Revisiting the Responsibility Gap in Human–AI Collaboration from an Affective Agency Perspective

Research Paper

Jonas Rieskamp, Annika Küster, Bünyamin Kalyoncuoglu, Paulina Frieda Saffer, and Milad Mirbabaie

University of Bamberg, Faculty Information Systems and Applied Computer Sciences, Bamberg, Germany
{jonas.rieskamp, milad.mirbabaie}@uni-bamberg.de
{annika.kuester, buenyamin.kalyoncuoglu, paulina-frieda.saffer}@stud.uni-bamberg.de

Abstract. This research investigates how responsibility evolves when artificial intelligence (AI) systems influence or dictate decision-making processes. As AI becomes integral in various domains, from healthcare to autonomous systems, its role challenges traditional notions of responsibility. The study highlights the "responsibility gap," where accountability diffuses among developers, users, and organizations due to AI's autonomous decision-making capabilities. By incorporating affective agency, we explore how responsibility emerges as a relational construct shaped by human, technological, and contextual interactions. Leveraging qualitative interviews with experts across sectors, this research examines responsibility emerging from relational agency in human–AI collaboration. Findings emphasize the dynamic interplay between human oversight, affective agency, and digital autonomy, contributing to mechanisms for clear accountability. By addressing this gap, the study supports the development of responsible and transparent AI systems aligned with societal values.

Keywords: Artificial Intelligence (AI), Responsibility Gap, Responsibility in Human–AI collaboration, Decision-Making, Sociomateriality, Affective Agency

1 Introduction

The use of AI systems for decision support and decision-making has increased exponentially in recent years and now permeates almost all areas of life—from everyday processes such as personalized recommendations in online retail to safety-critical fields such as healthcare and autonomous driving (Razmetaeva & Satokhina, 2022). AI systems promise to significantly increase the efficiency and accuracy of human decisions. At the same time, however, "human decision making and management […] is increasingly conditioned by AI and algorithms" (Fischer et al., 2023, p. 230), which underscores the growing reliance on and relevance of AI systems organisational contexts (Jöhnk et al. 2021).

20th International Conference on Wirtschaftsinformatik,
September 2025, Münster, Germany

Yet, the current discourse remains inconclusive regarding the ultimate responsibility if an AI system significantly influences a decision or even makes it autonomously. This question is discussed under the term "responsibility gap," as it is often difficult to assign blame clearly when no single person has full control over the outcome (Dong & Bocian, 2024). The responsibility gap refers to the difficulty of attributing moral or legal accountability to a single actor, given that AI systems often operate autonomously and beyond direct human control or knowledge (Smith et al., 2023; Berente et al., 2021). This phenomenon becomes particularly evident with autonomous systems, such as self-driving cars or autonomous weapon systems (Vallor & Bekey, 2017; Zeiser 2024), which can make decisions without the possibility of direct human intervention. In the healthcare sector, for instance, responsibility is shared among developers of AI systems, the medical staff, and the hospital as an institution, rendering problems regarding the clear determination of who is accountable for errors (Bleher & Braun, 2022; Taylor, 2024a). Despite these remaining issues of AI systems acting autonomously, scholars generally accepted that final responsibility lies with the human actor (Krügel et al., 2024).

In an attempt to bridge the responsibility gap, we need to further examine the relationship between users and AI systems. In this regard, we approach human–AI interactions and the responsibility problem by focusing on how agency emerges through relationships and interdependencies between humans and technologies. Specifically, we consider the role of affective agency since it was found to influence the relationship between humans and technology (Pignot & Thompson, 2024). Emotional attachments with technologies foster a sense of responsibility for one's own actions or their consequences and influence behaviour—for example, in environmentally conscious decisions (Antonetti & Maklan, 2014). This supports the assumption that responsibility is not purely cognitive but is strongly shaped by relational and affective dimensions. Papagiannidis et al. (2025) highlight that responsible AI governance directs attention to the role of human agency in the interpretation and use of AI systems. Relational and affective agency thus form the foundation for rethinking responsibility. Responsibility is directly tied to agency since the capacity to act is a prerequisite for the attribution of responsibility (Schatzki, 2025). In this regard, we focus on personal, moral, and legal responsibility related to organizational decision-making processes influenced by AI.

Building on these considerations, the present study examines the following research question: *How does responsibility change when AI systems influence or prescribe decisions?* To address this problem, we conduct qualitative interviews with experts from fields such as law, ethics, healthcare and technology. These perspectives will offer in-depth insights into the relationally affected responsibility attributed when AI systems shape or prescribe decisions. Our findings explain how affective positionalities toward AI (e.g., scepticism and trust) bridge the responsibility gap.

2 Background

2.1 Responsibility Gap in Human–AI Collaboration

The rapid integration of AI into organizational decision-making and service processes has sparked intensive debate about responsibility for AI-driven outcomes (Mikalef et al., 2022). As AI systems grow more autonomous, they do not merely function as passive tools but actively shape actions and decisions, sometimes in ways that even developers and users do not fully anticipate (Berente et al., 2021). This has led research to introduce the concept of a *responsibility gap*, highlighting the difficulty in assigning responsibility for actions that arise from AI systems operating beyond direct human control (Lüthi et al., 2023). Recent research on responsibility attribution in AI systems underlines how traditional models of accountability often fall short (Huang & Suang, 2024). Specifically, two unresolved normative debates are at play. First, whether AI itself should be considered a legitimate responsibility holder or viewed exclusively as a tool. Some argue that AI systems might be treated as "legal persons," analogous to corporate entities (Gervais & Nay, 2023), whereas others stress that AI lacks true moral agency if it does not fulfil fundamental conditions like full awareness of its actions (Coeckelbergh, 2020). Second, responsibility for an AI's behaviour may be distributed across multiple actors—developers, adopting organizations, frontline employees, and end users—leading to so-called "many-hands" problems when failure or harm occurs (de Sio & Mecacci, 2021). Beyond these theoretical debates, empirical studies show that human–AI collaborations in real-world scenarios produce considerable ambiguity in day-to-day work settings. As AI shapes decisions, both domain experts and data scientists may find themselves unsure how to interpret or justify AI outputs (Thuis et al., 2023). This uncertainty is compounded by the black box features of machine learning models (Kim et al., 2023), limiting the ability to trace or challenge algorithmic reasoning. Thus, "human decision-making and management […] is increasingly conditioned by AI and algorithms" (Fischer et al., 2023, p. 230), but corresponding responsibility structures remain ill-defined.

AI's capacity for adaptive learning and automation is the key driver of this responsibility gap. Autonomy, in this sense, does not imply complete independence from human oversight but rather the system's ability to act on complex data patterns in ways not fully pre-scripted by developers (Berente et al., 2021). Especially in safety-critical or high-stakes contexts like healthcare, finance, and autonomous driving, these autonomous decisions may carry ethical and legal implications (Li et al., 2024; Huang & Suang, 2024). As tasks formerly performed by human agents transition to AI, the locus of control and accountability often becomes blurred. Field studies illustrate that after such task shifts, organizational actors may struggle with the interpretability of AI recommendations and feel unable to "own" decisions, while data scientists may, in turn, feel they have produced only a model rather than a final judgment (Thuis et al., 2023). Such fragmentation intensifies the responsibility gap, as no single actor feels properly equipped or authorized to take responsibility for the AI's eventual behaviour. Possible approaches include establishing organisational responsibility for independent AI systems (Chen et al., 2023).

Research emphasizes that ethical and societal concerns related to AI-assisted decision-making. AI's involvement in critical decision contexts—such as healthcare diagnostics, judicial decisions, or autonomous driving—raises ethical challenges, notably issues of fairness, transparency, explainability, and potential biases inherent in the data (Mittelstadt, 2019). These concerns extend beyond individual accountability to include broader societal risks, such as reinforcing existing societal biases, increasing inequality, and diminishing trust in institutions (Floridi et al., 2020; Kellogg, Valentine & Christin, 2020). Thus, the ethics discourse calls for comprehensive governance frameworks and guidelines tailored explicitly to manage the societal risks and ethical implications associated with deploying AI systems, especially in high-stakes scenarios (Floridi & Cowls, 2019; Jobin, Ienca & Vayena, 2019).

In sum, the increasing autonomy of AI leads to a reconfiguration of decision authority and accountability. The responsibility gap emerges precisely because traditional frameworks assume a single identifiable decision-maker, whereas AI challenges those assumptions by adding new "actors" and redistributing the conditions for moral and legal responsibility (Coeckelbergh, 2020). Against this backdrop, it is imperative that future governance must explicitly account for both the technical complexity and the human–organizational interplay to foster ethically and legally sound AI adoption.

2.2 Affective and Relational Agency

Sociomateriality emphasizes the inseparability of technology and social practices, highlighting their mutual constitution in organizational life (Orlikowski & Scott, 2008; Niemimaa, 2016). Therefore, the sociomaterial perspective is well-suited to examine the responsibility at the fringe of human–AI collaboration. In the sociomaterial tradition, social and material actors are co-constitutive for phenomena in the world. This perspective is grounded in agential realism, which focuses on how agency emerges from dynamic entanglements of humans, artifacts, and broader contexts (Weber, 2020; Liang & Weber, 2024). Within such entanglements, the notion of 'relational agency' —the capacity for action that arises through relations and interdependencies rather than isolated actors—states that responsibility cannot be allocated simply by parsing human versus machine contributions. Agency arises collectively in the interaction of sociomaterial actors (Liang & Weber, 2024). Affective agency introduces an additional facet, demonstrating that decisions are often influenced by unconscious motives and affective positionalities (e.g., emotions or personal opinions). Furthermore, Pignot & Thompson (2024) complement this view with the concept of negative ontology, emphasizing that agency arises from a lack. In other words, decisions result from a web of interconnected influences, in which machines and humans alike co-condition outcomes. Individuals may also identify with artifacts or social structures to address experiences of 'lack' or misrecognition (Laclau, 2000; Epstein, 2018), imbuing technologies with moral significance, especially if the systems appear fair or autonomous. Schatzki (2025) expands this idea by describing agency as a nexus of actions, effects, and intentions across human and non-human actors. In such distributed networks, responsibilities become blurred since each participant's activity is inseparable from that of others. Digital technologies increasingly act as quasi-autonomous agents, taking on tasks and judgments

once limited to humans (Ågerfalk 2020; Fischer et al., 2023). Emotional factors are central to these dynamics: Decisions emerge not solely from reasoned deliberation but also from affective positionality—individuals' emotional orientations toward technology—and relational attachments to technology (Pignot & Thompson, 2024).

This relational focus resonates with findings that AI does not fully supplant human intelligence or moral decision-making. Rather, humans integrate or resist AI outputs based on perceptions of risk, trust, and the system's inscrutability (Baroni et al., 2022). Affective attachments, however, can strengthen a sense of responsibility, for instance, in environmentally conscious behaviour (Antonetti & Maklan, 2014) or foster blame-shifting if technologies seem too opaque or invasive. Consequently, responsibility in AI-assisted contexts hinges on sociomaterial entanglements, where humans, algorithms, and institutional structures jointly configure what counts as responsible action.

In short, relational agency argues AI-assisted decisions are co-constituted by technical artifacts and human subjectivities. As AI evolves beyond mere tools, clarifying responsibility demands close attention to how affect, subjectivity, and context shape the ways humans adopt, interpret, or challenge algorithmic outputs. This highlights why attributing accountability cannot be reduced to simple human-versus-machine dichotomies. Instead, it arises through the fluid interplay of people, practices, and technologies, each of which contributes to the ultimate moral, legal, and organizational consequences of AI-driven processes.

3 Research Design

To address the research question, we conducted 20 semi-structured interviews (I1- I20) with experts who are professionally involved with AI. The primary criterion for inclusion in this study was active engagement with AI systems in their respective work domains to ensure a diverse and pertinent sample of users. To avoid biases of a single industry, a wide range of sectors was included: automotive, healthcare, telecommunications, IT, and outdoor sports. Interviews were conducted online between December 2024 and January 2025, recorded, and transcribed with the interviewee's permission. We grouped the interview questions into six parts[1], including (1) demographics, (2) responsibility in general, (3) the influence of AI on the decision-making process, (4) affective dimensions of responsibility, (5) diffusion and attribution of responsibility, and (6) technical aspects and understanding of AI. The sample of interviewees includes 13 men and 7 women, aging from 24 to 47. The interviewees took 32 minutes on average. German interview transcripts were translated before analysis.

We used techniques of the grounded theory methodology (Gioia et al., 2013) to analyse the interviews while incorporating agency and responsibility as sensitizing concepts (Blumer, 1954; Goldkuhl & Cronholm, 2019). To this end, we first performed inductive open coding to gain an overview of all relevant data using *MAXQDA 2024*. The open codes were grouped into first-order concepts and then were successively grouped into second-order themes. The second-order codes were discussed among the

[1] Please, find a list of the interviewees and all questions in the online appendix: https://osf.io/4nsy6/files/osfstorage?view_only=39dbf36da87a47b18ed6e9bf6e8d5195

authors until agreement was reached. We then subsumed second-order themes into four aggregated dimensions, namely *(1) human-AI relationship and acceptance, (2) regulation and governance, (3) responsibility and control, and (4) processes and challenges* (see Table 1 for the data structure). While *responsibility* is our primary theoretical and literature-backed concept, notions of transparency, trust, ethics, and accountability were identified as subordinate or related to responsibility by the interviewees. That is, these notions are not grounded in literature but inferred from the data.

4 Findings

Table 1. Data structure.

First-order code	Second-order code	Aggregated dimension
Self-Assessment of Responsibility, Willingness to Assume Responsibility, Awareness of Responsibility, Feeling of Responsibility	Responsibility and Personal Attitude	Human-AI-Relationship and Acceptance
Perceived autonomy of AI, Emotional Interaction with AI, Human-Machine Symbiosis, Human-Machine Interaction, Change in the Type of Interaction, Autonomy vs Support	Human-AI-Interaction	
Emotional Interaction with AI	Acceptance and Trust	
Trust in AI, Uncertainty about AI Function, Transparency + Traceability, Transparency of AI Technology, Transparency Requirements	Transparency	Regulation and Governance
Risk Assessment, Normative Influences, Ethical Implications, Ethical Concerns	Ethics	
Liability Issues, Accountability, Regulatory Framework, Governance Structures	Regulation and Governance	
Situation of Responsibility, Responsibility in the Workplace	Responsibility at the Workplace	Responsibility and Control
Assumption of Responsibility, Attribution of Responsibility, Transfer of Responsibility, Degree of Responsibility	Duty of Care	
AI-Regulation, Controllability, Control Mechanisms	Control	
Technology Acceptance, Influence, Synergy Effect, Understanding of Roles, Role of AI	Understanding of Roles	Processes and Challenges
Barriers to Use, Technological Scepticism	Technological Challenges	
Decision-Making Process, Rationality vs Intuition, Adaptive Decision-Making Behavior, Rationality vs Emotionality	Decision-Making Processes	

4.1 Human-AI Relationship and Acceptance

Most interviewees indicate that their ideal scenario would relieve them of laborious data analysis while keeping final control in their hands. They observe two main shifts in responsibility following the use of AI systems: first, greater awareness of the need to validate and interpret AI outputs, and second, new grey areas concerning potentially severe consequences of incorrect inputs or unquestioned acceptance (I5, 189–193; I8, 168–171). As a result, personal responsibility grows since one cannot simply say, “The AI said so.” Only in scenarios where algorithms operate almost entirely on their own

responsibility would we undergo a more profound change. Despite AI's influence, participants widely agree that responsibility remains a primarily human affair. At the same time, interviewees call for more transparency, explainability, and ethical compliance—initiating a call for a shared yet not fully transferred responsibility among developers, users, and organizations.

Nevertheless, most respondents stay "accountable" for decisions, defining the AI's scope and input and facing any failures. Rather than a reduction of responsibility, this leads to what many call a recalibration—or sometimes an intensification—of accountability. Interviewees highlight that responsibility implies both acknowledging mistakes and contributing to successes: "living with the consequences" (I3, 19–27) and "standing up for one's actions" (I5, 15–22). In the words of I6, reliability is pivotal:

> "*For me, taking responsibility means that people can rely on me, and I can say with a clear conscience: 'This is my result' or 'give me the task, I can do it,' and I have the confidence to do so. […] I would say very similar to what I have already said: in my position, a lot about gathering and passing on information. Especially escalation is often a lot of pressure, a lot of stress. I try to pay attention to people and not just say 'just do it now', but also see what their needs are but also, to see what their needs are, how can I help?*" (I6, 29—37, telecommunications).

Meanwhile, I10 (33–36) underscores that to prove oneself responsible, one must work "cleanly and properly at the right time" to establish credibility. Many want to reaffirm their role as accountable actors in the face of AI support. Most perceive AI primarily as a tool for minor technical tasks or inspiration (I6, 42–49), with I5 (151–157) describing it as a "second opinion" rather than an autonomous agent. In I8 (66–74), AI is used to summarize papers, but the person ultimately decides which option makes sense. Only rarely do participants see AI taking the lead by itself. A recurring image is that AI acts like a "sparring partner" (I8, 113–116), providing data analyses or new ideas but leaving final authority to the human user.

4.2 Regulation and Governance

Most interviewees, especially in fields like healthcare or highly automated sectors, want maximum transparency from AI systems, highlighting transparency as key to assigning responsibility. I2 (126–127) calls for "the greatest possible transparency," I5 (223–227) approves that tools like ChatGPT now partially provide sources, and I6 (128–131) demands open data disclosure and explanations of how AI arrives at its output. I14 (342–358) stresses that people must at least grasp how data is weighted, noting that purely abstract formulas can breed scepticism. Lack of explainability can also undermine the sense of responsibility, as I2 (94–98) describes greater reluctance to use unchecked AI suggestions, and I4 (141–147) warns about the risk of blind trust. Others, like I8 (137–141) or I1 (137–140), see opacity as a spur to double-check results. The central insight is that AI usage often heightens critical scrutiny rather than diluting responsibility—particularly for individuals who understand the potential for AI's errors and remain cautious about relying on algorithmic outputs. I14 summarizes:

"*If an error occurs, it is up to us to ensure that the AI was used correctly, the data was accurate, and the decisions are ethical. This means that even when AI is used, humans still bear the ultimate responsibility. AI can support us, but we must ensure that its recommendations make sense and are in line with our goals and values […].*" (I14, 264–276, healthcare)

I5 (189–193) gives two reasons for this continuing human accountability: actively choosing to use AI and being responsible for how one queries (prompts) the tool. Only with fully autonomous systems does the question arise of whether developers or operators bear greater responsibility. I6 (114–118) criticizes a lack of ethics and data protection guidelines, while others argue that developers must help ensure reliability and ethics but cannot be liable for all use cases (I5, 194–199; I7, 388–397). I8 suggests that developers at least highlight limitations via a "warning notice":

"*I think the developers often don't want you to use AI as the ultimate decision-making authority because they themselves know that the system can make mistakes, of course it's a business model, but serious developers generally issue warnings that you shouldn't accept the results without checking them. They cannot take responsibility for every possible application because they do not know the context of use."* (I8, 180–183, environmental sector)

Although a few respondents worry about blame-shifting to AI in ethical lapses, most believe workplace norms and supervisory oversight would prevent excusing bad decisions on unverified outputs (e.g., I7; I3; I14).

4.3 Responsibility and Control

Although interviewees have varying focal points—product vs. team responsibility (I4, 19–23), team management (I4, 19–23), or customer focus (I7, 39–49)—all emphasize accountability for consequences (I2, 18–25; I18, 57–65). This holds true even with AI in the mix. Developers must uphold ethical standards and transparency but cannot be blamed for system misuse by others (I7, 388–397; I6, 114–118; I5, 194–199). Interviewees call for stronger legal regulations—especially for data protection and algorithmic transparency (e.g., I6; I16)—since the current regulatory gap means that, in practice, responsibility still lies with the humans who implement decisions.

AI does not absolve respondents from responsibility; rather, they incur an additional meta-level: they must not only account for decisions but also justify how AI was used, which data were input, and whether the algorithm's output was critically examined. This "expanded responsibility" emerges in statements like I10's:

"That I carry out the tasks assigned to me – especially because it is in the healthcare sector and can have a direct impact on patients as it involves X-ray equipment – in a clean, tidy, responsible manner and also handle them responsibly. […] For example, when I evaluate images and present the data, where the sources of error are, then you have to work very precisely and properly" (I10, 38–45, healthcare).

Although some foresee a future where advanced AI automates large data analyses with minimal human intervention (I16, 145–156), they note that errors remain the user's responsibility unless comprehensive new regulations emerge (I6, 114–118). I18 (264–

273) adds that “one could certainly hide behind AI,” but once mistakes occur, people ask why proper verification was not done. A shared awareness thus arises: AI can streamline workloads, but its recommendations only work as well as the data and the user’s critical review. Interviewees advocate assigning AI to repetitive, objective tasks while sensitive decisions remain under human judgment (e.g., I3; I8). A respondent in I15 (136–148) highlights AI’s value in supply chain management to make forecasts more efficient, yet management still decides how to implement or correct those forecasts.

In sum, responsibility in the workplace fundamentally rests on bearing the consequences of one’s decisions—AI-mediated or not. Because most rely on AI mainly as an analytical aid rather than a sole decision-maker, humans remain responsible. And while uncertainty about AI’s internal mechanisms persists, it often spurs caution and critical engagement rather than eroding responsibility.

4.4 Processes and Challenges

The interviews reveal a nuanced understanding of the roles that AI and humans play in decision-making processes. I4 recognizes that while AI is often seen as highly pragmatic and data-driven, it struggles to account for emotional values or empathy:

> “*I have the feeling that the decision made by AI is probably much more pragmatic or data-based than a decision that only I would make. So, I would assume that no emotions in that sense are involved at first.*” (I4, 60–62, energy)

This distinction becomes particularly relevant in contexts where political considerations or management decisions require balancing human sensitivities (I3, 59–62). In such cases, AI may support processes but cannot replace the human perspective. A key finding concerns the shifting of responsibility when AI is integrated into decision-making. The results indicate an ambivalence: on the one hand, AI can lead to a partial relinquishment of responsibility; on the other hand, it can reinforce users' sense of accountability. One interviewee (I3, 142–150) described how AI-assisted decision-making might enable her to "relinquish some of my own responsibility" because she could "rely on the machine." However, others (I8, 137–141; I4, 133–140) stated that precisely because of the uncertainty about how AI works, they feel an even greater need to scrutinize AI-generated recommendations, which ultimately increases their perceived responsibility. This dual effect aligns with answers of I15 (222–237), where respondents with higher technical expertise reported feeling more motivated to critically assess AI outputs. At the same time, they suspected that others, particularly those with less technical knowledge, might trust AI-generated results without thorough verification. These insights highlight the importance of human oversight in AI-assisted decision-making. The dataset consistently underscores the necessity of keeping a "human in the loop" (I7, 340–347): “Not for me personally now, because I am still responsible for my result. I can't say: ‘It was the AI.’ Then it would be: ‘Why didn't you test it beforehand?’ […]”. Many interviewees emphasized that AI systems should not operate entirely autonomously, even if they offer efficiency gains. Instead, they saw it as their duty to supervise AI-supported processes to ensure reliability and accountability. The notion that responsibility expands rather than diminishes with AI use is particularly evident in I8. One

respondent stated:

> "*Me, 100%. It's just a tool, just like an IDE or any other program. If I rely on it, it's my responsibility if something goes wrong. […] For me, it always solely lies with me. I can't make the AI tool responsible for anything because it only makes suggestions. […] It clearly lies with me or the person leading the project. AI is only a support tool.*" (I8, 169–179, environmental sector)

Beyond the decision-making process itself, some respondents stressed that responsibility also extends to input quality and the evaluation of AI-generated responses (I5, 189–193). This further reinforces the view that AI does not diminish human accountability but rather requires users to adopt a more reflective and active role in overseeing its use. Additionally, while AI is highly valued for its ability to process large amounts of data, its interpretation and the overall process remain firmly within the human sphere of responsibility (I14, 150–161). Particularly in areas where empathy, ethical considerations, or emotional aspects play a role, interviewees emphasized their duty to make the final decision (I3, 48–55; I18, 359–369). A recurring concern is that, despite its computational power, AI lacks contextual understanding, value-based reasoning, and an intuitive grasp of human complexities that are essential for responsible decision-making.

5 Discussion

Interviewees emphasize their need to understand how AI operates and critically assess its outputs, emphasizing their relational perspective of agency (Pignot & Thompson, 2024; Schatzki, 2025), captured by the aggregated dimension Human-AI Relationship and Acceptance. While AI can produce sophisticated forecasts or recommendations, these gains only materialize through human interpretation and judgment. Rather than feeling relieved of responsibility, participants report an amplified sense of duty when relying on opaque or black box AI models. This resonates with the negative ontology perspective (Pignot & Thompson, 2024), where uncertainty fuels rather than diminishes accountability, which is reflected in the aggregated dimension Responsibility and Control. In other words, the awareness of AI's complexity and potential biases drives people to reaffirm their professional ethos, exemplifying the affective agency (cf. Pignot & Thompson, 2024). Nonetheless, the lure to shift blame to technology is not absent. Some respondents noted that ambiguous AI outputs might invite scapegoating if errors arise. Yet, their organizations appear to implement clear protocols preventing staff from "hiding behind AI," reflecting the organization's commitment to Regulation and Governance structures. This aligns with broader frameworks positing that moral or legal responsibility remains with human actors (Verdicchio and Perin, 2022). Indeed, a fully autonomous AI that obviates human oversight was not desired by any interviewee, supporting calls for human-in-the-loop approaches to maintain human oversight (in line with Papagiannidis et al., 2025). In practice, respondents view AI tools as "sparring partners," not as independent decision-makers, highlighting again the aggregated dimension of Human-AI Relationship and Acceptance This upholds traditional responsibility models where a human or institution remains ultimately accountable (Champagne

& Tonkens, 2015). The potential for collective, AI-driven responsibility is not observed in our data. Instead, the data reveal how developers and vendors share some responsibilities—particularly around transparency, ethical standards, and traceability—but cannot be held liable for every misuse, reflecting the aggregated dimensions of Responsibility and Control as well as Regulation and Governance. This viewpoint aligns with arguments that manufacturers, operators, and end users collectively shape AI's ethical landscape (Bleher & Braun, 2022). Respondents particularly advocate clearer regulation and governance in critical domains like healthcare or autonomous driving (Müller, 2023; Widdau & Zednik, 2022). Current legal frameworks—such as the European AI Act—promise to help close these gaps, but interviewees find them insufficiently harmonized or binding.

The findings show that AI is widely perceived as a supportive and expanding people's decision-making capabilities rather than supplanting them, captured by the aggregated dimension of Human-AI Relationship and Acceptance. Most interviewees referred to AI systems such as large language models, such as ChatGPT, Perplexity, or Gemini. The hypothesized "responsibility gap" rarely surfaces in organizational practices because employees see themselves as answerable for both the choices they make and the ways they employ AI (aggregated dimension of Responsibility and Control). This outcome illustrates how technology broadens possibilities while leaving ethical and organizational accountability with human (Liang & Weber, 2024). Notably, the increased scrutiny, vigilance, and sense of accountability described by participants suggest that adopting AI often heightens awareness of one's professional and moral duties.

From a theoretical stance, these results highlight how relational and affective agency (Pignot & Thompson, 2024; Schatzki, 2025) unfold in everyday AI usage. First, agency is distributed: developers, users, and AI systems co-create decisions. Second, affect and emotional responses—ranging from trust to scepticism—shape how individuals perceive their responsibility. While many appreciate data-driven insights, they remain cautious about errors creeping in. This tension reinforces the proposition that agency is never purely rational; it is conditioned by affect and unconscious motives as well. Moreover, interviewees do not advocate for removing human judgment from the loop; on the contrary, they express that final decisions should remain the responsibility of people trained to interpret AI outputs critically. This preference echoes broader responsible AI frameworks, where technology serves as a powerful assistant but does not supplant the ethical and interpretive roles of human decision-makers (Papagiannidis et al., 2025), all within the aggregated dimension of Human-AI Relationship and Acceptance, though closely intertwined with elements of Responsibility and Control Regarding developers' roles, respondents typically consider them "co-responsible" rather than solely liable. Calls for explainable AI (Barredo Arrieta et al., 2020; Hulsen, 2023) suggest that users want transparency to understand how AI systems operate. Yet, our respondents recognize that developers cannot anticipate all misuse scenarios. Instead, a multilayered governance model—encompassing design, deployment, and usage—is seen as key to establishing trust and accountability (aggregated dimension of Regulation and Governance).

Figure 1 visualizes our theoretical contribution: We identify the negative ontology in AI's black box nature and the lack of the users' understanding. AI's inscrutable and

autonomous character has constituted the responsibility gap so far. We expand the prior understanding by explaining that users derive a sense of responsibility from the lack of understanding AI (i.e., negative ontology). The affective positionality toward AI that is conceptualized in the affective agency bridges the responsibility gap, explaining the change in responsibility. With respect to our RQ, the user's sense of responsibility shifts from covering only their own decisions to encompassing AI's output as well. The modified scope of responsibility to also include AI is driven by the affective positionality (i.e., scepticism and trust), which is engendered by the negative ontology imposed by the inscrutability of AI.

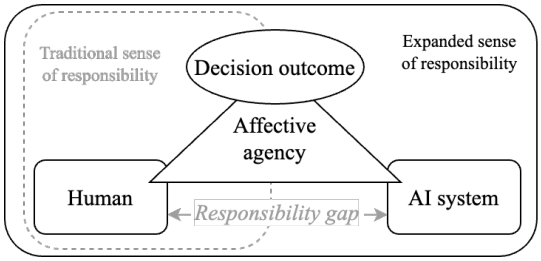

Figure 1. Bridging the responsibility gap.

Following the interviewees' call for sharing some responsibility with developers and vendors, we invite future research to attend to this matter. It aligns well with the genealogical approach proposed by Scott and Orlikowski (2025), which aims to uncover the sociomaterial origins of AI, potentially revealing responsibility in the development of AI systems. By studying how users, managers, and developers jointly navigate potential gaps, scholars and practitioners alike can better ensure responsible decision-making where the technological and the human remain in productive, accountable tension.

6 Conclusion

We explored how individuals assess responsibility in AI-assisted decisions, finding that responsibility does not simply shift away from humans but calls for new forms of critical engagement and oversight. Using a relational sociomaterial lens, our results show that (1) a negative ontology (emerging from uncertainty due to AI's inscrutability) and (2) affective positionalities (such as trust or scepticism) encourage deeper scrutiny of AI outputs and, in turn, bridge the responsibility gap. While agency is distributed across humans and technology, responsibility remains anchored in human oversight. In practice, organizations may establish clear role definitions and liability structures that enable this anchor across contexts. Fostering affective reflexivity through dedicated trainings enables critical engagement with AI outputs, encouraging practitioners to recognize and interrogate emotions such as trust or scepticism as integral to responsible human–AI collaboration. Our study, though limited in the number of respondents, underscores that a responsible use of AI systems requires critical engagement and challenging AI's outcomes.

References

Ågerfalk, P. J. (2020), Artificial intelligence as digital agency. *European Journal of Information Systems*, 29(1), 1–8. https://doi.org/10.1080/0960085X.2020.1721947.

Antonetti, P., & Maklan, S. (2014), Feelings that Make a Difference: How Guilt and Pride Convince Consumers of the Effectiveness of Sustainable Consumption Choices. *Journal of Business Ethics*, 124(1), 117–134. https://doi.org/10.1007/s10551-013-1841-9.

Baroni, I., Re Calegari, G., Scandolari, D., & Celino, I. (2022), AI-TAM: A model to investigate user acceptance and collaborative intention inhuman-in-the-loop AI applications. *Human Computation*, 9(1), 1–21. https://doi.org/10.15346/hc.v9i1.134.

Barredo Arrieta A, Díaz-Rodríguez N, Del Ser J, Bennetot A, Tabik S, Barbado A, Garcia S, Gil-Lopez S, Molina D, Benjamins R, Chatila R, Herrera F. (2020) Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion* 58: 82–115.

Berente, N, Gu, B, Recker, J & Santanam, R (2021), Managing Artificial Intelligence, *MIS Quarterly*, 45(3). https://doi.org/10.25300/MISQ/2021/16274.

Bleher, H, & Braun, M (2022), Diffused responsibility: attributions of responsibility in the use of AI-driven clinical decision support systems. *AI and Ethics*, 2(4), 747–761. https://doi.org/10.1007/s43681-022-00135-x.

Blumer H (1954) What is Wrong with Social Theory? *American Sociological Review* 19(1). 3–10. https://doi.org/10.2307/2088165.

Champagne, M., & Tonkens, R. (2015), Bridging the Responsibility Gap in Automated Warfare. *Philosophy & Technology*, 28(1), 125–137. https://doi.org/10.1007/s13347-013-0138-3.

Chen, A., Wang, C. and Zhang, X. (2023) 'Reflection on the equitable attribution of responsibility for artificial intelligence-assisted diagnosis treatment decisions', *Intelligent Medicine*, 3(2), pp. 139–143.

Coeckelbergh, M. (2020) 'Artificial intelligence, responsibility attribution, and a relational justification of explainability', *Science and Engineering Ethics*, 26(4), pp. 2051–2068.

Crompton, L. (2021), The decision-point-dilemma: Yet another problem of responsibility in human-AI interaction. *Journal of Responsible Technology*, 7–8, 100013. https://doi.org/10.1016/j.jrt.2021.100013.

Dong, M., & Bocian, K. (2024), Responsibility gaps and self-interest bias: People attribute moral responsibility to AI for their own but not others' transgressions. *Journal of Experimental Social Psychology*, 111, 104584. https://doi.org/10.1016/j.jesp.2023.104584.

Epstein, C. (2018), The productive force of the negative and the desire for recognition: Lessons from Hegel and Lacan. *Review of International Studies,* 44(5), 805–828. https://doi.org/10.1017/S0260210518000347.

Fischer, L. H., Wunderlich, N. and Baskerville, R. (2023), 'Artificial Intelligence and Digital Work: The Sociotechnical Reversal', in *Proceedings of the 56th Hawaii International Conference on System Sciences*. Manoa, HI. https://doi.org/10.24251/HICSS.2023.027.

Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*. https://doi.org/10.1162/99608f92.8cd550d1

Floridi, L., Cowls, J., King, T. C., & Taddeo, M. (2020). How to Design AI for Social Good: Seven Essential Factors. *Science and Engineering Ethics*, 26(3), 1771–1796. https://doi.org/10.1007/s11948-020-00213-5

Gervais, D.J. and Nay, J.J. (2023) 'Artificial intelligence and interspecific law', *Science*, 382(6669), pp. 376–378. https://doi.org/10.1126/science.adi8678

Gioia, D.A. et al. (2013), Seeking Qualitative Rigor in Inductive Research. *Organizational Research Methods*, 16(1). https://doi.org/10.1177/1094428112452151.

Goldkuhl G and Cronholm S (2019) Grounded Theory in Information Systems – from Themes in IS Discourse to Possible Developments. In: *Proceedings of the Fortieth International Conference on Information Systems*, Munich, Germany, 2019.

Gualdi, F. and Cordella, A. (2021), 'Artificial Intelligence and Decision-Making: The Question of Accountability', in *Proceedings of the 54th Hawaii International Conference on System Sciences*. Manoa, HI, 2297–2306. https://doi.org/10.24251/HICSS.2021.281.

Gürkan, N. & Suchow, J. W. (2024), 'Exploring Public Opinion on Responsible AI Through the Lens of Cultural Consensus Theory', In: *Proceedings of the 57th Hawaii International Conference on System Sciences*. Manoa, HI. https://doi.org/10.24251/HICSS.2024.085.

Hafner, A., Hammerschmidt, T. & Stolz, K. (2024), 'Analyzing the use of ethical theories within AI ethics research: A systematic scoping review', In: *Wirtschaftsinformatik 2024 Proceedings*, 46. https://aisel.aisnet.org/wi2024/46.

Huang, X. and Suang, H.C. (2024), 'A Systematic Review on Responsibility Attribution in AI Service and Human-AI Joint Service', In: *The Proceedings of the Thirty-Second European Conference on Information Systems*, Paphos, Cyprus. https://aisel.aisnet.org/ecis2024/track15_social_ict/track15_social_ict/11.

Hulsen, T. (2023), 'Explainable Artificial Intelligence (XAI): Concepts and Challenges in Healthcare', *AI*, 4(3), 652–666. https://doi.org/10.3390/ai4030034.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

Jöhnk J, Weißert M and Wyrtki K (2021) Ready or Not, AI Comes—An Interview Study of Organizational AI Readiness Factors. *Business & Information Systems Engineering,* 63(1): 5–20. https://doi.org/10.1007/s12599-020-00676-7

Kellogg, K. C., Valentine, M. A., & Christin, A. (2020). Algorithms at Work: The New Contested Terrain of Control. *Academy of Management Annals*, 14(1), 366–410. https://doi.org/10.5465/annals.2018.0174

Kim B (Raymond), Srinivasan K, Kong SH, et al. (2023) ROLEX: A Novel Method for Interpretable Machine Learning Using Robust Local Explanations. *MIS Quarterly* 47(3): 1303–1332. https://doi.org/10.25300/MISQ/2022/17141

Krügel, S., Ammeling, J., Aubreville, M., Fritz, A., Kießig, A., & Uhl, M. (2024), Perceived responsibility in AI-supported medicine. *AI & SOCIETY*, 40, 1485–1495. https://doi.org/10.1007/s00146-024-01972-6.

Laclau, E. (2000), Identity and Hegemony: The Role of Universality in the Constitution of Political Logics: Vol. J. P. Butler, E. Laclau, S. Zizek (Eds.), Contingency, Hegemony, Universality: Contemporary Dialogues on the Left.

Li, Z., Cui, X. & Hsieh, J. J. (2024), 'Whose Responsibility, Humans or Automation?', In: *Proceedings of the 45th International Conference on Information Systems*. Bangkok, Thailand. https://aisel.aisnet.org/icis2024/humtechinter/humtechinter/27/.

Liang T-P and Weber RA (2024), A Beginner's Guide To Agential Realism. *Communications of the Association for Information Systems*, 55(1): 978–003. https://doi.org/10.17705/1CAIS.05537.

Lüthi, N., Matt, C., Myrach, T. and Junglas, I. (2023), 'Augmented Intelligence, Augmented Responsibility?', *Business & Information Systems Engineering*, 65(4), pp. 391–401. https://doi.org/10.1007/s12599-023-00789-9.

Mikalef P, Conboy K, Eriksson Lundström J, and Popovič A. (2022) Thinking responsibly about responsible AI and 'the dark side' of AI. *European Journal of Information Systems*, 31(3): 257–268. https://doi.org/10.1080/0960085X.2022.2026621.

Mittelstadt, B. (2019). Principles alone cannot guarantee ethical AI. *Nature Machine Intelligence*, 1(11), 501–507. https://doi.org/10.1038/s42256-019-0114-4

Müller VC (2023) Ethics of Artificial Intelligence and Robotics. In: Zalta EN and Nodelman U (eds) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University. Available at: https://plato.stanford.edu/archives/fall2023/entries/ethics-ai/.

Niemimaa M (2016), Sociomateriality and Information Systems Research: Quantum Radicals and Cartesian Conservatives. *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, 47(4): 45–59.

Orlikowski WJ and Scott SV (2008), Sociomateriality: Challenging the Separation of Technology, Work and Organization. *The Academy of Management Annals,* 2(1): 433–474.

Papagiannidis, E., Mikalef, P., & Conboy, K. (2025), Responsible artificial intelligence governance: A review and research framework. *The Journal of Strategic Information Systems*, 34(2), 101885. https://doi.org/10.1016/j.jsis.2024.101885.

Pignot, E., & Thompson, M. (2024), Affect and relational agency: How a negative ontology can broaden our understanding of IS research. *Information and Organization*, 34(1), 100500. https://doi.org/10.1016/j.infoandorg.2023.100500.

Razmetaeva, Y., & Satokhina, N. (2022), AI-Based Decisions and Disappearance of Law. *Masaryk University Journal of Law and Technology*, 16(2), 241–267. https://doi.org/10.5817/MUJLT2022-2-5.

Santoni de Sio, F. and Mecacci, G. (2021) 'Four responsibility gaps with artificial intelligence: Why they matter and how to address them', Philosophy and Technology, 34, pp. 1057–1084.

Schatzki, T. R. (2025), Agency. *Information and Organization*, 35(1), 100553. https://doi.org/10.1016/j.infoandorg.2024.100553.

Schinagl, S., van den Hooff, B., Shahim, A. and Khapova, S. (2024), 'Governing Security in a Digital World Is Not a "One Size Fits All" Solution: A Multiple Case Study of Security Governance Modes and Their Effectiveness', in *Proceedings of the 57th Hawaii International Conference on System Sciences*. Manoa, HI, 6533–6542. https://doi.org/10.24251/HICSS.2024.783.

Scott SV and Orlikowski WJ (2025) Exploring AI-in-the-making: Sociomaterial genealogies of AI performativity. *Information and Organization,* 35(1) 100558. https://doi.org/10.1016/j.infoandorg.2025.100558.

Schulte-Derne, D. and Gnewuch, U. (2024) 'Translating AI Ethics Principles into Practice to Support Robotic Process Automation Implementation', *MIS Quarterly Executive*, 23(2), 205–221.

Smith, H., Birchley, G., & Ives, J. (2023), Artificial intelligence in clinical decision-making: Rethinking personal moral responsibility. *Bioethics*, 38(1), 78–86. https://doi.org/10.1111/bioe.13222.

Taylor, I. (2024a), Collective Responsibility and Artificial Intelligence. *Philosophy & Technology*, 37(1), 27. https://doi.org/10.1007/s13347-024-00718-y.

Thuis, T., Li, T. & van Heck, E. (2023), 'Who Takes Responsibility for AI? A Field Study on AI-Related Task Shifts, Explainability, and Responsibility Attributions', In: *Proceedings of the 44th International Conference on Information Systems*. Hyderabad, India. https://aisel.aisnet.org/icis2023/aiinbus/aiinbus/3

Vallor S and Bekey GA (2017) Artificial Intelligence and the Ethics of Self-Learning Robots. In: Lin P (ed.) Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence. Oxford University Press, pp. 338–353. https://doi.org/10.1093/oso/9780190652951.003.0022

Verdicchio M and Perin A (2022) When Doctors and AI Interact: on Human Responsibility for Artificial Risks. *Philosophy & Technology*, 35(1): 11. https://doi.org/10.1007/s13347-022-00506-6.

Weber R (2020), Taking the ontological and materialist turns: Agential realism, representation theory, and accounting information systems. *International Journal of Accounting Information Systems*, 39: 100485. https://doi.org/10.1016/j.accinf.2020.100485.

Widdau, C. S., & Zednik, C. (2022), Opake Systeme künstlicher Intelligenz und das Problem der Verantwortungslücke. In: Algorithmen und Autonomie: Interdisziplinäre Perspektiven auf das Verhältnis von Selbstbestimmung und Datenpraktiken (pp. 121–134). Verlag Barbara Budrich. https://doi.org/10.2307/j.ctv2xh53x1.9.

Zeiser, J. (2024), Owning Decisions: AI Decision-Support and the Attributability-Gap. *Science and Engineering Ethics*, 30(4), 27. https://doi.org/10.1007/s11948-024-00485-1.