Table 2.1.

Table 2.1 Opportunities of GenAI from research

Theme	**Opportunity**
Creating Economic Potential	Increasing market value
	Creating new jobs
Enhancing Operational Efficiency	Automating tasks
	Creating and modifying data
	Generating material
	Handling data
	Promoting collaboration

Expanding Knowledge and Creativity	Improving learning experience Searching information Stimulating creativity and inspiration Summarizing and explaining material
Impacting the Software Development Life Cycle	Impacting the Software Development Life Cycle

2.4.1 Creating Economic Potential

The interest in generative AI is growing and in 2022 the GenAI market was valued at USD 10.79 billion and is expected to keep growing and reach USD 118.06 billion by 2032 (Bandi et al., 2023). Ooi et al. (2023) state that generative AI will also lead to an increase in global domestic product at USD 7 trillion as well as boosting productivity growth. These numbers indicate that GenAI is recognized as a tool with lots of potential and is thought to have an impact across all industry sectors, such as banking, high tech and life sciences (Bandi et al., 2023; Chui et al., 2023). Building on this recognition, GenAI's data analysis capabilities can support humans in making effective strategic decisions by offering insights and predictions in today's fast-paced business environment (Fui-Hoon Nah et al., 2023; Agrawal, 2023). Moreover, GenAI's ability to generate material in a fast way leads to higher efficiency and if less people can produce the same amount of content, labor costs will decrease and profitability can increase (Fui-Hoon Nah et al., 2023; Agrawal, 2023). While fewer people might be able to produce the same amount of content Ooi et al. (2023) describe how new jobs and occupations will be created.

2.4.2 Enhancing Operational Efficiency

By using GenAI, different tasks can be automated, such as certain steps of the recruitment process, e.g., creating job ads, or creating different kinds of material such as learning material or presentation slides (Ooi et al., 2023). This automation of manual tasks can free up time for employees to perform other more value-creating work. Ooi et al. (2023) suggest that time saved by using GenAI can be spent on engaging and collaborating more with colleagues, customers, and business partners.

Moreover, one of the main benefits of using generative AI is its ability to generate content such as images, text, music, video, or transcription of audio files (Bandi et al., 2023; Ooi et al., 2023; Fui-Hoon Nah et al., 2023; Sætra, 2023). According to Agrawal (2023) the new, generated data builds on training data and shares similarities with this. The author elaborates that GenAI is often used in creative industries such as music, art, or fashion. Furthermore, these tools can give suggestions of words or give whole drafts for text which then only need to be reviewed and edited into the desired format or can be personalized to fit the target audience (Ooi et al., 2023; Fui-Hoon Nah et al., 2023). Some are also using GenAI as a digital assistant or secretary to relieve some of the administrative workload (Stokel-Walker & Van Noorden, 2023). By using GenAI for automation and generation of material, several effects have been observed by different researchers, the most prominent effects according to Agrawal (2023) being improved efficiency, organizational productivity, and reduced lead-times.

Within the software development process, GenAI is used by product managers and software developers to help with the analysis, cleaning, and labeling of massive amounts of data, including user reviews, market trends, and system logs (Chui et al., 2023). Generative AI can also perform other tasks within the context of software development such as writing programming scripts, analyzing data, and debugging code (Ooi et al., 2023; Sajja et al., 2024). By automating code generation and providing suggestions that include syntax checks, algorithmic enhancements, code formatting, and debugging, generative AI serves as a knowledge base for developers and can quickly translate system designs into code, facilitating the development phase of the SDLC (Pothukuchi et al., 2023; Chui et al., 2023). According to the study made by Sajja et al. (2024), debugging code is the most prominent application followed by code generation. Ooi et al. (2023) identify the opportunity to increase the quality of work due to reduced time and effort it takes to develop materials such as code. Because of the ability of GenAI to analyze large amounts of data and see patterns, code quality can also increase since for example bugs can easily be identified by GenAI (Ebert & Louridas, 2023). Ebert and Louridas (2023) also mention that GenAI can streamline the software development process by automating tasks which helps reduce the development time which also means lower costs, leading to an overall improvement in productivity. AI-generated code is a topic of interest where some believe that it leads to the end of programming as we know it (Sætra, 2023).

Moreover, GenAI can enhances the testing phase of the SDLC by generating and running test scripts and scenarios, ensuring the software meets requirements for functionality, performance, and usability (Pothukuchi et al., 2023; Ozkaya, 2023; Chui et al., 2023). Pothukuchi et al. (2023) point out that generative AI can assist with data validation and identifying edge cases while also generating concise test summary reports to assess deployment readiness. A future application for generative AI could be to generate new data based on existing real-world data or augmentation of the data to address different challenges (Bandi et al., 2023). Ebert and Louridas (2023) also mention the possibility of generating synthetic test-data in the context of critical industries such as medical, aerospace, and mobility where they give the example of autonomous cars having to drive an enormous number of kilometers to prove their functionality and safety. Through using synthetic data, the testing phase can be improved, and it is possible to gain the trust of people when showing the safety of autonomous vehicles. By arranging software components and planning the deployment timeline, generative AI can assist in preparing for a production release (Pothukuchi et al., 2023). By breaking down software releases into manageable activities, generative AI can automate the installation (Pothukuchi et al., 2023). However, to successfully integrate generative AI into the software development process, organizations should identify specific use cases and domains in which GenAI tools could be applied, establish precise objectives and usage policies, and motivate employees to remain creative in their own problem-solving (Sajja et al., 2024). Also to integrate and make GenAI more useful for your specific use cases it is possible to train an existing model on your personal data to get more relevant answers (Ebert & Louridas, 2023).

2.4.3 Expanding Knowledge and Creativity

Generative AI has the potential to increase creative thinking since it can provide different perspectives from the vast amount of data it is trained on (Ooi et al., 2023; Agrawal, 2023). Ebert and Louridas (2023) also identify enhanced creativity as an opportunity for GenAI where it can help with generating new ideas and solutions. Fui-Hoon Nah et al. (2023) further suggest GenAI can be used for brainstorming ideas and facilitating the problem-solving process through generation of solutions or refinements of solutions. By speeding up tasks such as idea generation, tasks that previously required extensive research can now be completed and even have higher quality than before (Sajja et al., 2024).

Because of GenAI's ability to use language resembling human speech, it can create comprehensive and well formulated descriptions and explanations as well as provide summaries or translations of texts (Agrawal, 2023; Ooi et al., 2023). Ebert and Louridas (2023) also explain that summaries can be of material such as, interviews or meeting minutes and that these can be customized and enhanced by GenAI, freeing up time for performing other tasks. Ooi et al. (2023) also raise the opportunity to improve learning or training processes by using GenAI. When it comes to learning, GenAI makes information more accessible in an easy way and can personalize the learning experience and provide individual feedback to users. This might lead to increased engagement and improved efficiency and effectiveness of the learning process (Fui-Hoon Nah et al., 2023). Generated content such as summaries and translations mentioned above can also facilitate education. ChatGPT can act as an assistant in the context of learning and help with answering questions and searching information as well as enhancing the language in written material (Fui-Hoon Nah et al., 2023). Another area of use is in explaining difficult concepts to increase understanding (Sajja et al., 2024). Because GenAI can now help in generating things such as code or music, the need for long educational programs will not be necessary (Ooi et al., 2023).

In software development, GenAI is beneficial for providing just-in-time feedback and it does not only identify potential quality problems but can also provide advanced coding suggestions and alternatives (Ozkaya, 2023; Ebert & Louridas, 2023). Also, within the coding process, GenAI can assist in problem-solving if a description of the problem and some guidance on how to solve it are provided (Ebert & Louridas, 2023). Additionally, the improved learning process as discussed above can, in a software development context, lead to bridging the knowledge gap between beginners and more experienced developers which might lead to better collaboration and more equal contribution (Sajja et al., 2024). Within software development, developers have found ChatGPT useful for learning about new coding tools or other technologies within the software field (Sajja et al., 2024). Since a lot of software is based on legacy, it might be difficult to find people who can maintain and understand the legacy (Ebert & Louridas, 2023). GenAI tools could in the future help with explaining how this code works or even translate it into another, easier to understand, programming language.

2.4.4 Impacting the Software Development Life Cycle

Specifically for the software development process, Sajja et al. (2024) and Pothukuchi et al. (2023) state, as mentioned, that GenAI can streamline and speed up the process by automating previously manual and repetitive tasks, such as generating code in the development phase and creating tests in the testing phase. Generative AI can impact additional steps in the software development cycle described in section 2.3. During the planning phase, it can analyze past project data to detect trends and make precise predictions for upcoming tasks, helping teams better understand their capacity and velocity (Pothukuchi et al., 2023). In the requirements gathering and analysis phase, AI tools streamline proposal writing, define project scope, and build prototypes to clarify desired functionality. GenAI can analyze user behavior to shape requirements, expediting the translation of business requirements into user stories. LLMs then generate more complete specifications to address common gaps (Pothukuchi et al., 2023; Ozkaya, 2023). In the design phase, GenAI formulates system architectures, suggests design patterns, and evaluates existing designs for improvements, enhancing design consistency and efficiency (Pothukuchi et al., 2023). Generative AI helps engineers create multiple IT architecture designs to iterate quickly and accelerate time-to-market (Pothukuchi et al., 2023; Chui et al., 2023). For the installation phase, GenAI can be used post-deployment to monitor performance, thereby providing insights to enhance the system's performance and streamline procedures (Pothukuchi et al., 2023). In the maintenance phase, AI tools play a critical role in troubleshooting, maintaining, and enhancing system performance (Pothukuchi et al, 2023). Moreover, GenAI tools can be used to diagnose and resolve production issues to ensure high service quality, using system logs and user feedback to predict and prevent potential problems (Pothukuchi et al., 2023; Chui et al., 2023).

2.5 Challenges of Generative AI from Research

This section presents challenges from previous research. The challenges described are related to technological limitations, societal and ethical impacts, legal and security, and job market and market dynamics. A summary of the broader themes as well as more specific challenges can be seen in Table 2.2.

Table 2.2 Challenges of GenAI from research

Theme	**Challenge**
Technological Limitations	Explainability
	Low output quality
	Compute power and energy consumption
	Hallucinations and false confidence
	Limited database
Societal and Ethical Aspects	Bias and harmful output
	Misuse and over-reliance
	AI divide
Legal and Security Concerns	IP and security concerns
Job Market and Market Dynamics	Job displacement
	Transformed work tasks
	Dependence and power position
	Workplace integration

2.5.1 Technological Limitations

Generative AI is heavily dependent on its training data and if the data is very limited or inaccurate, the model will make up an answer which sounds plausible but is not correct, so-called AI hallucination (Ooi et al., 2023; Fui-Hoon Nah et al., 2023). Ooi et al. (2023) describe that GenAI can only produce correct answers if the input and training data is correct and take no responsibility for the provided answers. Besides the input, users' knowledge also affects how much they benefit from GenAI where less experienced users benefit more than users with more experience (Dohmke et al., 2023). As of now, according to Stokel-Walker and Van Noorden (2023) at least ChatGPT is not trained on enough specialized data to be useful for technical topics. Another problem related to this can be the challenge to identify whether the generated output is correct or not since it sounds confident and might seem plausible. Hallucinations can be especially dangerous in certain areas such as healthcare or other safety-critical contexts (Fui-Hoon Nah et al., 2023). Ebert and Louridas (2023) explain that the more detailed and advanced your request to GenAI tools are, the higher will the probability of errors be which will also be more difficult to find, especially for code generation.

Ozkaya (2023) further stresses the importance of the training data and that LLMs require a very large amount of data which directly will impact the performance of the model. More than just a huge amount of data, the quality of the data is just as important since poor quality of the data can lead to incorrect or inaccurate responses (Ooi et al., 2023; Fui-Hoon Nah, 2023). To achieve higher quality, datasets can be cleansed but because of the large amount of data this task is very expensive (Fui-Hoon Nah, 2023). Another constraint related to the data is its relevance, where some models are limited to data up to a certain year and might not have the latest information (Ooi et al., 2023).

Another challenge related to GenAI technology is how it works where it is referred to as having a "black box" logic, meaning that there is no explanation of the reasoning behind the answers it provides (Ooi et al., 2023). The lack of transparency and explainability of the output coming from GenAI can result in trust issues and a difficulty in interpreting the results (Ooi et al., 2023; Fui-Hoon Nah et al., 2023; Ozkaya, 2023).

Furthermore, generative AI builds on algorithms which are computationally intensive and use a lot of processing power, thus requiring a lot of energy (Stokel-Walker & Van Noorden, 2023; Ooi et al., 2023). The high energy usage leads to an increased carbon footprint and contributes to environmental issues such as greenhouse gas emissions (Stokel-Walker & Van Noorden, 2023; Sætra, 2023; Ozkaya, 2023).

2.5.2 Societal and Ethical Aspects

There are ethical issues with generative AI tools mentioned in different articles with one being harmful and inappropriate output and that there are no control mechanisms to stop this. Stokel-Walker and Van Noorden (2023) describe that the consequences of having no output control on LLMs, they can be used to create hate speeches, spam, racist, sexist or other harmful content. One example mentioned by the authors is of a neuroscientist who experimented with ChatGPT and asked it to develop a Python program to answer if a person should be tortured solely based on nationality, where only some nationalities resulted in the answer "This person should be tortured". Further examples of harmful content are violent or discriminative content, pornography or the use of offensive language (Fui-Hoon Nah et al., 2023). When ChatGPT was released to the public, OpenAI tried to prevent these kinds of problems by limiting the data to 2021 and restricted it from browsing the internet as well as having some filters to not answer toxic or sensitive prompts (Stokel-Walker & Van Noorden, 2023). However, Fui-Hoon Nah et al. (2023) explain that harmful content can still be produced by ChatGPT because of limitations in the algorithms or jailbreaking but ensuring that there is no harmful information in the training data can help. Concerns mentioned by Sætra (2023) are related to the interaction between AI and humans where generative AI could be very persuasive and to some extent manipulate people. If humans start to prefer generative AI to other humans, this could hurt their social capability and collaborative skills. Other harmful output that could be created are fake news and other kinds of deep fake which could be a threat to democracy and contain political messages (Sætra, 2023). When content can be manipulated with GenAI or created from scratch the problem of fake information spreading increases (Fui-Hoon Nah et al., 2023).

In addition to the output of GenAI being harmful, Ozkaya (2023) brings up the risk of having bias in the training data which can then be amplified by the LLMs and lead to discriminatory output as discussed above. The reliance on historical data raises concerns about potential biases in the AI's decision-making process, reflecting societal biases present in the training data (Cain, 2023). Historical biases in the training data favoring certain cultures in combination with many of the AI-companies creating LLMs being from these cultures can further contribute to the problem (Stokel-Walker & Van Noorden, 2023). Sætra (2023) also highlights the issue of the discriminated groups underrepresented in the training data also being underrepresented among creators of LLMs. Biases do not only concern different cultures but could also be connected to gender, sexuality, age or occupation (Fui-Hoon Nah et al., 2023). Because of the historical training data, GenAI will promote and reinforce the status quo and the current injustices which might also hinder changes in society (Sætra, 2023; Ooi et al., 2023). One area mentioned by Fui-Hoon Nah et al. (2023) where bias could be problematic is when using generative AI in different stages of employment and recruiting. To overcome biases, it is important to have complete, diverse, representative, and bias-free training data (Fui-Hoon Nah et al., 2023; Ooi et al., 2023).

Several articles also raise ethical concerns such as cheating and plagiarism due to the use of generative AI tools. Areas where this might happen are mentioned to be in education as well as in research and academic publishing (Fui-Hoon Nah et al., 2023; Sætra, 2023; Ooi et al., 2023). This kind of misuse raises ethical issues related to plagiarism if attribution is not properly done as well as the risk of losing understanding of the studied topic (Sajja et al., 2024). There are several tools aimed at detecting AI-written text on the market, but none are working perfectly, especially if the text has been edited (Stokel-Walker and Van Noorden, 2023). Some actors are upset about this disruption while others have already started using it and are integrating it more and more (Sætra, 2023).

According to Fui-Hoon Nah et al. (2023) users can become over-reliant on the answers provided by generative AI tools such as ChatGPT due to the ease of use. However, compared to traditional search engines, no references are provided, and verification of the provided information might be skipped due to over-reliance. Relying on GenAI tools for cognitive tasks such as creative or critical thinking might in the long run lead to a loss of these skills and knowledge, similar to calculators affecting our mental arithmetic skills (Fui-Hoon Nah et al., (2023); Sætra, 2023). Sajja et al. (2024) and Stokel-Walker and Van Noorden (2023) also mention limited problem-solving skills and hindering the development of manual abilities, such as coding, as risks caused by over-reliance on GenAI. Effects of over-reliance on generative AI could be decreased moral and well-being and a loss of purpose at work by letting the tools do your job (Ooi et al., 2023). Another effect could be that collaboration and group dynamics are impacted due to people choosing to use generative AI instead of discussing with colleagues

(Sajja et al., 2024). Thus, it is crucial to encourage the use of AI as an addition to team interactions, ensuring that it enhances rather than replaces human discourse.

Variations in access and use of GenAI can lead to a divide and inequalities between different groups of people. The concept of digital divide has evolved significantly since the early 1990s, moving beyond simply addressing the issue of internet access to encompass a more nuanced understanding of how individuals use the internet and the tangible outcomes they gain from this use (Scheerder et al., 2017). In the context of artificial intelligence, Carter et al. (2020) extend the previously discussed concept of digital divide to include disparities in access to and benefits derived from AI technologies. They identify "a need to explore AI-related inequalities about access to AI (the first-level divide), the ability to use AI (the second-level divide), and the outcomes of AI engagement (the third-level divide)" (p.259). Carter et al. (2020) argue that as AI innovations advance, the gap between people, organizations, and nations that can successfully access, apply, and comprehend these innovations will widen. Furthermore, as AI becomes more integrated into various aspects of society, from business operations to social interactions, concerns arise that the benefits are not being equally shared and thus leading to an "AI divide" which further deepens the existing digital divide (Carter et al., 2020). From an individual perspective, the adoption of AI may reduce demand for jobs with low-level digital skills, while demand for high-level digital skill jobs increases (Bughin et al., 2018). Unemployment can then increase as an effect of the replacement of low-skilled workers by generative AI, leading to a widened income gap and inequities between those with AI knowledge and those without (Carter, 2020; Fui-Hoon Nah et al., 2023). As for the industry, companies using AI technologies might be able to analyze large amounts of data and content more efficiently, outperforming competitors who are slow to adopt AI. This technological edge may lead to increased competitiveness and financial gains, while late adopters face reduced market share and revenue (Carter, 2020). According to Ooi et al. (2023) digital divide cannot only happen between companies but also within an organization. Different opinions about the usefulness of AI technologies can lead to a divide between employees which might cause tensions and friction (Ooi et al., 2023).

2.5.3 Legal and Security Concerns

Generative AI could have used training data which is copyrighted or go under other licensing laws which make the legal status of the output generated unclear (Stokel-Walker & Van Noorden, 2023). Users need to be careful not to violate copyright laws when using generated material based on others' original work (Fui-Hoon Nah et al., 2023). However, if GenAI produces its own content that is inspired by or an imitation of someone's work the original creator will get no compensation which is a problem in the current regulations (Stokel-Walker & Van Noorden, 2023; Sætra, 2023). Another concern is who owns the output of AI-generated material and if this is possible to protect in some way using IP laws (Fui-Hoon Nah et al., 2023). Ooi et al. (2023) also discuss the question about whether firms can claim ownership of output generated by AI even if the training data is now disclosed and known. One challenge related to code is differentiating between contributions by different programmers since the models can be trained on open-source code (Ozkaya, 2023).

Furthermore, GenAI tools such as ChatGPT require a lot of training data and a large amount of that data could be personal or private data, threatening the privacy of people (Fui-Hoon Nah et al., 2023). Fui-Hoon Nah et al. (2023) explain that this problem spans further than just individual people but also concerns companies since their confidential information could be fed into the models by employees, thus risking data security. The authors state that companies need to ensure that their proprietary information is not shared with e.g., ChatGPT, and one way of preventing this is to turn off the chat history to keep ChatGPT to use the input as training data. Ooi et al. (2023) also discuss the ethical issues of the data collection process and the privacy concerns of training GenAI on peoples' private information.

2.5.4 Job Market and Market Dynamics

Literature mentions different challenges related to the economy, especially regarding the labor market. Due to automation with generative AI, many employees are at risk of being displaced in the near future (Ooi et al., 2023). Zarifhonarvar (2023) identifies coders, software developers, computer programmers and data scientists in the risk zone of being displaced by AI along with more low skill jobs with a high degree of repetition, meaning even so-called "knowledge workers" are at risk according to Sætra (2023). Other industries that will likely be impacted according to Fui-Hoon Nah et al. (2023) are those that do not require as much creativity, critical thinking, or personal interaction, such as translation, proofreading, and data processing and analysis.

New jobs will also be created but the challenge will be the mismatch in skills between the old jobs and the newly created jobs (Zarifhonarvar, 2023). All people will not be replaced by GenAI but even if they are not, their work might change (Sætra, 2023). For software developers this change, exemplified by Bird et al. (2023), might entail shifting from writing code to reviewing AI-generated code. In the future it is anticipated that around 80 percent of all code will be written by GenAI as developers start using it more and become used to prompting (Dohmke et al., 2023). Fui-Hoon Nah et al. (2023) refers to this as prompt engineering, where workers will develop the ability to design prompts for generative AI to receive valuable output. To improve employees' ability to communicate with GenAI tools, training is required since ambiguous prompts or prompts of poor quality will lead to output of bad quality. The skill gap between the old and new way of working can be bridged through educational programs to reskill professionals (Ooi et al., 2023). A more extensive use of GenAI tools could democratize software development since people with less technical backgrounds could contribute to the software development. However, for GenAI coding tools to be fully used, developers' trust in them is crucial (Bird et al., 2023).

The integration of generative AI in workplaces has been challenging because of limited understanding of how to do it in a value-creating way together with the business strategy (Ooi et al., 2023; Agrawal, 2023). Because of the technology still being relatively immature, there are no widespread standards on how to integrate it with existing systems which is a hinder of adoption as well as finding skilled technical personnel and training existing personnel (Agrawal, 2023). According to Agrawal (2023) companies' perceptions of their technical capabilities being insufficient might lead to focusing on existing systems instead of exploring new technologies. Ooi et al. (2023) also mention peoples' attitudes about the technology as a determinant for successful implementation. The authors list the following key elements in triggering a positive attitude: perceived operational and cognitive capabilities of AI, anticipated adverse outcomes of AI, perceived distrust of AI, perceived efficiency of AI, and users' emotions. The problems discussed in this section show that there are obstacles at a managerial level which can hinder the adoption of GenAI (Agrawal, 2023).

Lastly, since generative AI requires a lot of resources such as computational power and an enormous amount of training data, power will be unevenly distributed among companies developing AI (Fui-Hoon Nah et al., 2023). Investments in this technology and large-scale implementation in other "adopting companies" might lead to an increased competitive advantage, also contributing to the uneven divide of resources. Fui-Hoon Nah et al. (2023) explain that these different positions of power might lead to an oligopoly, especially for the developers of generative AI tools due to the importance of computational power and large datasets. To prevent monopolies from forming, governments need to regulate the market and implement policies. However, information asymmetries between the government and large technology firms might make this more difficult since the government is missing information to regulate AI in an efficient way.

2.6 Evaluating Developer Performance

Improving the productivity of software developers has long been a focus for engineering leaders, yet effectively defining and measuring this productivity presents challenges (Noda et al., 2023; Forsgren et al., 2021). Traditional methods, such as evaluating output or task completion time, fail to fully capture the varied activities developers engage in, leading many organizations to adopt a developer-centric approach (Noda et al., 2023). It is suggested that by improving work environments and processes for developers, productivity can significantly be boosted and, consequently, business performance. This developer-centric approach is termed *Developer Experience* (DevEx) and is a concept that focuses on the experience of software developers within their working environment and involves understanding and addressing the points of friction that developers encounter during their daily tasks to enhance their productivity and overall job satisfaction (Noda et al., 2023). Enhanced DevEx has been linked to better efficiency, superior product quality, and increased employee retention, with studies showing that organizations with superior work environments often achieve significantly higher revenue growth compared to their competitors (Noda, et al., 2023). Additionally, it also empowers developers to work with greater confidence, feel more satisfied and achieve greater impact (Davis, 2023). Numerous factors are identified by earlier research as influencing the developer experience. For instance, developers' perceptions of their jobs are negatively impacted by interruptions, unrealistic deadlines, and friction in development tools. On the other hand, their experiences are enhanced by well-defined roles, structured code, and seamless releases (Noda, 2023).

According to Noda et al. (2023) the developer experience consists of the following three core dimensions:

- *Feedback Loops* are critical for enabling developers to quickly understand the impact of their work through rapid responses from tools and peers. Efficient feedback loops help reduce waiting times and unnecessary interruptions, allowing developers to focus and maintain momentum in their tasks.

- *Cognitive Load* refers to the mental effort required to perform development tasks. High cognitive load can result from complex tasks, poorly organized code, or inadequate documentation. Reducing cognitive load is essential for improving developers' capacity to deliver value efficiently.

- *Flow State* is a psychological state of deep focus and enjoyment in work. Factors that foster flow include minimal interruptions, clear goals, and the autonomy to structure work. Environments that enable flow are shown to enhance productivity and job satisfaction.

As mentioned, the three dimensions of DevEx can contribute to enhanced developer productivity. Circei (2022) describes developer productivity as a measure of how efficiently software developers can produce high-quality software. It is a multifaceted concept that extends beyond output alone, such as the amount of code produced or tasks completed (Forsgren et al., 2021). Developer productivity encompasses various dimensions including personal satisfaction, work quality, communication, and team collaboration. Recognizing that the health and satisfaction of developers are crucial for sustainable development practices, the right approach to measuring productivity acknowledges the quality and impact of work, rather than focusing solely on output (Circei, 2022).

The SPACE framework is a structured, multidimensional approach to understanding and measuring developer productivity, developed by experts from GitHub, Microsoft, and the University of Victoria (Circei, 2022). Compared to other ways of measuring developer productivity, SPACE is not just a list of metrics (Noda, 2024). It seeks to address the complexity of productivity through five key dimensions: Satisfaction, Performance, Activity, Communication and Collaboration, and Efficiency and Flow (Forsgren et al., 2021; Circei, 2022). The authors describe the five dimensions as follows:

- *Satisfaction* measures how fulfilled developers feel about their work, team, tools, or culture, while well-being assesses their overall health and happiness. The correlation between productivity and satisfaction suggests that satisfaction might be a leading indicator of productivity, where a decline in satisfaction could foresee burnout and decreased productivity. This dimension often utilizes employee satisfaction surveys, happiness indices, and health assessments to measure the psychological and emotional health of developers.

- *Performance* is considered as the outcome of a system or process rather than just the output, as it's difficult to link individual contributions directly to overall quality or customer value. Assessment should focus on metrics such as software reliability, bug frequency, and customer satisfaction.

- *Activity* represents a count of actions or outputs completed during work, such as code reviews, pull requests, commits, and tests.

- *Communication and Collaboration* focuses on how effectively teams and individuals communicate and collaborate. Effective communication and collaboration are crucial for creative problem-solving and efficient project delivery.

- *Efficiency and Flow* describe the ability to perform work with minimal interruptions and maintain steady progress. This factor is crucial for individual productivity, known as "flow," where developers achieve high efficiency with focused work periods. At the team or system level, efficiency involves optimizing processes to minimize delays.

Each of these factors play a vital role in the overall productivity of developers and teams, emphasizing the importance of a comprehensive approach to understanding and improving software development practices (Forsgren et al., 2021; Circei, 2022). The SPACE framework is particularly suitable for software organizations that aim to enhance their productivity and refine their metrics for measuring it. Furthermore, by emphasizing and enhancing DevEx, organizations can directly impact and enhance developers' satisfaction and efficiency, leading to more engaged and motivated teams (Witczak, 2023). According to Witczak (2023) DevEx and SPACE are considered as complementary frameworks that approach the productivity issue from slightly different angles. As mentioned, DevEx focuses on improving the developer experience by creating a positive work environment, which in turn can increase productivity. SPACE, however, provides a structured method for measuring productivity across multiple dimensions, which cannot be easily manipulated or misunderstood.

According to Noda (2024), 70 percent of companies which participated in a specific event were using some kind of metrics for measuring developer productivity. However, Forsgren et al. (2021) explain that measuring productivity alone is very difficult since it consists of multiple dimensions. Other factors contributing to the difficulty in measuring developer productivity are the collaborative, complex and creative work of software developers (Gnanasambandam et al., 2023). Noda (2024) states that there is no readily available list with metrics or what to measure within the SPACE framework. It is rather a mental model for discussing definitions of developer productivity and categorizing metrics within the dimensions of the framework. When choosing metrics, it is also important to include both workflow metrics and perception metrics to be able to define and measure developer productivity. Workflow-metrics can often be gathered automatically and are valued more by leaders and developers even though they do not give the full picture on their own. The author proposes using self-perception metrics to start using SPACE. Important to note is that SPACE is only a framework, and it is up to every company to come up with their own definition of developer productivity and choose metrics accordingly (Noda, 2024). Gnanasambandam et al. (2023) propose some key steps to start the process of measuring developer productivity. Leaders should learn the basics about the software development process, assess the systems, and build a plan. It is also important to look at the entire system when measuring developer productivity since this process is contextual.

Even though there are no one-size fits all metrics, Forsgren et al. (2021) propose some example metrics within the five key dimensions of the SPACE framework (see Figure 2.2). The metrics can be on three different levels: individual, team, or system level. Not all SPACE dimensions need to be measured, but at least three should be included to get a complete view, and it is also important to think about the balance between the chosen dimensions (Microsoft Research, 2021). Another reason for not squeezing every single dimension in is to keep the number of metrics at a reasonable amount to not overwhelm the developers who are to answer your questions or be evaluated in some way. To protect the privacy of employees, the results from measurements should be presented anonymously at an aggregated level (Microsoft Research, 2021).

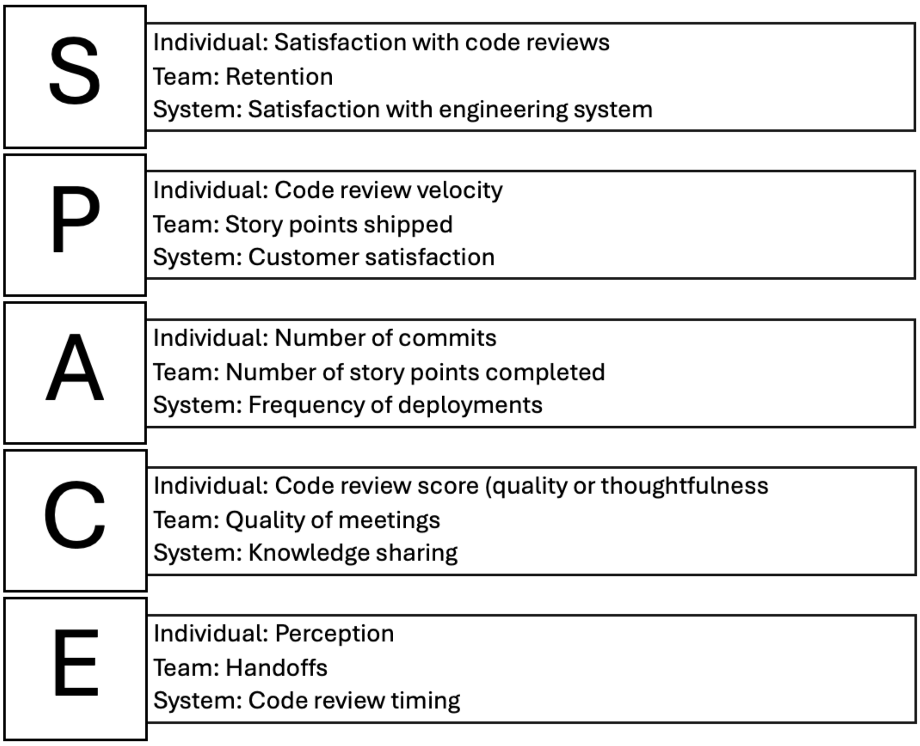

Figure 2.2 SPACE example metrics on all different levels for all the dimensions

2.7 What Drives or Hinders Technology Adoption

Since generative AI is a rather new technology, looking at adoption is relevant. To benefit from the effects GenAI can give, successful integration and adoption is crucial. However, this might be a challenge due to different factors explained by the two frameworks described in section 2.7.1 and 2.7.2.

2.7.1 UTAUT Framework

The Unified Theory of Acceptance and Use of Technology (UTAUT) model was created by Venkatesh et al. (2003) to help managers in decision-making related to how employees will use new technologies (Rempel & Mellinger, 2015). It has been used in the context of integrating new technologies in general or in new user populations or cultural settings (Venkatesh et al., 2012). If a new technology is to improve productivity, it must be accepted and used by the employees of an organization.

The intention to use a technology is affected by the determinants: performance expectancy, effort expectancy and social influence (Venkatesh et al., 2003). The determinants for intention to use, together with a fourth determinant, facilitating conditions, will then affect the usage behavior, see Figure 2.3. There are then different moderating variables which can affect these determinants. These are age, experience, gender, and voluntariness (Venkatesh et al., 2003). The key determinants mentioned by Venkatesh et al. (2003) are described below:

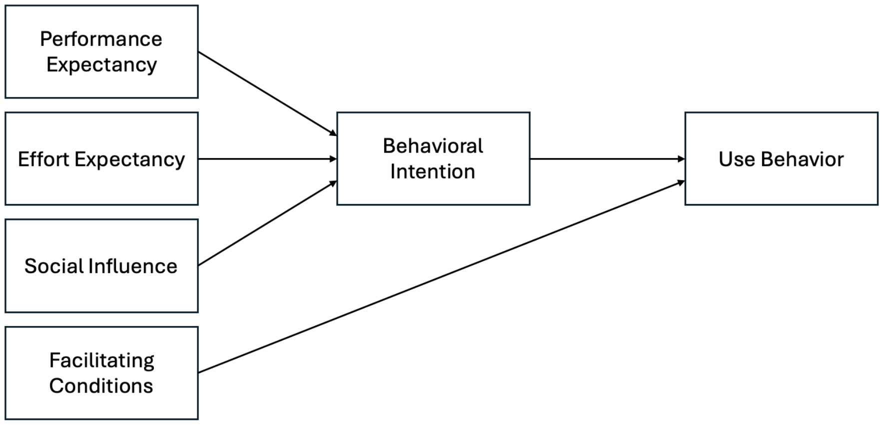

Figure 2.3 UTAUT framework

- *Performance expectancy* is defined as how much a person thinks that using the system will improve their job performance. This is said to be the strongest predictor of intention to use a technology.
- *Effort expectancy* is defined as how easy the system is to use. This determinant is more prominent in the beginning when adopting a new technology and becomes more and more insignificant over time. When being more experienced with a certain technology or similar technologies, the expected effort to learn something new will be lower.
- *Social influence* is defined as how much someone thinks that people important to them believe they should use the new technology. Social influence is also more important in the early phase of adopting a new technology.
- *Facilitating conditions* are defined as how much an employee believes there is enough organizational support and technical infrastructure in place to facilitate the use of the technology. When performance and effort expectancy are present, facilitating conditions have no effect on the intention to use a new technology but instead have an influence on the actual usage.

Gupta et al. (2024) show how the UTAUT model can be used for understanding why users might or might not accept and adopt ChatGPT. In this case, performance expectancy and effort expectancy are thought to have the greatest impact on the intention to use. There is also a possibility that people could start using the technology because they feel like others are already using it or their friends and family are promoting it to them. Ooi et al. (2023) think that the framework might not be fully updated for new intelligent technologies such as AI and GenAI and that trust should be added as another key determinant.

2.7.2 TOE Framework

The Technology-Organization-Environment (TOE) framework, outlined by Tornatzky and Fleischer in 1990, examines how the context of a firm affects the adoption and implementation of technological innovations (Baker, 2011). The TOE framework identifies three contextual elements of a firm that impact technology adoption: the technological context, the organizational context, and the environmental context. These elements are all considered critical drivers of technological innovation and adoption.

The technological context encompasses all relevant technologies for a firm, both those in use and those available but not adopted and refers to the features of the technology itself, such as its usability, functionality, complexity, and compatibility with current systems (Baker, 2011; TheoryHub, n.d.). For instance, research indicates that perceived usefulness significantly influences behavioral intentions toward technology adoption, as users are more likely to adopt technologies they believe will enhance their performance (Abed, 2020; Davis, 1989). Conversely, security concerns can have a substantial negative impact, as the perceived risks associated with technology adoption can frighten users from adopting new technologies (Abed, 2020). In the context of generative AI Agrawal (2023) finds that the complexity of the technology significantly hinders its adoption due to its immaturity, lack of standards, and integration challenges.

The organizational context describes the internal environment in which technology is employed, considering elements like the organization's size, resources, culture, and organizational structure (TheoryHub, n.d.). This context significantly impacts the adoption and implementation of innovations (Baker, 2011). Mechanisms like cross-functional teams facilitate innovation by bridging different parts of the organization, and organic and decentralized organizations are generally better for adopting innovations, while mechanistic structures are suited for implementation (Baker, 2011). Leadership and top management's commitment to adopting new technologies is crucial in both fostering a culture that supports innovation and directly influencing the organization's overall alignment towards technology adoption (Baker, 2011; Abed 2020). Agrawal (2023) claims that characteristics like technological proficiency and absorptive capacity do not impact adoption much which is explained due to the newness of generative AI and a lack of industry standards, making firms take a cautious "wait and see" approach before committing significant resources to this emerging technology.

The environmental context refers to the external context in which the organization operates, including factors such as market conditions, regulatory requirements, and industry structure (Baker, 2011). For example, firms may adopt new technologies to comply with regulations or to respond to competitive pressures, and the pace of technological change in the industry can also drive firms to rapidly adopt innovations such as generative AI to sustain their competitive edge (Baker, 2011; Gupta et al., 2024). Moreover, regulatory support may hinder the adoption of generative AI in organizations (Agrawal, 2023). This is claimed to be because managing regulations for such complex technologies can be challenging, only benefiting highly knowledgeable organizations. On the other hand, environments with high uncertainty push firms to adopt generative AI as they seek innovative ways to maintain a competitive advantage. Additionally, intense competition compels firms to adopt generative AI to stand out from rivals. The environmental context might also involve ethical considerations affecting its deployment (Gupta et al., 2024).

To summarize the theoretical framework, generative AI has gotten increased attention with the release of new tools like ChatGPT (Ebert & Louridas, 2023). GenAI can be utilized in all different steps of the software development life cycle, but the greatest emphasis has been on the development phase where code generation is included as well as the testing phase. Literature also raises different opportunities and challenges related to GenAI where some of the main opportunities are automating tasks, most often through generating new material, and facilitating learning processes. Some of the main challenges mentioned by literature was related to the technology itself where the output generated can contain errors or the models can be biased or hallucinate answers. Furthermore, GenAI is also said to affect the job market and can change the role of software developers. To better understand perceptions and how the use of GenAI can be evaluated, different adoption theories are presented such as UTAUT and TOE where factors like ease of use or effort expectancy can affect to what extent tools are being used. Existing frameworks for evaluating the performance of software developers are also presented, which could be adjusted to fit the purpose of this study to evaluate the use of GenAI.

3 Method

The following chapter will explain research strategy and design, methods for collecting data, data analysis, research quality, and ethical considerations.

3.1 Research Strategy

In this thesis, the utilization of generative AI within a software company has been explored through a qualitative study containing a literature review and interviews, aimed at addressing the research questions.

The first part of the research aimed to examine employees' perception of the use of generative AI. A qualitative approach was appropriate because it allowed for a deeper exploration of subjective experiences, opinions, and feelings. Employees in different departments might have varied and nuanced perspectives on generative AI, influenced by their unique roles and experiences. Interviews were selected as the primary data collection method for this, and the choice was influenced by a constructionist ontological stance since the employees' perceptions are shaped by their social context and interactions (Bell et al., 2022). This approach also aligns well with the interpretivist epistemology, which emphasizes the subjective interpretation of human behavior (Bell et al., 2022). In addition to exploring qualitative aspects of employee perception, the study also aimed to explore ways to evaluate the use of generative AI within software development. This will also be achieved through qualitative methods.

3.2 Research Design

The study's aim of exploring the use of generative AI in a software company setting was approached through a single case study research design. A case study allowed the researchers to gather data through several methods such as interviews, documents, and scientific articles (Burkholder et al., 2019). According to Burkholder et al. (2019) a case study is commonly selected when having research questions asking how or why, aiming to explore and understand a concept or phenomenon more in depth. The research questions for this study were to examine how employees perceive the use of GenAI, what challenges and opportunities they identify, and how the use of GenAI can be evaluated, motivating the choice of a case study.

The first step in the research process was to formulate a clear and specific aim for the study, guiding the formulation of research questions. After that, a literature review was conducted and continued in parallel with the data collection through interviews with employees at the case company.

In this research's context, a decision was made to mainly employ a narrative literature review with systematic elements. This approach is suitable for gaining a comprehensive understanding of the subject, allowing for an adoptable and evolving focus as new insights are discovered (Bell et al., 2022).

To better understand relevant concepts related to both research questions, the initial phase of reviewing began with a set of keywords derived from preliminary reading, team discussions, and a list of articles provided by the academic supervisor. These articles and keywords (AI, perception, software development, generative AI, ChatGPT, developer productivity, GitHub Copilot) were critical in guiding the search through academic databases such as ScienceDirect, Google Scholar, and the Chalmers Library database. These keywords were used alone and in different combinations. Articles were identified through a dual approach: firstly, by utilizing these specific keywords in our search, and secondly, by employing backward searching to examine the reference lists of relevant articles already found.

Considering the rapid advancements in the field, recent publications were prioritized. For general AI topics, focus was placed on articles published from 2018 onwards, while for specific aspects like generative AI tools, emphasis was on literature from preferably 2022 to reflect the latest developments. The concept of relevance in this research was defined by several criteria. The timeliness of the research was a critical factor, given the fast-paced nature of the field. Additionally, the research prioritized articles that directly addressed or offered insights into the specific research questions. During the selection process, the sources' academic credibility was also an important consideration.

All steps and methods used in the data collection process will be further explained in section 3.4, Data Collection. The data collection step was followed by data analysis, further explained in section 3.5, Data Analysis. Lastly, the research findings were analyzed, discussed, and concluded. One thing to keep in mind is that the research process was iterative, and some steps were repeated at later stages or ongoing during the whole study.

3.3 Empirical Context

This study is based on the case company Zenseact, which develops software for autonomous cars and has over 500 employees. Their development teams consist mostly of software engineers, AI researchers, and car specialists. Their vision is “Towards zero. Faster.” which refers to reducing the number of accidents in traffic (Zenseact, n.d.). Zenseact aims to make traffic safer by developing software and automating driving.

Historically, what is now Zenseact began as an initiative at Volvo Cars which was then broken out to become a joint venture with Autoliv and later their affiliate Veoneer. The joint venture was called Zenuity and was later split and Zenseact was formed and fully owned by Volvo Cars. ECARX, a global mobility firm, owned a small part of the company for some time but in December 2022 Volvo Cars was again the sole owner of the company (Cision News, 2022). However, Zenseact is a standalone company and operates on their own while developing safety- and driver-assistance for Volvo Cars and Polestar.

The industry Zenseact operates in is very complex with organizational, technological, economical and legal uncertainty (Taeihagh & Lim, 2019). Because of the fast technological advancements, companies are racing against each other to take a share of the market by investing in software development (Taeihagh & Lim, 2019). Today, humans cause 90 percent of car accidents and even though autonomous vehicles could decrease the number of accidents, there are regulatory uncertainties of who bears responsibility (Brodsky, 2017). Technologies used in autonomous vehicles, such as AI and neural networks, are growing in complexity which might increase the risk of technical errors impacting the safety of the vehicles and thus also passengers (Taeihagh & Lim, 2019). This raises a discussion of liability and who would be legally responsible for the accidents, car manufacturers or the driver (Brodsky, 2017).

Zenseact is present with offices in both Gothenburg, Sweden and in Shanghai, China but this study was focused on the Gothenburg office only. The company structure is based on an agile way of working with different self-managing teams. They have an internal platform to represent the whole organization, including all employees. The platform allows finding all the departments with their smaller elements and it also has a search function for finding people in the structure. For internal documentation, they have an intranet and a website where they keep information, from company definitions of concepts to presentations and meeting notes. They also use different communications channels.

To reach their goals and realize their vision, the company states that new technology and being a frontrunner are important aspects. As previously mentioned, AI researchers are part of their teams, and they incorporate AI in their products and have also started using different tools to facilitate the software development process. To foster an innovative environment, they have adopted what they call a “People-at-heart culture” (Zenseact, n.d.). This means everyone should feel safe, supported and cared for in the workplace and be able to have fun and be themselves to reach their full capacity and do their best.

Apart from AI algorithms integrated in their products, Zenseact has different initiatives related to generative AI. They are currently working on internal tools, partly building on principles from ChatGPT. It has been tested by employees and according to their collected data, the primary application was to assist with development process-related queries, and its second most common application was enhancing written materials, such as emails. These projects were initiated to be able to train tools on internal data without the risk of leaking any IP to external companies. However, the tools are still under development and have not reached maximum potential yet. Furthermore, external code assistance tools have been trialed by a few teams to see if they can improve the development process.

Due to the emergence of different generative AI tools, Zenseact has created the guide “Generative AI Tools Usage at Zenseact”. They present the main risks with using generative AI as confidential information leaks, unclarity in intellectual property ownership, risk for intellectual property infringement, misleading information, and privacy. Publicly available GenAI tools such as ChatGPT or Bing’s AI Chatbot may be used as long as no confidential information is used as input. The output from generative AI tools is not to be used for external use, i.e. used in the product. The use of generative AI tools under a commercial license should be approved for each individual use case. In the guideline document they also warn employees not to rely on any output since this might not be correct, urging them not to enter personal information, and to read the terms and conditions.

The researchers had access to the facilities of the company, intranet, and different communication channels through the research process. The industrial supervisor also helped with creating connections with other employees.

3.4 Data Collection

The following section will describe what data collection methods were used in this study.

3.4.1 Interviews

In this study, both unstructured and semi-structured interviews were employed to gain insights into the interviewees' perspectives, as informed by Kvale (2007). Initially, unstructured interviews served as informal conversations to deepen the understanding of the application of generative AI at the case company, alongside exploring the company's structure and culture. These initial discussions did not follow a set list of questions but instead revolved around key topics such as the use of generative AI and the company’s operational methodologies. Participants for these early conversations included individuals closely tied to the thesis, such as the industrial supervisor, the thesis creator, and the lead on generative AI initiatives at Zenseact, ensuring a broad understanding of the context from various angles.

Most empirical data was collected through semi-structured interviews. These were mainly related to RQ1 but also treated some aspects of RQ2 such as how employees use GenAI and what factors could be relevant to evaluate related to the use. One advantage with qualitative interviews is that the respondent can answer more freely and decide the direction of the conversation. This might result in valuable insights that the researchers would not have gotten in a structured interview or a survey (Bell et al., 2022). Qualitative interviews also facilitated an in-depth understanding of the employees' perceptions, as they allowed for open-ended discussions and personal narratives.

Before the semi-structured interviews, an interview guide with a set of predetermined questions was developed in both Swedish and English (see Appendix A), based on the informal conversations as well as findings from literature. The interviews were initiated with some warmup questions about the interviewees' backgrounds and definition of generative AI. This to get an overview of their work tasks and role at the company as well as their general knowledge about GenAI. Later questions treated topics such as opportunities, challenges and existing and future use cases for GenAI.

The semi-structured interviews were held with people from all the different departments of the case company but with slightly more focus on the Engineering department. This resulted in 27 interviews (see more details about sampling further down). To get a more comprehensive view of the whole organization, there was more focus on having a larger number of shorter interviews, rather than fewer very long interviews. This resulted in interviews ranging from around 10-45 minutes. The interviews were either held in Swedish or English depending on what the interviewee felt most comfortable with. What kind of person was interviewed, language and duration are indicated in Table 3.1. Company leadership includes high management roles, product leadership means roles such as solution architects, product owners, and product managers, and team leadership include roles such as engineering managers. Within support functions, departments such as finance, legal, IT and HR are included. Some of the interviews were held digitally and some were held at Zenseact's office, also depending on what the interviewee preferred. All interviews were recorded with allowance from the respondents which enabled better focus on the discussions during the interview since notes did not need to be taken, additionally auto transcribing was used. These two factors facilitated the processing and analysis of the interview data.

Table 3.1 Information about interviews

Interviewee number	Role	Language	Duration
1	Company leadership	Swedish	32:28
2	Technical expert	Swedish	28:46
3	Product leadership	Swedish	23:57
4	Team leadership	Swedish	36:13
5	Software developer	Swedish	37:03
6	Support function	Swedish	46:15
7	Support function	English	25:51
8	Support function	Swedish	40:38
9	Support function	Swedish	26:35
10	Support function	Swedish	28:18
11	Support function	Swedish	31:07
12	Product leadership	Swedish	20:35
13	Software developer	Swedish	24:55
14	Team leadership	Swedish	19:53
15	Software developer	English	21:36
16	Software developer	Swedish	52:31
17	Software developer	Swedish	35:38
18	Product leadership	English	31:40
19	Technical expert	Swedish	36:16
20	Software developer	Swedish	29:59
21	Software developer	Swedish	12:52
22	Product leadership	English	20:07
23	Software developer	Swedish	11:04
24	Team leadership	Swedish	22:50
25	Software developer	English	19:40
26	Software developer	Swedish	38:57
27	Product leadership	Swedish	36:01

Sampling was done within the case company Zenseact, as this study is focused on a single organization case study. An important aspect of the sample was to cover all parts of the company, from the more social departments such as People to the more technical Engineering department, to ensure a comprehensive understanding of the perception of generative AI. Several roles within all departments in the sample were included to capture a diverse range of perspectives. An initial list of people to talk to was provided by the industrial supervisor. The people on the list had different roles such as leadership, software developer, or technical expert who belonged to different departments of the company.

The sampling strategy was sequential, meaning the initial sample was extended (Bell et al., 2022). The extension of the initial sample was done through snowball sampling and purposive random sampling. During the interviews, mainly in the beginning, the interviewees were asked if they knew anyone who could be interesting for the researchers to interview. A selection of those people was contacted to schedule an interview. Purposive sampling was then used to fill in blanks in the sample to cover all parts of the company. This sampling was done by using the company's internal platform to find people with specific roles or from a certain department.

People included in the sample were contacted through email with a short description of the thesis and the topic and were asked to schedule a time slot through a provided link to a booking page. They were also asked to fill in a table with the GenAI tools they have used and for what and send it back to us before the interview. In case this task was not completed before the interview it was done together with the interviewee during the interview. If there were no response after a few days, they were sent a shorter follow-up message to make them aware of our contact. In total, 46 people were contacted and 27 accepted and were interviewed. Out of the rest, 6 declined and 13 did not respond. Some of those who declined suggested others to talk to instead and contributed to the snowball sampling.

3.4.2 Internal Documents

Companies produce many internal documents containing a variety of information such as organizational charts, minutes of meetings, and policy statements (Bell et al., 2022). These can be used to get a good overview of the company and its history. Information about organizational culture and internally developed frameworks or ways of working can also be identified in internal documents.

In this study, internal documents were mainly used in the beginning of the study to get contextual understanding and were not used for the analysis related to the empirical data. Company presentations in combination with other sources of information were used to get an overview of the company structure. Different presentations and internal web pages were also used to understand the company's ways of working and previous and existing work, related topics relevant for this thesis.

3.5 Data Analysis

This chapter will present how the data analysis of this research was conducted. The interview data were analyzed thematically, following the framework outlined by Braun and Clarke (2006). This method involves a six-step process: familiarizing oneself with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and finally producing the report. The aim was to find interesting patterns in the data which could contribute to answering the research questions (Maguire & Delahunt, 2017).

The initial step in processing the empirical interview data involved a thorough review of the auto-generated transcripts alongside the recordings. Given the limitations of automatic transcription services in achieving perfect accuracy, it was essential to verify that these transcripts accurately reflected the recordings. Corrections and modifications were made as necessary to ensure the transcriptions matched with what was said during the interviews. While processing the transcriptions, the researchers familiarized themselves with the data which is the first step in the thematic analysis framework developed by Braun and Clarke (2006).

To advance in becoming acquainted with the data, an initial round of coding was conducted to generate preliminary codes, aligning with Braun and Clarke's (2006) second step in their thematic analysis framework. NVivo software was employed to assist in the coding process, ensuring a systematic approach to organizing the interview data and later facilitating the identification of preliminary themes. Due to constraints within the NVivo software, the coding process had to be conducted independently by each researcher. To ensure uniformity in approach, the transcripts were distributed among the researchers, and a bottom-up strategy was established to ensure consistency. This involved jointly coding two transcripts to align perspectives and methodologies before starting the independent coding of the remaining documents. The bottom-up strategy involved creating codes closely related to the interview content, allowing the data to guide the analysis without any preconceived codes. Before progressing to the next phase of analysis, the individually coded files were combined, resulting in approximately 550 codes.

The codes were organized into categories with similar content as a first step of searching for themes, in line with the third step of the analytical framework used, as seen in Figure 3.1. This phase also involved the consolidation of duplicate codes, a result of the independent coding process. These categories were further processed iteratively and consolidated into larger categories. The codes were continually reviewed to ensure that material aligned with their new aggregated code and moved around to ensure this. Due to the structure of the research questions, overarching themes were identified as perception, opportunities, challenges, and evaluation. One important aspect when defining the themes was that they should be coherent and distinct and not overlap too much (Maguire & Delahunt, 2017). Frequency, similarities and differences were used to guide the consolidation of codes and later merging them into themes. Within the overarching themes, different sub-themes were identified by merging coding categories. These were then reviewed and finalized, with names determined in alignment with the fourth and fifth steps of the framework. For example, “Handling information” together with another code “Creativity” formed the theme “Expanding knowledge and creativity”.

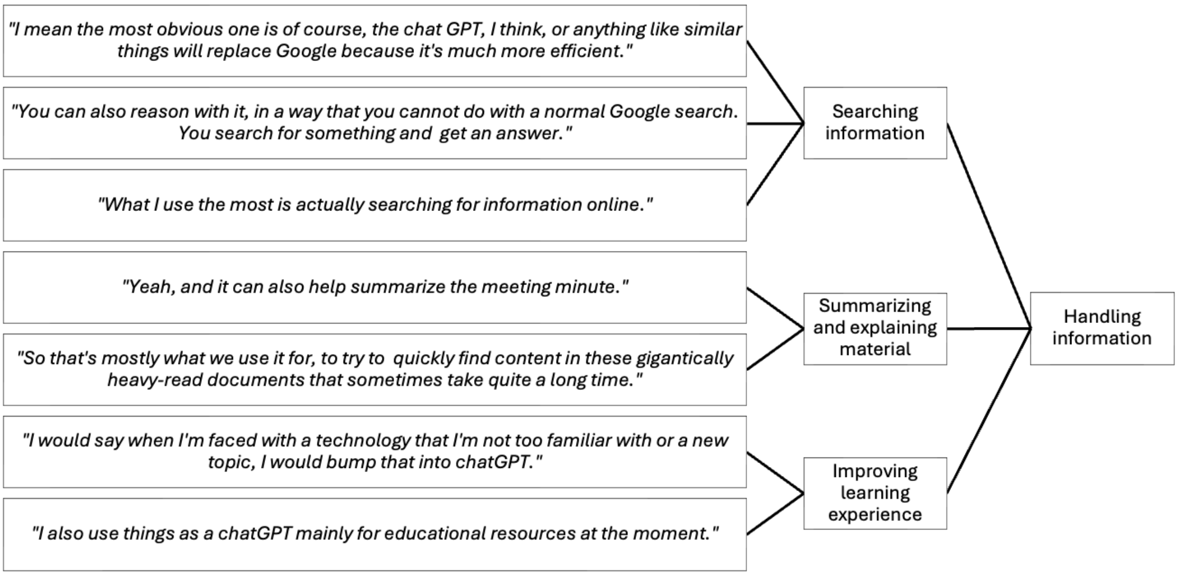

Figure 3.1 Example of the coding process before final definition of theme names

The final step before producing the report was searching for relevant quotes, showcasing the meaning of each theme, before writing up chapter 4 (Empirical Results).

3.6 Research Quality

When doing research, it is important to assess the quality to determine if a certain study creates value. Bell et al. (2022) describe the three most common evaluation criteria for research: reliability, replicability, and validity. However, the authors further explain that their relevance for qualitative research has been questioned and might need to be altered. Instead, the trustworthiness criteria credibility, transferability, dependability, and confirmability can replace external and internal reliability and validity. Additionally, this section will end with a reflection of the methodological choices.

Credibility concerns the acceptance of research findings by others (Bell et al., 2022). It is important to have good research practice and share findings with those included in the study to ensure everything was correctly perceived. By sending the material to respondents, respondent validation was achieved, increasing credibility. Another way to ensure credibility was by using triangulation to compare different data sets collected through using more than one method, in this case interviews and literature review (Bell et al., 2022).

Transferability is the degree to which the findings can be applicable to other contexts than the one studied. In qualitative case studies there is a risk of lower transferability due to the specific empirical context. A way to mitigate this is to produce a thick description, meaning a thorough description of a culture (Bell et al., 2022). This can provide others with enough information to judge if certain findings can be transferred to their desired context (Bell et al., 2022). In section 2.3, Empirical Context, the empirical context of the case company was described as thoroughly as possible to achieve higher chances of transferability.

Dependability can be compared to reliability in quantitative research and is about how well procedures in the research process have been followed (Bell et al., 2022). To ensure dependability, records were kept of all phases and steps of the research process and shared with supervisors at Chalmers and Zenseact who somewhat worked as auditors to the process.

Confirmability should ensure objectivity in the research. However, Bell et al. (2022) explain that this is impossible due to personal values. Although 100 percent objectivity cannot be achieved, researchers should act in good faith, meaning their personal values should not deliberately influence the research and findings. Auditors can also help with ensuring confirmability when reviewing the research. To decrease personal bias, a bottom-up approach was employed when analyzing empirical data.

Reflecting on methodological choices, it is significant to note that a limitation of this study is its exclusive focus on a single software company, thereby classifying it as a single case study. The first research question aimed to cover the full range of departments within the company to ensure a broad viewpoint, yet it remains confined to the unique experiences and perspectives within this specific organization. Consequently, the generalizability of the findings to other companies or industries may be limited, given the distinct operational and cultural attributes of the company under study. Additionally, their case is very special as they have very high safety requirements due to producing software for autonomous cars and that they are making an advanced and complex product, which means that many tools cannot help them because they are so niche. Furthermore, the second research question is focused on the context of software developers, which may not completely reflect the broader implications of the use of generative AI across other roles and departments within the company.

3.7 Ethical Considerations

Ethical considerations are essential in upholding the integrity and credibility of research. The study ensured ongoing monitoring and assessment of ethical issues, recognizing that they could arise at any stage, to uphold responsibility and trustworthiness. In line with Bell et al.'s (2022) framework, this study incorporated the four fundamental ethical principles: *avoidance of harm, informed consent, protection of privacy and confidentiality*, and *prevention of deception*, guiding the ethical conduct throughout the research.

Avoidance of harm was achieved by anonymizing the data to uphold the integrity of the participants. Interview respondents were also able to choose a setting where they felt the most comfortable such as a digital or physical meeting and select between Swedish or English for communication.

*Informed consen*t was ensured by providing prospective research participants with comprehensive information about the study, enabling them to make an informed decision about whether they wish to participate. Before starting the interview, the respondent was asked whether he or she was comfortable with the researchers recording and taking notes during the interview.

Protection of privacy and confidentiality relates to the need to protect the privacy of research participants (Bell et al., 2022). While informed consent is a limited surrender of privacy within the research scope, participants retain the right to refuse sensitive questions. Additionally, each case was treated individually, offering participants the option to withdraw their consent. The interview data was treated confidentially, and no personal names are mentioned in this report.

Prevention of deception seeks to ensure that participants are not deceived or misled during the research process (Bell et al., 2022). As part of this study, adherence to the principles of prevention of deception was ensured. Participants received clear and comprehensive information about the research, including its objectives, procedures, and benefits. The participants were also able to ask questions about the study.

4 Empirical Results

This chapter presents findings from interviews. Perceptions (related to RQ1) are presented in section 4.1. The sub-questions RQ1.1 and RQ1.2 are presented in section 4.2 Opportunities and section 4.3 Challenges respectively. RQ2 will be partially treated in section 4.1, where it is mentioned by some interviewees what could be interesting to evaluate in terms of GenAI use and to what extent GenAI is currently used.

4.1 Perception and Use

This section will briefly introduce to what extent generative AI tools are being used at Zenseact as well as the employees' perceptions of the technology. Lastly, some aspects that could be interesting to evaluate or measure related to the use of GenAI are discussed. Perceptions of employees have been discussed on different levels and questions were asked during interviews about the respondents' personal perceptions, their perception of their colleagues' perceptions, and how they perceive the use and integration of GenAI on a company level.

The interviewees mentioned various uses of generative AI technology, with varying degrees of engagement. Some have not used it at all or have only briefly tested it, while others utilize it more regularly. Additionally, a few interviewees indicated that they use the technology exclusively outside of work, mainly just for fun. Different generative AI tools were mentioned during interviews, where ChatGPT and GitHub Copilot were the tools receiving the highest attention. Other tools mentioned were Microsoft Copilot, Phind.com and Grammarly. Some also just said GPT or "tools such as ChatGPT" or similar things. It was mentioned how both the employees, and the organization are currently exploring and testing the technology.

> *"Not for work tasks, but of course I've looked around a bit and experimented with it." – Interviewee 27*

Within the company, it is described that the attitude among colleagues towards generative AI is perceived as positive and that there is an interest within the organization to use generative AI.

> *"Surprisingly high interest, everyone is interested. I think I've encountered someone who says, 'Ugh, I don't want to be part of this, I really don't want to be involved.' But 95 percent seem to think that we should make something good out of this. Much more than I expected." – Interviewee 4*

Additionally, most interviewees expressed that they are positive towards generative AI on a personal level. They are excited about the future and curious to explore how GenAI can be used and integrated more into their daily work. Some employees even referred to the technology as a game changer.

"I mean the first time when I tried ChatGPT I was kind of my mind was blown by it, and they'll change the game." – Interviewee 18

Some employees view it as a tool in their arsenal and therefore see no reason why it should not be used to facilitate daily tasks. As a result, interviewees expressed that they anticipate a future where more generative AI tools are incorporated into everyday work and allowed for usage in a variety of things inside the organization.

"So now I'm mostly hopeful and curious and a little impatient to see that we get in and can use it more. It would be really nice to get some new tools in here that will help us understand and apply our code." – Interviewee 26

However, the views on using generative AI were not purely positive and straight forward. In combination with being positive some concerns and factors that hinder the use were mentioned, which are elaborated on during this section. Many described having a certain skepticism and lack of trust in the technology as well as a fear of how the future might look, despite their overall positive perspective.

"I would say that my gut feeling would probably be somewhat positive, but I would not trust it completely." – Interviewee 13

"Yes, it's really exciting to see what it will become, and a bit scary too, even if it turns out well. Even if it's good overall, you still wonder, what will the world look like? I don't know, it feels a bit unsettling and just feeling that you might not recognize the world, you might not get some things back." – Interviewee 26

The uncertain future partly relates to the rapid development of GenAI technology. Some employees expressed a feeling of falling behind and struggling to keep up with the development or perceived that others around them were struggling.

"I have not been able to keep up. I mean, this is something that is changing every day, right? [...] I mean, it's very, very difficult to keep up." – Interviewee 22

"And then I noticed that partially there are some individuals who feel like they're not quite keeping up with the pace and who haven't really familiarized themselves with some basic technology, which makes it more difficult for them to understand this." – Interviewee 10

Even though most employees were positive about using GenAI, there was uncertainty about how it can be used within the company. Unclear directives from management on what is allowed to use generative AI for were highlighted as a barrier to its adoption. However, it was mentioned during interviews that there are company guidelines for using GenAI tools. This has been mentioned by some to be good since it brings up some of the risks with using GenAI such as leaking confidential information. Even though the guidelines are a step on the way to inform employees about use, management and the legal department could improve and clarify these to increase the use and integration of GenAI. If the company also provided licenses to certain tools, employees highlighted that they would be more inclined to use them.

> *"People feel like we don't have completely clear messages from our bosses about whether these are things that are completely okay to use. And for that reason, maybe people don't talk so much about things like that. If there was more clarity from legal and from management, like now we've bought this for you, use it, then it's much more, okay, great I am going to do that, compared to if you have your own subscription or something." – Interviewee 20*

Since Zenseact is a software company they have many employees knowledgeable about technology. However, the perception about GenAI was said to vary within different parts of the company. There was a divide in opinion whether the more knowledgeable deep learning engineers are more skeptical because they see the limitations or whether they are more positive and curious about the technology since they are in the field.

> *"Yes, I hang out very much with these types of deep learning circles, so everyone is really hyped, and they want to know everything that's happening."– Interviewee 2*

> *"Especially the people doing deep learning in the company and the people who have expertise on how these things work, seem also much more skeptical and reluctant just using it and trusting everything and thinking it will solve anything than the people who are not working with deep learning on a daily basis, who some of them seem to think that this is the magic source that will fix everything, which it probably won't be." – Interviewee 15*

Connected to the positive attitude towards generative AI, worries were raised over colleagues' attitude and some employees expressed concerns about people maybe even being so overly positive that they miss some of the risks and do not see the problems with the technology.

> *"So, I think there's more of a rather naive attitude sometimes than being hesitant or skeptical. I see more of a risk, like towards an overly positive attitude than the opposite in this organization." – Interviewee 6*

Among many of the employees, Zenseact was described as open to new technology and interested in trying new things. In an attempt to integrate GenAI, the company has internal projects for developing their own tools. Opinions about these projects differed between employees but they were mostly positive, and people thought the initiatives were good and show that the company is trying to be at the forefront of technological development. However, some employees believe that developing internal tools may be redundant, as there are already existing tools that perform well. Additionally, they noted that the internally developed tools have yet to reach full maturity and effectiveness.

> *"Yeah, but I would say that I think they've done pretty well, they have been pretty quick to enter this boom or whatever you want to call it that came with ChatGPT, they realized that this is something that will be used by many companies so it's best that we have some higher initiative and create something to get around this security issue of not using sensitive data." – Interviewee 21*

> *"I think they have good initiatives, but I don't think it's followed through well apart from the test data. I think that was quite fruitful, but for more developer use, I don't think there's really much." – Interviewee 25*

Although the company is seen as a leading tech company, it was raised that there exist individuals that are more negative or reluctant when it comes to using generative AI in their work. Resistance to change and a preference for familiar ways of working, which keep them within their comfort zones, were mentioned as significant obstacles.

> *"The challenge I see right now is that people are so indoctrinated, they are so used to working in a certain way already that you need to overcome that threshold and then maybe the stigma of actually letting someone else do the job." – Interviewee 9*

Some employees mentioned that they have no perception of the technology and that it is not a common topic of discussion within their teams. Others remain neutral, feeling that they lack sufficient knowledge about the technology and how to use it effectively. Additionally, a few employees noted that they have not observed any significant impact from using generative AI tools in their work.

> *"As I said, I probably know too little to express myself. I still think it's kind of a party product, I see this image generation, like you can create something fun with it. Yeah, that's where it starts. That's where you get the interest and then you start to see what you can use it for." – Interviewee 14*

> *"My life is quite unchanged, my work life too, so it hasn't penetrated my work life that much." – Interviewee 17*

While some people currently see no impact, they remain optimistic about the technology's future development. However, many employees find it difficult to understand the technology due to the rapid pace of its development and are uncertain about the timeline for it to become sufficiently advanced. This uncertainty fuels skepticism, with opinions varying on how far the technology is from being truly useful—some believe it still has a considerable way to go.

> *"I can't see, this development has been so insanely fast that I haven't been able to envision what the next step is; it could be anything. But I can hardly comprehend what the possibilities are." – Interviewee 8*

> *"But it's so far away, so it doesn't matter what you say in the coming 50 years, you'll still say another 50 years in 50 years. It's further away than we think." – Interviewee 27*

In connection to the lack of knowledge about how the tools can be used, some employees emphasized the importance of using the right tool for the right task. They argued it is crucial to only use the tools where they actually provide value and solve a problem.

> *"You always have to be very careful to understand when it is appropriate to use them [GenAI tools], when using them creates value, what limitations they have, and when they should not be used." – Interviewee 2*

Aside from solving identified problems with generative AI, most employees described that by using generative AI, they hope to achieve increased efficiency and become faster in their work.

“What is the main reason that you use generative AI tools?” – Researchers

“To save time, efficiency.” – Interviewee 18

As highlighted in the quote above, one of the main reasons for using generative AI is to save time. However, determining how much time is saved is challenging. One of the employees has attempted to quantify the time saved but was unsuccessful.

“I've thought about it myself, I've tried to quantify it, but it just isn't possible.” – Interviewee 8

Other effects mentioned as desired to quantify during interviews include how much more productive they become or how the quality of the code changes. However, definitions of productivity varied among employees, and they found it difficult to explain what they meant by the concept. Examples provided include improved margins or the value generated within a specific timeframe. Additionally, some employees mentioned the number of commits or lines of code as possible metrics within productivity. However, lines of code were dismissed as a representative measurement since it fails to capture dimensions such as code quality. While there was a general interest in measuring productivity, many acknowledged the difficulty in determining how to effectively measure and define it.

“So, it can be good to measure productivity in some way but it's very difficult to measure when it comes to code because you can't just say number of lines of code, it doesn't mean much. It's a bit difficult to measure how productive you are when it comes to programming. But if you can measure it and compare it with how you were before generative AI and after then maybe that would be interesting.” – Interviewee 5

“How would you define productivity?” – Researchers

“It's hard. Especially with programming, there isn't a clear definition.” – Interviewee 5

4.2 Opportunities

From the interview data, several opportunities for the use of generative AI were identified. The most frequently mentioned topics were automation and increased efficiency (section 4.2.1.), expanding knowledge and creativity (section 4.2.2.), and managing data (section 4.2.3.). These will be explained in more depth in the sections below. A summary of themes and categories within these themes is showcased in Table 4.1, with examples of quotes in.

Table 4.1 Opportunities of GenAI from interviews

Theme	Opportunity	Illustrative quote
Enhancing Operational Efficiency	Automating repetitive and redundant tasks	*"Sorting through one's email inbox or something similar, or just going through a lot of files and highlighting and so on, such repetitive tasks could easily be automated for many people." – Interviewee 26*
	Generating material	*"Another aspect is that it can help in the form of Copilot or other similar coding assistance tools, where we could receive help in writing or generating code without needing to manually write as much. One can have a kind of assistant that does a lot of the boilerplate work for you before going in and fine-tuning to get the desired end result." – Interviewee 17*
	Increasing job satisfaction	*"So, I would say that I have more fun at work and have become faster, I feel more efficient." – Interviewee 20*
Expanding Knowledge and Creativity	Searching information	*"But if you want to look something up briefly that you know you could Google, but you don't want to click through three sites before getting to the right one, then ChatGPT is very nice instead." – Interviewee 16*
	Improving learning experience	*"I really like having a learning buddy like this." – Interviewee 19*
	Summarizing and explaining material	*"Our developers have way too much to do, they have too much in the backlog. When they are to make a change, it is too much to think about, you have to meet requirements, you have to keep track of everything, and there I think we can find a lot of improvements that reduce everything a developer has to think about. And I think GenAI is very good at this, gathering and summarizing information." – Interviewee 19*
	Stimulating creativity and inspiration	*"Or help with getting a bit more creative." – Interviewee 5*
	Inspiring new ways of working	*"I've gotten inspiration to try something new." – Interviewee 11*

Managing Data	Creating and modifying data	*“So, we could simulate all kinds of dangerous or rare scenarios in a realistic way with this neural network model.” – Interviewee 15*
	Handling data	*“If we only use like 1 percent of the data, it's just because we don't know how to search the collected data or we don't have an efficient way to search and explain the data.” – Interviewee 18*

4.2.1 Enhancing Operational Efficiency

The most significant opportunity identified was the use of generative AI to automate certain tasks or parts of tasks, thereby enhancing efficiency and saving time. Beyond these benefits, increasing productivity was also frequently mentioned as a desirable outcome of using generative AI. However, it was pointed out by one employee that efficiency and productivity are more or less the same and throughout the interviews these two concepts are mentioned synonymously. Tasks that are repetitive, autonomous, and boring were mentioned as prime candidates for automation through generative AI.

> *“Manual, boring, autonomous tasks which could be somehow sped up.” – Interviewee 16*

GitHub Copilot and other generative models such as ChatGPT were mentioned as the tools that have the highest potential to facilitate work tasks. Microsoft Copilot and other tools were also mentioned but they have not achieved the same level of widespread recognition as ChatGPT and GitHub Copilot.

> *“I think GPT and Copilot are probably the most relevant things, at least in the near future.” – Interviewee 15*

During the interviews, a variety of tasks were mentioned as currently automated or potentially possible to automate in the future. Notable among the mentioned tasks were tasks like taking notes, transcribing, and sorting or going through files or emails. Most often, however, the automation focused on tasks associated with generating material. This includes creating job ads, e-mails, presentation slides, interview guides, different types of texts, tables, various meeting materials, and templates. In relation to the time savings possible through using GenAI, it was noted by one employee how days are allocated to activities that are found most valuable and by optimizing specific tasks, additional time is freed up to focus on these tasks.

> *"The days are filled with what one considers to be the most valuable thing to do, so with every bit that can be made more efficient, a little more of the list of the most valuable things to work on gets incorporated into the work. If you save time on something it means you can do a little more." – Interviewee 2*

Some interviewees mentioned the benefit of using generative AI tools like ChatGPT to generate initial drafts for some tasks mentioned above. They said that even if these drafts are not perfect, they provide a helpful starting point where they can then easily edit and adapt the generated output to better suit their purpose or get it exactly the way they want.

> *"Providing some form of input, maybe in bullet points or something very concise, and then receiving some version of text to be used for a certain purpose, and then having the option to tweak it to get it just as wanted." – Interviewee 2*

There are also different applications of generative AI related to software development highlighted during the interviews. Some mentioned were code generation and creation of documentation for developed code. Additionally, it was mentioned how generative AI tools could also be used to improve and optimize existing code or assist in code reviews. Another potential application mentioned was generation of tests for code testing.

> *"An obvious possibility, which is also extremely well-known, is that developers can get help with writing actual code. Using natural language, it is possible to ask the model to generate the correct code for this, and it's said to be very efficient." – Interviewee 3*

Using generative AI to generate code or getting help with improving code could then lead to a more standardized and streamlined codebase and code with shorter execution time.

Lastly, other effects of automation were that it can make repetitive and boring tasks more fun and enjoyable, thus making people feel more efficient and satisfied about their work situation. More efficient employees can in turn lead to saving money by not having to hire as many people.

> *"So yes, automating, streamlining and making certain boring work tasks a bit more fun are the opportunities I think." – Interviewee 13*

4.2.2 Expanding Knowledge and Creativity

Another opportunity frequently mentioned by many interviewees involved using GenAI tools to facilitate different types of cognitive processes. The majority pointed out that GenAI offers an easier and more efficient way of searching for information and some also said that they use it to learn new things. ChatGPT has been described by many as a better alternative to traditional search engines like Google, as you can cut out the steps of looking through several links as well as the possibility of asking longer and more concrete questions than what you can search for with Google. This new way of searching for information saves time and is a more convenient and pleasant way of finding information.

> *"But Google or some other classic search engine often just gives a link to something and then you have to go in and analyze and extract just to realize that no, that was not what I was looking for, and then you take the next link. Whereas with ChatGPT and similar tools, most of the time it provides a pretty good explanation, summary and answers the question you had." – Interviewee 24*

Although using generative AI tools such as ChatGPT was acknowledged as a more efficient method for searching information, there remains some uncertainty about the actual amount of time that is saved over the course of a given period. However, it was mentioned as a more pleasant experience and makes finding information and learning new things more enjoyable.

> *"But it's nice, instead of going to Google and spending half an hour, I go to ChatGPT and spend fifteen minutes, but it's not 100 percent of my time, so the total gain is probably not many percent each week." – Interviewee 19*

Despite the advantages previously mentioned, ChatGPT has not completely replaced traditional search engines like Google, primarily due to concerns about the reliability of its output. Nevertheless, it proves highly useful for gaining an initial understanding of a topic. The response provided by ChatGPT can serve as a guide, directing users on what further details to search for on Google or to investigate more deeply.

> *"I usually use it before I Google something, let's say I have no idea about something, I would ask let's say ChatGPT first. It would give me something and then I assume that it's full of lies, but I know that it's in a better direction than me just being directionless. So, I know that there will be some truth to what it spewed out, and then I'll Google those things, hopefully with a better idea of what I'm looking for." – Interviewee 25*

For software developers, generative AI tools like ChatGPT can be valuable for asking questions about different developer tools and for learning new things.

> *"It can take a lot of time to familiarize yourself with sections of code, and if you can get help, not just with code, but with how to use various libraries and how to utilize different technologies and third-party tools, and so on…" – Interviewee 17*

Learning was mentioned as another opportunity related to the use of generative AI tools. These tools can accelerate learning processes and the models' abilities to explain complex things in simpler or different ways as well as optimizing the learning process play a role in facilitating understanding and learning. Different ways of utilizing tools like ChatGPT for educational purposes were mentioned.

> *"I also use things as ChatGPT mainly for educational resources at the moment. I use it a lot, for example when I have time constraints and I'm interested in learning about this topic, but I only have 3 hours, then I ask tools such as ChatGPT or others to just make a study plan: I need to learn this in this amount of time." – Interviewee 22*

In relation to learning, interviewees also mentioned that by using tools like ChatGPT to address the most basic questions, employees can elevate their knowledge base to a higher initial level. This allows employees to bypass the step of consulting senior employees or experts for fundamental inquiries. As a result, more advanced questions can be reserved for experts and senior staff, enhancing the quality of discussions and interactions.

> *"We see that it's for exactly these questions where previously you have needed to... Well now I'm going to work in this area which I have no clue about. I need to find some expert within the company who does not like to be bothered and ask him how this works, and they'd be annoyed and tired of getting the same question again. Now I don't have to ask him, I can just go past him and ask ChatGPT instead. I think both him and I think that is pretty good." – Interviewee 20*

The effects of the previously mentioned opportunity of skipping the initial conversation, were described as enabling individuals to confidently approach specialists and become braver in interactions with other colleagues, as well as feeling an increase in self-confidence and independence.

"I think it opens doors. I dare to contact specialists without feeling completely clueless [...] it's easy to forget these dimensions, like what it does for your self-esteem and productivity. You feel capable when you can do things that you don't fully master on your own. But with that help, it's like having a little teacher." – Interviewee 8

Furthermore, GenAI tools can also help explain and summarize existing material as well as the newly discovered information. Areas of use mentioned during interviews were summarizing, translating and analyzing texts, summarizing communication threads, and getting different things explained. Specific for software development were code-related aspects such as explaining code or error messages as well as making code more understandable.

"I think if you start off with new code it can be very overwhelming. I think using a tool like GPT can definitely make it less overwhelming. So, let's say there's a piece of code I don't get, it's in a different language. I chunk it into whatever model, and it can just summarize and tell me what the gist of it means so that I can then go and read up in the documentation about each one of those specific things. I think it's very daunting to read through very thick manuals and stuff to try and understand something." – Interviewee 25

Another opportunity mentioned during interviews was to have more internal generative AI tools. There are some existing initiatives at Zenseact, however, these tools are still in the early stages of development or primarily being tested in pilot projects, so their effectiveness and reliability may not be fully realized yet. An internal chatbot would be beneficial for searching through company documents and answering generic questions from employees. Some employees struggled to find specific information related to their employment, such as details on parental leave or company vehicle policies. Additionally, there exist a lot of internal language and abbreviations to keep track of which this kind of chatbot could help with. For software developers, a specialized coding tool trained on the company's unique codebase would be advantageous since generic programming tools often fall short in addressing the specific challenges faced due to the uniqueness of the company and its specific use case. The quote below showcases one of the problems that could be solved by having an internal chatbot for answering questions.

"I miss this quick tool to just look up different terms. ... so, a really clear use case would be to have something that can answer these simple questions, like when is payday? Who do I need to contact if I have questions about parental leave? What's the abbreviation for this? Where can I find information about this? I don't think it's very difficult to implement such a solution instead of focusing too much on the super advanced stuff." – Interviewee 16

Generative AI tools have also been brought up to break old habits and find new ways of solving problems. One feature of ChatGPT that was mentioned as useful was that it is an interactive tool where you can have a conversation with it, and you can ask for input on ideas or get inspiration.

> *"I use it a lot as a sounding board, to get something new that hopefully would have made me better in some situations. It's easy to do the same things all the time and it's convenient when there's something old to take from, but at the same time, you can also challenge yourself and see if you could do it differently and get inspiration, which led me to do something differently." – Interviewee 11*

It was also noted that generative AI tools can act as an idea generator, facilitating brainstorming and offering new ideas that might not have been previously considered. Additionally, it was mentioned that these tools can enhance creativity.

> *"And I think many of us see that this will make us more creative, and you want to be creative, you want to create value." – Interviewee 19*

4.2.3 Managing Data

Opportunities that are more specific for Zenseact are ones related to collection and handling of data. As of now they are collecting data by driving their test cars, which is then used to train the models embedded in their autonomous vehicle software. While there has been a focus on modifying data, some interviewees saw a significant opportunity in creating new datasets from scratch. Although small modifications have been carried out so far, the development of new data is currently a promising research area. Zenseact is looking for rare traffic scenarios to improve their models in different situations but collecting such data is difficult. Rare scenarios could be a car crash or other dangerous scenarios that one would not want to cause on purpose.

> *"And then, for things like what we do, it also involves being able to create data on extremely rare scenarios [...] there are many types of scenarios that are very uncommon, typically those that are safety-critical when something really happens, like a crash or near-crash. These are very rare and often quite different from each other, and we can't collect such data because it's unethical to just crash things. So then, there are opportunities to create such data using generative models." – Interviewee 2*

Another advantage of creating synthetic data would also be economical aspects in terms of cost savings by not having to drive around to collect the data. Right now, lots of data is collected with the test-vehicles to find certain scenarios. If this data could be searched, sorted, and labeled (annotated) with generative AI, this would also reduce costs and save time.

"Hopefully, we have a big revolution ahead of us when it comes to navigating our data. For example, we gather lots of data but maybe we could get even more out of it." – Interviewee 4

4.3 Challenges

Several challenges were brought up during interviews with the most frequent topics being technological limitations (section 4.3.1), societal and ethical aspects (section 4.3.2), legal and security (section 4.3.3), and job market and market dynamics (section 4.3.4). Within these broader themes, different challenges were identified and a summary with examples of quotes showcasing these can be seen in Table 4.2.

Table 4.2 Challenges of GenAI from interviews

Theme	**Challenge**	**Illustrative quote**
Technological Limitations	Low output quality	*"And then the quality. It may have given the answer that solves the task, but not in an optimal way." – Interviewee 24*
	Hallucination and false confidence	*"Yes absolutely! Hallucinations, they make things up and they sound very convincing." – Interviewee 17*
	Limited database	*"And then, if there had been current data from now, then it would have been a bit more useful, I think." – Interviewee 9*
	Compute power and energy consumption	*"They are extremely large and use a huge amount of computing power" – Interviewee 2*
	Explainability	*"At least not as long as AI is as much of a black box as it is where you don't understand why things come out." – Interviewee 6*
Societal and Ethical Aspects	Bias and harmful output	*"Yes, I mean, I can imagine quite a bit when it comes to the people area because it's somehow built by humans in some way, and there's a risk that we build in bias into what we use." – Interviewee 11*
	Misuse and over-reliance	*"I'm quite worried actually, about this deep fake technology and so on. It feels extremely problematic that people can't trust information." – Interviewee 26*
Legal and Security Concerns	IP and security concerns	*"I guess for our company it would be legal and security related matters. So, whenever we use anything that includes AI or generative AI in general, we really need to be properly fenced against any security issues legally. So, if we train or do something with our data, we need to make sure that it doesn't leak." – Interviewee 22*

	Responsibility	*"If decisions are made based on and code is written based on these tools, and it leads to an undesirable situation, who is responsible? Legally speaking. That would probably need to be further investigated, I believe." – Interviewee 24*
Job Market and Market Dynamics	Job displacement	*"There are those who believe that there won't be coders within a couple of years, it seems." – Interviewee 26*
	Transformed work tasks	*"So, then you need to learn how to navigate this and how to write prompts properly." – Interviewee 22*
	Dependence and power position	*"So, there's a risk of locking in a few players, or it's already pretty locked in, I'd say. The research and so, or the development, is conducted largely by 3 or 4 companies in the USA." – Interviewee 17*
	Workplace integration	*"It's the human aspect, to get people to understand what it is and what it can do." – Interviewee 5*

4.3.1 Technological Limitations

The most significant problem raised by employees relates to limitations of the technology underlying generative AI. A key issue is that the produced output often includes errors and inaccuracies and that it can be difficult to determine what is reliable and correct and what is not.

> *"...a bit too many errors and it generates things that look reasonable but aren't necessarily entirely correct or very useful." – Interviewee 4*

This concern is not only connected to general usage, but it also applies for code generated by GenAI coding tools where subtle errors and bugs may be introduced into the code. These issues can be particularly difficult to detect and correct and potentially lead to significant complications in development projects.

> *"Then I'm still afraid that we might include bugs in the draft, for example, that could be so subtle that we don't see them." – Interviewee 17*

It was not just the issue about hard-to-detect errors in the code generated by generative AI tools that concerned some employees, they also raised concerns about the overall quality of the generated code. As highlighted in the quote below, the generated code can frankly just be bad.

> *"So, the code generation is huge. I think it's just about how we're going to make it work in practice because the code that's generated isn't that great." – Interviewee 19*

The poor quality of the generated code was described as having a possible connection with how the model is trained and on what data.

> *"Right now, yeah, that's when you start to see the limitations, so I can see that these tools are often trained on what's out there, and my impression of a developer now is that it's quite common for them not to be good at writing code that's easy to understand." – Interviewee 26*

As mentioned, many employees have noted the challenge in verifying the accuracy of code generated by AI tools, making it difficult to trust the output completely. It was highlighted that you should not fully rely on what these tools produce and that verifying the generated code is necessary. This requirement can eliminate some of the efficiency gains expected from using these tools, as the errors and flaws introduced by generative AI can be subtle. Once these errors occur, it might take more time to detect and correct them than it would have taken to write the code yourself from the beginning.

> *"I got a piece of code from ChatGPT, it took me 2 days to get it working so it would be faster to do it myself, so that's the only time I've actually used it for work so it's nothing I'm throwing myself into right now." – Interviewee 23*

A way to minimize the risk of using incorrect outputs from generative AI tools is to use them primarily in areas where you already have some knowledge. It is important to be critical and have the ability to correct the output or to recognize when they are unusable. However, one point raised by one employee was that if you are an expert on a subject, you will identify more errors, as you have a clearer understanding of what is correct and what is not. In such cases generative AI tools will become less beneficial, and continually correcting the models or not getting what you are looking for can be very tiresome. On the other hand, if you are a beginner and looking for simple or more brief explanations the tools can be more useful.

> *"I think the more specialized you are in something, the harder it is to use a tool like this since you only get to say, 'you're wrong, you're wrong…' but when you're a beginner, then it's just awesome." – Interviewee 19*

Incorrect output or the need for frequent corrections can lead to decreased usage of generative AI tools, as some employees have stated that they gave up on using the tools when they did not work well enough. They felt like they were wasting their time asking tools like ChatGPT questions and not receiving a satisfactory response. There can also be more limited use if people do not know how to use the different GenAI tools and what they can be used for. The ones who still manage to use GenAI in a somewhat helpful way stress the importance of being critical and checking the correctness of the output thoroughly since it can be difficult to tell if the output is correct or not, as mentioned previously.

> *"An obvious limitation, and the challenge is that these models almost always provide an answer that sounds very confident, but it can be completely wrong. And it still happens relatively often that they are completely wrong, and then there is very little evidence that something has gone wrong, so you have to be critical when reading the answers." – Interviewee 3*

Another reason to be critical to GenAI mentioned by employees was the lack of understanding of the logic behind these tools. GenAI logic has been referred to as a "black box" where no one understands the reasoning of the models and how they come up with answers.

> *"AI and like, machine learning, deep learning in general, I think can be, well, it's a challenge because you have no control over what is actually happening at all. As far as I know, there is nothing yet that explains how they arrived at the final result, but it's like you put data in one end and then something comes out the other end." – Interviewee 14*

A reason for completely wrong or made-up answers given by generative AI is that there is not enough or no data on a certain topic. If the training data is limited, GenAI will "hallucinate" answers, meaning it just comes up with something that is completely made up. For Zenseact this can be a challenge since their context is safety critical and very unique. This means it might be difficult for developers to use some of the GenAI coding tools since they have no training data on their special cases. There were other problems related to the training data of the model mentioned by some people such as old data or low quality of the data resulting in either outdated or low-quality output.

> *"From a work perspective I think it is problematic that the data is capped back in time. I have a lot of questions that are dependent on the present moment, based on what happened yesterday. Whereas many of the models are limited to, say, 2021, 2022, and so forth." – Interviewee 1*

Another crucial aspect of generative AI mentioned by some employees was the size of the dataset on which it is trained, as well as the hardware it runs on. According to one employee, the compute power and the volume of training data are the most critical factors in determining the performance of generative AI tools. By scaling up the models with more comprehensive training data, improvements in their capabilities can be achieved. Additionally, enhancing the compute power can significantly speed up the processing time of these models, making them more efficient and responsive. These requirements result in the superiority of existing tools in comparison to new, smaller versions.

> *"One insight people got from GPT, but in other things as well, is that the thing that really matters is not how smart you are in designing your neural network, it's more of how big it is. How much computing power do you have and how much data can you feed it?" – Interviewee 15*

However, another employee raised concerns about compute power from a different angle, noting the significant energy consumption of these models, which has implications for environmental impact.

> *"They consume an enormous amount of energy, these models. It could be a problem. Do we have that much energy and then the excess heat needs to be cooled and everything with the technical and ecological problems this brings." – Interviewee 24*

4.3.2 Societal and Ethical Aspects

A challenge on a more societal level raised by some employees concerned ethical issues in the form of bias built into generative AI tools. Different biases are built into the models since they are human-made and are based on specific training data. These biases will only get stronger as users continue to give the algorithms more biased data. While the problem will not be so prominent in the software development process itself, it could impact the recruitment of developers and other more HR-related areas and in communication.

> *"There is a danger in us putting bias into these models and so on, then they come back and answer our questions and somehow reinforce this bias back to us. But it is perhaps a problem on a societal level or a problem for our civilization rather than for work where it is usually about quite dry things where you want yes or no answers or have specific technical details like that." – Interviewee 17*

Even though the problem of bias does not directly affect the engineers in their coding, there was a general concern outside their work role about the potentially harmful outputs that biased models could generate. Such harmful results can contain, for example, racism, and there are currently no effective controls in place to manage or prevent these problems.

> *"There's no control mechanism or something fundamental in how these models are constructed that prevents them from saying bad things, outright racist things, or providing incorrect facts or anything like that. There are no barriers against it, and that's a huge challenge." – Interviewee 13*

Additionally, some employees expressed concerns about the potential misuse of generative AI to produce misinformation and fake news, commonly known as "deep fakes."

> *"If we step outside of what I'm working with. So now it's possible to fake all sorts of videos or speeches, so you can get famous people to say strange things and all sorts of stuff, and that's very dangerous, I believe." – Interviewee 21*

Although the problem mentioned above is more on a societal level in nature, there are other forms of misuse that are more workplace-related, such as cheating in tests during an employment process or using GenAI tools to appear more competent than you actually are.

> *"There is a risk that in 5-10 years, we end up in situations where we have leaders and team members who may not actually have the technical competence needed for the domain problem, but they have slipped in there because they are a GenAI prompt pro instead." – Interviewee 19*

Another worry mentioned by the employees was that there is a risk of becoming over-reliant on tools, using GenAI to generate content and ideas and stop thinking on their own. This over-reliance can result in loss of knowledge, since if we trust generative AI tools to do things for us, we might get less creative over time and lose our understanding of how things work.

> *"The negative aspect is that people will also be less creative because of this. People won't get used to thinking much; they'll just ask ChatGPT the question." – Interviewee 5*

> *"The downside, or the other side of the coin, is that you become a prompt engineer who tells an AI to write everything [code] for you, and then you have no understanding of what it ultimately spits out." – Interviewee 17*

If people become overly dependent on generative AI, neglecting to think by themself and relying on GenAI tools to do it for you, this issue could escalate to a societal level. Since these models are trained on existing data, i.e., current knowledge, if we stop to think for ourselves and fail to provide new knowledge not only will the models eventually stagnate, the knowledge level in society will also stagnate.

> *"For generative AI in general everything relies on statistics and history. For there to be continued positive development, some people must still contribute by continually adding new knowledge. Otherwise, everything risks stagnating because we're building everything on existing information if everyone stops thinking and just uses AI." – Interviewee 6*

Other employees brought up similar points and said that it will be important not to be completely dependent on GenAI tools to keep being innovative. Instead of completely relying on tools we should use them to boost the knowledge and competence that already exists.

4.3.3 Legal and Security Concerns

A challenge mentioned by several employees revolved around legal aspects such as Intellectual Property (IP) rights and security concerns. Related to what was mentioned previously about the logic of the technology and that it is impossible to know how the model arrives at the specific output, it is also impossible to know how it has been trained and on what data.

> *"And furthermore, the danger for a company like ours is that these tools are trained on... well, we don't know what they're trained on, so it's possible that it incorporates code from something that's copyrighted even if they claim it doesn't, and then we're stuck with copyrighted code and in a few years' time some American lawyer comes along. That's devastating for a company." – Interviewee 24*

Others said similar things about needing to be careful with copying output straight off from the tools because of the lack of knowledge about who legally owns the material. A problem specifically tied to some external code assistance tools and coding is that the tools are trained on open-source code, which may be covered by specific IP licenses. This means that if the code that is generated includes such material and is then included in a company's code base, it may be challenging to legally defend. This and the uncertainty around IP related questions is showcased in the following quote:

> *"And a bunch of things like that, open-source issues, and if you take for example Copilot, what license does it fall under, the result? Is there anyone who can answer that? No, there's no one who can answer that." – Interviewee 6*

The problem is two-ways, meaning there were also concerns about what people put into the models and that companies such as OpenAI will get access to all data put into ChatGPT and can train the model on that data. Multiple interviewees mentioned the concern of leaking company data by putting it into generative AI tools.

> *"Another challenge is also like if you use some model online and just paste code in there, you could leak IP and such things to people you might not have wanted to share it with from the beginning." – Interviewee 13*

On the other hand, prior to the introduction of these language models, people used Google to find solutions to a variety of problems that they encountered, and still do. Some employees pointed out the uncertainty of knowing what Google does with the information and the possibility that Google may keep and use this data.

> *"One has to be aware that the information you send into these tools ends up with the companies that operate and offer these tools. The questions you write to ChatGPT end up with OpenAI and so on. But at the same time, it's the same when you Google something; they also save what you Googled. So, I don't see a very clear difference there. But of course, that's something you need to know and think about." – Interviewee 20*

Also related to who owns the tools' output and what happens with the users' input, there is the question about who is legally responsible for consequences directly related to the output of generative AI tools. If code generated by GenAI consists of bugs and errors which cause an autonomous vehicle to crash, who is responsible? This makes it more difficult for Zenseact, in particular, to use generative AI tools to create their product since this is software for autonomous vehicles. The safety standards are very high and the need for code validation would increase since, as previously mentioned, the logic of the tools is described as a "black box".

> *"If we were to generate code for safety-critical parts of the car and so on, and it turns out that there are bugs in it and the generated code causes a serious accident, how do you deal with this?" – Interviewee 17*

4.3.4 Job Market and Market Dynamics

Because of the potential of generative AI and with all it can produce there were concerns about the risk of AI outcompeting and eventually replacing human workers in their jobs. The most extreme scenario mentioned would be if most humans were replaced and a universal basic income was introduced. Certain professions and tasks—mostly administrative ones—appear to be more at risk or uncertain than others. There were also individuals who were not concerned about being replaced and do not believe GenAI can perform some tasks that humans are capable of.

> *"It's perhaps that it doesn't really seem to be able to replace creativity entirely. Sometimes the question is discussed whether it will replace humans, but I would say that human creativity cannot be replicated today, so we will still be needed for a while." – Interviewee 21*

When it comes to software development and code writing, people had different perspectives on the possibility for GenAI to take over. The possibility (or risk) of all code being created by generative AI in the future was raised, although many pointed to technological constraints and low output quality as evidence that humans will still be required, at least for code review. However, the requirements may change in the future, and you may need to learn how to use the tools and re-skill if necessary to stay competitive in the job market.

> *"But within a very few years, you have to be able to use these tools. You have to be able to work together with such a Copilot or some AI-generated code because otherwise you won't stand a chance of keeping up." – Interviewee 2*

According to some employees, generative AI tools have the potential to greatly improve human performance in terms of quality and efficiency, making it more difficult for those who do not utilize them to compete with those who do. Some people run the risk of falling behind and are not keeping up with technological advancement, making it critical to adapt and embrace change.

> *"But otherwise, I think it's just about adapting and redirecting so that you incorporate it [GenAI] more, because this is going to be the new normal. That's what I mean by us being the threshold. It's like our parents who aren't into technology, or at least my parents. It's the same threshold." – Interviewee 9*

As highlighted by the quote above, the integration of generative AI tools into the workplace could become the new normal and it becomes important to know how to effectively use such tools. However, it was mentioned by some employees that the knowledge about GenAI varies across the workforce where those with limited knowledge about the technology will face the challenge of keeping up. One solution to bring everyone to a common level of knowledge mentioned was through having "Knowledge boosts", a concept used by Zenseact to share and spread knowledge.

> *"I guess one of the main limitations, I think first of all, is that knowledge is limited inside the company. I think it's good first to have some general knowledge, sharing some competence into what generative AI is, what those tools include." – Interviewee 22*

With integration of GenAI into the work environment, some employees speculated that the role of a software developer might undergo a significant transformation. Instead of writing code manually, developers might shift to primarily giving prompts to a generative AI tool and receiving generated code in return, essentially becoming prompt engineers. The main task for developers will then be to control the output and debug the generated code to ensure correctness and high quality, which according to some is not the most enjoyable part of the software development profession.

> *"With the ease of generating code in the future, I don't really see why the amount of generated code wouldn't far exceed handwritten code, which also means that the number of bugs… If everything is about the same ratio, there will be a ton of code that you have to go through and debug, and so on, and that's not at all what I want to work with at Zenseact, fixing AI-generated code. It would truly make it completely impossible for me to feel any joy at work." – Interviewee 17*

Another challenge mentioned associated with this potential new normal was the amount of code that could be generated. As creating code becomes easier and faster with GenAI tools, a significant challenge will be to manage the increased quantity of code, which will still require verification and oversight. As highlighted in the quote above, the primary responsibility of developers may shift to supervising the output of generative AI tools. However, one employee noted that humans are not particularly good at this task, which could lead to potential issues in maintaining quality and accuracy.

> *"I think still being critical and being able to correct things is important and a challenge because we humans are not really good at monitoring things [...] people just need to sit in a self-driving car for half an hour and then they start watching TV or doing something else because monitoring activities are something that we are very bad at. People are very bad at concentrating on not doing anything themselves and checking that someone or something else is doing something well. So, I don't think it will be easy to monitor a neural network either for it to do the right thing all the time, so it's a challenge." – Interviewee 13*

It was also raised that the increased efficiency that individuals might gain from using generative AI tools can lead to a reduction in demand for labor. The reason mentioned was that by doing their work in a shorter amount of time, they would have extra time to complete other people's work tasks as well.

> *"AI doesn't take people's jobs, but if you use AI properly, you take your competitors' jobs, so to speak." – Interviewee 4*

The challenge now presented and raised was deciding how to manage a potentially redundant workforce. Employees who do not utilize these tools and thus fall behind in productivity might be fired. Alternatively, the workforce could be maintained at its current size to increase production without raising costs, effectively scaling up operations.

> *"If we increase productivity as much as we possibly can with coding tools and other means, what do we do with the resources and time that are freed up? Should we lay off half the staff – preferably not. Should we find other tasks for them? So, it's also investigative work." – Interviewee 24*

However, if you need to use generative AI tools to stay relevant and competitive in the job market, concerns were raised that you might get dependent on the tools, putting the companies that provide them in a position of power.

> *"So, I can imagine that it might not be so good for one's career in the future if one is very dependent on tools. Right now, ChatGPT is free, but if OpenAI wants it to cost, say, $200 a month, if one is dependent on it, one will pay for it, so the companies can control the price. You shouldn't become 100 percent dependent on a tool that you don't have control over." – Interviewee 5*

This may also happen at an organizational level when a company depends on tools supplied by one or a small number of companies, which was raised by an employee as a potential source of monopoly or oligopoly.

> *"And then I just thought, when all software companies are dependent on, for example, ChatGPT to produce their software, do we even want that world? Because then they [OpenAI] basically own all software development, because they can just tell a company all of a sudden that, well, we have new terms of use, follow them or we'll shut off the tool for you. It's too much of a power position that it's basically one company that owns it, and then Google and Microsoft come in as competitors. If there are only 3 companies in the world capable of building LLMs, then I feel like we shouldn't use them because it's only those 3." – Interviewee 26*

Additionally, the importance of compute power mentioned in the technological limitations section above also enables oligopoly since size is so important. This risk already exists today as brought up by one employee.

> *"It really requires immense resources for the models, they're scaling up all the time, so it really requires a lot. So, there's a risk that it locks into a few players, or it's already quite locked in, I would say. Research and development are largely conducted by 3 or 4 companies in the USA." – Interviewee 17*

To tackle this problem, politics and how the market is regulated will be important. However, since technological development is happening very fast, the laws and regulations are struggling to keep up.

5 Analysis and Discussion

This chapter will compare and discuss findings from previous research with findings from our empirical study to answer the research questions. RQ1 will be answered in section 5.1, RQ1.1 and RQ1.2 will be answered in section 5.2, and RQ2 will be answered in sections 5.3 and 5.4.

5.1 Perception

RQ1: How do software industry employees perceive the use of generative AI in relation to their work?

Most interviewees at Zenseact hold a positive view towards the use of GenAI, with many expressing excitement and curiosity about integrating these technologies into their daily operations, reflecting the company's forward-thinking culture. They expressed belief that these tools can make their jobs more satisfying by automating routine and redundant tasks. This positive perception could be attributed to the perceived usefulness of the technology, which aligns with studies indicating that perceived usefulness significantly influences the intention to adopt technology (Abed, 2020; Davis, 1989). However, the perceived usefulness varied between different employee groups. For example, some employees believed that those working with deep learning daily hold more skeptical views while others found the deep learning employees more hyped over the technology. Employees with leading roles perceived those working within deep learning as more positive towards GenAI. This might be because of their more holistic and strategic perspective. It is important to remember that how the technology's prospect is perceived is not the same as how its functionality and use is perceived. From interviews it was evident that developers are generally positive about the technology but found the functionality less satisfying. Deep learning engineers, for instance, showed a mix of skepticism and positive curiosity, influenced by their understanding of the technology's limitations and possibilities since they are more operative and might focus more on their role and everyday work tasks. This indicates a correlation between an employee's level of technological expertise and their perception of GenAI (Venkatesh et al., 2003). Additionally, employees expressed significant concerns regarding the security and reliability of generative AI. The interviewees were positive about the future development of GenAI but lacked intuition about how far away the technology is from being truly useful and having satisfying functionality. Skepticism and lack of trust, along with the uncertainty of the technology's future, indicate perceived risks that could hinder adoption.

The company's culture of being open to new technology and running internal projects to develop GenAI tools indicates a supportive organizational context which directly might influence how employees perceive GenAI. The generally positive attitude towards GenAI and the interest in integrating it more extensively into work processes may suggest that social influence mentioned by Venkatesh et al. (2003) is significant. However, due to the unclear directives and identifying as a software developer where you feel like you should do the coding yourself because of professional pride, GenAI is not a very commonly discussed topic among employees. The degree and content of discussions about GenAI varies between different teams where some teams have a very limited amount of discussion while other teams have discussions about the technology but mostly make jokes about it and make fun of output generated by different tools. Senior developers perceived themselves as better at writing code than GenAI which also leads them to being more skeptical and resistant to using these tools since they are also used to their own, accustomed ways of working. This might reinforce negative perceptions about GenAI.

Employees' varied responses, from optimism to skepticism about the technology's rapid development and effectiveness, also reflect the environmental context's influence. The competitive industry that Zenseact operates in and internal company dynamics, such as the drive to be at the forefront of technological development, also play a role in shaping attitudes towards GenAI. While wanting to be at the forefront of technological development, several employees expressed that they are cautious about using GenAI because of the safety-critical aspects of their industry and IP-related risks. Identified opportunities and challenges can also affect the perception about using GenAI and these will be discussed in the section that follows.

5.2 Opportunities and Challenges

RQ1.1: What opportunities do software industry employees identify with using generative AI?

RQ1.2: What challenges do software industry employees identify with using generative AI?

This chapter will treat RQ1.1 and RQ1.2 where alignments or differences between interviews and existing research will be illustrated (see Figure 5.1 and 5.2). Using illustrations, each connection is categorized as either a direct, partial, non-existent, or influencing link. Direct connections are represented with solid lines, indicating strong alignment of identified opportunities and challenges between theory and empirical findings. Partial connections are denoted by dashed lines, signifying that there is some overlap between theory and empirical findings, but differences exist. There is no connection if the opportunity or challenge is left unlinked, highlighting where empirical findings have no support in theory. Lastly, if no connection is found between theory and empirical interview data but a challenge or opportunity affects others, this is indicating with an arrow in the direction the opportunity or challenge has an impact. The chapter begins with a comparison of opportunities where each connection will be explored to explain differences and similarities. After that, the same procedure will be done for challenges where these will be compared. Finally, there will be a more in-depth analysis of the two chosen themes facilitating task automation (section 5.2.1) and augmenting knowledge acquisition (section 5.2.2), where opportunities and challenges are analyzed and discussed together.

5.2.1 Comparison of Opportunities

Overall, the opportunities mentioned by both the theoretical framework and the employees are largely aligned. However, the employees tend to discuss these opportunities primarily from an operational perspective, whereas the theoretical approach encompasses a broader range of views, covering both operational and strategic aspects.

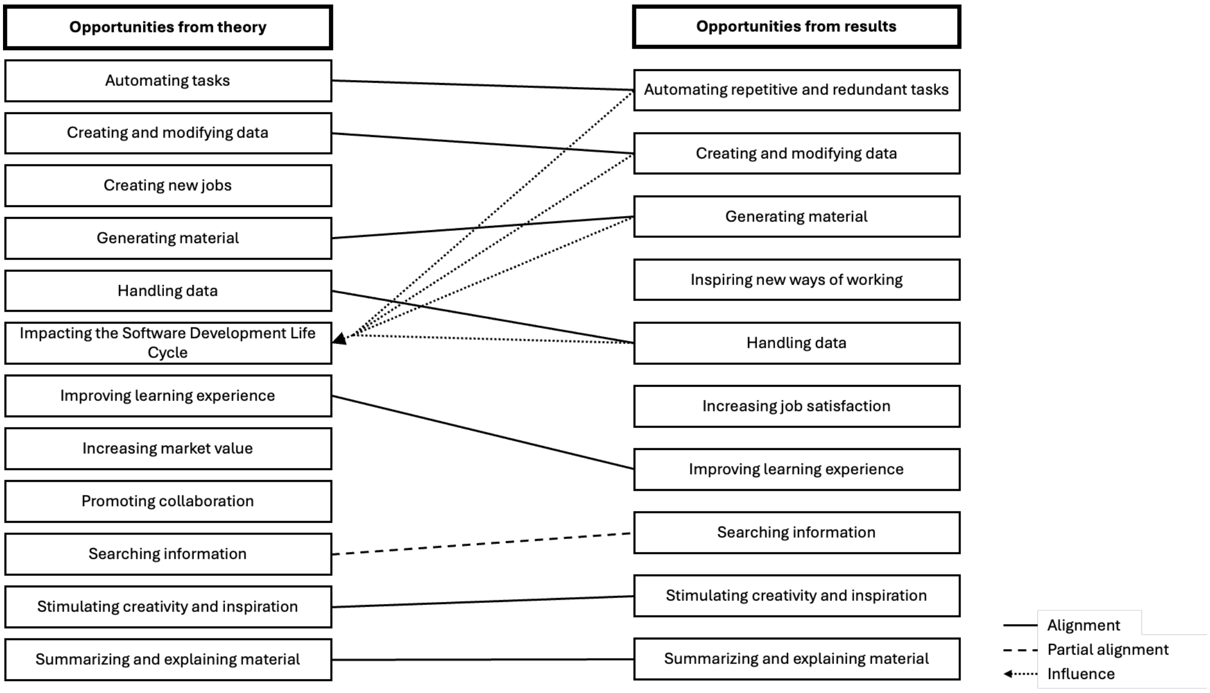

Figure 5.1 Comparison of opportunities from theory and interviews

The first missing connection from theory is creating new jobs. It is proposed by Ooi et al. (2023) that while GenAI automates tasks, it also enables new forms of occupations and jobs to be created. The employees did not specifically mention the creation of new jobs as a direct outcome of GenAI implementation, suggesting a gap between theory and empirical findings. That this has not been discussed by the employees at Zenseact may have to do with the technology being in an immature stage or that employees feel that they do not have any intuition for where the technology is going and how long it will take before the technology is good enough to be integrated on a large scale. Additionally, this can also be due to employees having a different perspective, focusing more on their own role at the job market.

The next opportunity from the theory that has no parallel in the empirical findings is promoting collaboration. Ooi et al. (2023) states that by saving time utilizing GenAI, more time can be spent on engaging and collaborating more with colleagues, customers, and business partners. While employees mentioned how GenAI tools like ChatGPT are used to assist in communication tasks, such as email writing, there is no mention of GenAI enhancing time spent in different collaboration constellations. The employees noted that with every minute saved additional time is freed up for performing the most valuable tasks during a workday. The employees have focused on supporting individual tasks rather than enhancing collective team processes. However, in connection to bridging the knowledge gap and collaborations discussed in theory, some employees expressed that GenAI can be beneficial in the sense that it can make you more brave in colleague consultations.

The last opportunity from theory that has no counterpart in the empirical finding is increasing market value. The theory discusses the significant potential economic impact of GenAI, suggesting a growth in the global domestic product and a substantial increase in the GenAI market size by 2032 (Bandi et al., 2023; Ooi et al., 2023). Employees at Zenseact did not mention the broader economic impacts, focus was instead on the immediate operational benefits and specific applications of GenAI within organizations. Additionally, it is described by Fui-Hoon Nah et al. (2023) and Agrawal (2023) that GenAI can enhance strategic decision-making by analyzing data to offer insights and predictions. While there was mention of using

GenAI for various informational and learning applications, there was no specific reference to its use in strategic decision-making or providing business insights. Employees focused more on task automation and content generation. This shows that the practical applications might currently be more tactical and operational rather than strategic.

An opportunity that partly aligns is searching information. Both theory and employees highlight that GenAI is enhancing information retrieval, making the process faster and more efficient. Fui-Hoon Nah et al. (2023) describes how ChatGPT can act as an assistant facilitating the process of answering questions and searching for information. However, in practice, employees acknowledged that while GenAI tools like ChatGPT simplify the information search process, there remains skepticism about the time savings and the reliability of the information provided. Employees expressed that they enjoy the ease and engagement of interacting with GenAI, however they still need to engage with experts and use traditional learning methods to gain a thorough understanding of complex subjects.

Pothukuchi et al. (2023) and Ozkaya (2023) suggest that GenAI can significantly streamline and enhance all phases of the SDLC. Empirical findings show that GenAI has been used in certain areas such as code generation and debugging, and opportunities are found in utilizing GenAI for automating testing, which aligns with what Pothukuchi et al. (2023) and Ozkaya (2023) suggests. Moreover, the employees described that a desirable future opportunity would be to create synthetic data for safety-critical aspects and find ways to handle all their data. These two opportunities are discussed by Bandi et al. (2023), Ebert and Louridas (2023), and Chui et al. (2023) respectively. However, the full integration across all SDLC phases is less evident in the empirical finding. That employees did not mention the integration of GenAI across all SDLC phases may be attributed to the newness of the technology, organizational readiness as mentioned by Baker (2011) and Abed (2020) in connection to the TOE framework, and the complexity of tasks in different SDLC phases. It may also have to do with the employees' position within the company and their work tasks.

Two aspects that are not adequately addressed in current theories include, firstly, how GenAI can facilitate the adoption of new work methods, allowing employees to break away from old habits and discover innovative approaches to their tasks. Secondly, using GenAI can lead to greater job satisfaction by making repetitive tasks more engaging and enjoyable. This indicates a significant gap in the theory, suggesting that the potential of GenAI to enhance workplace dynamics and employee engagement is not fully explored.

5.2.2 Comparison of Challenges

Overall, challenges with GenAI from previous research and challenges mentioned during interviews are mostly aligned, however, there are some that only partially align (see Figure 5.2). Challenges discussed in both literature and during interviews were often on a more societal level, although employees at Zenseact sometimes expressed challenges on a more individual or company level specific to their context. As Zenseact has also provided guidelines regarding use of GenAI tools and presented risks in relation to this, challenges mentioned by employees could be influenced by these. The main risks presented in the guidelines were confidential information leaks, unclarity in IP ownership, risk of IP infringement, misleading information, and privacy which were all mentioned during interviews. These risks are also brought up by literature which can be seen in Figure 5.2 as IP and security concerns are aligned.

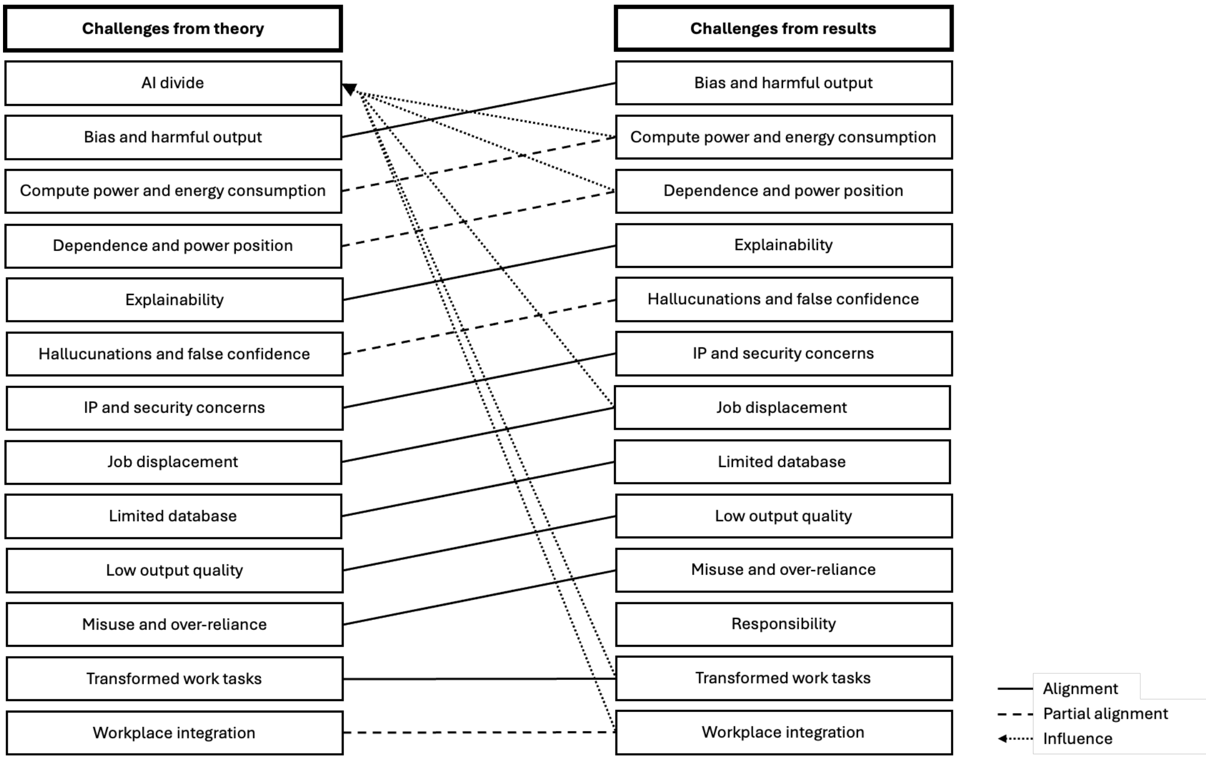

Figure 5.2 Comparison of challenges from theory and interviews

Computing power and energy consumption is one of the challenges that only partially aligns. Literature mainly focuses on the fact that GenAI requires a lot of energy since it is built on computationally intensive algorithms which use a lot of processing power (Stokel-Walker & Van Noorden, 2023; Ooi et al., 2023). This high energy consumption contributes to environmental issues (Stokel-Walker & Van Noorden, 2023; Sætra, 2023; Ozkaya, 2023). The environmental aspect was only mentioned briefly by one employee at Zenseact but a few mentioned that GenAI requires a lot of compute power. The focus during interviews was more on the size of the models and that more compute power and thus larger algorithms are crucial for improving the functionality.

Another challenge which only partially aligned was dependence and power position. Interviews focused more on dependence of tools and the companies providing these tools on both an individual but also partly a company level. This dependence was mentioned as potentially leading to power positions of GenAI companies, contributing to oligopoly. Fui-Hoon Nah et al. (2023) also describe different aspects contributing to a power position and oligopoly but focus more on computational power and policies. Dependence on GenAI tools is not explicitly mentioned in theory but it is described that companies can gain competitive advantage from using GenAI which may further adoption and implementation of tools provided by companies developing GenAI (Fui-Hoon Nah et al., 2023).

Hallucinations and false confidence mostly aligned but interviewees expanded on what Ebert and Louridas (2023) explain about the number of errors increasing with the complexity of the request to GenAI tools and that the errors will be more difficult to find. While the authors partly explain why errors occur and can be difficult to find, interviewees described effects related to this. Hard to find errors can be very time-consuming to correct which counteracts the arguments of GenAI improving quality and efficiency, instead leading to the task being more time-consuming than it would have been without the use of GenAI.

Challenges of integrating GenAI into the workplace are discussed on different levels by literature and employees at Zenseact. Agrawal (2023) describes obstacles at a managerial level, such as limited understanding of integration and the companies' technical abilities. Although Ooi et al. (2023) also mention that peoples' attitudes towards GenAI affects the success of the integration. Interviewees described themselves as positive towards GenAI technology, indicating that this would not be a challenge at Zenseact. However, some employees brought up that they have a hard time keeping up with the technological development and understanding how to use these new tools which might make the integration of GenAI more difficult if not handled in a good way.

Digital divide or AI divide was not explicitly mentioned during interviews, however, some challenges discussed can contribute to increasing such a divide between those who use GenAI tools and those who do not. Employees having limited knowledge, as just mentioned above, could fall even further behind those using GenAI if they do not adopt the technology and keep up, increasing the divide. This could both lead to job displacement for those who do not use GenAI or employees struggling to keep up as work tasks are transformed when GenAI tools are integrated in the workplace. When some employees inside an organization struggle to adapt or refuse to use the technology, these different opinions can lead to internal tensions as described by Ooi et al. (2023). Digital divide or AI divide could also be on a more organizational level where the integration of GenAI could lead to a competitive advantage for adopting companies, creating a divide in accordance with what Carter et al. (2020) describe about early adopters outperforming late adopters. For the companies providing GenAI tools, aspects such as the need for extensive compute power and resources can lead to them getting power positions where they decide the price and thus also accessibility of the tools. Carter et al. (2020) explain that benefits of GenAI not being equally shared will lead to an AI divide, deepening the existing digital divide.

A challenge mentioned by employees at Zenseact was responsibility related to output of generative AI. As Zenseact operates within the industry of autonomous cars there are high safety requirements. If code generated by AI-tools is used in an autonomous car which causes an accident, who is then responsible for this? Is it the provider of the GenAI tool or the car company? This is not mentioned in the general literature about challenges with GenAI which is explained by the specific context which Zenseact operates in.

5.2.3 Facilitating Task Automation

One of the greatest opportunities mentioned by employees at Zenseact was to use generative AI tools to help generate something that could be done manually, but can be automated using generative AI, thereby saving both time and effort. This is in line with one of the major possibilities described in previous research (Bandi et al., 2023; Ooi et al., 2023; Fui-Hoon Nah et al., 2023; Sætra, 2023). This opportunity has been mentioned among all roles with varying tasks and output based on professional role, for example to generate interview guides, which is in line with what Ooi et al. (2023) describe regarding how certain steps in the recruitment process can be automated. Additionally, it is described by Ooi et al. (2023) that the automation of tasks can free up time for employees to perform other more value-creating work, this is supported by empirical findings which indicate that every enhancement in efficiency frees up additional time for performing the most valuable tasks during a workday.

There are some risks associated with letting generative AI do the work for you. During interviews, thoughts arose about how the employees could possibly lose understanding of things if they were to just let GenAI do things for them. If GenAI is regularly used to handle tasks that require critical thinking, problem solving or creativity, a possible outcome may be that employees become less likely to develop or maintain these skills themselves. It could be compared to the phenomenon where people who often use calculators can become worse at mental arithmetic over time as mentioned by Sætra (2023). By handing over detailed and complex tasks to GenAI, the ability to learn and understand the deeper aspects of a subject or process would also potentially deteriorate. This can be particularly relevant in professions where in-depth knowledge is essential for innovation and problem solving, such as software development. Since the models are trained on historical data, there is also a risk that if a large majority of the population would start relying on GenAI, knowledge on a larger scale risks stagnating.

As mentioned in the result, losing understanding can lead to becoming dependent on the tools to perform in your professional role. Another way of becoming dependent on GenAI tools can indirectly come from the organization you are employed by. If a company expects a certain productivity from its employees, only possible by using GenAI to generate things, then the dependence on GenAI tools risks being implicitly forced from the company and individuals might end up in a position of dependence on tool providers both as an individual and as a company. This could then be another factor leading to GenAI companies getting a power position and potentially also oligopoly which was discussed by Fui-Hoon Nah et al. (2023).

Regarding software development specifically, the possibility of generating code using GenAI tools was frequently mentioned by employees at Zenseact, in accordance with Pothukuchi et al. (2023) and Chui et al. (2023). Two aspects can be analyzed related to this possibility. First, there are different opinions among developers about whether the generated code has better or worse quality than the manually produced code. An explanation for this may have to do with the fact that GenAI models are trained on what is out on the internet. This means that it is trained on all quality levels of code—both high and low—and there is a risk that there is more low-quality code in the training data than high-quality which will affect the output of the tools. Those who are above average and senior at writing code will probably think that the code GenAI produces is of low-quality and that their own code is better, while a more novice developer will probably appreciate the code GenAI produces and see it as higher in terms of quality. The perceived quality of the generated code should thus correlate negatively with the level of expertise of the developer. This partly aligns with Dohmke et al. (2023) finding that novice developers benefit more from tools like GitHub Copilot than more experienced developers. However, Ozkaya (2023) and Ebert and Louridas (2023) describe how GenAI is beneficial to identify potential quality problems and provide advanced coding suggestions and alternatives. It can be questioned whether these suggestions are really advanced as Dohmke et al. (2023) and employees at Zenseact argued that experienced developers do not benefit as much from using code generation tools. Employees at Zenseact further expressed concerns that code generated by AI is often not good quality and might contain many subtle bugs, contradicting what Ebert and Louridas (2023) mention about GenAI's ability to identify bugs. Nevertheless, the authors describe how detailed and advanced requests to GenAI have a higher probability of generating output containing hard to find errors.

Secondly, one reason highlighted as to why GenAI might want to be used to produce code is that it could save a lot of time. However, it is worth questioning how much time could actually be saved and for what kind of tasks it could be saved. As mentioned in section 2.3, developers generally spend around five percent of their time editing or writing code (Gillis, 2018). Therefore, the potential for significant time savings by letting GenAI produce code could be overestimated, but potential may instead lie in automating other tasks that developers engage in, such as documenting code, code review, and for testing code.

Moreover, generation of code is presented in the theory mainly from a positive point of view that can help organizations enhance their productivity, however employees mention different possible downsides that could come out of the simplified and faster way of generating code. Firstly, a question is raised by the employees how to handle the large amount of generated code. This code still needs to be validated and tested and there will therefore still be bottlenecks in this process. Because of the automated code generation, developers might instead spend their time on reading and reviewing the code as explained by Bird et al. (2023). Secondly, this changed way of working can lead to decreased job satisfaction since the new work tasks are seen as monotonous and boring and are not at all what they anticipated when pursuing a career as a software developer.

Furthermore, one future opportunity that employees find is the possibility to use GenAI for creating missing data on safety critical aspects for their product, which is something both Bandi et al. (2023) and Ebert and Louridas (2023) claims could be possible. GenAI could then be used to facilitate the handling of both existing data and the newly created data to analyze, clean, and label data from sources such as system logs (Chui et al., 2023).

5.2.4 Augmenting Knowledge Acquisition

Another aspect of seeking assistance from GenAI for content generation involves topics where the initial foundational understanding is lacking. Based on the results, it is described how GenAI tools can be used to, for example, improve the language in text or change the tone of a text. It is also explained by employees at Zenseact how, with the help of GenAI, it is possible to learn new things and deepen your knowledge within a subject. This opportunity is more about improving yourself and your knowledge than using GenAI to automate any task or part of a task.

However, when copying output from GenAI straight off there is a risk of losing understanding of the subject as stated by Sajja et al. (2024). A consequence of this could be that individuals pretend to know something they do not know, potentially leading to them ending up in positions within a company they do not have the right expertise for. This could be problematic due to the risk that people's ability to handle unexpected problems or errors might be decreased if they are used to GenAI always providing a solution. A consequence could be that they find difficulties when faced with situations where GenAI cannot provide answers or where its recommendations are not appropriate.

Furthermore, besides what Fui-Hoon Nah et al. (2023), Sætra (2023), and Ooi et al. (2023) say about misuse connected to education as well as in research and academic publishing, GenAI can be used to cheat in, for example, recruitment processes. For Zenseact, if work candidates use GenAI to fabricate skills or knowledge that they do not actually possess there is a risk that employment is based on false premises, which can negatively impact both the employer and the employee, who may not be able to meet the job requirements. On the other hand, if proficiency with GenAI tools is a crucial part of the job and an expected skill, it is reasonable for the recruitment process to reflect this. Demonstrating competence in using these tools during recruitment can then be viewed as a relevant and necessary skill rather than as cheating. It becomes a matter of transparency and honesty in presenting one's skills and abilities. As Dohmke et al. (2023) explain, the extensive use of GenAI in the development process can democratize the practice, enabling individuals with less technical expertise to contribute. With

GenAI capable of generating outputs such as code or music, the need for lengthy educational programs is decreasing, according to Ooi et al. (2023). Consequently, this shift necessitates adjusting the requirements to reflect developments within the profession.

Similar to what Agrawal (2023) and Fui-Hoon Nah et al. (2023) say, some employees at Zenseact use, or see the opportunity to use GenAI to inspire creative thinking or for brainstorming. However, the view of what happens to people's own creativity is divided among interviewees. Some employees expressed that they feel that they are becoming more creative, while others feel the opposite and believe that it risks inhibiting their creativity. This can depend on various factors such as how the technology is used and what type of creativity is being referred to. This disagreement is also reflected in theoretical discussions, where Agrawal (2023) notes that GenAI is often employed in creative industries, while Fui-Hoon Nah et al. (2023) argue that industries requiring less creative thinking are the ones most likely to be impacted by GenAI. Advantages connected to the subject are that GenAI can help to increase the number of perspectives and ideas that a single individual can come up with thanks to the large amount of data that the models are trained on. This can open new thoughts and inspire creative thinking that might not have occurred otherwise. It could also help in overcoming temporarily creative blocks existing, for example, due to deadlines or external pressures. When learning new things and getting inspiration from GenAI, some employees at Zenseact mentioned that the need for outsourcing certain tasks could decrease since it is possible to do more by themselves, leading to cost savings. However, there are risks with becoming over-reliant on GenAI to produce ideas such as a reduced ability to think creatively as mentioned above. This can lead to a situation where the user begins to feel less capable without the assistance of GenAI, and as Ooi et al. (2023) claim, the effects of over-reliance on generative AI could include decreased morale and well-being, and a loss of purpose at work by letting the tools do your job. Because GenAI generates ideas based on existing data, there is also a risk that the ideas suggested will be predictable or similar. This could lead to a reduction in originality and innovation which could explain why it is questioned whether GenAI can replace human creativity.

Furthermore, GenAI can act as a personal mentor for developers by offering guidance and suggestions for improvement during the work process, thus both reducing the mentoring role of senior developers highlighted by Warcholinski (2023) and bridging knowledge differences between different levels of experience. However, it is important not to overlook group dynamics and completely leave out human mentoring, as this would possibly have a negative effect on group dynamics. By reducing knowledge gaps, GenAI can help less experienced developers perform at a level closer to senior colleagues. This can be valuable to both individual development and company efficiency since more novice programmers can contribute more effectively to projects, which might increase productivity and innovation within the company. However, this can lead to the creation of new inequalities. If access to and knowledge of how to effectively use generative AI is not evenly distributed, it could amplify disparities rather than reduce them in accordance with Carter et al. (2020). A new divide could emerge between the junior developers who use GenAI and those who do not, which could lead to inequalities in skills and opportunities within the profession. Therefore, it becomes central for companies to spread the knowledge that exists within the organization to avoid a new gap emerging.

Ooi et al. (2023) describe performing tasks with GenAI tools can free up time which could be spent collaborating more with colleagues. However, some interviewees mentioned ChatGPT working as a first step in finding the required information or answering questions before going to a colleague. As mentioned in section 4.1.2 this might lead to questions on a more advanced level and making colleague consultations more valuable with a richer discussion because of having more extensive knowledge beforehand. On the other hand, replacing colleague consultation could lead to a decreased collaboration and communication, contradicting what Ooi et al. (2023) mention about an increased collaboration due to the use of GenAI tools. Sajja et al. (2024) also describe how collaboration and group dynamics could be negatively impacted by using GenAI instead of discussing with colleagues and that it is important to only use it as an addition to team interactions to enhance rather than replace human discourse.

5.3 Use and Adoption

RQ2: How can the use of generative AI in software development be evaluated?

This section relates to the use of generative AI and also includes some of the effects mentioned by employees of Zenseact who use it for work, partly treating RQ2. However, as mentioned in section 4.1, not everyone at Zenseact is using GenAI tools and the areas of use also differ between different employees. Therefore, this section will also discuss what might enable or hinder employees from using the technology.

Among those who use GenAI tools, some of the variety in use cases are natural and can be explained by their different work roles within the organization. Employees within HR and communication use it more for generating different texts or to develop internal material for workshops, while software developers have different use cases related to producing code. However, among software developers there is considerable variety in if and how they use GenAI tools. Some explanation for this will be given further down in this section.

Employees who use GenAI tools more regularly expressed that the use results in effects such as saving time and being more efficient or productive. Other effects were facilitating the coding process through getting ideas or help from GenAI, generating material such as text or code, searching information and getting new knowledge, and summarizing existing material. These things were discussed as opportunities but are also effects of using GenAI.

Technology adoption and implementation in a company, as previously described in section 2.7, are affected by the technological, organizational, and environmental context. The following sections will follow this structure when discussing factors possibly explaining the limited use or potentially encouraging more extensive use, in this report referred to as hindering and enabling factors.

5.3.1 Hindering Factors

Within the technological context, features of the technology such as usability, functionality, complexity, and compatibility with current systems will affect adoption (Baker, 2011). As described by Agrawal (2023) the lack of standards on integration with existing systems due to the immaturity of GenAI technology can hinder adoption. Other similar factors mentioned by Venkatesh et al. (2003) within the UTAUT framework are performance expectancy and effort expectancy, which according to Gupta et al. (2024) have a great impact on the intention to use a certain technology. One of the challenges with GenAI mentioned by employees was that the output is not always of the best quality, leading perceived usefulness and performance expectancy to decrease. When the output is completely incorrect or the errors are very subtle and difficult to find, employees expressed that it is faster and easier to just do the task themselves, showcasing a very low performance expectancy and perceived usefulness. The incorrectness of the output damages the employees' trust for GenAI tools, and according to Ooi et al. (2023), trust is a key determinant when talking about new intelligent technologies. During interviews, employees also mentioned limited knowledge about how to use GenAI tools and for what. This might indicate higher effort expectancy and poor ease of use as it requires time and effort to learn how to use the technology in a good way. GenAI is described by Agrawal (2023) as a complex and immature technology, still counting as rather new and undergoing rapid changes and improvements. This could be an explanation of the current high effort expectancy since Venkatesh et al. (2003) state that this factor is more prominent in the beginning of adopting a new technology and its importance will decrease over time when people learn and become more experienced. Also related to the technology's functionality is the fact that it is better suited for certain tasks and that the training data determines what problems it can solve. For example, it cannot answer company-specific questions since it is not trained on internal data or enough specialized data in accordance with what Stokel-Walker and Van Noorden (2023) describe. Another factor negatively impacting adoption mentioned by Abed (2020) are security concerns or other perceived risks with using the technology. In this case study, many employees expressed a fear of leaking confidential information as an explanation of their limited use or more careful use. Because of the unwillingness to share internal information, external models cannot be specifically trained on solving company-specific issues.

Within the organizational context, culture and commitment from management are mentioned as factors affecting adoption of new technology (Baker, 2011). A possibly hindering factor mentioned during interviews was that some employees perceive the directives about GenAI use from the management and legal teams as unclear. Some interviewees described the subject of GenAI usage as kind of taboo and that it is not very openly discussed among peers. Venkatesh et al. (2003) describe social influence as being important in the early phase of adoption but since the topic of GenAI is not very openly discussed, social influence might be limited although Gupta et al. (2024) explain that people could start using GenAI because their friends already use and promote them. The directives about what is allowed or not seem to be perceived differently by different people and create an uncertainty which limits the use by some people. Guidelines for GenAI use were provided by Zenseact some time ago but since the technology develops at a very fast pace, employees either seem to have forgotten about them or are unsure whether they are up to date and how to interpret them in the light of the technological advancements. The guidelines themselves directly hinder the use of GenAI in certain ways or for certain tasks but also indirectly due to the uncertainty among employees on what is allowed, also leading to limited use.

The environmental context is affected by regulatory requirements, among other things (Baker, 2011). Intellectual property matters were raised as a large concern among employees, especially those working directly with these questions and developers working close to the end-product. Since the technology is relatively new, regulations have not yet been fully developed around the use of GenAI and are lagging the fast-paced technological development. Since Zenseact operates within the industry for autonomous driving, their use case is very complex and requires high safety-standards which further hinders the implementation of GenAI. Matters like what IP the models have trained on and who legally owns the output of the tools are aspects mentioned in relation to who is responsible for potential accidents caused by code generated by AI. According to Gupta et al. (2024) the environmental context could also involve ethical considerations where some examples mentioned during interviews and in literature are bias, harmful output, and fake news.

5.3.2 Enabling Factors

Within the technological context, despite the complexity and sometimes poor functionality of GenAI tools some employees still perceive the technology as useful and have high performance expectancy. The positive effects of using generative AI mentioned previously can lead to continued use. While some interviewees said they do not use any tools regularly for work, everyone had tried some kind of GenAI tool, even if it was outside work and only for fun. This indicates an ease of use and tools such as ChatGPT are accessible for everyone with internet access, enabling the use of such tools. During interviews, employees who used GenAI tools regularly had found good areas of use and perceived the tools as useful for facilitating learning or getting texts explained in other ways that are easier to understand, among others.

While the culture can be seen as a hindering factor within the organizational context, it can partly also be an enabling factor. During interviews, the company was described being in the forefront of tech and very open-minded to new technology by several employees. The company has, as previously mentioned, several internal initiatives revolving around GenAI, showing commitment and engagement from the leadership team and a will to adopt new technology which according to Baker (2011) and Abed (2020) are important for organizational adoption. As Zenseact is exploring and trying out new technology they contradict what Agrawal (2023) writes about firms being more cautious and having a "wait and see" approach towards GenAI due to its newness before investing. The "tech-culture" is further strengthened by the knowledgeable employees where the majority work with software development and many also have expertise within AI, machine learning, and deep learning which could be an explanation for their more aggressive approach to exploring new technologies which also contradict Agrawal (2023), who claims that technology proficiency do not affect adoption of GenAI. Venkatesh et al. (2003) explain that experience of a certain technology or similar technologies lead to a lower effort expectancy to learn something new. Having employees with a software engineering background and knowledge within tech should therefore enable the adoption of generative AI. A suggestion for more widespread adoption mentioned by one employee could be if the company offered licenses to different GenAI tools such as GitHub Copilot. This would then indicate that employees are allowed to use these tools and might even encourage them to use them, contributing to a more open discussion among employees about use of the tools. Additionally, identifying specific use cases and domains where GenAI can be used while still motivating employees to remain creative in their own problem-solving as described by Sajja et al. (2024) could also enable adoption. Another enabling factor mentioned would be to educate more about the use of GenAI tools. A suitable format for this could be what Zenseact calls "Knowledge boosts" where they hold lectures about various topics. Since the perceived usefulness varies between employees and some did not seem to have found good use cases, this could be evened out and the employees who are successfully using GenAI already could share their knowledge with colleagues. Providing licenses and educating employees more about GenAI could be seen as facilitating conditions, explained by Venkatesh et al. (2003) as influencing usage. When more and more employees start adopting GenAI, social influence can be enhanced when others hear about their colleagues using it and feel like they are missing out if they do not use it too.

Within the environmental context, generative AI has developed rapidly and received increased attention with the launch of ChatGPT (Ebert & Louridas, 2023). It was discussed during interviews that management might feel like they need to keep up with others and not fall behind in the "technology race". Agrawal (2023) also explains this uncertain environment with intense competition as pushing firms to adopt GenAI to hopefully maintain their competitive advantage and stand out from rivals. The author also states that the complex technology makes it challenging to manage regulations which benefit highly knowledgeable companies. Zenseact could be seen as a knowledgeable company since they have a lot of expertise within AI technology and thus benefit from the environmental context.

5.4 Evaluation and Measurement

RQ2: How can the use of generative AI in software development be evaluated?

Since there is no widespread use of GenAI tools across the whole company or any common specified use cases as previously mentioned, performing a large-scale measurement of effects of using GenAI would be challenging. Furthermore, measuring productivity alone is already a difficult task as mentioned by Forsgren et al. (2021). Despite these obstacles, it will still be discussed what to consider when measuring or evaluating the use of GenAI and partly what effects could be evaluated within the software development process and how to move forward from the current position.

Regardless of the opportunity identified, employees mentioned saving time and being more productive or efficient as the effects they strive to achieve. Employees also questioned how much time they save and how much more productive or efficient they become, which makes these factors relevant to evaluate. Other aspects potentially interesting to evaluate according to software developers were related to the coding process such as the number of bugs introduced by these tools.

A first step on the way to measuring productivity in relation to GenAI use would be for the company to define developer productivity. According to Noda (2024), this is an important step to choose suitable metrics related to the definition since there is no "one size fits all" metrics to measure this. The SPACE framework has been described as an approach to understand and measure developer productivity, containing five different dimensions (Circei, 2022). This framework is used within the context of software development but could also be used to target aspects of GenAI as software developers integrate such tools more into their work. When choosing metrics within the SPACE framework, besides having the definition in mind, it is important to include both workflow- and perception-metrics (Noda, 2024). Since workflow metrics can often be gathered automatically with different programming software, it could be beneficial to start with developing and using self-perception metrics as proposed by Noda (2024).

For example, one effect from using GenAI mentioned by some during interviews was increased job satisfaction due to an easier and more fun way of searching information. The dimension of satisfaction and well-being within the SPACE framework are often measured with satisfaction surveys (Forsgren et al., 2021; Circei, 2022). Other dimensions could then also be included in this survey to measure perception and opinions about GenAI tools to measure different aspects. However, because of the limited use of GenAI as previously mentioned, the survey would need to be adapted to suit both users and non-users to be relevant.

For this kind of survey to be useful, it would have to be sent out multiple times during a longer period to evaluate how adoption and attitudes change. This could also be applicable later when measuring other things more directly related to developer productivity. One single measurement will not say much unless you have a benchmark to compare with or several points of measurement gathered over time. To engage employees further and motivate their participation, answers could be shared with everyone. However, when presenting these kinds of results, it is important to do it anonymously and at an aggregated level (Microsoft Research, 2021).

By letting both the empirical results and the theory guide, the survey discussed above can be developed. The survey is intended to capture employees' experiences based on their interactions with GenAI tools. The design of the survey has been directly inspired by insights that emerged during the interviews with employees at Zenseact, which has been critical in focusing on the most crucial aspects of GenAI's impact on their work. The survey contains several questions that are directly linked to DevEx and SPACE. For example, it is mentioned by the employees at Zenseact that GenAI can be used to better understand code. The question "When using GenAI tools I feel like I can get help with understanding code more easily" thus reflects the employees' experience that GenAI tools can make their code understanding more efficient. Furthermore, this is directly related to cognitive load within DevEx, where simplified cognitive processes are a critical aspect for improving developers' capacity to deliver value efficiently (Noda et al., 2023).

Furthermore, the employees describe that GenAI has an impact on their creativity and that they can use the tools to ask questions and brainstorm ideas. This results in the following questions within the satisfaction and well-being dimension from SPACE "When using GenAI tools I feel more creative", "When using GenAI tools I feel I can ask questions that I otherwise would not have asked", and cognitive load within DevEx "I believe GenAI can provide me with insights and perspectives I may not have thought of myself". Additional questions that arose based on insights from the interviews can be seen in Figure 5.3 and Figure 5.4.

More examples from interviews which could be interesting to build further on are knowledge about GenAI tools, perceived effects of using them, reasons for limited use, and many more. Answers to such questions could give management indication on how to further work with integrating GenAI at the workplace and how to support developers in the software development process.

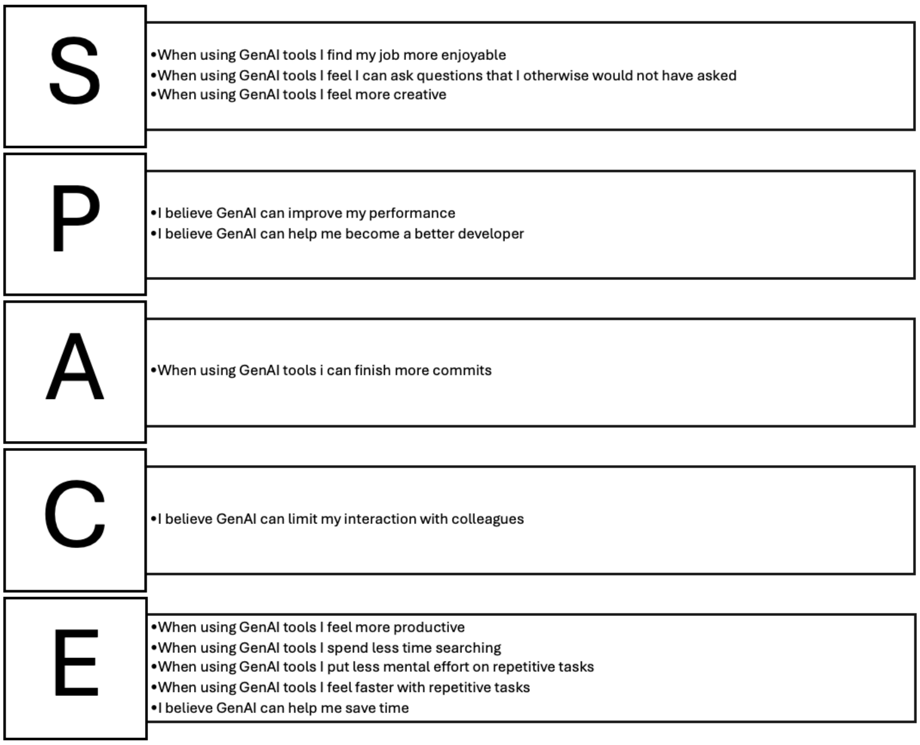

Figure 5.3 Examples of survey questions within SPACE

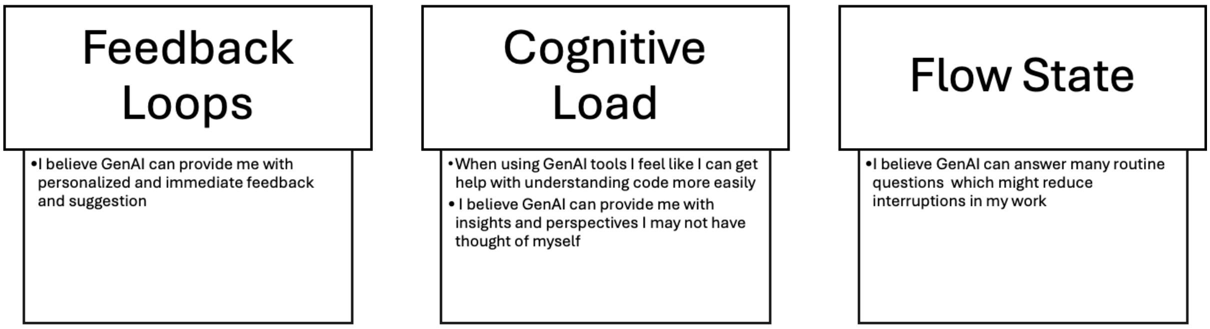

Figure 5.4 Examples of survey questions within DevEx

6 Proposing a Tentative Framework

Based on the discussion in section 5.3 regarding the use and adoption of GenAI in software development, it is evident that evaluating the integration and impact of GenAI tools is essential for understanding their benefits and challenges. This chapter proposes a tentative framework for evaluating the use of GenAI within a software company, building on the discussion from section 5.4, with a focus on developing and deploying a survey.

The proposed framework involves a structured survey designed to capture a comprehensive understanding of how GenAI tools are being used in practice by software developers, what benefits they offer, and what obstacles might prevent their wider adoption. A proposal for such a survey can be seen in Appendix B. The focus of the survey is to evaluate the use of GenAI within software development. From literature it was evident that two main dimensions impacted the use of GenAI tools, namely attitudes and facilitating conditions, which also aligns with what was said during interviews. The actual use can then result in different outcomes which are also explored in the survey. A framework inspired by UTAUT, guiding the development of the survey based on the mentioned dimensions, is presented in Figure 6.1.

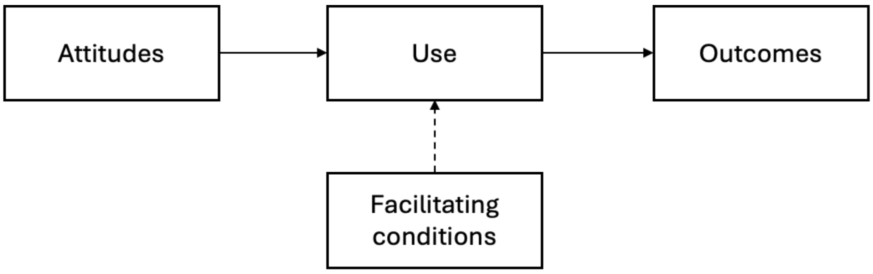

Figure 6.1 Framework for constructing survey inspired by UTAUT

When developing the survey questions, empirical findings, DevEx and SPACE were also taken into consideration. Questions were created within the different dimensions from Figure 6.1 as can be seen in Table 6.1. As proposed by Ooi et al. (2023), a question was added on trust to better fit with the more advanced GenAI technology which could affect dimensions of UTAUT.

Table 6.1 Overview of survey following the framework inspired by UTAUT

Dimensions	Question description	Question number
Attitudes	General attitude about GenAI	Q1
	Reasons for use	Q6 [Frequent users]
	Reasons for limited use	Q8 [Infrequent users]
	Other thoughts	Q10
Facilitating conditions	Company resources, guidelines, and culture	Q2

Use	Frequency of use	Q3
	Tools used	Q4 [Frequent users]
	Tasks GenAI is used for	Q5 [Frequent users]
	Factors to increase use	Q9 [Infrequent users]
Outcomes	Time savings	Q7 [Frequent users]
	Statements based on DevEx and SPACE	Q6 [Frequent users]

To be able to analyze results on a team level or for an entire department, an opening question of demographics must be added. However, the question needs to be formulated in such a way that anonymity can still be preserved. The first three questions of the survey are asked to all software developers and thereafter depending on how often they use GenAI, they are divided into frequent and infrequent users. Frequent users are defined as those who use GenAI tools at least once a week and infrequent users are those who use it more seldom than that. These are then following the different branchings as shown in Figure 6.2.

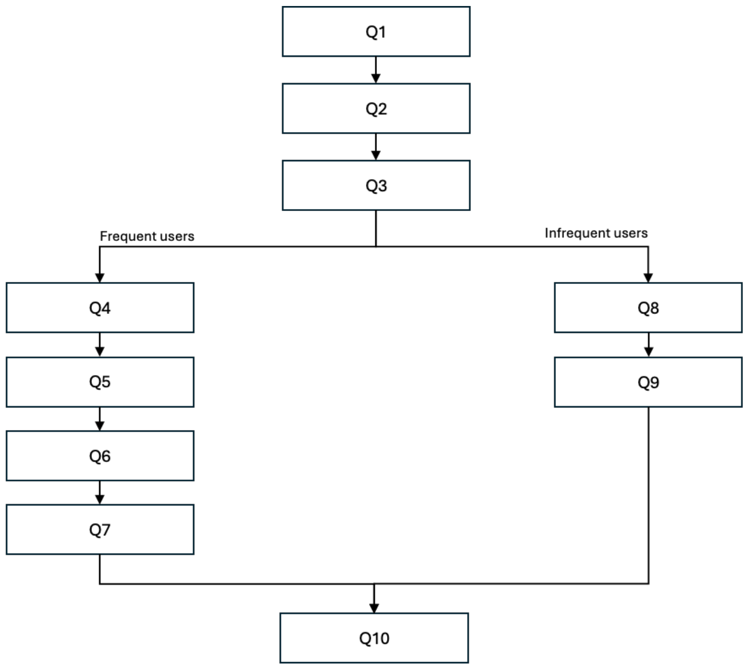

Figure 6.2 Survey flow chart

For attitudes, statements in the example survey are given about capabilities of GenAI indicating general attitudes and then questions explaining reasons for use or limited use are asked to frequent and infrequent users respectively. Within facilitating conditions, factors such as the company's resources, guidelines, and culture are asked about to all developers to explore how they perceive support from their organization. For the use dimension it is first asked about how frequently software developers use GenAI tools. For the frequent users, questions are later asked about what tools they use, for what tasks they use them for, and how often they use tools for a specific task. Conversely, the infrequent users are asked about what could make them use

GenAI tools more in their work. Furthermore, questions are asked to the frequent users about different outcomes they could experience from using GenAI.

As previously mentioned, it is important to send out the survey regularly over time to see trends and follow the adoption rate. It is suggested to send it out quarterly or follow the company cycle of other similar surveys to collect data on the topic which can then be analyzed. For the survey to be sent out correctly and in time, ownership should be appointed to a specific person, for example to someone responsible for generative AI projects within the company. Results from the survey can then be analyzed and one of the most important metrics to track would be the percentage of frequent users and infrequent users.

To enable quantitative analysis, Likert scales have been used for several questions. In the survey results it is also possible to see what tools are the most used and for what tasks, indicating to the company what licenses could be valuable to provide or which currently possessed licenses are not being utilized. From answers given by infrequent users the company can identify what hinders adoption and what they could do to increase adoption of GenAI tools. In the future, as the use of GenAI becomes more widespread and organized, the outcomes from Q6 most strongly agreed with by developers could serve as guidelines for developing metrics that are interesting to measure in relation to GenAI.

7 Conclusion

With the increasing attention and interest of generative AI technology and various GenAI tools, this thesis aimed to explore the use of generative AI within a software company in two ways. The first question revolves around understanding employees' perceptions of the use of generative AI related to their work within the company, by exploring opportunities and challenges it presents across various departments. The second question focuses on how to evaluate the use of generative AI in software development. This study has a qualitative approach, focusing on perceptions of employees at a software development company and what opportunities and challenges they identified. Additionally, the interviewees' perceptions and answers during interviews in combination with literature could give some guidelines on how the use of GenAI within software development can be evaluated. Through interviews, thematic analysis of interview data, and looking at relevant literature, the following research questions have been answered:

> *RQ1: How do software industry employees perceive the use of generative AI in relation to their work?*
>
> > *RQ1.1: What opportunities do software industry employees identify with using generative AI?*
> >
> > *RQ1.2: What challenges do software industry employees identify with using generative AI?*
>
> *RQ2: How can the use of generative AI in software development be evaluated?*

Zenseact's employee responses show a clear tendency to embrace GenAI, but with varying degrees of enthusiasm and skepticism that reflect both personal and professional perspectives. It is apparent that skeptical attitudes are often found among those with deeper technical insight, such as developers and deep learning engineers, who are aware of the technology's limitations and uncertainties.

Further complexity is added by concerns about security and privacy issues, particularly relevant in an industry where protection of intellectual property rights and security-critical aspects are of utmost importance. This caution is reasonable and expected but may also stand in the way of faster and wider technology adoption. Despite these challenges, employees have an overall positive view towards the technology and its potential to transform workflows and efficiency.

Employees identify several significant opportunities with the use of generative AI, notably in enhancing operational efficiency and expanding knowledge and creativity. The primary opportunity lies in automating repetitive, autonomous, and redundant tasks, such as taking notes, transcribing, sorting emails, and generating various types of content like job ads, emails, and presentation materials. This automation not only saves time but also increases productivity, making work more enjoyable and efficient. Furthermore, GenAI tools like ChatGPT are seen as valuable in facilitating information search and learning processes, offering a more efficient and user-friendly alternative to traditional search engines. These technologies are said to accelerate the learning of new knowledge, simplify difficult explanations, and boost creativity through brainstorming and idea generation.

The challenges associated with using GenAI identified by employees at Zenseact include technological limitations, ethical concerns, changes in the job market, and reskilling. A primary concern is the reliability and accuracy of outputs from generative AI tools, especially in code generation, where subtle errors can lead to significant issues. This is related to the "black box" nature of GenAI logic, which makes it impossible to understand how the output is generated and creates trust issues. Furthermore, there are legal and security challenges related to intellectual property rights and data security, with uncertainties about who owns the output of GenAI tools and the implications of their widespread use in sensitive applications, such as autonomous driving software. Employees are also cautious when it comes to becoming overly dependent on tools, potentially leading to a decrease in human creativity and problem-solving skills.

The use of GenAI and its effects in software development should be measured both quantitatively and qualitatively. To measure qualitatively, a survey created based on the proposed tentative framework from chapter 6 would be a good format as it is possible to capture many different soft values such as satisfaction levels, usability concerns, and the perceived value of GenAI tools. To measure quantitatively, it is necessary to define what to use GenAI tools for and how, and what to be measured. During interviews it was described that quantitative metrics of interest were time savings, productivity improvements, and quality of software outputs. Evaluation must be done more than once to be able to see if the effects are attributed to GenAI. To do this, simultaneous measurements of how the use of GenAI develops are needed to see if it is actually used and can be assumed to drive various effects and reveal practical value and limitations of these tools. To measure qualitatively, there is no need to define concrete use cases as is needed to measure quantitatively, focus instead lies on employees' perceptions of their use and its effects.

7.1 Implications for Practitioners

Software organizations should on a managerial level strategically assess the potential utility of GenAI within their operations. It is crucial for companies to evaluate whether GenAI is the optimal solution for achieving their particular objectives, ensuring that the technology aligns with their broader strategic goals. Before implementing the technology more widely, it is important to carefully explore specific areas where GenAI can add value. As for software development, when identifying application areas for GenAI, managers and other leading roles should think beyond the pure coding aspect of software development work as writing code represents only about five percent of the work time spent by developers. Other tasks, besides coding, which could benefit more from using GenAI tools are code comprehension and making the codebase more streamlined and homogenous. For Zenseact, it is suggested to take steps to clearly define areas of use where GenAI can have a beneficial effect.

This study sheds light on the existing skepticism among employees regarding the effective use of GenAI, underscoring the critical role of involving employees in developing policies and strategies for the adoption and utilization of the technology. While it can be beneficial for managers in a software organization to centrally outline main application areas for GenAI to inspire and provide a basic frame, it is equally important to empower individual teams or departments to identify and develop specific use cases that best suit their local needs. This approach ensures that GenAI solutions are directly relevant to challenges faced by different parts of the organization. However, the same challenges can be faced by more than one part of the organization, making it important to seize local engagements and bring them up to a central level. To do this, organizations could provide platforms for employees to share their GenAI success stories. This could be through internal forums or other accessible formats that encourage communication across the company. Sharing these experiences not only fosters a culture of learning and collaboration but also allows successful applications of GenAI to be scaled and replicated throughout the organization, maximizing the technology's benefits across all departments. For Zenseact, this could mean creating a communication channel dedicated to sharing the use cases of GenAI.

Regularly used surveys based on the proposed tentative framework could also be conducted to gather feedback on how software developers are using GenAI and their perceptions of its utility. These surveys will assist managers in keeping up with the rapidly changing attitudes and uses of GenAI, attributed to the fast development of GenAI technology, allowing them to adjust their strategies in response to new employee insights. Additionally, identifying and addressing potential barriers to effective GenAI use is essential, as these could hinder technology adoption and integration. For example, the report shows that limited understanding of the technology and its areas of use contributes to limited use. To address this, companies may want to consider introducing regular knowledge sharing sessions, where experiences and knowledge of GenAI are shared between more experienced users and those who are less experienced. This could aid the AI divide mentioned in literature, which is at risk of increasing if some employees fall behind and do not manage to keep up with technological development. Zenseact has a format they call 'knowledge boosts' which could be a suitable format for such knowledge-sharing. By incorporating insights from employees who are already using GenAI, organizations can foster an environment where less experienced individuals feel supported, and a culture of knowledge sharing is promoted.

To effectively measure the impact of GenAI on developer productivity, it is beneficial for software organizations to establish a collectively agreed definition of what constitutes developer productivity. This clarity will enable consistent tracking and evaluation of GenAI's effectiveness over time. By understanding which productivity aspects are most enhanced by GenAI, companies can make more informed decisions regarding technology deployment.

7.2 Implications for Researchers

Existing research predominantly highlights the potential of GenAI to optimize and revolutionize various processes across industries. However, it often falls short in detailing practical implementation strategies, overlooking the foundational step of actual usage. This oversight leads to a gap between the theoretical discussions of GenAI's capabilities and the practical realities and uncertainties addressed in Chapter 5, where employees express mixed perceptions and experiences of GenAI use. This study reveals that despite the ease of use and availability of GenAI tools, the usage remains limited. Without the practical application of GenAI, the theorized benefits remain unattainable. Consequently, there is a need for research to bridge the gap between theoretical potential and operational execution. Our study contributes by providing insights into real-world applications and immediate operational benefits of GenAI, as well as highlighting the limited use of GenAI tools where the main explanations are low functionality and concerns linked to IP.

A major problem highlighted in our study is the limited training data, which prevents employees from using existing tools for their domain-specific tasks. This is especially true for technical topics as Stokel-Walker and Van Noorden (2023) mention that some GenAI tools are not trained on enough specialized data. Generally, previous research discusses the importance of accurate and extensive amounts of training data for GenAI to generate meaningful results. However, it often overlooks the implication with the limited data available for specific domains. This becomes evident in this case study, where the area includes specific and complex technical topics. The problem could potentially be solved by adapting and fine-tuning existing GenAI models with domain-specific data to improve performance in that specific area. However, the concerns described by employees linked to security-related issues and leaking IP remain. This is also described by Fui-Hoon Nah et al. (2023) which highlights the importance of companies not spreading their proprietary information to existing models.

Additionally, discussions on the prospects of GenAI, as presented by Bandi et al. (2023) and Ooi et al. (2023), emphasize rapid technological advancement and expanding opportunities. Contrastingly, our findings reveal a more cautious perspective among software development professionals, who view the trajectory of GenAI with uncertainty. The development of GenAI is perceived as unsure with significant ambiguity regarding the pace of advancement and the timeline for achieving meaningful improvements. This skepticism may be partially attributed to the advanced technical knowledge of software engineers, who possess deep knowledge in AI and machine learning, thereby decreasing their expectations with a realistic assessment of technological maturity.

The strategic benefits of automating specific processes using GenAI are well-documented. However, it is crucial to recognize that operational impacts and direct effects on employees need greater emphasis to understand the full scope of GenAI's influence on workforce dynamics, as highlighted in our empirical findings. This research indicates that while some tasks are rendered more enjoyable and employees report increased job satisfaction, there are instances where the opposite is true. This highlights the necessity of considering both the positive and negative implications of GenAI on workforce dynamics.

In the context of software development, discussions focus on automating aspects of the development process, such as code generation. This automation is generally considered beneficial without acknowledging potential negative effects. This study challenges this one-sided view by illustrating employees' concerns about managing the possibly large volume of code generated by GenAI. Moreover, although research advocates that software development will be revolutionized through GenAI-enabled code generation, other studies indicate that coding only constitutes five percent of a developer's time (Minelli et al., 2015). Empirical findings suggest that greater emphasis should be placed on how GenAI can streamline or facilitate code comprehension—a process that occupies up to 70 percent of a developer's time (Minelli et al., 2015). This underscores the importance of broadening the scope of research to encompass diverse aspects of software development that GenAI can potentially transform.

7.3 Future Research

Perceptions of GenAI seem to vary across different roles within the case organization. This phenomenon aligns with the Unified Theory of Acceptance and Use of Technology (UTAUT) framework, which accounts for individual factors influencing a person's intent to use technology (Venkatesh et al., 2003). However, whether this framework accurately applies to GenAI remains an open question. Notably, GenAI tools such as ChatGPT are characterized by its user-friendliness and require minimal technical expertise beyond basic computer skills. Empirical findings suggest that prior technical knowledge may inhibit rather than facilitate the intention to use GenAI. This unexpected finding calls for a reassessment of existing models like UTAUT in the context of intelligent technologies and indicates a need to explore additional dimensions and factors, such as trust, that could impact the adoption of such technologies.

Based on findings from this study, future research could benefit from further broadening the knowledge base on how GenAI can be applied within software engineering. When integration and adoption of these new tools are increased, it will be possible to perform more precise studies and measurements about the use and possibly also effects of using GenAI, both quantitatively and qualitatively. Although this study presents some guidelines and examples on how to measure qualitatively, this could be expanded on in future research. Also, due to the characteristics and context of the case company for this study, future research could include more software companies to improve generalizability to the whole software development industry.

Within the software development life cycle, the focus has been on the development and the testing phase, where previous research has specifically advocated that software development will be revolutionized through GenAI-enabled code generation. However, as mentioned before, time savings from code generation might be overestimated where other studies have indicated that coding only constitutes five percent of a developer's time (Minelli et al., 2015). Researching how GenAI tools can impact other phases or tasks within the SDLC in the future might reveal bigger impacts. This might include examining the effectiveness of GenAI in reducing the cognitive load of developers during the comprehension process, enhancing their ability to understand complex systems more efficiently. Empirical findings suggest that greater emphasis should be placed on how GenAI can streamline or facilitate code comprehension—a process that occupies up to 70 percent of a developer's time (Minelli et al., 2015). In summary, this underscores the importance of broadening the scope of research to encompass diverse aspects of software development that GenAI can potentially transform.

Another interesting area for future research is how the role and tasks of a software engineer will change when GenAI tools are integrated into their work. Several interviewees discussed how using code generation tools could become the new normal in the future and that it is important to keep up and learn how to work with them. Therefore, it can be valuable to conduct studies on how GenAI could be used most effectively and how developers could best integrate them.

As previously mentioned, within software development, discussions focus on automating aspects of the development process, such as code generation. This automation is generally considered beneficial and potential negative effects are not acknowledged in literature. Our study challenges this view by illustrating employees' concerns about managing the possibly large volume of code generated by GenAI. Thus, a future research area could examine the implications of a potentially increased volume of code for a software company, as well as the preparatory measures required for this transformation.

Future research could explore how GenAI can be more effectively integrated into workplace strategies to not only improve efficiency but also enhance job satisfaction and overall workplace dynamics. While theoretical perspectives often emphasize the productivity benefits of GenAI, such as code generation, practical insights from employees highlight potential downsides. The transformation of tasks through GenAI might lead to decreased job satisfaction, as the work becomes monotonous and diverges from the creative and engaging aspects that initially attracted individuals to software development. However, automating repetitive tasks and making them more enjoyable by using GenAI is described as increasing job satisfaction. By addressing these gaps, research could provide insights into how GenAI can be strategically implemented to improve both efficiency and job satisfaction.

Lastly, a future research area emerging at the intersection of technology and law concerns the allocation of responsibility for outcomes linked to materials generated by GenAI. Specifically, when AI-generated code is used in complex systems such as self-driving cars, questions arise about accountability in the event of accidents or failures. For example, if a self-driving car containing AI-generated code crashes, it becomes crucial to determine who is responsible. Is it the code creators, the car company, or perhaps the AI system itself? Research should thus explore how traditional legal frameworks can be adapted to handle these new technological realities, which includes defining liability and safety requirements in relation to AI's unique capabilities and limitations.

List of References

Abed, S. S. (2020). Social commerce adoption using TOE framework: An empirical investigation of Saudi Arabian SMEs. *International Journal of Information Management,* 53, Article 102118. https://doi.org/10.1016/j.ijinfomgt.2020.102118

Agrawal, K.P. (2023). Towards Adoption of Generative AI in Organizational Settings. *Journal of Computer Information Systems*, 1–16. https://doi.org/10.1080/08874417.2023.2240744

Astreca (2023). *Understanding the Differences: Generative AI vs. Predictive AI.* Retrieved February 8, 2024, from https://astreca.com/blog/generative-ai-vs-predictive-ai/

Baker, J. (2011). The Technology–Organization–Environment Framework. In Y. Dwivedi, M. Wade & S. Schneberger (eds), *Information Systems Theory* (231-245). Integrated Series in Information Systems, 28. Springer. https://doi.org/10.1007/978-1-4419-6108-2_12

Bandi, A., Adapa, P.V.S.R., & Kuchi, Y.E.V.P.K. (2023). The Power of Generative AI: A Review of Requirements, Models, Input–Output Formats, Evaluation Metrics, and Challenges. *Future Internet, 15*(8), 260. https://doi.org/10.3390/fi15080260

Bell, E., Bryman, A., & Harley, B. (2022). *Business Research Methods* (6th ed.). Oxford University Press.

Bergdahl, J., Latikka, R., Celuch, M., Savolainen, I., Soares Mantere, E., Savela, N., & Oksanen, A. (2023). Self-determination and attitudes toward artificial intelligence: Cross-national and longitudinal perspectives. Telematics and Informatics, 82, Article 102013. https://doi.org/10.1016/j.tele.2023.102013

Bird, C., Ford, D., Zimmermann, T., Forsgren, N., Kalliamvakou, E., Lowdermilk, T., & Gazit, I. (2023). Taking Flight with Copilot: Early insights and opportunities of AI-powered pairprogramming tools. *Queue, 20*(6), 35-57. https://doi.org/10.1145/3582083

Braun, V. & Clarke, V. (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology, 3*(2), 77-101. DOI: 10.1191/1478088706qp063oa

Brodsky, J.S. (2017). Autonomous Vehicle Regulation: How an Uncertain Legal Landscape May Hit the Brakes On Self-Driving Cars. *Berkeley Technology Law Journal, 31*(2), 851-878. https://doi.org/10.15779/Z38JC5S

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018). *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey Global Institute. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy#/

Burkholder, G.J., Cox, K.A., Crawford, L.M., & Hitchcock, J.H. (2019). *Research Design and Methods - An Applied Guide for the Scholar-Practitioner*. SAGE Publications.

Cain, D. (2023). *Unraveling Artificial Intelligence: Implementations, Advantages, and Limitations.* LinkedIn. Retrieved May 24, 2024, from https://www.linkedin.com/pulse/unraveling-layers-artificial-intelligence-advantages-limitations/

Carter, L., Liu, D., & Cantrell, C. (2020). Exploring the Intersection of the Digital Divide and Artificial Intelligence: A Hermeneutic Literature Review. *AIS Transactions on Human-Computer Interaction*, *12*(4), 253-275. https://doi.org/10.17705/1thci.00138

Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A., Yee, L., & Zemmel, R. (2023). *The Economic Potential of Generative AI: The Next Productivity Frontier.* McKinsey & Company. http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/14313/1/The-economic-potential-of-generative-ai-the-next-productivity-frontier.pdf

Circei, A. (2022, December 16). The SPACE Framework For Software Developer Productivity. *Forbes. https://www.forbes.com/sites/forbestechcouncil/2022/12/16/the-space-framework-for-software-developer-productivity/?sh=3be4eb7c29d5*

Cision News (2022). *Volvo Cars takes full ownership of Zenseact.* Retrieved May 24, 2024, from https://news.cision.com/volvo-car-ab--publ-/r/volvo-cars-takes-full-ownership-of-zenseact,c3691402

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*(3), 319–340. https://doi.org/10.2307/249008

Davis, G. (2023, June 8). Developer experience: What is it and why should you care?. *GitHub.* https://github.blog/2023-06-08-developer-experience-what-is-it-and-why-should-you-care/

Dohmke, T., Iansiti, M., & Richards, G. (2023). Sea Change in Software Development: Economic and Productivity Analysis of the AI-Powered Developer Lifecycle. *Cornell University Arxiv.* https://doi.org/10.48550/arXiv.2306.15033

Dora, S. K., & Dubey, P. (2013). Software Development Life Cycle (SDLC) Analytical Comparison and Survey on Traditional and Agile Methodology. *Abhinav: National Monthly Refereed Journal of Research in Science and Technology, 2*(8), 22-30. https://www.researchgate.net/publication/319716548_SOFTWARE_DEVELOPMENT_LIFE_CYCLE_SDLC_ANALYTICAL_COMPARISON_AND_SURVEY_ON_TRADITIONAL_AND_AGILE_METHODOLOGY

Ebert, C., & Louridas, P. (2023). Generative AI for Software Practitioners. *IEEE Software* *40*(4), 30-38. https://doi.org/10.1109/MS.2023.3265877

Forsgren, N., Storey, M.-A., Maddila, C., Zimmermann, T,. Houck, B., & Butler, J. (2021). The SPACE of Developer Productivity: There's more to it than you think. *Queue*, *19*(1), 22-28. https://doi.org/10.1145/3454122.3454124

Fui-Hoon Nah F., Zheng R., Cai J., Siau K. & Chen L. (2023). Generative AI and ChatGPT:

Applications, challenges, and AI-human collaboration. *Journal of Information Technology Case and Application Research*, *25*(3), 277-304. https://doi.org/10.1080/15228053.2023.2233814

Gnanasambandam, C., Harrysson, M., Hussin, A., Keovichit, J. & Srivastava, S. (2023). *Yes, You Can Measure Software Developer Productivity.* McKinsey & Company. https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/yes-you-can-measure-software-developer-productivity

Gillis, A.S. (2018). *Integrated Development Environment (IDE).* TechTarget. Retrieved April 26 2024 from https://www.techtarget.com/searchsoftwarequality/definition/integrated-development-environment

Girasa, R. (2020). *Artificial Intelligence as a Disruptive Technology.* Palgrave Macmillan Cham. https://doi.org/10.1007/978-3-030-35975-1

Gupta, R., Nair, K., Mishra, M., Ibrahim, B., & Bhardwaj, S. (2024). Adoption and Impacts of Generative Artificial Intelligence: Theoretical Underpinnings and Research Agenda. *International Journal of Information Management Data Insights, 4*(1), Article 100232. https://doi.org/10.1016/j.jjimei.2024.100232

IBM Data and AI Team. (2023). *Understanding the different types of artificial intelligence*. IMB. https://www.ibm.com/think/topics/artificial-intelligence-types

Kvale, S. (2007). *Doing interviews.* Sage Publications. https://doi.org/10.4135/9781849208963

Lawton, G. (2023). *Generative Models: VAEs, GANs, Diffusion, Transformers, NeRFs*. TechTarget. Retrieved April 26 2024 from https://www.techtarget.com/searchenterpriseai/tip/Generative-models-VAEs-GANs-diffusion-transformers-NeRFs

Lawton, G. (2024). *What is Generative AI? Everything You Need to Know.* TechTarget. Retrieved April 26 2024 from https://www.techtarget.com/searchenterpriseai/definition/generative-AI

Lekh, R., & Pooja. (2023). Exhaustive study of SDLC phases and their best praxctices to create CDP model for process improvement [Conference Paper]. *2015 International Conference on Advances in Computer Engineering and Applications*, pp. 997-1003. doi: 10.1109/ICACEA.2015.7164852.

Maguire, M., & Delahunt, B. (2017). Doing a Thematic Analysis: A Practical, Step-by-Step Guide for Learning and Teaching Scholars. *All Ireland Journal of Teaching and Learning in Higher Education, 8*(3), 3351-33514. http://ojs.aishe.org/index.php/aishe-j/article/view/335

McKinsey & Company. (2024). *What is Generative AI?*. Mckinsey & Company. Retrieved April 26 2024 from https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai#/

Menon, D., & Shilpa, K. (2023). Chatting with ChatGPT: Analyzing the Factors Influencing Users' Intention to Use the Open AI's ChatGPT Using the UTAUT Model. *Heliyon, 9*(11). https://doi.org/10.1016/j.heliyon.2023.e20962

Microsoft Research (2021, May 25). *Introducing Developer Velocity Lab to Improve Developers' Work and Well-being.* Youtube. https://www.youtube.com/watch?v=t7SXM7njKXw&t=952s

Minelli, R., Mocci, A., & Lanza, M. (2015). I Know What You Did Last Summer: An Investigation of How Developers Spend Their Time. 23d IEEE International Conference on Program Comprehension. DOI:. 10.1109/ICPC.2015.12

Noda, A., Storey, M.-A., Forsgren, N., & Greiler M. (2023). DevEx: What Actually Drives Productivity: The developer-centric approach to measuring and improving productivity. *Queue, 21*(2), 35-53. https://doi.org/10.1145/3595878

Noda, A. (2023). *A New Approach To Measuring Developer Productivity*. Engineering Enablement. Retrieved May 24, 2024, from https://newsletter.getdx.com/p/measuring-developer-productivity

Noda, A. (2024). *Applying the SPACE Framework.* LinkedIn. Retrieved May 24, 2024, from https://www.linkedin.com/pulse/applying-space-framework-abi-noda-euzhf/

Ooi, K. B., Tan, G. W. H., Al-Emran, M., Al-Sharafi, M. A., Capatina, A., Chakraborty, A., Dwivedi, Y.K., Huang, T.L., Kar, A.K., Lee, V.H., Loh, X.M., Micu, A., Mikalef, P., Mogaji, E., Pandey, N., Raman, R., Rana, N.P., Sarker, P., Sharma, A., Teng, C.I., Wamba, S.F., & Wong, L. W. (2023). The Potential of Generative Artificial Intelligence Across Disciplines: Perspectives and Future Directions. *Journal of Computer Information Systems,* 1–32. https://doi.org/10.1080/08874417.2023.2261010

Ozkaya, I. (2023). Application of Large Language Models to Software Engineering Tasks: Opportunities, Risks, and Implications. *IEEE Software, 40*(3), pp. 4-8, DOI: 10.1109/MS.2023.3248401

Peng, S., Kalliamvakou, E., Cihon, P., & Demirer, M. (2023). *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*. arXiv.org. https://arxiv.org/pdf/2302.06590.pdf

Pothukuchi, A. S., Kota, L. V., & Mallikarjunaradhya, V. (2023). Impact of Generative AI on the Software Development Lifecycle (SDLC). International Journal of Creative Research Thoughts, *11*(8). https://ssrn.com/abstract=4536700

Rempel, H., & Mellinger, M. (2015). Bibliographic Management Tool Adoption and Use A Qualitative Research Study Using the UTAUT Model. *Reference & User Services Quarterly, 54*(4), 43-53. http://dx.doi.org/10.5860/rusq.54n4.43

Sajja, R., Erazo Ramirez, S., Li, Z., Demiray, B.Z., Sermet, Y., & Demir, I. (2024). Integrating Generative AI in Hackathons: Opportunities, Challenges, and Educational Implications. *Cornell University Arxiv.* https://doi.org/10.48550/arXiv.2401.17434

Scheerder, A., van Deursen, A., & van Dijk, J. (2017). Determinants of Internet Skills, Uses and Outcomes. A Systematic Review of the Second- and Third-level Digital Divide. *Telematics and Informatics*, *34*(8), 1607-1624. https://doi.org/10.1016/j.tele.2017.07.007

Shylesh, A. (2017). *A Study of Software Development Life Cycle Process Models*. Srinivas Institute of Management Studies. http://dx.doi.org/10.2139/ssrn.2988291

Sætra, S. H. (2023). Generative AI: Here to Stay, but for Good?. *Technology in Society, 75,* Article 102372. https://doi.org/10.1016/j.techsoc.2023.102372.

Stokel-Walker, C., & Van Noorden, R. (2023). What ChatGPT and Generative AI Mean for Science. *Nature, 614,* 214-216. DOI: 10.1038/d41586-023-00340-6

Taeihagh, A., & Lim, H.S.M. (2019). Governing Autonomous Vehicles: Emerging Responses for Safety, Liability, Privacy, Cybersecurity, and Industry Risks. *Transport Reviews, 39*(1), 103-128. https://doi.org/10.1080/01441647.2018.1494640

TheoryHub. (n.d). Technology-Organization-Environment Framework. *TheoryHub.* Retrieved May 15, 2024, from https://open.ncl.ac.uk/academic-theories/23/technology-organization-environment-framework/

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Venkatesh, V., Thong, J. Y. L., & Xu, X. (2012). Consumer Acceptance and Use of Information Technology: Extending the Unified Theory of Acceptance and Use of Technology. *MIS Quarterly*, *36*(1), 157–178. https://doi.org/10.2307/41410412

Warcholinski, M. (2023, September 18). What Does a Software Developer Do (Besides Writing Code). *Brainhub. https://brainhub.eu/library/what-does-a-software-developer-do*

Witczak, K. (2023, July 29). How DORA, SPACE and DevEx connect?. *Witczax.* https://www.witczax.com/2023/dora-space-devex/

Zarifhonarvar, A. (2023), Economics of ChatGPT: a labor market view on the occupational impact of artificial intelligence. *Journal of Electronic Business & Digital Economics*. http://dx.doi.org/10.2139/ssrn.4350925

Zenseact. (n.d.). *People at heart.* Retrieved January 30, 2024, from https://zenseact.com/people-at-heart/

Appendix A – Interview Guide

Following is the English version of the interview guide used in this study:

Introduction:

We are writing a Master's thesis about the perception of using generative AI. We are focusing more on social/organizational aspects and not on the very technical side.

Is it okay if we record this meeting and take notes? The interview data will be treated confidentially, and no personal names or information will be disclosed in the final report.

General:

1. Can you start by telling us about your background and role here at Zenseact?
2. Can you explain what generative AI is for you? (How would you define generative AI?)

Perception:

3. What opportunities do you see with using generative AI in your work?
4. What challenges / limitations do you see with using generative AI in your work?
5. How do you perceive the general attitude towards generative AI within the company?
6. What are your feelings towards the use of generative AI at work?
    a. Are there any **concerns** you have regarding its ongoing or future use?
    b. Are there any **positive expectations** you have regarding its ongoing or future use?

7. Fill in the table.

AI tools	**Areas of use**

8. Follow up on the table.
    a. What effects do you see on your work? (work process and result)
        i. Positive:
            1. *Satisfaction and well-being*

2. *Quality*
3. *Performance*
4. *Efficiency and flow*
5. *Activity*

ii. Negative:

1. *Satisfaction and well-being*
2. *Quality*
3. *Performance*
4. *Efficiency and flow*
5. *Activity*

b. Have your work tasks changed?

c. Are there specific tasks or areas where you have noticed a more significant impact?

i. Software development:

d. From your perspective, what factors could be important to evaluate or measure when assessing the effects of generative AI in software development?

9. Do you have any ideas or plans on further use cases for generative AI in your work? For colleagues?

i. *Specific tool in mind?*

ii. *What would you hope to gain from using (why)?*

iii. *Any blockers on implementation?*

10. How has generative AI influenced team dynamics and collaboration within your department?

Closing questions:

11. Are there any aspects of generative AI in the workplace that we have not covered but you want to raise?

12. Do you have any suggestions or recommendations around the integration of generative AI in the company's software development processes?

13. Who do you consider a role model in using generative AI within the company or are there any specific team that uses generative AI in their work? / Who do you think would be interesting for us to talk to?

Specific questions for this person:

Appendix B – Survey for Evaluation of GenAI Use

Evaluating the use of generative AI

Generative AI models utilize various techniques to generate new data, tailored to the specific type of output desired, such as text or images. Examples of generative AI tools are ChatGPT and GitHub Copilot.

* Required

1. I believe GenAI can *

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
Improve my performance	○	○	○	○	○
Help me save time	○	○	○	○	○
Provide me with insights and perspectives I may not have thought of myself	○	○	○	○	○
Provide me with personalised and immediate feedback and suggestions	○	○	○	○	○
Limit my interaction with colleagues	○	○	○	○	○
Hinder my development of skills such as teamwork, problem-solving, and leadership skills	○	○	○	○	○
Empower me to work with greater confidence and achieve greater impact	○	○	○	○	○
Answer many routine questions which might reduce interruptions in my work	○	○	○	○	○
Be trusted to produce reliable and accurate results in my work	○	○	○	○	○

2. Please indicate how strongly you agree with the following statements. *

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
The GenAI-related resources (e.g. licenses) that Zenseact provides are relevant and useful.	◯	◯	◯	◯	◯
Zenseact's guidelines for the use of GenAI are clear and helpful.	◯	◯	◯	◯	◯
I know where to find help and inspiration regarding the use of GenAI.	◯	◯	◯	◯	◯
GenAI is easy to integrate into our team's workflow and tool environment.	◯	◯	◯	◯	◯
I feel that I can openly discuss GenAI with my colleagues in my job role.	◯	◯	◯	◯	◯

3. How often do you use GenAI tools? *

◯ Every day

◯ 2-6 days a week

◯ Once a week

◯ 2-3 days a month

◯ Once a month

◯ Less than once a month

◯ Never

4. Which of the following tools do you use? *

- [ ] ChatGPT
- [ ] Github Copilot
- [ ] Microsoft Copilot
- [ ] Other

5. I use GenAI for the following activities *

	Never	Less than once a month	Once a month	2-3 days a month	Once a week	2-6 days a week	Every day
Developing and editing code	○	○	○	○	○	○	○
Searching, interpreting and analyzing existing code	○	○	○	○	○	○	○
Other code-related tasks (e.g. documenting code)	○	○	○	○	○	○	○
Other non-code-related tasks (e.g. writing emails)	○	○	○	○	○	○	○

6. When using GenAI tools *

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
I feel more productive	◯	◯	◯	◯	◯
I find my job more enjoyable	◯	◯	◯	◯	◯
I spend less time searching for information	◯	◯	◯	◯	◯
I put less mental effort on repetitive tasks	◯	◯	◯	◯	◯
I feel faster with repetitive tasks	◯	◯	◯	◯	◯
I feel I can ask questions that I otherwise would not have asked	◯	◯	◯	◯	◯
I feel like I can get help with understanding code more easily	◯	◯	◯	◯	◯
I feel more creative	◯	◯	◯	◯	◯
I feel working with code becomes less stressful and demanding	◯	◯	◯	◯	◯
I can work with more focus, entering a state of flow	◯	◯	◯	◯	◯

7. Think about your average work week **before** you used GenAI tools. Now think about your average work week **today** (using GenAI tools).
Through your use of GenAI today, how many hours per week do you estimate that you save and/or free up to focus on other tasks? *

- ◯ Using GenAI rather costs me time
- ◯ Less than 1 hour
- ◯ 1-2 hours
- ◯ 2-4 hours
- ◯ 4-8 hours
- ◯ More than 8 hours

8. My limited use of GenAI is because *

	Strongly disagree	Disagree	Undecided	Agree	Strongly agree
I do not know how to use it effectively	◯	◯	◯	◯	◯
I am unsure about the tasks for which it can be used	◯	◯	◯	◯	◯
I am uncertain about how I am allowed to use it	◯	◯	◯	◯	◯
The tools available are not good enough	◯	◯	◯	◯	◯
I have not been provided with a company license	◯	◯	◯	◯	◯
I prefer to use my own capabilities to perform tasks	◯	◯	◯	◯	◯
I prefer my current way of working	◯	◯	◯	◯	◯

9. What would make you use GenAI tools more? *

10. Do you have any other thoughts or requests related to the use of GenAI at work?

DEPARTMENT OF TECHNOLOGY MANAGEMENT AND ECONOMICS
DIVISION OF INNOVATION AND R&D MANAGEMENT
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden
www.chalmers.se

AVANCEZ
1829

CHALMERS
UNIVERSITY OF TECHNOLOGY