

When LLM Governance Falls to Employees: Informal Practices of Responsible AI Use at Work

Tarini Saka^{1*}, Sandy Hong², Veelasha Moonsamy²

¹ Max Planck Institute for Security and Privacy

² Faculty of Computer Science, Ruhr University Bochum

tarini.saka@mpi-sp.org, sandy.hong@ruhr-uni-bochum.de, veelasha.moonsamy@ruhr-uni-bochum.de

Abstract

Large Language Model (LLM) tools are rapidly entering workplaces, raising concerns around accountability, privacy, security, and responsible use. While organizations increasingly respond through appropriate usage policies, training, and governance structures, less is known about how employees experience these efforts in everyday work. This paper examines the gap between formal organizational guidance and the practical realities of workplace LLM use. Drawing on semi-structured interviews with 20 employees across diverse industries in Germany, we investigate how employees use LLMs at work, how they perceive organizational readiness and support, and how they manage risks when policies or training are absent, unclear, or difficult to apply. Our findings show that responsible LLM use is often governed informally by employees themselves. Participants relied on self-directed learning, peer exchange, personal judgment, prompt anonymization, output checking, selective disclosure, and ad hoc decisions about what information was sensitive. Although these practices helped employees address various concerns related to LLM use, they also resulted in inconsistent approaches shaped by individual confidence, expertise, and workplace culture. We argue that responsible adoption requires moving beyond formal policies toward situated governance that combines clear, actionable guidance with role-specific training, technical safeguards, open communication, and employee participation in governance design.

1 Introduction

Large Language Model (LLM) tools, such as OpenAI's ChatGPT and GitHub Copilot, are increasingly becoming part of everyday work. Employees use these tools to automate tasks and support writing, coding, research, communication, and decision-making across a range of workplace contexts (Ferino et al. 2025; Balayn et al. 2025). According to the 2024 Work Trend Index Annual Report, 75% of knowledge workers already use AI at work (Microsoft and LinkedIn 2024), but 60% of leaders report that their organization lacks a clear plan for AI implementation (Microsoft and LinkedIn 2024). This suggests a growing gap between LLM adoption and organizational preparedness.

The general-purpose nature of LLMs makes their workplace adoption a distinctive governance challenge. Unlike workplace technologies introduced for a specific department, workflow, or expert group, LLMs can be used by employees across many tasks, roles, and organizational contexts (Balayn et al. 2025; Wagman, Dearing, and Chetty 2025). Some risks associated with LLM use, such as supply-chain vulnerabilities, are primarily technical and require engineering defenses (OWASP 2025; Sebastian 2024). Other risks are use-centric: they emerge from how employees prompt, interpret, and rely on LLM outputs, and therefore require employee-facing governance measures such as education, training, and clear guidance (Prümmer, van Steen, and van den Berg 2024; Grill et al. 2025). When such support is limited or difficult to operationalize, responsibility may shift toward practitioners, who rely on informal learning and situated judgment to interpret and enact responsible-use expectations (Gagua 2025; Madaio et al. 2024).

Responsible AI governance should not just be a static checklist of principles, but a dynamic process embedded in institutional, organizational, and sociotechnical contexts (Gagua et al. 2025). Accordingly, we define *workplace LLM governance* as the formal and informal organizational arrangements through which LLM use is guided, supported, monitored, and held accountable. Such governance may include acceptable-use policies, practical guidelines, approved tools, training programs, technical safeguards, oversight teams, audit logs, responsibility assignments, support contacts, and protocols for responding to failures or misuse (Balayn et al. 2025; Tavasoli, Sharbaf, and Madani 2025). These mechanisms address what employees may or may not do, how responsibilities are distributed, how employees learn to recognize and manage risks, and what happens when mistakes or uncertain situations arise (Jeong et al. 2026; Zhang and Lee 2025).

Organizations have adopted various measures to address LLM-related risks, ranging from policies and approved tools to training programs and outright bans on applications such as ChatGPT (Cardon et al. 2023; Tavasoli, Sharbaf, and Madani 2025; Jöhnk, Weißert, and Wyrski 2021). However, the existence of such measures does not necessarily ensure responsible use. Organizational support remains uneven, and employees often receive limited guidance on how to manage the security, privacy, and context-specific risks

*Work done while at Ruhr University Bochum.

Copyright © 2026, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

they encounter in practice (Kimbel, Glas, and Pernul 2024; Li et al. 2024; Zhang and Lee 2025). More broadly, responsible AI research shows that translating formal governance into everyday work is difficult and can result in fragmented implementation, improvised accountability, informal learning, and communication gaps (Rakova et al. 2021; Jafari Meimandi et al. 2025; Gagua 2025; Madaio et al. 2024; Fischer et al. 2025). Yet existing research has focused primarily on the governance mechanisms organizations should provide, rather than on how employees experience and interpret them. Building on workplace studies of AI policies, training needs, and context-dependent risks (Jeong et al. 2026; Zhang and Lee 2025; Wagman, Dearing, and Chetty 2025), we therefore examine whether employees are aware of organizational guidance, find it useful and trustworthy, and can apply it when rules are unclear or workplace norms remain informal. This addresses an important gap in understanding how formal LLM governance translates, or fails to translate, into everyday practice.

To address this gap, we conducted semi-structured qualitative interviews with employees from diverse industries in Germany ($n = 20$) to examine how these forms of governance are received, interpreted, and put into practice by employees. Specifically, we study how employees use LLMs in everyday work, how they perceive organizational preparedness, and how they navigate unclear, missing, or difficult-to-apply guidance. By foregrounding their perspectives, we address the following research questions:

- RQ1** *How do employees incorporate LLMs into everyday work, and how do they navigate the practical, social, and risk-related conditions of responsible use?*
- RQ2** *What organizational governance mechanisms currently exist to support responsible LLM use, and what challenges do employees perceive in their implementation?*
- RQ3** *What organizational support do employees consider necessary to enable more responsible and confident LLM adoption?*

Our findings show that LLM use is already embedded in everyday work across diverse roles, often through a mix of official tools, personal accounts, and informal experimentation. However, organizational governance remains uneven: policies and training are absent or unclear in many workplaces, and even when they exist, employees often perceive them as too general, inaccessible, or insufficiently connected to daily work. In response, employees develop *informal governance practices* around privacy, security, reliability, disclosure, and professional legitimacy. These practices include self-directed learning, peer exchange, personal judgment, prompt anonymization, output checking, selective disclosure, and ad hoc decisions about what information is safe to enter into LLM tools. Our findings further show that organizational culture, including open communication, leadership support, and the normalization of responsible LLM use, shapes employees' confidence and willingness to discuss risks. Based on employees' recommendations, we argue that responsible LLM adoption requires moving beyond policies on paper toward actionable guidance, continuous and

role-specific training, technical safeguards, and participatory governance processes that incorporate employee feedback.

2 Background

LLM tools are increasingly becoming part of everyday work. Unlike specialist workplace software designed for a specific task or user group, LLMs can be used by employees across many roles and levels of technical expertise. Developers may use them for coding, researchers for information synthesis, and employees in legal, finance, or administrative roles for drafting, summarization, translation, and routine communication (Shan and Shan 2024; Chkirbene et al. 2024; Ferino et al. 2025; Liao et al. 2024). Their variety therefore concerns both *who* uses them and *how* they are used. This broad user-task landscape makes organizational adoption more difficult than introducing a specialized tool with a clearly defined audience. Organizations must provide training and guidance that account for different roles, skills, use cases, and levels of risk, while also managing access, privacy, compliance, and acceptable use (Uren and Edwards 2023; Tadimalla and Maher 2024; Diro et al. 2025).

Risks of Workplace LLM Use

This general-purpose nature of LLMs makes workplace governance challenging. Responsible workplace use requires attention to privacy, security, reliability, and ethical risks. Some risks are primarily technical, including prompt injection, sensitive information disclosure, and supply-chain vulnerabilities (OWASP 2025). Others arise through everyday use: employees may enter confidential business information, customer data, proprietary code, or internal documents into external tools, creating risks of data leakage or inappropriate third-party processing (Vaishnav, Singh, and Cornell 2024; Yan et al. 2024). LLM outputs can also be incorrect, fabricated, biased, or overly persuasive, making verification and human judgment essential (Sood, Zeadally, and Hong 2025; Diro et al. 2025). However, prior work shows that users do not always respond appropriately to erroneous or uncertain AI advice: such failures can miscalibrate trust and reliance, lead users to follow poor recommendations, and ultimately impair decision quality or task performance (Saka et al. 2025; Klingbeil, Grützner, and Schreck 2024; Kim, Yang, and Zhang 2023).

These risks are difficult to govern through technical safeguards alone. Emerging AI-specific regulations emphasize accountability, transparency, human oversight, and protection of rights (Pathan and Shah 2025; Radanliev et al. 2024), but they do not determine how employees actually use LLMs in daily work. Because many risks emerge at the point of use, responsible adoption also depends on practical employee-facing support: employees need to understand what information should not be entered into tools, how to evaluate outputs, when to avoid LLM use, and how to respond to uncertainty or misuse. Thus, workplace LLM governance must combine technical safeguards with clear guidance, training, and support that employees can apply in everyday practice.

LLM Governance Mechanisms

Recent work has proposed governance mechanisms such as policies, guidelines, training programs, approved tools, technical controls, oversight structures, and, in some cases, restrictions or bans on public generative AI tools (Cardon et al. 2023; Tavasoli, Sharbaf, and Madani 2025). Responsible AI governance may include assigning roles and responsibilities, maintaining audit logs, defining approved tools, specifying what data may be used, and establishing procedures for failures or misuse (Tavasoli, Sharbaf, and Madani 2025). Organizational AI readiness research similarly emphasizes upskilling, ethical awareness, and governance as prerequisites for effective adoption (Jöhnk, Weißert, and Wyrski 2021).

Li et al. study AI tool adoption in software development and find that organizational rules, training, peer support, and perceived risk shape whether individuals and teams adopt AI tools (Li et al. 2024). They found that employees report greater willingness to use AI tools when such provisions exist. Such support may include instructions on approved tools, data-handling rules, prompting practices, known limitations, bias, and ethical issues (Li et al. 2024; Estaiteyeh and McQuirter 2024). Yet existing evidence also suggests that concrete guidance is often missing or uneven. Kimbel et al. provide closely related interview-based evidence on workplace ChatGPT use, showing that employees recognize security and privacy risks but that few organizations provide specific guidance (Kimbel, Glas, and Pernul 2024). While they focus on risk perceptions and self-protective practices, we examine multiple LLM tools and a broader governance context, including organizational rollout, policies, training, output reliability, openness about use, and employee recommendations. Together these studies suggest that employees often want clearer support, while organizational guidance remains incomplete or uneven.

This points to a broader challenge in responsible AI: *governance mechanisms do not automatically translate into practice*. Responsible AI principles often remain difficult to operationalize, requiring changes to incentives, processes and everyday decision-making (Rakova et al. 2021), and can become fragmented across organizational units and local contexts (Jafari Meimandi et al. 2025). Gagua presents preliminary findings on how accountability is improvised around two purpose-built AI systems within a single public-sector agency (Gagua 2025). We broaden this perspective by examining employees across organizations who use general-purpose LLMs in varied everyday work contexts, allowing us to explore whether similar governance dynamics arise across a wider and less bounded setting.

An essential research gap is understanding how workplace LLM governance operates in practice. Examining whether employees are aware of organizational provisions, how they interpret and use them, and how they respond when formal guidance is incomplete can reveal the gap between governance as designed and governance as experienced. Our study addresses this gap by examining employees as recipients and everyday enactors of workplace LLM governance.

3 Methodology

To gain an in-depth understanding of employees' experiences with organizational LLM adoption, we adopted a qualitative research design (Myers and Newman 2007; Creswell and Poth 2016). Semi-structured interviews were appropriate for eliciting detailed accounts of how employees use LLM tools, how they perceive organizational support, and how they navigate privacy, security, and ethical concerns in everyday work.

Interview Guide

The interview guide was developed iteratively based on prior literature and industry reports on LLM use, organizational policies, training, and employee experiences. The final guide included 24 main questions organized into five thematic sections: *Usage of LLM Tools in Daily Work*, *Policies & Training*, *Privacy & Security Risk Awareness*, *Perceived Preparedness for Responsible Use*, and *Suggestions for Improvement*. These sections were designed to elicit participants' experiences with LLM tools, their awareness of associated risks, their perceptions of organizational preparedness, and their recommendations for improving support structures. Follow-up questions were used where appropriate to clarify responses and explore emerging topics in greater depth. To accommodate participants' language preferences, the interview guide was prepared in both English and German, and interviews were conducted in the participant's chosen language. Before the main data collection, we piloted the interview process with two participants, one in English and one in German. Based on the pilot interviews, we reworded unclear questions, removed redundancies, and added follow-up prompts to support more detailed responses. The English version of the full interview guide is provided in the supplementary material.

Participants and Recruitment

Participants were eligible if they were at least 18 years old, worked in Germany, and were current or former users of LLM tools in a workplace context. Recruitment took place between July and August 2025. We distributed an invitation email via companies, universities, and professional networks, and encouraged recipients to forward the call within their organizations. In addition, we shared the call via LinkedIn and personal and institutional networks. Interested individuals registered through an online sign-up form, where they consented to be contacted. They subsequently completed a pre-interview survey collecting demographic information and background on their LLM use, and provided informed consent. Interviews were scheduled via Calendly and conducted remotely via Zoom.

Data Collection and Piloting

Data collection involved a two-step recruitment and screening process. Interested participants first completed a sign-up survey that provided study information, collected contact details and preferred interview language, and assessed eligibility. Eligible participants were then invited to complete a

pre-interview survey, which included informed consent, demographic information, professional background, and current experience with LLM tools. This process supported participant screening, scheduling, and communication in either English or German.

In total, 20 semi-structured interviews were conducted (17 in German, 3 in English). All interviews were conducted remotely via Zoom and lasted approximately 30–45 minutes (mean = 31:40 minutes). Participants selected an interview time through an online scheduling tool and received a meeting link by email. Interviews were conducted in the participant's preferred language. With participants' consent, interviews were recorded, securely stored, transcribed, pseudonymized, and anonymized prior to analysis.

Before the main data collection, we piloted the interview guide and study procedure with two participants, one in English and one in German. Based on the pilots, we clarified wording, removed redundant questions, and added follow-up prompts to support deeper responses. The revised guide was reviewed and approved by the supervisors before recruitment began.

Data Analysis

We analyzed the interview data using thematic analysis following Braun and Clarke's six-stage framework (Braun and Clarke 2021; Ahmed et al. 2025). The analysis involved familiarization with the data, generating initial codes, reviewing and refining themes, defining and naming themes, and selecting illustrative quotations for reporting. This approach allowed us to identify recurring patterns in employees' experiences while remaining attentive to variation across organizational contexts.

All interviews were transcribed using noScribe¹, a local transcription tool, and all transcripts were manually checked and corrected. German transcripts were first thoroughly anonymized and then translated into English using DeepL². The transcripts were then coded in MAXQDA³. Coding began with four randomly selected interviews, which were coded iteratively to develop an initial codebook. This preliminary codebook was refined as additional interviews were analyzed, with new codes added when relevant and overlapping codes merged during regular reviews. The coding process combined deductive categories informed by the interview guide with inductive codes that emerged from the data.

As analysis progressed, themes were organized around the three research questions, including employees' LLM use and perceived preparedness, organizational policies and training, privacy and security risk awareness, and recommendations for improvement. Around the sixteenth interview, no substantially new codes or themes emerged, indicating thematic saturation (Hennink and Kaiser 2022). All interviews were coded by a single researcher; to support consistency and reduce interpretive bias, the codebook and emerging themes were discussed regularly with the supervising researchers, and transcripts were reviewed iteratively

before finalizing the analysis. The full codebook is shared in the supplementary material.

Study Limitations

Despite efforts to ensure methodological rigor throughout the study design, data collection, and analysis, some limitations remain. First, the sample includes 20 participants. Although no substantially new themes emerged toward the end of the analysis, indicating thematic saturation within this sample (Hennink and Kaiser 2022), the study cannot represent the full range of workplace roles, industries, and organizational settings. Second, all participants worked in Germany, where regulatory expectations, data-protection norms, employee protections, and workplace cultures may have shaped their experiences. The findings may therefore not transfer directly to other national or institutional contexts. Third, the study relies on self-reported experiences rather than observed behavior, making the findings vulnerable to recall and social-desirability biases, especially for sensitive topics such as hidden LLM use, policy violations, and incident handling. Finally, most interviews were conducted in German and translated into English. Although transcripts and translations were carefully reviewed, some linguistic or cultural nuance may have been lost (Verma 2024).

4 Study Findings

This section presents participants' demographic summary and our interview findings, organized by the three research questions.

Demographic Summary

Table 1 summarizes the demographic composition of the 20 participants. The sample included employees across different age groups, levels of work experience, and organizational sizes. Participants were predominantly women (70%), with most aged between 20 and 39. Their experience in their current roles ranged from less than one year to more than 15 years, and they worked in organizations ranging from small companies to large organizations with more than 5,000 employees. The sample included technical, managerial, administrative, educational, compliance, and public-sector roles, enabling us to capture diverse workplace perspectives on LLM adoption. A detailed participant-level overview, including professional roles and self-reported LLM use, experience, and confidence, is provided in the supplementary material.

RQ1: How do employees appropriate LLMs in everyday work?

Our findings show that employees' use of LLMs in the workplace cannot be understood only through the types of tasks for which they use these tools, but rather through how employees incorporate LLMs within their everyday organizational contexts. Our findings show that workplace LLM use was shaped by three interrelated processes. First, employees often developed their readiness through self-directed

¹<https://github.com/kaixxx/noScribe>

²<https://www.deepl.com/en>

³<https://www.maxqda.com/>

Gender	
Woman	14 (70%)
Man	6 (30%)
Age Group [Years]	
20–29	11 (55%)
30–39	4 (20%)
40–49	3 (15%)
50–59	2 (10%)
Experience in Current Job [Years]	
Less than 1	2 (10%)
1–3	9 (45%)
4–7	6 (30%)
8–15	2 (10%)
More than 15	1 (5%)
Company/Organization Size [Employees]	
1–10	1 (5%)
11–50	1 (5%)
51–200	4 (20%)
201–500	3 (15%)
501–1000	0 (0%)
1001–5000	5 (25%)
More than 5000	6 (30%)

Table 1: Demographic Summary of the 20 Participants

and informal learning. Second, LLM use was socially negotiated: participants described supportive, skeptical, or ambiguous workplace cultures, and some hesitated to disclose their use because they feared negative judgments. Third, some employees engaged in everyday risk mitigation practices, such as anonymizing prompts, checking outputs, relying on domain expertise, and discussing uncertain cases with colleagues. Together, these findings suggest that employees were not merely adopting a new tool, but actively constructing the conditions under which LLM use could be experienced as competent, legitimate, and safe.

LLM use as employee-led appropriation: Participants reported different patterns regarding their access to and use of LLMs. Many participants used official company accounts, while just as many used personal accounts. A few accessed LLMs via free or open web platforms; some used both official and personal accounts, depending on the tool; and one participant used a combination of personal and fake accounts. Adoption was not always initiated by the organization. About half of the participants reported that they had adopted their first LLM independently, while others encountered LLMs through company-provided tools or experienced a mixture of self-initiated and organizational adoption.

Participants used LLMs for a wide range of tasks, including research, coding, debugging, summarization, writing, structuring information, generating ideas, creating tables, seeking explanations, and supporting administrative or decision-making work. Yet task type alone did not explain their experiences. More important was how participants *selectively integrated* LLMs into their routines, shaped by the task, their topic familiarity, and their evolving understand-

ing of the tools’ capabilities. Many reported using LLMs more over time as they became more comfortable with them. However, they remained selective, turning to LLMs mainly for complex, unfamiliar, or exploratory tasks, while relying less on them for work they felt confident completing independently.

Self-governed readiness: Participants’ self-assessed readiness revealed a tension between confidence and uncertainty. Many felt generally prepared to use LLMs responsibly and were confident in their ability to protect sensitive data. Several stated that they saw little need for formal training, believing effective LLM use could be self-taught. However, this confidence was often built through informal, self-initiated learning rather than formal organizational support. Participants described learning through trial and error, online resources, webinars, external training, peer exchange, and personal judgment or “gut feeling,” while formal training and company guidelines were mentioned less frequently.

This produced a form of *self-governed readiness*: employees often felt capable, but their preparedness depended largely on individual experimentation. This became especially visible in relation to sensitive data and security-related decisions, where some participants wished for clearer organizational rules. In contrast, those who had access to company guidelines or information classification systems described them as a source of confidence. Thus, readiness was not simply an individual skill, but was shaped by the employees’ own learning efforts and the clarity, or absence, of organizational guidance.

Negotiating the Social Legitimacy of LLM Use Participants described LLM use as a socially situated and, at times, socially risky practice. Organizational cultures varied substantially: while some participants perceived their workplaces as supportive, especially when management framed LLMs as tools for efficiency and productivity, others described more skeptical or critical environments. Such skepticism was linked to generational differences, conservative organizational cultures, concerns about data leakage, and limited familiarity with digital technologies.

Beyond general support or resistance, participants raised a more specific concern: whether LLM use would be recognized as legitimate work. Some worried that using LLMs could be interpreted as a lack of competence, effort, or ownership. For instance, participants referred to the fear of “looking stupid” (P15) or the perception that using an LLM meant, “I didn’t do anything, I had someone else do the work for me” (P7). These concerns suggest that LLM use was evaluated not only in terms of utility, but also in relation to professional identity, authorship, and accountability.

This ambiguity also shaped disclosure. Many participants reported that LLMs were not discussed in depth within their organizations. Some believed management was unaware of employees’ LLM use, while others hesitated to disclose their own use because they anticipated negative reactions. Thus, employees did not only consider whether LLMs were useful or formally allowed; they also assessed whether their use would be viewed as acceptable, competent, and profession-

ally appropriate. Where open discussion was limited, employees had to infer informal norms and decide whether disclosure felt safe.

Everyday risk mitigation as informal governance Participants expressed high awareness of risks related to LLM use. Privacy, data protection, security, and output reliability were the most prominent concerns. Almost all participants could identify concrete risks, including accidental disclosure of sensitive data, unsafe or malicious code, external data processing, and the use of AI meeting assistants by third parties. Ethical concerns also extended beyond immediate workplace risks. Participants mentioned environmental sustainability, bias and discrimination, skill loss, overreliance, academic integrity, copyright issues, the replacement of human judgment, and the use of LLMs as substitutes for human interaction or human labor. Environmental sustainability was one of the most frequently mentioned ethical concerns.

However, the most important pattern was not only that participants perceived risks, but that they actively developed everyday mitigation practices. Most participants anonymized or rewrote prompts to avoid entering sensitive information. Others limited LLM use to public data, adjusted privacy settings, changed workflows, or avoided particular tools or use cases. These practices were often self-directed rather than prescribed by formal organizational policy.

Participants also developed strategies for managing accuracy risks. Many had encountered inaccurate outputs, including faulty code, outdated information, hallucinated facts, or responses that initially appeared trustworthy but were later found to be incorrect. To address this, participants compared LLM answers with external sources, relied on their own expertise, tested code, checked sources suggested by the LLM, asked the LLM to provide references, refined prompts iteratively, or discussed outputs with colleagues.

These examples show that employees' risk mitigation practices were not limited to technical safeguards. They also involved professional judgment, role-taking, contextual evaluation, and social review. In this sense, employees were engaging in a form of informal governance: they translated broad concerns about privacy, reliability, and ethics into practical decisions about when, how, and whether to use LLMs.

The informality of this governance was also visible in how incidents were handled. Most participants had no direct knowledge of LLM-related incidents in their organizations. Among those who did, incidents included sensitive data being entered into an LLM, concerns about AI meeting assistants in external meetings, and data breaches related to shared organizational accounts. Responses to such incidents were often informal, ranging from peer warnings and brief discussions to cases where superiors were not informed and no significant consequences followed. Risk management often depended on individual caution, peer communication, and ad hoc judgment, making it a very subjective process.

RQ2: Organizational Provisions and Support

Among the 20 participants, there are clear differences in the availability of formal policies and training. Larger organizations, especially those with more than 1000 employees in fields like IT, education, and energy, tend to have both policies and structured training programs. On the other hand, smaller organizations or those in sectors such as logistics, media, and public administration often report not having these formal provisions, highlighting a significant organizational gap. Additionally, a few participants reported training without corresponding policies or guidelines, or vice versa. The supplementary material provides an overview of the existence of organizational policies, guidelines, and training initiatives among the study participants.

Uneven Policy Visibility and Limited Actionability

About half of the participants were unaware of any official guidelines on LLM use in their workplaces. Although many participants stated that official guidance could be important, many also never inquired about them, either because they saw no immediate need, felt disengaged from formal documentation, or perceived such policies as irrelevant or obvious. However, a few who actively sought clarification reported receiving no clear answer or being told that another department was responsible for developing guidance.

By contrast, about half of the participants confirmed that policies or guidelines existed in their organizations. These documents most often addressed data processing, data protection, and general usage instructions, while issues such as bias, transparency, prompting, incident response, and broader risks were mentioned less frequently. Further, even though existing guidelines are generally understandable and helpful for broad tasks, they often cover only basics, lack depth and comprehensiveness, and are designed more for non-technicians. Many other participants stated that the documents are lengthy, too general, or hard to understand, with some describing them as “dry” or “not helpful”. A few participants even expressed uncertainty about where they can access the guidelines. Consequently, awareness and engagement remain rather low, as some participants reported not having read the policies.

Where policies were more explicit, they often provided clear “dos and don’ts,” particularly around data protection, such as prohibiting the entry of private, confidential, internal, or personal data into LLM tools. Some organizations also required employees to use only approved platforms, and a few mentioned tools compliant with external regulations such as the EU AI Act. However, a few organizations treat these policies as optional, explicitly allowing unrestricted use of provided versions, with employees being “*allowed to freely use the version that is made available to [them]*” and with no or minor restrictions.

Our findings also highlight that keeping policies current and communicating changes are challenging. Some participants are unaware of recent updates, and others report that updates are infrequent or poorly communicated, as they “*haven’t noticed anything*” (P19), and are often shared only through intranet articles or email distributions that few engage with.

Training Formats, Content, and Effectiveness Because policies and guidelines are often consumed passively, training represents an important complementary mechanism for translating organizational policies into practice. Unlike static documentation, training can engage employees more actively and support responsible behavior. In our study, organizations used a wide range of training formats, including online courses, webinars, meetings, quizzes in which participants “*have to achieve a certain number of points*” (P4), workshops, gamified “serious games”, and hackathon-style events such as “prompt-a-thons”. However, half of the participants reported that their organizations offered no formal training at all, indicating a substantial gap in support.

Even where training existed, participation was often uneven. Most programs were voluntary, and only a few required or strongly encouraged attendance. In some cases, organizations used reminders to encourage participation, with employees receiving email alerts to “*complete this e-learning course*” (P4). However, some employees had not engaged with any training because they felt they had “*managed quite well*” (P10) and did not expect to learn anything new. Thus, training availability did not necessarily translate into uptake, particularly among employees who already perceived themselves as competent users.

Training content typically covered appropriate usage, potential risks, safety considerations, prompting techniques, and, in some cases, compliance with regulatory frameworks such as the EU AI Act. However, participants frequently described these offerings as too basic, overly general, or insufficient for developing deeper understanding. Some saw little added value in the available formats, stating that they “*might as well just watch a YouTube video*” (P6). Others reported that formal training was limited in scope, frequency, and availability, with restricted places or sessions that did not reliably take place. As a result, many employees remained responsible for developing their own competence.

Finally, organizations rarely assessed whether employees had understood LLM-related risks. Only a few participants reported risk checks through seminar discussions or e-learning quizzes, and even these were criticized as too simple to test meaningful understanding. Overall, training provisions often lacked the depth, reach, and evaluative mechanisms needed to reinforce responsible LLM use in practice.

Technical Controls and Human Support Systems Technical support mechanisms play a crucial role in encouraging responsible use of LLM tools, complementing written policies and guidelines. Some organizations used automated solutions such as pop-up warnings, URL blocking, data masking, and access controls to prevent inappropriate use and reinforce rules around safety, privacy, and ethical standards. For example, one participant described data masking as a system “*where AI already recognizes what the employee is entering*” (P2).

Interestingly, human support systems were less developed. Although some participants reported having a clear contact point for LLM-related questions, many indicated that their organizations lacked a dedicated LLM contact person or support team. This suggests that technical safe-

guards were sometimes more established than the human support structures needed to help employees interpret and apply responsible-use expectations in practice.

RQ3: Closing the Gap: Policy, Guideline, and Training Improvements Proposed by Employees

Participants provided a wide range of concrete recommendations for organizational support, policy improvements, and technical solutions to facilitate a more responsible and confident use of LLM tools in the workplace.

Make Policies Understandable and Actionable Almost all participants emphasized the need for clear, accessible, and actionable policies to support responsible LLM use. Rather than lengthy or abstract documents, participants wanted guidance that was easy to find, concise, and understandable, so they could apply it without “*this flood of information*” (P4). They recommended written policies that are regularly updated and include concrete risk scenarios, explicit examples of permitted and prohibited use, and negative examples of misuse.

Participants also suggested embedding LLM guidance into existing organizational, legal, and operational frameworks. Several recommended checklist-style formats covering data protection, security, and procedures for accidental uploads or other mistakes. Others called for more context-specific guidance, including department-level policies, clearer rules around external LLM tools, and, in some cases, bans on non-approved systems. A few also proposed external standardization, such as EU-level certification, to help organizations rely on vetted LLM tools without conducting all assessments internally.

Design Practical and Effective Training Most participants called for practical and engaging training formats that move beyond passive information delivery. They recommended diverse online formats, including videos, short courses, and demonstration-based sessions focused on concrete workplace use cases. Many also emphasized the value of in-person workshops, interactive sessions, scenario-based exercises simulating real incidents, and quizzes “*just to deepen people’s knowledge and ensure that they don’t just listen to the information*” (P4). Participants also requested more advanced and regularly updated content covering sensitive data disclosure, emerging risks, data protection, and responsible prompting.

Risk awareness was a central training need. Some participants suggested live Q&A sessions, while others wanted training to include hallucinations, negative use cases, and examples of organizational mistakes, finding “*it quite interesting when other entrepreneurs made mistakes, that [they] could point out what consequences that can have*” (P10).

Participants differed on whether training should be voluntary or mandatory. Some preferred voluntary participation, while others argued that mandatory training was needed to ensure a shared baseline of responsible LLM use. A few suggested a hybrid model, where training is required only for employees who use LLM tools at work. About half also recommended integrating risk-awareness checks into training,

with some proposing random checks imilar to phishing simulations.

Embed Technical Safeguards into LLM Workflows Beyond policies and training, participants recommended technical safeguards that would make responsible LLM use easier to enact in practice. Most supported automated data sanitization to prevent sensitive information from being entered into LLM systems. However, some cautioned that such tools should not block all potentially sensitive content, since employees may intentionally use fictional names or anonymized examples. Instead, they suggested pop-up warnings that prompt users to verify whether sensitive data is being disclosed.

Participants also expressed mixed trust in automated safeguards. A few worried that data-sanitization tools would themselves need to process sensitive content, noting that *“this tool also recognizes this sensitive data and can see it”* (P3). Others preferred secure-by-design LLMs, including self-hosted models, task-specific functions, embedded documentation references, and systems developed through trusted institutional or European partnerships. One participant argued for alternatives to free commercial tools, explaining that *“if it’s for free, [they’re] not the client, [they’re] the product”* (P1).

Additional recommendations included firewalls or access controls to prevent unauthorized external LLM use, multiple output options to support comparison, test-driven LLM code generation to improve reliability, centralized information hubs, and tighter integration of LLM tools into existing organizational workflows.

Foster an Open and Transparent LLM Use Culture Participants emphasized that responsible LLM adoption requires not only policies, training, and technical safeguards, but also an organizational culture in which employees can discuss LLM use openly. Several called for greater awareness and open-mindedness, *“instead of just demonizing it”* (P14), and encouraged regular exchanges about both practical use and associated risks. Such communication was seen as important for reducing uncertainty and helping employees feel *“that they don’t have to be afraid to use these tools”* (P18). A few also suggested that officially introducing and structurally rolling out LLM tools would help normalize their use and reduce hidden or informal practices.

Participants also highlighted the need for continuous support infrastructures, including expert guidance, mentorship, and designated contact persons who could answer questions and help employees navigate responsible use. While a few emphasized individual responsibility, most argued that organizations should actively support employees, since companies have *“a responsibility to train, educate, and support its employees so that [LLMs] can be used safely in everyday work”* (P8). Transparency and open communication were therefore viewed as central to both confidence and accountability.

Finally, participants differed in how strongly LLM use should be promoted. About half supported active promotion to encourage innovation and competence, noting that these tools are increasingly present in everyday life and can sup-

port even simple tasks. Others preferred a balanced approach that combines encouragement with appropriate restrictions. Overall, participants called for a culture that neither stigmatizes nor uncritically promotes LLM use, but instead enables transparent, informed, and responsible adoption.

5 Discussion

This study examined workplace LLM adoption from the perspective of employees who are expected to interpret and enact responsible use in practice. Our findings show that responsible LLM governance can break down at the point of everyday use. When policies, training, and communication are absent, unclear, or difficult to apply, employees fill this gap through informal governance practices. These practices help employees manage privacy, security, reliability, and legitimacy concerns, but they also make responsible use uneven, dependent on individual judgment, and sometimes hidden from organizational view. In the following sections, we discuss three implications of this finding. First, the general-purpose nature of LLMs makes workplace governance difficult because organizations cannot assume a single user group, workflow, or risk profile. Second, organizations need to move from implicit reliance on employee self-governance toward actionable guidance and role-sensitive training. Third, responsible adoption requires workplace cultures where LLM use can be discussed openly without normalizing overreliance.

LLMs Are Not Just Another Workplace Technology

A key implication of our findings is that LLMs should not be treated as a conventional workplace tool. Many enterprise technologies are introduced for specific departments, workflows, or expert groups. LLMs, by contrast, are general-purpose systems that can be used across many tasks, roles, and organizational contexts (Eloundou et al. 2024; Calvino, Haerle, and Liu 2025; Tabassi 2023). This makes their adoption qualitatively different from the rollout of a specialized software system: employees can use the same tool for writing, coding, summarizing, research, communication, decision support, or administrative work, often with different types of data and different levels of risk.

This general-purpose character makes governance more complex. Organizations cannot assume a single user group, a stable workflow, or a bounded set of use cases. Instead, they need to understand how different employees interact with LLMs, what data they may enter, what outputs they rely on, and what risks arise in different work contexts. Our findings show that participants used LLMs through both official and personal accounts, often experimenting before or alongside formal organizational adoption. This bottom-up pattern aligns with reports of *bring-your-own-AI*, where employees adopt AI tools even when organizational strategies remain incomplete (Microsoft and LinkedIn 2024).

Responsible adoption therefore requires more than granting access to a tool. Organizations need to assess the risks associated with different user groups, tasks, data types, and deployment contexts. This aligns with emerging expecta-

tions around AI literacy, which emphasize that organizations should consider users' technical knowledge, experience, education, training, and the context in which AI systems are used (European Union 2024; European Commission 2025). In this sense, LLM governance must be organization-wide, but not generic: it must be specific enough to address the varied ways employees actually use these tools.

From Subjective Self-Governance to Actionable Organizational Guidance

Our findings show that employees were often left to govern their own LLM use. In the absence of clear or accessible guidance, participants relied on trial and error, personal judgment, peer exchange, anonymization, output checking, and their own sense of what seemed safe or appropriate. These practices show that employees were actively managing risk, but also that responsible use was often treated as an individual burden. This aligns with prior work showing that responsible AI learning often happens informally on the job, through information seeking, interpersonal learning, and situated workplace practice (Madaio et al. 2024). It also extends prior studies of workplace generative AI use, which find that employees want clearer guidance and that organizational rules, training, and support shape how AI tools are adopted in practice (Cardon et al. 2023; Li et al. 2024; Kimbel, Glas, and Pernul 2024).

This reliance on subjective judgment creates an accountability gap. What one employee considers sufficiently anonymized, ethically acceptable, or low risk may differ from another's interpretation, depending on their technical knowledge, role, experience, and access to sensitive data. This echoes broader responsible AI research showing that high-level principles and governance structures are difficult to translate into organizational routines (Rakova et al. 2021; Jafari Meimandi et al. 2025). In decentralized or complex organizational settings, governance may become fragmented across local contexts, producing variation in how responsibilities are interpreted and enacted (Jafari Meimandi et al. 2025). Related work on everyday responsible AI practice similarly shows that when formal governance structures are limited, accountability may be improvised through day-to-day work rather than applied through stable procedures (Gagua 2025). Our findings suggest that a similar dynamic can occur among ordinary workplace LLM users: employees become informal enactors of governance, even though they are not formally responsible for designing or overseeing it.

Organizations should therefore move from implicit reliance on employee judgment to actionable guidance. Policies and guidelines should be released alongside LLM tools, clearly communicated, easy to access, and specific enough to cover realistic workplace use cases. Rather than only stating broad principles, they should explain what is allowed, what is prohibited, what is ambiguous, and what employees should do when mistakes occur (Kawakami, Wilkinson, and Chouldechova 2024). This is particularly important because several participants were unaware of whether policies existed, did not know where to find them, or perceived existing documents as too vague, superficial, or detached from every-

day work. Governance cannot be effective if employees do not know that it exists or cannot translate it into action.

Training Must Move Beyond Generic Awareness

Training was another major gap. About half of the participants reported no formal training, and where training existed, it was often described as too basic, superficial, or insufficiently engaging. Some participants felt they already knew the content, while others wanted more practical and interactive formats. This suggests that training cannot be treated as a one-time introduction to LLMs or as a generic awareness exercise.

The need for better training follows directly from the general-purpose nature of LLMs. Because employees use these tools for different tasks, with different data, risks, and levels of expertise, training must be contextual rather than one-size-fits-all. A baseline module may be useful for all employees, but it should be complemented by role- and task-specific training for higher-risk or higher-frequency use cases. For example, developers, educators, administrators, managers, and employees handling customer or confidential data may require different examples, risk scenarios, and verification practices.

Participants' recommendations point toward more engaging forms of learning, including scenario-based exercises, live Q&A sessions, quizzes, workshops, gamified formats, and concrete examples of policy violations or harmful outcomes. Such formats are valuable because they move training from passive information delivery to practiced judgment. Personalized and adaptive training approaches from usable security similarly suggest that content is more effective when it reflects users' prior knowledge, work context, and risk exposure (Schöni et al. 2025; Šeda et al. 2021). For LLM adoption, this means training should not only teach employees how to use the tools, but how to decide when not to use them, how to verify outputs, and how to respond when something goes wrong.

Usage Stigma, Selective Disclosure, and Governance Blind Spots

A further implication of our findings is that concerns about professional legitimacy and stigma surrounding LLM use can create a governance blind spot. Participants worried that disclosing their use of LLMs could make them appear lazy, less competent, or less accountable for their work. These concerns are not necessarily unfounded: recent work shows that AI assistance can influence how others judge the person behind a message, making disclosure reputationally sensitive (Khadpe et al. 2025). When employees perceive LLM use as socially risky, they may disclose it selectively, use these tools "secretly," or avoid reporting mistakes. This is a significant governance concern because stigmatized use can become hidden use, limiting organizational oversight and opportunities to identify risks or learn from incidents.

This finding relates to Zhang et al.'s distinction between passive non-disclosure and active concealment of LLM use (Zhang et al. 2025). Their study shows that people are more likely to hide LLM use when they expect negative

judgments from others. Our study extends this idea to organizational governance. When employees fear reputational harm, they may disclose their use only selectively or keep it outside formal organizational view. This creates a governance blind spot. Organizations may underestimate how often and in what ways LLMs are used, remain unaware of risky practices, and miss opportunities to learn from uncertainty, mistakes, or near misses. In our study, some incidents were handled through peer warnings or informal discussions, while supervisors were not always informed. A culture that discourages disclosure may therefore weaken incident reporting, organizational learning, and the improvement of policies and support.

Organizations should treat disclosure as an important part of responsible governance, rather than only as an individual duty. Employees need clear guidance on when disclosure is expected, accessible channels for asking questions and reporting mistakes, and reassurance that honest disclosure will not automatically be treated as incompetence or misconduct. This does not mean that organizations should promote LLM use without limits. Instead, they should distinguish responsible assistance from inappropriate delegation or overreliance. Employees should be able to discuss how they used an LLM, what uncertainties arose, and when human review was needed.

One practical mechanism is the introduction of designated AI champions, similar to security or privacy champions. These employees could act as local contact points, support peer learning, encourage responsible experimentation, and help reduce stigma around LLM use. They would not replace formal governance, but could make it more visible and usable by connecting policies, training, and everyday practice. More broadly, our findings suggest that responsible LLM adoption requires a shift from generic governance to situated enablement: organizations must not only define rules, but also create the cultural, educational, and technical conditions that allow employees to follow them in practice.

6 Conclusion and Future Work

This study examined responsible LLM use in the workplace from employees' perspectives. Drawing on semi-structured interviews with 20 employees across diverse industries in Germany, we investigated how employees use LLMs in everyday work, how they experience organizational policies, training, and support, and what forms of guidance they consider necessary for responsible adoption. Our findings show that responsible LLM governance currently breaks down at the point of everyday use. When policies, training, and communication are absent, unclear, or difficult to apply, employees fill this gap through informal governance practices: they anonymize prompts, check outputs, seek peer advice, decide what information is safe to enter, and judge whether LLM use is professionally acceptable. These practices help employees manage privacy, security, reliability, and legitimacy concerns. However, they also make responsible use uneven, dependent on individual judgment, and sometimes hidden from organizational view.

We argue that responsible LLM adoption requires moving beyond policies on paper. Organizations need to make

governance visible, actionable, and integrated into everyday work. This includes clear and practice-oriented policies, role-sensitive training, technical safeguards, accessible support channels, and open communication cultures where employees can discuss uncertainty, disclose mistakes, and share responsible practices. Importantly, employees should not only be treated as recipients of governance, but as key stakeholders whose experiences can help shape more realistic and usable policies and training.

Future work should evaluate which support mechanisms most effectively improve responsible LLM use in practice. Comparative studies could test different training formats, such as short e-learning modules, scenario-based workshops, role-specific training, live Q&A sessions, gamified exercises, or phishing-style risk simulations, and examine their effects on employee confidence, risk awareness, disclosure behavior, and actual workplace practices. Such work would help organizations move from providing governance in name only toward evidence-based approaches for building meaningful and sustained AI literacy.

Ethical Considerations Statement

This study was conducted as part of a master's thesis, and hence participants received no financial compensation. They were informed of the study's academic context and participated voluntarily. At the time of the study, our institute did not have an ethics review board. However, participants were provided with a consent form outlining the study's scope, data usage, and retention policies, approved by our data protection officer. We adhered to General Data Protection Regulation (GDPR) guidelines, informing participants how their data is processed and stored. Participants were also made aware that they could withdraw their data at any point during or after the study without any consequences. We assured them that only anonymized data and quotes would be analyzed and published. Before starting the study, participants were required to consent and were encouraged to download a copy of the consent form for their records. Interviews were recorded only with explicit consent, camera use was optional, and transcripts were pseudonymized before analysis, and only anonymized data and quotations are reported.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2092 CASA - 390781972. Tarini Saka (lead author) was partially funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 538873437.

References

- Ahmed, S. K.; Mohammed, R. A.; Nashwan, A. J.; Ibrahim, R. H.; Abdalla, A. Q.; Ameen, B. M. M.; and Khdir, R. M. 2025. Using thematic analysis in qualitative research. *Journal of Medicine, Surgery, and Public Health*, 6: 100198.
- Balayn, A.; Yurrita, M.; Rancourt, F.; Casati, F.; and Gadri, U. 2025. Unpacking Trust Dynamics in the LLM Supply Chain: An Empirical Exploration to Foster Trustworthy

- LLM Production & Use. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–20.
- Braun, V.; and Clarke, V. 2021. Thematic analysis: A practical guide.
- Calvino, F.; Haerle, D.; and Liu, S. 2025. Is Generative AI a General-Purpose Technology? Implications for Productivity and Policy. *OECD Artificial Intelligence Papers* 40, OECD Publishing, Paris.
- Cardon, P. W.; Getchell, K.; Carradini, S.; Fleischmann, C.; and Stapp, J. 2023. Generative AI in the workplace: Employee perspectives of ChatGPT benefits and organizational policies. *SocArXiv papers*.
- Chkribene, Z.; Hamila, R.; Gouissem, A.; and Devrim, U. 2024. Large language models (llm) in industry: A survey of applications, challenges, and trends. In *2024 IEEE 21st International Conference on Smart Communities: Improving Quality of Life using AI, Robotics and IoT (HONET)*, 229–234. IEEE.
- Creswell, J. W.; and Poth, C. N. 2016. *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.
- Diro, A.; Kaisar, S.; Saini, A.; Fatima, S.; Hiep, P. C.; and Erba, F. 2025. Workplace security and privacy implications in the GenAI age: A survey. *Journal of Information Security and Applications*, 89: 103960.
- Eloundou, T.; Manning, S.; Mishkin, P.; and Rock, D. 2024. GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384(6702): 1306–1308.
- Estaiteyeh, M.; and McQuirter, R. 2024. Generative or degenerative?! Implications of AI tools in pre-service teacher education and reflections on instructors’ professional development. *Brock Education Journal*, 33(3): 75–98.
- European Commission. 2025. AI Literacy: Questions & Answers. Shaping Europe’s Digital Future.
- European Union. 2024. Article 4: AI Literacy. Regulation (EU) 2024/1689, Artificial Intelligence Act.
- Ferino, S.; Hoda, R.; Grundy, J.; and Treude, C. 2025. Junior Software Developers’ Perspectives on Adopting LLMs for Software Engineering: a Systematic Literature Review. *arXiv preprint arXiv:2503.07556*.
- Fischer, R.; Wischniewski, M.; van der Staay, A.; Poitz, K.; Janiesch, C.; and Liebig, T. 2025. Bridging the Communication Gap: Evaluating AI Labeling Practices for Trustworthy AI Development. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(1): 926–939.
- Gagua, A. 2025. Improvising Accountability: The Everyday Governance Work of Responsible AI in the Public Sector. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(3): 2869–2871.
- Gagua, A.; van der Voort, H.; Goyal, N.; and Verbraeck, A. 2025. Responsible AI Governance in the Public Sector: Explaining Contextual Dynamics through a Realist Synthesis Review. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, volume 8, 990–1002.
- Grill, M.; Sommestad, T.; Karlzén, H.; and Pousette, A. 2025. Training for improved information security culture: a longitudinal randomized controlled trial. *Information and Computer Security*, 33(5): 690–718.
- Hennink, M.; and Kaiser, B. N. 2022. Sample sizes for saturation in qualitative research: A systematic review of empirical tests. *Social science & medicine*, 292: 114523.
- Jafari Meimandi, K.; Reuel, A.; Aranguiz-Dias, G.; Rahama, H.; Ayadi, A.-E.; Boullier, X.; Verdo, J.; Montanie, L.; and Kochenderfer, M. 2025. An Adaptive Responsible AI Governance Framework for Decentralized Organizations. *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, 8(2): 1324–1334.
- Jeong, K. J.; Zhang, Z.; Dong, J.; Chung, A.; and Shih, P. C. 2026. Everyday Practitioner Experiences of AI-First Policies Adopted by U.S. Big Tech Companies. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- Jöhnk, J.; Weißert, M.; and Wyrski, K. 2021. Ready or not, AI comes—an interview study of organizational AI readiness factors. *Business & information systems engineering*, 63(1): 5–20.
- Kawakami, A.; Wilkinson, D.; and Chouldechova, A. 2024. Do Responsible AI Artifacts Advance Stakeholder Goals? Four Key Barriers Perceived by Legal and Civil Stakeholders. In *Proceedings of the 2024 AAAI/ACM Conference on AI, Ethics, and Society*, AIES ’24, 670–682. AAAI Press.
- Khadpe, P.; Wenzel, K.; Loewenstein, G.; and Kaufman, G. 2025. Explaining the reputational risks of ai-mediated communication: Messages labeled as ai-assisted are viewed as less diagnostic of the sender’s moral character. In *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society*, volume 8, 1413–1423.
- Kim, A.; Yang, M.; and Zhang, J. 2023. When Algorithms Err: Differential Impact of Early vs. Late Errors on Users’ Reliance on Algorithms. *ACM Transactions on Computer-Human Interaction*, 30(1).
- Kimbel, A.; Glas, M.; and Pernul, G. 2024. Security and Privacy Perspectives on Using ChatGPT at the Workplace: An Interview Study. In *International Symposium on Human Aspects of Information Security and Assurance*, 184–197. Springer.
- Klingbeil, A.; Grützner, C.; and Schreck, P. 2024. Trust and reliance on AI—An experimental study on the extent and costs of overreliance on AI. *Computers in Human Behavior*, 160: 108352.
- Li, Z. S.; Arony, N. N.; Awon, A. M.; Damian, D.; and Xu, B. 2024. AI tool use and adoption in software development by individuals and organizations: a grounded theory study. *arXiv preprint arXiv:2406.17325*.
- Liao, Z.; Antoniak, M.; Cheong, I.; Cheng, E. Y.-Y.; Lee, A.-H.; Lo, K.; Chang, J. C.; and Zhang, A. X. 2024. LLMs as Research Tools: A Large Scale Survey of Researchers’ Usage and Perceptions. *arXiv preprint arXiv:2411.05025*.
- Madaio, M.; Kapania, S.; Qadri, R.; Wang, D.; Zaldivar, A.; Denton, R.; and Wilcox, L. 2024. Learning about Responsible AI On-The-Job: Learning Pathways, Orientations, and Aspirations. In *Proceedings of the 2024 ACM Conference*

- on Fairness, Accountability, and Transparency, FAccT '24, 1544–1558. New York, NY, USA: Association for Computing Machinery. ISBN 9798400704505.
- Microsoft and LinkedIn. 2024. 2024 Work Trend Index Annual Report: AI at Work Is Here. Now Comes the Hard Part. <https://www.microsoft.com/en-us/worklab/work-trend-index/ai-at-work-is-here-now-comes-the-hard-part>.
- Myers, M. D.; and Newman, M. 2007. The qualitative interview in IS research: Examining the craft. *Information and organization*, 17(1): 2–26.
- OWASP. 2025. OWASP Top 10 for Large Language Model Applications. <https://owasp.org/www-project-top-10-for-large-language-model-applications/>. [Accessed 12-11-2025].
- Pathan, M. K.; and Shah, A. 2025. Ethical Considerations and Responsible Governance of Generative AI: A Systematic Review. *Premier J Artif Intell*, 4: 100016.
- Prümmer, J.; van Steen, T.; and van den Berg, B. 2024. A Systematic Review of Current Cybersecurity Training Methods. *Computers & Security*, 136: 103585.
- Radanliev, P.; Santos, O.; Brandon-Jones, A.; and Joinson, A. 2024. Ethics and responsible AI deployment. *Frontiers in Artificial Intelligence*, 7: 1377011.
- Rakova, B.; Yang, J.; Cramer, H.; and Chowdhury, R. 2021. Where responsible AI meets reality: Practitioner perspectives on enablers for shifting organizational practices. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1): 1–23.
- Saka, T.; Vakali, K.; Jenkins, A. D. G.; Kokciyan, N.; and Vaniea, K. 2025. Judging Phishing Under Uncertainty: How Do Users Handle Inaccurate Automated Advice? In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25. New York, NY, USA: Association for Computing Machinery. ISBN 9798400713941.
- Schöni, L.; Roch, N.; Sievers, H.; Strohmeier, M.; Mayer, P.; and Zimmermann, V. 2025. It's a Match - Enhancing the Fit between Users and Phishing Training through Personalisation. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, CHI '25. New York, NY, USA: Association for Computing Machinery. ISBN 9798400713941.
- Sebastian, G. 2024. Digital Shadow Risk Theory (DSRT): Assessing Organizational Risks of AI in the Workplace. Available at SSRN 5038258.
- Šeda, P.; Vykopal, J.; Švábenský, V.; and Čeleda, P. 2021. Reinforcing Cybersecurity Hands-on Training With Adaptive Learning. In *2021 IEEE Frontiers in Education Conference (FIE)*, 1–9.
- Shan, R.; and Shan, T. 2024. Enterprise LLMOps: Advancing Large Language Models Operations Practice. In *2024 IEEE Cloud Summit*, 143–148. IEEE.
- Sood, A. K.; Zeadally, S.; and Hong, E. 2025. The Paradigm of Hallucinations in AI-driven cybersecurity systems: Understanding taxonomy, classification outcomes, and mitigations. *Computers and Electrical Engineering*, 124: 110307.
- Tabassi, E. 2023. Artificial Intelligence Risk Management Framework (AI RMF 1.0).
- Tadimalla, S. Y.; and Maher, M. L. 2024. AI literacy for all: adjustable interdisciplinary socio-technical curriculum. In *2024 IEEE Frontiers in Education Conference (FIE)*, 1–9. IEEE.
- Tavasoli, A.; Sharbaf, M.; and Madani, S. M. 2025. Responsible innovation: A strategic framework for financial LLM integration. *arXiv preprint arXiv:2504.02165*.
- Uren, V.; and Edwards, J. S. 2023. Technology readiness and the organizational journey towards AI adoption: An empirical study. *International Journal of Information Management*, 68: 102588.
- Vaishnav, L.; Singh, S.; and Cornell, K. A. 2024. Transparency, security, and workplace training & awareness in the age of generative ai. *arXiv preprint arXiv:2501.10389*.
- Verma, S. J. 2024. Lessons learned from qualitative fieldwork in a multilingual setting. *Qualitative Research in Medicine & Healthcare*, 8(1): 11554.
- Wagman, K. B.; Dearing, M. T.; and Chetty, M. 2025. Generative AI Uses and Risks for Knowledge Workers in a Science Organization. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 1–17.
- Yan, B.; Li, K.; Xu, M.; Dong, Y.; Zhang, Y.; Ren, Z.; and Cheng, X. 2024. On protecting the data privacy of large language models (llms): A survey. *arXiv preprint arXiv:2403.05156*.
- Zhang, A.; and Lee, M. K. 2025. Knowledge Workers' Perspectives on AI Training for Responsible AI Use. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery.
- Zhang, Z.; Shen, C.; Yao, B.; Wang, D.; and Li, T. 2025. Secret Use of Large Language Model (LLM). *Proceedings of the ACM on Human-Computer Interaction*, 9(2): 1–26.