DEPARTMENT OF APPLIED IT

UNIVERSITY OF
GOTHENBURG

Between Automation and Expertise:

Employee Sensemaking in Early-Stage Agentic AI Implementation

By Frida Piscator Pettersson

A Thesis submitted for the Master's Degree
in Human-Centered Artificial Intelligence.
May 21st, 2026

Internal Supervisor: Dina Koutsikouri
External Supervisors: Viktor Ström & Johan Gärdebo
Examiner: Marisa Ponti

Abstract

As organizations increasingly invest in AI, understanding the gap between implementation and actual adoption remains a significant challenge. While technical considerations have received considerable attention, employee-level dynamics and the influence of agentic AI remain underexplored in AI adoption research. This thesis examines how employees within a management consultancy firm make sense of the introduction of an agentic AI system during the early stages of organizational implementation, and how this sensemaking process shapes their engagement with the technology.

Drawing on sensemaking theory, eight sensemaking mechanisms, and a human-centered AI perspective, a qualitative case study was conducted through six interviews with employees at a Swedish management consultancy firm that recently implemented a locally hosted, agentic AI-powered chatbot.

The analysis reveals that employees made sense of the system by drawing on available resources such as familiar generative AI tools, future-oriented beliefs and delegated safety claims, to construct plausible but not necessarily accurate or organizationally aligned interpretations. Employees also turned to colleagues to make sense of the system collectively, which emphasizes the social nature of such sensemaking processes. Furthermore, the specific context of knowledge-intensive work surfaced how employees were not only interpreting a new tool but negotiating their professional identity in relation to a technology that challenges the same capabilities that their professional value is built on.

The findings contribute to research on employee-level sensemaking in AI implementation and they offer practical implications for leaders seeking to facilitate the adoption of agentic AI through a human-centered approach.

Keywords

Human-Centered Artificial Intelligence, Agentic AI, AI Implementation, AI Adoption

Foreword

This thesis marks the end of my master's in human-centered AI, and it reflects my genuine belief in preserving the human perspective in a world where AI is increasingly seen as a one-fits-all solution. My hope is that this thesis, even if just a little, can contribute to conversations about why keeping human values and needs at the center of AI development matters, and that it can inspire someone to view AI from a different angle.

First, I would like to thank my supervisor Dina for your amazing support and for showing genuine belief that I would always figure things out, even in moments of doubt. And thank you for giving us students the best start of the programme last year, by welcoming us with such warmth and enthusiasm.

I would also like to thank Viktor and Johan at Thalamind for giving me the opportunity to work together with you, and to be a part of such an inspiring and hard-working team. When I said yes to writing a thesis with you this fall, I could never have imagined what a unique and valuable situation I would find myself in, one that made this thesis possible in the first place.

Finally, thank you to my classmates for the early mornings in Kuggen and for two amazing years together!

Table of contents

1 Introduction 1
2 Related work 3
2.1 Agentic AI 3
2.2 Knowledge-intensive firms 4
2.3 AI adoption and AI implementation 4
2.4 Sensemaking and AI Implementation 6
3 Theoretical framework 8
3.1 Sensemaking mechanisms 8
3.2 Human-Centered Artificial Intelligence (HCAI) 10
4 Method 11
4.1 Research design 11
4.2 Research setting 11
4.3 Data collection 12
4.3.1 Interviews 12
4.3.2 Selection process 13
4.4 Data Analysis 13
4.5 Researcher positionality 14
4.6 Ethical considerations 15
4.7 Use of AI 15
5 Findings 17
5.1 Translating the technology through familiar tools 17
5.2 Using professional identity to navigate AI use 18
5.3 Anchoring meaning in future potential 19
5.4 Understanding system safety 20
5.5 Collective sensemaking through peer learning 21
6 Discussion 22
6.1 Practical implications 24
6.1.1 Translate 24
6.1.2 Converse candidly 25

6.1.3 Look closely and update 25
6.1.4 Provide a direction 26
6.1.5 Facilitate learning 26
6.2 Limitations 27
7 Conclusion 29
8 References 30

1 Introduction

Almost every organization is concerned with artificial intelligence (AI). It has rapidly become a frequent subject of discussion across industries, with the potential to fundamentally reshape how organizations operate and compete. Organizations are investing heavily to capture that potential, with enterprise spending on generative AI reaching billions of dollars, driven largely by expectations of cost decreases and revenue growth (Singla et al., 2024). Yet, the scale of investment has not translated into equivalent organizational impact. Many organizations remain stuck in stages of experimentation and pilot phases rather than achieving widespread adoption, and a recent industry report indicates that 95% of organizations achieve no measurable outcomes from their AI initiatives (IBM, 2024; Challapally et al., 2025).

Organizational AI initiatives often fail to deliver their intended benefits, not only due to technical considerations but also because they encounter barriers to AI adoption related to social and organizational dynamics, such as unrealistic expectations and misaligned understandings of AI procedures (Westenberger et al., 2022). While technology can be implemented relatively quickly, changing the way people think and work is harder, and research shows that most successful digital change initiatives start with a change of mindset at both employee and organizational levels (Kane, 2019). This is further shown in practice, as industry reports show patterns of leaders pushing AI adoption faster than employees are comfortable with, which widens the gap between what AI can do and what employees are willing to integrate in their working life (IBM, 2024; Challapally et al., 2025). Conditions for successful AI adoption are thereby constructed through human and organizational processes and not through technology alone.

If the limiting factor is not only technical considerations but people and processes in the context to which AI is introduced, understanding how individuals experience and respond to AI implementation becomes essential. AI systems present distinct organizational challenges compared to traditional information systems, when properties such as autonomy and inscrutability generate ongoing uncertainties that must be navigated through the interpretative work of the individuals who encounter them in practice (Berente et al., 2021; Sagodi et al., 2023). These challenges become even more emphasized with the introduction of agentic AI, where increased autonomy and goal-directed behaviors are achieved through architectural approaches and training techniques (Hosseini & Seilani, 2025). Despite the significance of these individual-level dynamics, research within technology transformation has focused predominantly on business and strategic levels of implementation, while employee experiences and behavior remain underexplored in AI adoption research (Trenerry et al., 2021: Gottumukkala & Prasad, 2025).

Within this underexplored area, the specific characteristics of agentic AI make certain contexts particularly relevant to examine. Knowledge-intensive firms (KIFs) represent one such

context, where professional value is closely linked to human expertise and autonomy, independent judgment and established legitimacy through professional identity (Alvesson, 2001; Starbuck, 1992; Drucker, 1999). Introducing agentic technology into such an environment may create certain tensions related to professional identity, autonomy and perceptions of work.

What employees are doing when encountering a complex technology such as agentic AI is essentially a process of meaning construction. Sensemaking theory (Weick, 1995) offers a valuable lens for understanding how individuals construct meaning around AI technologies during implementation. How individuals make sense of new technology during early implementation further shapes whether engagement and adoption follow, and this process is particularly critical in the context of AI when those interpretative demands are intensified (Engström et al., 2024; Sagodi et al., 2023). Understanding this process of meaning construction and the organizational conditions that support or constrain them, can therefore provide essential information on how employee-level processes in AI implementation affects AI adoption initiatives.

Research has highlighted the need for greater attention to employee-level aspects and the impact of agentic systems on AI adoption (Gottumukkala & Prasad, 2025), while emphasizing the impact of employee-level sensemaking in early-stage AI implementation on adoption. The specific characteristics of knowledge-intensive contexts further suggest that they provide a particularly relevant setting in which to examine these dynamics. Therefore, this study aims to explore how employees within management consultancy make sense of the introduction of agentic AI during early-stage organizational implementation, and how this shapes the engagement with the technology. The study addresses the following research question:

RQ: How do employees within management consulting firms make sense of agentic AI systems during early-stage implementation?

By answering this question, the study seeks to contribute to a human-centered perspective of early-stage AI implementation by surfacing employee sensemaking processes and translating these into practical insights for organizational leaders managing AI implementation with the goal of facilitating AI adoption.

2 Related work

The following section reviews research relevant to the focus of the study. First, the characteristics of agentic AI are outlined, followed by an overview of defining features of knowledge-intensive firms. Thereafter, previous research on AI implementation and AI adoption is examined, before sensemaking is introduced as the study's theoretical lens. Together these perspectives provide the conceptual foundation that guides the framing and analysis of the study.

2.1 Agentic AI

Over the past years, the field of AI has moved beyond systems designed for well-defined tasks in static environments, giving rise to a new class of technologies commonly referred to as agentic AI. Unlike traditional AI agents and chatbots that respond to explicit user inputs, these systems are characterized by their ability to make autonomous decisions and operate within environments with limited or no human intervention (Hosseini & Seilani, 2025).

Recent advances in generative models and large language models have enabled the development of AI agents capable of performing specific tasks, operating with high autonomy and adaptability within their specific domain (Sapkota et al., 2025). Agentic AI systems can further combine multiple AI agents that collaborate or compete to achieve a common goal, often through modular or hierarchical architectures such as multi-agent systems, hierarchical reinforcement learning, and goal-oriented modular designs (Acharya & Kuppan, 2025). Rather than acting as isolated assistants, these systems of AI agents are thereby explicitly designed to collaborate through goal-directed behavior, combining autonomy, task specificity, and adaptability (Sapkota et al., 2025).

This marks an important transition in the role of AI within organizations. Where previous generations of AI tools assisted users with defined tasks, agentic AI systems can act independently across workflows, perceiving their environment and initiating actions without user instructions. Holmström (2022) notes that AI technologies with more human-like capabilities such as acting and perceiving independently may introduce new and consequential challenges for different actors when introduced into organizations. Hosseini and Seilani (2025) further argue that despite agentic AI's potential to revolutionize industries through efficiency and improved capabilities, it must be adopted with strategies to mitigate risks associated with the technology. One such challenge mentioned is the transformation of the workforce, highlighting new questions about how these systems interact with human workers and organizational structures.

2.2 Knowledge-intensive firms

Knowledge-intensive firms are organizations in which knowledge constitutes both the primary input and output of work (Starbuck, 1992). These organizations are characterized by workers applying intellectual expertise and professional judgment to complex, non-standardized tasks, with examples including management consultancy firms, law firms and investment firms (Alvesson, 2004).

A defining feature of KIFs is the high degree of autonomy afforded to their employees. Drucker (1999) argues that knowledge workers are largely responsible for their own productivity and therefore require autonomy to perform effectively. Alvesson (2004) reinforces this need for autonomy by noting that individuals within KIFs are often considered to have the best insights into problem areas, and the most familiarity with specific challenges because of their professional experiences. This is reflected in the structure of these organizations, where authority tends to follow expertise, rather than any formal position. Starbuck (1992) further argues that at least a third of employees in KIFs can be considered experts, referring to their high degree of education and experience.

Related to this is the significance of professional identity in knowledge-intensive work. Alvesson (2001) argues that KIFs are inherently ambiguity-intensive, since the value of knowledge, what the employees do, and the result of their work, are difficult to evaluate objectively. In the absence of clear performance criteria, professional identity and the perception of one's own expertise therefore become central to how knowledge workers establish their value and legitimacy. Starbuck (1992) notes that the value of professional identity and expertise in KIFs may also influence employees' tendency to resist change, since the need to learn can be considered evidence of lack of sufficient knowledge.

Because the value of expertise is difficult to objectively verify in these contexts, knowledge workers are continuously engaged in negotiating meaning, managing impressions and accounts of their own competence and contributions (Alvesson, 2001). These characteristics suggest that the introduction of new, autonomous and knowledge-based technology such as agentic AI lands in an environment already defined by ambiguity and ongoing negotiation of individual meaning. Therefore, understanding the dynamics emerging within these contexts when agentic technology is implemented can be considered crucial to form the conditions for organization-wide adoption.

2.3 AI adoption and AI implementation

AI adoption presents organizations with characteristics that distinguish it from the adoption of traditional information systems. Berente et al. (2021) argue that managers within organizations must treat AI as changing sociotechnical systems, where concepts such as autonomy, learning and inscrutability represent core challenges of managing AI. Inscrutability refers to the difficulty of making AI systems intelligible due to the complex nature of the technology,

creating a fundamental challenge for users who cannot rely on consistent or transparent outputs to guide their understanding (Berente et al., 2021). These characteristics are not merely technical properties but management challenges that must be navigated organizationally, generating new uncertainties that shape how individuals engage with these technologies.

AI implementation refers to the organizational process through which a technology is introduced, and these early stages of adoption shape how members of an organization come to understand it (Engström et al., 2024). This process is particularly important for AI-based technology, as the unique characteristics of these systems generate new uncertainties that must be interpreted through the organization before adoption can take hold (Sagodi et al., 2023). Implementation is therefore not merely a precondition, but an active state where the conditions for adoption are constructed.

Research has increasingly shown that employee-level factors play a significant role in whether such processes succeed. Trenerry et al. (2021) identify five individual-level factors shaping how employees engage with new technologies, including their perceptions and attitudes towards technological change, their skills, access to training, and their resilience and adaptability. Importantly, Trenerry et al. (2021) note that existing research has focused largely on business and strategic levels of digital transformation, with less attention given to these individual factors. This is echoed in AI adoption research, where Gottumukkala and Prasad (2025) identify the need for extended research on individual behavior within the context of AI adoption. This despite initiatives such as Tursunbayeva and Gal (2024), who developed the Technology, Organization and People (TOP) framework to address individual levels and people within organizations in the process of AI adoption. Although the individual level is addressed in the framework, it is rather framed as something to manage than to explicitly explore or understand. The social and individual aspects of AI adoption are further highlighted by Cubric (2020), who argues that social considerations such as lack of knowledge, trust issues and safety concerns represent main barriers for AI adoption. In a related study on clinicians adopting machine learning tools, Henry et al. (2022) emphasize how individuals understood AI systems through experience, and how that experience and communication with colleagues defined trust in the system, highlighting the importance of employees' experiences in interaction with complex systems. This points to a broader pattern in the literature emphasizing the role of employee-level factors in the context of AI adoption, and the need to develop deeper understanding for these factors within the research area.

Overall, the complexity and uniqueness of AI implementation create new challenges for both employees and managers within organizations hoping to introduce and adopt this technology. With agentic AI, challenges around inscrutability, trust and safety become even more pressing, as increased autonomy and goal-directed behavior achieved through architectural designs further intensify the complexity of these systems. Gottumukkala and Prasad (2025) highlight this by arguing that current AI adoption literature calls for more research towards generative and agentic AI. Additionally, research shows that individuals within these organizations play a crucial role in determining whether implementation leads to organization-wide adoption or

not. Understanding how individuals understand and construct meaning around novel AI technology during implementation phases therefore becomes essential.

2.4 Sensemaking and AI Implementation

The concept of sensemaking offers a valuable theoretical lens for understanding how individuals within organizations come to interpret and engage with agentic AI. Importantly, sensemaking is not purely an individual process, but operates across multiple levels simultaneously, including the personal sensemaking of employees, the shared sensemaking of teams and groups and broader sensemaking processes across the organization (Maitlis & Christianson, 2014; Weick et al., 2005). These levels are further shaped by the broader environment in which actors of the organization are embedded, such as industry pressure and media narratives, which affect how individuals and organizations make sense of new technologies (Maitlis & Christianson, 2014). By examining how individuals construct meaning around AI during implementation, it becomes possible to trace consequences that reach the organizational level, while also revealing how macro-level influences translate into individual-level processes.

More recent work has begun to address this by bringing sensemaking into the field of AI implementation specifically. Engström et al. (2024) demonstrate that how employees perceive AI technology shapes the sensemaking processes they activate, which in turn determines what kind of learning and engagement with the technology that follows. Goto (2022) similarly shows that professionals in service firms construct their understanding of AI during early stages of technology adoption through an ongoing cycle of monitoring, experimenting, and developing views of the future. This sensemaking process is further argued to directly shape their willingness to adopt AI in professional practice, while interactions in professional settings determine which future possibilities they consider viable. Taken together, research shows that individual sensemaking processes are specifically important during early stages of AI implementation and that they have a crucial impact on whether the technology is adopted at organizational levels.

Sensemaking, as conceptualized by Weick (1995), refers to the process through which individuals interpret and construct meaning when faced with novel or unfamiliar situations. More specifically, the process involves relating unfamiliar events to prior experiences and existing understanding, through both cognitive and behavioral processes (Gioia et al., 1994). The process is set in motion when individuals encounter something they cannot already explain with existing understanding, forcing them to seek new interpretations (Maitlis & Christianson, 2014). Sagodi et al. (2023) argue that the complex characteristics of AI function as triggers for such sensemaking processes, creating demands for interpretations which are not present with conventional information systems. Weick (1995) identifies seven characteristics which together explain how this process of sensemaking construction unfolds:

Grounded in identity construction – Sensemaking is shaped by how individuals understand who they are in each context. As individuals encounter new situations, their sensemaking both reflects and reshapes their identity.

Retrospective **–** Individuals make sense of events after they occur, by reflecting on previous experience and actions.

Enactive of sensible environments **–** Individuals partly create the environments they interpret through action, meaning that interpretation and action are mutually important for the process.

Social **–** Interpretations are constructed through interaction, communication and shared practices, rather than in isolation.

Ongoing **–** Sensemaking does not have a clear beginning or end, since individuals continuously update their interpretations of the environment as new events occur.

Focused on and by extracted cues **–** Individuals rely on selected cues from complex environments to construct meaning.

Driven by plausibility rather than accuracy **–** Sensemaking aims for interpretations that are coherent and believable rather than objectively correct.

Weick's (1995) characteristics provide the foundational properties of sensemaking, but more recent initiatives emphasize the prospective dimension of the process, which is particularly relevant for novel technologies like agentic AI. Sandberg and Tsoukas (2015) argue that sensemaking is not only retrospective but also involves prospective imagination, where individuals project possible futures and test potential interpretations against these imagined realities. Similarly, Goto (2022) raises the prospective aspect of sensemaking by demonstrating how professionals construct future scenarios around AI practices, to evaluate the role of the technology in their own work. Hence, the forward-looking dimension becomes a critical addition to the seven characteristics, as employee's early projections about novel technology such as agentic AI may shape their willingness to engage and adopt.

Together, these characteristics demonstrate why sensemaking is particularly relevant for understanding agentic AI implementation in knowledge-intensive firms. First, it plays a crucial role in the early stages of technology implementation, shaping preconditions for organization-wide adoption (Engström et al., 2024). Second, agentic AI's high autonomy and unpredictability intensify interpretive demands, while the centrality of professional expertise signals that employees may interpret it as potentially reshaping their roles given how the process of sensemaking is intertwined with identity construction. Third, external pressures from industry reports provide macro-level cues that influence these processes (Maitlis & Christianson, 2014). For leaders, understanding these dynamics is essential, as they represent a determining factor to whether implementation gets stuck at experimentation stages or progresses toward adoption.

3 Theoretical framework

This study is guided by two complementary theoretical perspectives, presented in the following section. The eight sensemaking mechanisms proposed by Iveroth and Hallencreutz (2015) provide an interpretative lens to interrogate what employee sensemaking reveals about organizational practice, while the human-centered AI perspective frames the study's orientation toward implementation that places human experiences and needs at the center.

3.1 Sensemaking mechanisms

Iveroth and Hallencreutz (2015) argue that "leaders have to become aware of, understand and find meaning in new things." (p. 4). To do so, leaders must acknowledge emerging interpretations and understand how meaning is constructed within their organization through ongoing communication. Hence, understanding the current narrative on both an individual and organizational level is crucial for facilitating sensemaking processes.

Iveroth and Hallencreutz (2015) use a landscape metaphor, framing organizations as landscapes that leaders must map by understanding diverse employee perspectives. Each individual actor within this landscape will have different perspectives depending on where they are located, just like individuals within an organization will have different meaning and preferences depending on, for example, skills and hierarchical position. Thus, all these characteristics together will form the whole landscape, which the leaders of organizational change should map (Iveroth & Hallencreutz, 2015).

To navigate such a landscape, Iveroth and Hallencreutz (2015) propose eight mechanisms that describe how sensemaking processes can be facilitated in organizational change contexts, as shown in table 1. In this study, these mechanisms are used as a sensitizing and interpretative lens in the analysis, supporting the interpretations of how employees make sense of agentic AI during early-stage implementation. Specifically, the mechanisms are used to connect the observed patterns of employee sensemaking to potential leadership practices that may support or constrain these processes. In this way, they enable the translation of empirical findings into practical insights for leaders navigating the early stages of AI implementation in knowledge-intensive firms.

Table 1. Eight sensemaking mechanisms, adapted from Iveroth and Hallencreutz (2015).

Guidelines	Description

Change through logic of attraction	• Emphasizes the importance of leaders changing their own practices first, to inspire others to engage in transformation. • Demonstrate change in practice, and allow others emulate new ways of working.
Provide a direction	• Emphasizes the importance of articulating a clear and aligned direction which can strengthen confidence in others, reduce confusion and show the way forward. • Help people understand where they are and where they are going next.
Translate	• Emphasizes the important role of translating visions and strategic plans into concrete plans of action. • Strengthen the engagement of people through translation of complex strategies into smooth plans.
Stay in motion	• Emphasizes the importance of generating activities and make visible decisions early in the change process. • Praise initiatives to move ahead instead of getting stuck in processes and policies.
Look closely and update	• Emphasizes the importance of being involved in the change situation on the ground and pick up cues in the front line. • Improve people's situational awareness, to understand the current situation and act based on that.
Converse candidly	• Emphasizes the importance of communicating throughout the whole change process, rather than only at early stages. • Make people talk and listen carefully to understand their meanings and truths.
Unblock improvisation	• Emphasizes the importance of allowing early adopters to experiment and test new ideas. • Encourage themselves and others to deviate from plans to work with material at hand and get rid of organizational hurdles which may stifle improvisation.
Facilitate learning	• Emphasizes the importance of facilitating learning to guide future actions and interpretations. • View mistakes as an opportunity for learning and reinforce new behavior through feedback and reward.

3.2 Human-Centered Artificial Intelligence (HCAI)

According to Shneiderman (2022), human-centered AI (HCAI) emphasizes the development of technologies that "amplify, augment, empower and enhance human performance" (p. 4). Rather than positioning AI as a replacement for human capabilities, this perspective frames AI systems as supportive tools designed to extend human competence and agency. From this viewpoint, the design and implementation of AI technologies should aim to reduce uncertainty, foster user control, and enable meaningful engagement with technology in everyday life (Shneiderman, 2022).

Shneiderman (2022) primarily frames HCAI as a design and governance perspective, offering principles intended to guide the development and design of AI systems that are reliable, safe, and aligned with human needs. Central to this view is the framing of AI systems as "super-tools", which support users in achieving goals while maintaining meaningful control over technological processes. Beyond its design-oriented focus, HCAI also provides a valuable lens for understanding the individual experiences related to human-AI collaboration, by emphasizing individual goals such as self-efficacy, creativity and responsibility. Shneiderman (2022) highlights the importance of designing highly automated systems in ways that preserve human control, thereby challenging earlier assumptions that increased automation necessarily implies reduced human involvement. This emphasis on balancing autonomy and control is particularly relevant in contexts where highly autonomous systems are introduced into organizations characterised by high levels of individual responsibility and self-management.

Overall, HCAI highlights the importance of the human perspective in AI adoption by drawing attention to goals of creating reliable, safe, and trustworthy AI systems. In this study, HCAI serves as a complementary sensitizing device, with the aim to better understand how a human-centered approach to AI implementation can be achieved. This complements the sense-making mechanisms, which provide the analytical structure for understanding how such technologies are interpreted and experienced by employees in relation to organizational practices.

4 Method

The following section describes the methodological approach used for the study, including the selected research case, data collection methods and the data analysis process. It also describes the positionality of the researcher and ethical considerations relevant to the study.

4.1 Research design

The study adopts a qualitative case study approach, motivated by the nature of the research question under investigation. Rather than seeking to measure variables or causal relationships, the aim is to explore how employees within a management consulting firm make sense of agentic AI in early-stage implementation. In other words, the focus is to capture the meaning and lived experience of employees, which calls for an approach capable of capturing the depth and complexity of that experience (Denzin & Lincoln, 2011).

The rationale for selecting the case organization was that it provided access to a context where it was possible to examine the early phases of agentic AI implementation at a knowledge-intensive firm. Furthermore, a case design was motivated by the broad definition of agentic AI systems, implying that the characteristics of these systems may differ across organizations. A cross-organizational approach could thereby risk compromising the depth of insight.

4.2 Research setting

The setting of the studied case was a small management consultancy firm in Sweden, which recently introduced an agentic AI-powered chat bot into the organization. The firm employs approximately 20 people and operates across various sectors and client contexts. Access to the firm was established through prior contact with the organization, which enabled the researcher to study the implementation process at an early stage.

In the fall of 2025, the firm introduced a locally hosted chatbot powered by multiple AI agents, constituting an agentic AI system. The chatbot enables employees to access and query internal documents and knowledge, while being tailored to organizational needs through fine-tuning and continuous development. Furthermore, the local hosting of the system offers sovereign data control through a secure, air-gapped architecture. At the time of data collection, the introduction was still in its early stages, as employees had had access to the system for approximately two months. Employees were therefore actively engaged in making sense of what the system meant for their work and how it had or would affect their everyday work practices.

Prior to the initiation of the study, the researcher had established contact with the firm through an internship at the startup responsible for developing and implementing the system, which facilitated access to the organization.

For the purposes of this study, agentic AI is operationalized mainly through two characteristics drawn from the related work and in accordance with what the implemented system specifically offers. The first characteristic is high autonomy, referring to the system's capacity to act through multiple agents operating in the background without continuous human instruction. The second is goal-directed behavior, referring to the system's ability to pursue objectives across workflows rather than responding to isolated inputs. These characteristics are thereby treated as analytically relevant throughout the study, as they represent the properties most likely to shape employee sensemaking, given the specifics of the implemented system.

4.3 Data collection

The primary source of data for this study consists of semi-structured interviews conducted with employees within the organization. The following sections describe the interview method, the participants and how they were selected, and the practical aspects involved in the data collection process.

4.3.1 Interviews

Six in-depth interviews with employees at the case organization were carried out via Google Meets. Following Hennink et al. (2020), in-depth interviews seek to gain information on an individual level regarding a specific experience or topic. Further, these interviews were conducted in a semi-structured manner, using an interview guide that was developed with the aim to capture the employee sensemaking around agentic AI in the early implementation stages. The focus was thereby to design questions that would allow employees to freely talk about their experiences and interpretations.

The interview guide consisted of open questions related to work procedures, the technology introduced, the implementation process, and potential changes in workflows. Iveroth and Hallencreutz's (2015) eight mechanisms for sensemaking were used as a guide to develop the interview guide, with the goal to capture the sensemaking processes happening within the organization. However, the questions were used to guide the interviews rather than dictate them.

Each interview was recorded and lasted between 40 and 45 minutes.

4.3.2 Selection process

The participants were selected in collaboration with contacts at the organization, to allow for individuals with deeper knowledge about the organization to identify employees with different roles and experiences. Contact information was provided through these organizational contacts, and each potential participant was approached individually via email, asked about their interest to participate. In total, six employees took part in the study, all listed in table 2.

Table 2. Overview of interview participants

Pseudonym	**Role**
P1	Consultant
P2	Consultant
P3	Consultant
P4	Former CEO/Partner
P5	Consultant
P6	Partner/Consultant

The roles of participants varied from consultants to partners and a former CEO, allowing for multiple perspectives on the implementation process. This includes both those who may have been more closely involved in the introduction of the system and those relatively distant from it. Such variation in roles and experiences is theoretically grounded, as Iveroth and Hallencreutz (2015) argue that individuals with different positions within an organization perceive the organizational landscape differently.

4.4 Data Analysis

The interview data were analyzed using thematic analysis as described by Braun and Clarke (2006), a method for identifying and analyzing patterns across a dataset to report recurring themes. Thematic analysis emphasizes the active role of the researcher in constructing themes through interpretation and reflection, which is consistent with the interpretative approach of the study, motivated by the complex nature of the experiences under study.

The analysis moved back and forth between six phases in an iterative process, as defined by Braun and Clarke (2006). In the first phase, interview data were transcribed with support from a generative AI transcription tool and read through carefully to establish familiarity with the data. In the second phase, codes were generated inductively from the data, without reference to the mechanisms proposed by Iveroth and Hallencreutz (2015). The inductive starting point was motivated by the limited research on individual sensemaking in the context of organizational implementation of agentic AI, where an open approach was considered necessary to generate new knowledge rather than mirroring the previous research. However, as Braun and Clarke (2006) note, researchers are never free of theoretical commitments, and the theoretical foundation of the study inevitably shaped this part of the process as well.

The first two phases resulted in 295 individual codes, organized in a coding sheet. In the third phase of searching for themes, frequently occurring codes were identified and grouped by similarity into initial candidate themes. In the fourth phase of reviewing themes, the themes were then evaluated against the context of the study and the assigned research question, and those that were either infrequent, inconsistent across the data, or not analytically relevant given the context of the study were discarded. Examples included themes with codes only related to a few participants or themes that were off topic given the focus on employee sensemaking related to the implemented technology. This resulted in the five themes presented in the findings.

Theory began influencing the process during the fifth phase of refining and naming themes, rather than the initial coding stage, which preserved the inductive grounding of the analysis, as motivated in earlier sections. This step was theoretically motivated by the intertwined nature of employee and leader sensemaking, where employee experiences during implementation inevitably reflect the organizational conditions that leaders create. Similarly, leaders must be aware of and responsive to employee sensemaking processes to effectively facilitate sensemaking during implementation. Iveroth and Hallencreutz (2015) sensemaking mechanisms were applied as an interpretative lens to interrogate what each theme revealed about organizational practice, rather than as predefined categories into which data was sorted. This distinction was considered important to avoid simply mapping findings onto the framework and instead allowed the researcher to interrogate each theme by asking which of them were most relevant, present or absent in relation to what employees described. This step enabled the translation of empirical findings into practical implications for leaders rather than shaping the entire analytical process.

In the sixth phase of producing the report, the most relevant quotes were selected to represent each theme in the written analysis, and the movement between data and framework shaped the final articulation and naming of the themes.

4.5 Researcher positionality

According to Hennink et al. (2020), the background and identity of the interviewer may influence the quality of the data collected, primarily because participants may censor, perform or withhold information depending on how they perceive the role or intentions of the interviewer. Power relations between interviewer and participants are also relevant, as differences in expertise, partnerships or perceived authority may constrain or enable certain kinds of responses.

In the context of this study, the researcher's prior involvement in the implementation process represents both a potential source of bias and a prerequisite for the research itself. Access to this environment, including participants, the organization and the ongoing implementation of agentic AI was made possible precisely because of that prior relationship. Many contexts of

this kind remain inaccessible to research; hence, the researcher's prior engagement created the conditions under which the study could be conducted.

At the same time, this positioning introduced a risk that participants might perceive the interview as an evaluation of the technology introduced, rather than an open conversation about their own experiences. To mitigate this, the researcher explicitly informed each participant of their current role as an independent master's student and clarified that the purpose of the interviews was unrelated to the quality of this technology. The participants were also encouraged to treat the interview as an independent research study and answer based on what they experienced rather than giving feedback on the technology or introduction process initiated by the providing tech company.

4.6 Ethical considerations

This study was conducted in accordance with ethical guidelines for research involving human participants as proposed by the Swedish Research Council (2024), specifically with regards to informed consent and handling of personal data. Prior to data collection, participants received a written consent form informing them about the purpose of the study, the anonymity of their participation, data handling and storage procedures, their right to access, change and delete their data and that the interviews would be recorded. They were also informed that participation was completely voluntary and that they could withdraw from the study at any given time without consequences. Verbal consent was collected at the start of each interview, at which point the conditions for participation were briefly summarized to ensure that participants had read and understood the consent form. To protect the participant anonymity, the organization is not named in the study, and participants are referred to by numerical IDs throughout.

Data was stored on local devices protected by personal passwords, and only available to the researcher and supervisors. The generative AI device used for transcription was a fully local model, and the transcriptions were thereby only available to the researcher and supervisors. All recordings and transcriptions will be deleted after the completion of this thesis.

4.7 Use of AI

Generative AI tools were used during the study for three different purposes, reported here in the interest of transparency.

First, a fully on-premises generative AI tool (Thalamind, n.d.) was used for transcribing the recorded interviews. As the models for the tool were downloaded and run locally with no third-party dependencies, this was considered appropriate for handling interview data while preserving participants' privacy.

Second, Claude AI (Anthropic, 2025) was used during the analysis as a brainstorming tool. While the full analysis was conducted by the researcher, the tool was used to explore ideas related to theme naming and the structure of the findings.

Third, Claude AI was used to support the writing process in relation to grammar and suggestions on reformulation of written text sequences. The tool was thereby not used to generate new text, but only to correct or restructure text already formulated by the researcher.

5 Findings

In this section, the findings from the data analysis are presented across five different themes. Each theme illustrates aspects of how employees make sense of the early-stage implementation of agentic AI in practice. These themes are interpreted using the sensemaking mechanisms proposed by Iveroth and Hallencreutz (2015), to highlight how employees interpret and respond to the introduction of agentic AI in relation to the organizational implementation process.

5.1 Translating the technology through familiar tools

The analysis highlights that the respondents made sense of the implemented system by translating it through familiar tools they already knew, especially ChatGPT. Rather than describing the technology in terms of autonomy, multi-agent functionality or agentic capabilities, they interpreted it through existing reference points from everyday AI-use.

ChatGPT functioned in this sense as a descriptive shorthand that employees reached for to explain what the new system was. One respondent expressed expectations about the solution matching ChatGPT in functionality when asked to describe the implemented system, illustrating the role of familiar tools when describing the new tool and setting expectations:

> *“My expectations were probably that we would get something that functions just like ChatGPT… And I don’t know what those expectations came from.” – P1*

For other respondents, the implemented system was situated within a broader set of generative AI tools currently in use at the company. When asked to describe the solution, they did not refer exclusively to the implemented system, but instead broadened the frame by mentioning other models used in practice:

> *“The other things implemented are still a lot of free GPT or, I was about to say Cloud, but other LLM-models, not just ChatGPT. CoPilot and yeah. I was using that yesterday. There are still other models that we use. – P3*

The analysis also showed how ChatGPT was not only used for description, but also as a performance benchmark as respondents did not evaluate the system on its own terms, but against what they were currently using. Several respondents explain how ChatGPT was seen as the lowest threshold of performance, and that the new system should match what they already use:

> *"I think the minimum hygiene factor, the minimum requirement needs to be that it matches the solution I use today. Which in my case is kind of exclusively ChatGPT. So that was probably my minimum expectation. – P4*

This shows how employees used ChatGPT and other generative AI to translate functionalities and capabilities of the new system into something they already knew, and through that process made sense of the implemented system.

5.2 Using professional identity to navigate AI use

The analysis also illustrates how employees made sense of the system through their professional identity. Without shared organizational guidelines for how to use the system, employees built individual norms for what legitimate AI use looked like to them. These norms were shaped by concerns about what the technology meant for their critical thinking and credibility as consultants.

A first aspect of how professional identity shaped this norm-construction concerns the boundary between legitimate task delegation and problematic outsourcing of thinking. Several respondents expressed concerns that the convenience of turning to AI could gradually reduce their own cognitive capacity, since repeatedly reaching for an answer before thinking independently might over time weaken their ability to think for themselves:

> *"You don't even bother to activate your brain before you go to AI, can I just get the answer? Over time, everything may become the same. Where will the new elements come in? Maybe it will be more like you abandon your own thinking capacity" - P3*

One respondent even explicitly distances themselves from using AI without thinking when talking about the general risks related to excessive AI use:

> *"But yes, It's also kind of scary how many people could easily just stop thinking completely. I am not talking about myself here though, I have not stopped thinking completely." – P1*

At the same time, respondents drew a distinction between this kind of cognitive outsourcing and the deliberate delegation of tasks they considered low value, such as formatting and structuring presentation material to define own thoughts and ideas.

The analysis suggests that norms for AI use were not only constructed at an abstract level but also translated into individual rules for practice. Respondents described how their confidence in AI output was closely tied to their own expertise. In areas where respondents had expertise, they felt capable of evaluating and verifying what the AI system produced, while in areas where domain knowledge was limited, output became harder to evaluate:

> *"...Since it hallucinates at times, or gives you answers to things that aren't really true. That's what you often need, knowledge. You either need to ask for sources so you can double check the original source as well. Or you really need to have your own understanding to verify."- P5*

A further dimension of this process is shaped by the external pressures employees navigate as management consultants. Across interviews, AI competence was consistently framed not only as personally useful but as professionally necessary. Respondents described how AI competence was considered something consultants needed to demonstrate to maintain credibility and guide clients through their own AI processes:

> *"So we have to be on it here too. We must adapt our way of working to be able to educate customers on how to use AI in their workflows. So, they've been talking about it all the time. We're going to test, we're going to be at the forefront, we're going to work with this so that it's constantly becoming a natural part of our everyday work." – P2*

This points to a sensemaking process where professional identity played an important role to navigate what the system might mean for employees' role as consultants, and to figure out how to use the system to protect that role.

5.3 Anchoring meaning in future potential

Another theme that emerged from the analysis concerns how respondents made sense of the new system through its future potential rather than its current state. In this way, employees constructed meaning around the system through a forward-oriented perspective.

Several respondents described the current version of the system as incomplete but still meaningful, as it represented progress toward a more developed system. One participant noted that the solution already met expectations for a first step, suggesting that present limitations were interpreted as a part of a natural development process rather than absolute flaws:

> *"I think there are visions of what it will be able to do in the future. I think it meets the expectations that I had for a first step" – P5*

Beyond normalizing current limitations, respondents also constructed visions of what the system could eventually enable for the organization. Several respondents emphasized the potential value of accessing the organization's accumulated knowledge through the system. Here, the system was envisioned as a tool to connect internal project material and organizational expertise in ways that could enhance future work situations and eventually deliver organizational value:

> *"The idea is brilliant and I think it is the future. Being able to find our own knowledge bank, which is our structural capital, and combine it with other things. I think that is a huge asset for us, and I believe in that even more." – P4*

Although this vision of the system eventually outperforming current reality was widely shared, it was not fully consistent. One respondent stands out by claiming to not have used the system when asked, because of its basic performance, while also showing limited belief in the systems' ability to improve with time:

> *"Well nothing. And why is that? I think it is too rudimentary, too hard to navigate. I think the effort I put in does not match the results that I am looking for. And I have no reason to believe that it will change just because I give it another try." – P6*

The quote suggests that while most respondents valued the system through its future potential, that valuation was not automatic but dependent on beliefs that the technology would continue to improve over time.

This shows how beliefs of future potential functioned as an important resource in the sense-making process, allowing employees to understand an early-stage agentic system.

5.4 Understanding system safety

The analysis also highlights how safety played an important role in how employees made sense of the system, as well as how they came to use it. Although the system was consistently framed by the organization as safe and secure, there was no way for employees to independently verify these claims. Instead, employees constructed their understanding of safety around the system through what they had been told.

The analysis illustrates an organizational framing of the implemented system as secure, referring to the on-premises location and independence from third parties. However, respondents' everyday practices revealed a cautious stance to these security claims. Even when system security was acknowledged, several respondents described hesitancy about using it for customer specific documents, to avoid the risk of breaching contracts. In practice, safety was not simply accepted as a promise, but the system was still used with caution:

> *"Even if [THE SYSTEM] is supersafe etc, I still haven't felt that I want to feed it any information. Based on NDA's and such. And if someone in my team have done that, that's on them." – P1*

This caution was not limited to professional context and handling of customer data. Respondents also described resistance to using the system for personal queries, concerned that colleagues might have visibility into the conversations:

> *"For me personally, I'm cautious about, I don't want colleagues who are potentially in the system to see, they don't have anything to do with what I want to search for… Those kinds of questions, I'd rather send to GPT, because I pay for that myself" – P4*

A fundamental challenge underlying the understanding of the safety of the agentic system is shown in how employees had no way of independently verifying the security claims given by the organization. Sense of safety was therefore constructed through what had been communicated, which left the system difficult to act on with confidence, without specific competence:

> *"And from a safety perspective, I have no opinion at all. I assume that the solution we have implemented is much more secure both for us as an organization and our*

customers, in relation to if it is external. I am just basing this on what I am told, I have no answer for that myself." – P3

This shows how safety shaped employees' sensemaking of the system, where it was accepted in theory through trust in organizational safety claims, but not acted upon in practice.

5.5 Collective sensemaking through peer learning

Finally, the analysis shows how employees made sense of the system collectively, through peer discussions, shared experimentation and common team identity.

A first aspect of this collective sensemaking concerns how employees framed themselves as a team. Across interviews, respondents consistently described their group as self-going and technically capable. This collective identity was used to make an unstructured implementation process feel consistent with how they already work, rather than alarming:

"We are generally a very self-going group. So, I think the expectation was that we can just give this to you…So, I think [NAME OF THE COMPANY] is probably the easiest place to implement a new AI system. Or any AI for that matter." – P1

In the absence of structured learning opportunities from the organization, the analysis reveals how employees turned to each other to discuss best practices on AI use. Having normalized the limited plan around the implementation through team identity, employees then built their own learning collectively through word of mouth, peer experimentation and shared reflection. Although not explicitly connected to the new AI system, employees described how meaning was constructed collectively by learning about best practices together:

"We haven't really had any structured education. This is how you prompt in a good way? How have you done this to learn more? I prompt by myself. Then I go to a colleague and brainstorm ideas. This was a good. Then we use that together. You learn together" – P4

However, as the learning was initiated through individual employees, not everyone was involved in the process. Alongside statements of AI learning being driven by curiosity, experience and interest, the analysis reveals that not everyone was equally involved in this collective learning process:

"In any case, I'm not involved in those discussions. Maybe they do, but I don't know about them, or because I didn't sit in on it. But I don't really know if that's our… I don't really know if that's how we should do it…" – P3

This shows how colleagues and team identity played an important role in how employees made sense of the implemented system, where these aspects filled the gap left by the absence of organizational structure.

6 Discussion

The aim of this study has been to explore how employees within a management consultancy firm make sense of the introduction of an agentic AI system during early-stage organizational implementation, and how that shapes the engagement with the technology. The findings reveal that employee sensemaking in this early-stage implementation was largely individually initiated and shaped by existing resources such as prior AI usage, professional identity and peer relationships. The themes emerging from the findings are discussed in detail in the following section, to develop an answer to the research question.

A central insight of this study is that the autonomous and multi-agent characteristics of the agentic AI system were largely invisible to employees during the early stages of implementation, and that the system was not framed or described as "agentic". This invisibility does not appear random but reflects the inscrutability of AI systems which is in line with prior research by Berente et al. (2021). In other words, these underlying AI mechanisms are largely inaccessible to less technical users. In this specific case, and as emphasized by Acharya and Kuppan (2025), the high autonomy and goal-directed behavior that define the implemented system operate through multiple agents in the background, making them largely architectural and therefore not perceived directly by employees interacting with a familiar chat interface. Because of the technology itself not offering a specific frame for interpretation through interaction, employees reach for something accessible, which in this case was ChatGPT and other generative AI tools. These references consequently shape not only how employees understood the system, but also what they noticed during their initial interaction and what they expected the system to do.

Inscrutability further extends beyond system capability to safety. Employees struggle to independently verify the safety of the system, meaning that this sense of safety had to be borrowed from organizational claims. This produces a fragile outcome where employees acknowledge the system as secure, but that acknowledgement does not translate into confident use, as seen in how they avoid uploading confidential data and turn to public solutions for personal queries. This further extends the inscrutability argument as defined by Berente et al. (2021), from opacity related to understanding internal processes and verifying outputs, to employees' ability to verify the specifications that determine whether an AI system is safe to use in practice.

When faced with the interpretative challenges of AI in early-stage implementation, which Sagodi et al. (2023) argue are particularly complex and specific, employee sensemaking defaults to available resources such as familiar AI tools to understand what the system was, future visions to justify its current state, and delegated claims to navigate safety. This aligns with Weick´s (1995) argument that sensemaking is driven by plausibility rather than accuracy, where employees seek interpretations that are believable and coherent enough to act on, rather than objectively correct. The findings show that employees did exactly this and did it actively.

What is analytically interesting is thereby what the gap between a coherent and an accurate interpretation means for employees' sustained engagement. When sensemaking defaults to familiar tools as reference points and the resulting sense of the system reflects the logic of those tools rather than the logic of the system, it risks producing expectations that the implemented system cannot or was never intended to meet. Furthermore, while this result may enable action, it is not necessarily accurate or organizationally aligned.

Similarly, when future beliefs function as a resource to make sense of a current system, this sense is dependent on individual conviction rather than organizationally grounded visions, and those convictions might not be equally shared. As Goto (2022) argues, these early-stage projections shape individuals' willingness to engage with AI professionally, illustrating that the stability of these initial visions matters for organizations that aim towards sustainable AI adoption. This argument was further motivated in the findings, as limited beliefs in future improvement already led to disengagement in at least one case.

The same applies to the social process through which meaning was constructed. As sensemaking happens through interaction rather than in isolation (Weick, 1995), it is a process that should be equally available to everyone. The findings show that collective sensemaking did occur, as employees turned to colleagues to build understanding and normalized the implementation process through shared team identity. However, when this social process is left unorganized and self-initiated, it may not reach everyone equally and does not necessarily address the implemented system specifically, according to the findings. This highlights that although collective sensemaking plays an important role in how employees make sense of the implemented system, it cannot be assumed to emerge effectively on its own. Without organizational facilitation, it risks becoming uneven and unfocused, with potential consequences for organization-wide adoption.

A further dimension of this sensemaking process is specific to the context of implementing agentic AI in KIFs, grounded in how employees make sense of who they are in relation to the newly implemented system. Concerns around cognitive deskilling and a lack of unique knowledge production in relation to AI use, as suggested by the findings, illustrate a possible challenge to the identity of employees as management consultants. As knowledge-intensive work is inherently dependent on individual expertise and knowledge retrieved through professional experiences (Alvesson, 2004; Starbuck, 1992), professional identity in this context is deeply entangled with cognitive ability. To be a management consultant is to be someone who thinks independently and critically.

When agentic AI offers abilities to shortcut the exact cognitive tasks that largely define their professional self, employees are not faced only with adjusted workflows, but they are confronted with a question of who they are or will become in relation to their work. According to Weick (1995), identity construction is one characteristic of sensemaking where individuals ask themselves what certain events might mean to their own identity to make sense of those events. The findings show that the increase of AI use in daily work tasks implies an

experienced risk of deskilling in cognitive capabilities, which introduces a challenge to the professional self into the sensemaking process.

This tension is further sharpened by the three motives Weick (1995) identifies as central to individual sensemaking: the need for self-enhancement, the self-efficacy motive and self-consistency. On one hand, employees may feel motivated to use AI as an act of self-enhancement, demonstrating AI knowledge and staying ahead in a rapidly developing market. On the other hand, over-reliance on AI risks undermining self-efficacy derived from critical thinking. Lastly, the findings illustrate a challenge related to self-consistency, referring to employees' sense of remaining the knowledgeable and capable employees they have become. The findings suggest that employees navigated this conflict by individually shaping norms for what legitimate AI use looks like in practice, to potentially protect their professional identity. As these norms or guidelines for AI use were not organizationally facilitated, they may be fragile and pose a risk that may affect future employee engagement. This illustrates the multi-level nature of sensemaking as proposed by Maitlis and Christianson (2014), where industry pressure shapes micro-level identity concerns, and where individual negotiations of professional self-concept may carry consequences for organizational adoption outcomes.

To answer the research question, this study shows that employees make sense of agentic AI during early-stage implementation by drawing on available resources to construct interpretations they can understand and act on, rather than seeking accurate and complete understanding of the characteristics which define the implemented system. Furthermore, colleague relationships play an important role in the sensemaking process, as employees turn to each other to collectively build meaning around the system, even when such initiatives are not organizationally facilitated. Lastly, the process is shaped by the specific stakes of the KIF context, where employees are not only interpreting a new tool but negotiating their professional self in relation to a technology that challenges the cognitive capabilities their value is built on.

6.1 Practical implications

To understand what the findings implies for organizational leadership, they are discussed in relation to the sensemaking mechanisms proposed by Iveroth and Hallencreutz (2015), with the aim of deriving practical insights for leaders navigating the early stages of agentic AI implementation in knowledge-intensive firms. These implications are further informed by human-centered AI principles from Shneiderman (2022), which provide an additional lens for understanding and interpreting what a human-centered implementation could entail.

6.1.1 Translate

The findings illustrate how employees built their own understanding of the implemented system based on other generative AI tools such as ChatGPT, shaping expectations that the implemented system was not necessarily designed to meet. This suggests that limited shared organizational explanations or shared vocabulary around the new system had been established, leaving employees to construct their own frames of reference. This points to the value of active

translation as a leadership practice, as emphasized through the *translate* mechanism proposed by Iveroth and Hallencreutz (2015). By drawing on employees' established AI experiences with other tools, leaders could build a shared vocabulary that clarifies what the new system is, how it differs from familiar tools and what it uniquely offers. Thereby, leaders may utilize the potential AI literacy at the organization to reduce some of the complexity around the implemented agentic system, by translating it into terms employees are familiar with.

From a human-centered perspective, this translation is an ongoing practice rather than a one-time communication. As Shneiderman (2022) emphasizes, human-centered AI requires systems to be interpretable, and to preserve human control even in highly autonomous systems. Actively building this understanding through ongoing translation in familiar terms therefore becomes a way for leaders to preserve human control and interpretability in practice.

6.1.2 Converse candidly

The findings suggest that safety was communicated as a key objective at the start of the implementation process, but was not further justified or revisited, leaving employees to individually make sense of a system they could not independently verify as safe. This points to the value of ongoing and candid communication in leadership, as emphasized through the *converse candidly* mechanism proposed by Iveroth and Hallencreutz (2015). Rather than assuming initial safety claims are sufficient, leaders could create ongoing space for dialogue around system safety, allowing employees to surface doubts and develop a genuine rather than delegated sense of safety. In relation to the findings, this could allow employees to unlock the full potential of the system and potentially increase overall confidence in use patterns.

From a human-centered perspective, Shneiderman (2022) places an active responsibility on leaders to develop sufficient knowledge of how safety has been achieved within a system and to communicate this actively to users, to facilitate a safe culture around AI. This means not only committing to safety as a principle but actively engaging with the technical basis of that safety, creating room for sensemaking of those claims through credible conversations.

6.1.3 Look closely and update

The findings show that employees were individually constructing norms for AI use, navigating a tension between developing AI competence and preserving their ability to think independently. Since the findings suggest that these norms were individually constructed with limited organizational input, they risk being inconsistent across the organization. This points to the value of close observation as a leadership practice, as emphasized through *the look closely and update* mechanism proposed by Iveroth and Hallencreutz (2015). By actively observing how the system is landing in the organization, including worries and individual norms emerging as a result, leaders could let those observations guide the development of shared guidelines for sustainable use of the system. These guidelines could thereby address the potential concerns observed, while being tailored to specific organizational dynamics.

According to Shneiderman (2022), preserving individual goals by supporting self-efficacy and creativity are core principles of human-centered AI. The individual norm construction observed in this study reflects employees actively attempting to preserve those individual attributes. This implies that from a human-centered perspective, leaders should translate these principles into organizational practices, creating shared guidelines that support preservation of the professional self, rather than leaving individuals to meet those questions alone.

6.1.4 Provide a direction

The findings show that visions related to specific functionalities and potential capabilities had been communicated around the implemented system, which employees used to legitimize current limitations by anchoring engagement in future potential. However, as shown by at least one employee who had already disengaged, this orientation was fragile and dependent on individual belief rather than an organizationally grounded understanding. This points to the value of providing a clear and continuously reassessed direction as a leadership practice, as emphasized through the *provide a direction* mechanism proposed by Iveroth and Hallencreutz (2015). Rather than communicating a vision once and assuming it holds, leaders could reassess that direction to ensure it remains aligned with what employees are actually experiencing.

Shneiderman (2022) argues that AI implementation should actively support individual goals of self-efficacy and personal development and thereby extend beyond organizational goals such as revenue growth and cost reduction. Given that the findings show how future-oriented beliefs played an important role in sustaining engagement with the system, the direction leaders communicate and how it resonates personally with employees may carry real consequences for continued engagement. From a human-centered perspective, this suggests that leaders should incorporate individual perspectives when communicating an organizational vision, rather than framing it solely in terms of organizational benefit.

6.1.5 Facilitate learning

The findings show that employees facilitated their own learning both independently and through peer dialogue, but this process was unevenly distributed and largely oriented around general AI use rather than the specific capabilities of the implemented system. This points to the value of structured and inclusive learning as a leadership practice, as emphasized through the *facilitate learning* mechanism proposed by Iveroth and Hallencreutz (2015). By directing learning initiatives specifically toward the implemented system and ensuring they reach everyone in the organization, leaders could build a more consistent and shared understanding rather than leaving it to emerge unevenly through individual curiosity and specific groups.

From a human-centered perspective, Shneiderman (2022) emphasizes that AI systems should be interpretable and transparent to their users. Facilitating learning specifically toward the implemented system thereby becomes a concrete way for leaders to build that interpretability in practice, utilizing the collective sensemaking already happening within the organization.

Table 3. Summary of the practical implications in relation to findings.

Mechanism	Theme in findings	Leadership implication
Translate	Translating the technology through familiar tools	Leaders can decode complexity of agentic AI systems by understanding and using current AI use at the organization as a frame for translation.
Converse candidly	Understanding system safety	Leaders play an important role in providing space for surfacing uncertainties around safe use of agentic systems, which can help employees understand and act accordingly.
Look closely and update	Using professional identity to navigate AI use	Leaders can create organizational guidelines for good AI use in practice, by picking up on discussions and individual anxieties.
Provide a direction	Anchoring meaning in future potential	Leaders can provide organizationally aligned visions of AI implementation by including individual perspectives and ensuring everyone is moving in the same direction.
Facilitate learning	Collective sensemaking through peer learning	Leaders can facilitate collective learning initiatives specifically designed to the implemented agentic system, and make sure these initiatives reaches everyone equally.

6.2 Limitations

The limitations of this study relates primarily to the sample size, specifics of the investigated case and the methodological challenges in translating individual sensemaking experience into broader organizational implications.

Fifteen employees within the studied organization were contacted and invited to participate, resulting in a final sample of six interview participants. Given the size of the organization, this represents approximately one third of all employees. Hence, the findings should be interpreted as indicative rather than exhaustive, as it reflects individual experiences within a larger organizational context.

Additionally, the study was conducted within a single knowledge-intensive organization implementing a specific agentic system, and the ability to generalize the findings to other organizational and technological contexts can thereby be questioned. The architectural and capability-related features of agentic systems entail that such technologies can differ considerably in functionality and scope across cases. Direct interaction through interface design, as well as the organizational framing surrounding a system, can arguably represent a greater influence

on sensemaking than the system's underlying technical architecture. A system with identical capabilities but a different interface or framing may therefore have been interpreted differently from the system examined in this study. Similarly, knowledge-intensive organizations, and management consultancy firms in particular, vary in structure and size in ways that arguably shape sensemaking processes in relation to organizational leadership. Nevertheless, a single case study design was chosen precisely in response to this variation across both technology and organizational context.

A fundamental challenge in this study concerned the transition between individual sensemaking experiences and organizational leadership implications. This transition was motivated by the goal of translating employee perspectives into practical guidelines, resulting in a research question focused on how employees made sense of the system, while practical implications were directed towards leaders. The sensemaking mechanisms by Iveroth and Hallencreutz (2015) were chosen as a theoretical framework to bridge these two levels, although they were originally designed as leadership tools rather than a lens for analyzing individual experiences. This meant that the findings related to leadership practices were based on how employees perceived and described the implementation process, rather than on direct accounts from leaders. Whether leadership practices were absent or simply not visible to employees can therefore not be confirmed by the data, and the leadership implications of this study rest on an interpretive step based on what employees described in relation to organizational practices. Furthermore, moving back and forth between individual and organizational levels also introduced a degree of abstraction that may have affected the clarity and focus of the research question itself. While the research question centered on individual sensemaking, the analysis was also interpreted through an organizational leadership lens, which at times pulled in different directions.

7 Conclusion

Despite investing heavily in AI, organizations continue to experience the gap between AI initiatives and organization-wide adoption. This study offers a glimpse into how employees receiving such initiatives make sense of the implemented technology, and how these interpretations may shape their continued engagement with the system.

By exploring how employees within management consultancy made sense of the early-stage implementation of agentic AI, the study found that employees made sense of the system by drawing on available resources – translating it through familiar AI tools, anchoring meaning in future potential and trusting organizational safety claims they could not verify. Each of these resources shaped engagement in different ways, as familiar tools set expectations the system was not designed to meet, future potential was dependent on individual conviction and safety claims were accepted but not acted upon with confidence. The findings also illustrate the social aspect of this process, as employees turned to colleagues to construct meaning collectively, though this process was not specifically targeting the implemented agentic system and did not involve everyone equally. Lastly, in the specific context of management consultancy, professional identity emerged as a central aspect of this process. Employees were not only making sense of what the new tool was, but also what a technology that challenges the cognitive abilities that their professional value is built on might mean for their professional self. This influenced engagement through the need to individually develop norms for AI use which answered to the external pressure of showing AI competence while preserving the ability to think individually.

The study makes three contributions, primarily through the integration and application of existing theoretical perspectives within a novel empirical context. First, it responds to an identified need for research on agentic AI and individual-level behavior in AI adoption research, offering new insights into how employees make sense of agentic AI during the early-stage implementation. Second, it offers a practically oriented starting point for leaders seeking to facilitate agentic AI implementation with a human-centered approach, grounded in employee perspectives from an underexplored context. Lastly, it surfaces the stakes specific to agentic AI within knowledge-intensive work, showing the influence of professional identity in the sensemaking process when agentic AI is introduced.

As the interest in successful AI adoption grows, future research should focus on investigating these dynamics at larger scale, with more developed agentic systems. Specifically, more attention should be given to investigate how these early-stage sensemaking processes related to professional identity shape long-term adoption, since it surfaces a dynamic where agentic AI risks challenging individual values that are specifically emphasized in human-centered AI literature. Through this, research can continue to show how implementing AI with human values and needs at its core should be of interest to everyone seeking sustainable AI adoption.

8 References

Acharya, D. B., Kuppan, K., & Divya, B. (2025). Agentic AI: Autonomous intelligence for complex goals—A comprehensive survey. *IEEe Access*, 13, 18912-18936. 10.1109/ACCESS.2025.3532853

Alvesson, M. (2001). Knowledge work: Ambiguity, image and identity. *Human Relations,* 54(7), 863-886. https://doi.org/10.1177/0018726701547004

Alvesson, M. (2004). Knowledge work and knowledge-intensive firms. OUP Oxford.

Anthropic. (2025). Claude [Large language model]. https://claude.com/product/overview

Berente, N., Gu, B., Recker, J., & Santhanam, R. (2021). Managing artificial intelligence. *MIS quarterly*, *45*(3), 1433-1450. https://doi.org/10.25300/MISQ/2021/16274

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.

Challapally, A., Pease, C., Raskar, R., & Chari, P. (2025). *The GenAI divide: State of AI in business 2025*. MIT Project NANDA / MLQ. https://mlq.ai/media/quarterly_decks/v0.1_State_of_AI_in_Business_2025_Report.pdf

Cubric, M. (2020). Drivers, barriers and social considerations for AI adoption in business and management: A tertiary study. *Technology in Society*, *62*, 101257. https://doi.org/10.1016/j.techsoc.2020.101257

Denzin, N. K., & Lincoln, Y. S. (Eds.). (2011). *The Sage handbook of qualitative research*. Sage.

Drucker, P. F. (1999). Knowledge-worker productivity: The biggest challenge. *California Management Review*, *41*(2), 79–94.

Engström, A., Pittino, D., Mohlin, A., Johansson, A., & Edh Mirzaei, N. (2024). Artificial intelligence and work transformations: integrating sensemaking and workplace learning perspectives. *Information Technology & People*, *37*(7), 2441-2461. https://doi.org/10.1108/ITP-01-2023-0048

Gioia, D. A., Thomas, J. B., Clark, S. M., & Chittipeddi, K. (1994). Symbolism and strategic change in academia: The dynamics of sensemaking and influence. *Organization Science*, 5(3), 363–383. https://doi.org/10.1287/orsc.5.3.363

Goto, M. (2022). Accepting the future as ever-changing: Professionals' sensemaking about artificial intelligence. *Journal of Professions and Organization*, *9*(1), 77-99. https://doi.org/10.1093/jpo/joab022

Gottumukkala, S., & Prasad, S. S. (2025). Decoding the Organisational AI Adoption: What Do Theories and Frameworks Reveal?. *Journal of International Commercial Law and Technology*, 6, 744-754. https://doi.org/10.61336/Jiclt/26-01-110

Hennink, M., Hutter, I., & Bailey, A. (2020). *Qualitative research methods* (2nd ed.). SAGE Publications.

Henry, K. E., Kornfield, R., Sridharan, A., Linton, R. C., Groh, C., Wang, T., ... & Saria, S. (2022). Human–machine teaming is key to AI adoption: clinicians' experiences with a deployed machine learning system. *NPJ digital medicine,* 5(1), 97. https://doi.org/10.1038/s41746-022-00597-7

Holmström, J. (2022). From AI to digital transformation: The AI readiness framework. *Business Horizons*, 65(3), 329-339. https://doi.org/10.1016/j.bushor.2021.03.006

Hosseini, S., & Seilani, H. (2025). The role of agentic ai in shaping a smart future: A systematic review. *Array*, 26, 100399. https://doi.org/10.1016/j.array.2025.100399

IBM. (2024, May 16). IBM study: As CEOs race towards Gen AI adoption, questions around workforce and culture persist. IBM Newsroom.https://newsroom.ibm.com/2024-05-16-IBM-Study-As-CEOs-Race-Towards-Gen-AI-Adoption,-Questions-Around-Workforce-and-Culture-Persist

Iveroth, E., & Hallencreutz, J. (2016). Effective organizational change: leading through sensemaking. Routledge.

Kane, G. C. (2019). The technology fallacy: People are the real key to digital transformation. *Research Technology Management,* 62(6), 44–49. https://doi.org/10.1080/08956308.2019.1661079

Maitlis, S., & Christianson, M. (2014). Sensemaking in organizations: Taking stock and moving forward. *Academy of management annals*, *8*(1), 57-125. https://doi.org/10.5465/19416520.2014.873177

Sagodi, A., Dremel, C., & van Giffen, B. (2023). Sensemaking in AI-based digital innovations: insights from a manufacturing case study, *ECIS 2023 Research Papers*, 390, available at: https://aisel.aisnet.org/ecis2023_rp/390

Sandberg, J., & Tsoukas, H. (2015). Making sense of the sensemaking perspective: Its constituents, limitations, and opportunities for further development. *Journal of organizational behavior*, *36*(S1), S6-S32. https://doi.org/10.1002/job.1937

Sapkota, R., Roumeliotis, K. I., & Karkee, M. (2025). Ai agents vs. agentic ai: A conceptual taxonomy, applications and challenges. *Information Fusion*, 103599. https://doi.org/10.1016/j.inffus.2025.103599

Shneiderman, B. (2022). *Human-centered AI*. Oxford University Press.

Singla, A., Sukharevsky, A., Yee, L., Chui, M., & Hall, B. (2024, May 30). *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value*. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-2024

Singla, A., Sukharevsky, A., Yee, L., Chui, M., Hall, B., & Balakrishnan, T. (2025, November 5). *The state of AI in 2025: Agents, innovation, and transformation*. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Starbuck, W. H. (1992). Learning by knowledge-intensive firms. *Journal of Management Studies*, *29*(6), 713–740.

Swedish Research Council. (2024*). Good research practice* (New rev ed.). Swedish Research Council.

Thalamind. (n.d.). *Thalamind.* https://www.thalamind.com

Trenerry, B., Chng, S., Wang, Y., Suhaila, Z. S., Lim, S. S., Lu, H. Y., & Oh, P. H. (2021). Preparing workplaces for digital transformation: An integrative review and framework of multi-level factors. *Frontiers in psychology*, *12*, 620766. https://doi.org/10.3389/fpsyg.2021.620766

Tursunbayeva, A., & Gal, H. C. B. (2024). Adoption of artificial intelligence: A TOP framework-based checklist for digital leaders. *Business Horizons*, 67(4), 357-368. https://doi.org/10.1016/j.bushor.2024.04.006

Weick, K. E. (1995). *Sensemaking in organizations* (Vol. 3). Sage.

Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. (2005). Organizing and the process of sensemaking. *Organization science*, *16*(4), 409-421. https://doi.org/10.1287/orsc.1050.0103

Westenberger, J., Schuler, K., & Schlegel, D. (2022). Failure of AI projects: understanding the critical factors. *Procedia computer science*, *196*, 69-76. https://doi.org/10.1016/j.procs.2021.11.074