

AI Accountability: Navigating Legal, Ethical, and Transparency Challenges

Priyanka Sajnani¹

¹School of Computer Science Engineering and Application, D Y Patil International University Akurdi pune 411035

Abstract - Artificial Intelligence (AI) continues to revolutionise industries, presenting immense opportunities while simultaneously introducing intricate challenges of accountability, transparency, and ethics. This review explores the multifaceted landscape of AI accountability by synthesising insights from diverse domains, including legal liability, intellectual property, ethical considerations, and regulatory frameworks. Key themes include the complexities of ascribing accountability in AI-driven systems, especially in scenarios involving criminal liability and decision-making failures, and the proposal of electronic legal personhood for autonomous systems. The interplay between human biases, such as self-serving attribution in generative AI use, and the sociotechnical nature of AI governance is examined to reveal gaps in current accountability frameworks. Further, this paper delves into the regulatory implications of transparency, particularly under frameworks like the General Data Protection Regulation (GDPR), and the emerging need for relational transparency that extends beyond information provision to meaningful communication. Ethical considerations in sectors like financial services and intellectual property law are critically analysed, emphasising the need for clear guidelines to navigate algorithmic bias, fairness, and trustworthiness. Through this synthesis, the review proposes a novel framework integrating legal, ethical, and societal dimensions to redefine AI accountability. This approach underscores the importance of proactive governance mechanisms, such as enhanced transparency measures, context-sensitive accountability architectures, and industry-wide ethical standards, to foster responsible AI innovation.

Key Words: AI accountability, legal liability, transparency, ethical AI, intellectual property, GDPR, criminal responsibility, governance mechanisms, algorithmic fairness.

1. INTRODUCTION

Artificial intelligence (AI) has become an integral part of modern life, driving innovation and reshaping industries ranging from healthcare and education to business and governance. Among the most transformative advancements within AI is the development of Generative AI (GAI), which leverages deep learning and machine learning algorithms to produce novel content, including text, images, and audio, based on user input. Tools like ChatGPT, GitHub Copilot, and DALL-E have garnered significant attention due to their potential to enhance productivity, creativity, and efficiency

in tasks such as writing, coding, and designing. This rapid adoption of GAI tools highlights their transformative role in improving problem-solving and decision-making capabilities in the workplace.

However, the integration of GAI systems into human activities brings forth challenges related to ethics, reliability, and accountability. Instances such as biased outputs in recommendation letters or the spread of misinformation underscore the risks associated with GAI misuse or overreliance. These issues highlight the importance of ensuring accountability in AI systems to mitigate potential harms. Accountability in AI encompasses the obligation of actors involved in the design, deployment, or use of such systems to explain their actions, justify decisions, and take responsibility for adverse outcomes. Without well-defined accountability frameworks, addressing the root causes of AI failures becomes challenging, thereby increasing the likelihood of repetition and escalating risks.

A key obstacle in establishing accountability within GAI is the "many hands" problem, wherein responsibility is diffused among multiple stakeholders, such as developers, organisations, and users. This complexity is exacerbated by the opaque nature of GAI systems, which often lack transparency in decision-making processes. For example, when GAI-generated outputs result in errors or damages, it becomes difficult to pinpoint responsibility. This diffusion of accountability can lead to scenarios where individuals underestimate their role in the proper use of these systems, fostering overreliance and potentially detrimental consequences, such as automation bias or data breaches.

Understanding how accountability is perceived and distributed among stakeholders is critical for fostering ethical and sustainable AI usage. Previous research has predominantly focused on the accountability of single actors, such as developers or organisations, and has explored how accountability influences trust, adoption, and behaviour. However, less attention has been given to multi-actor accountability frameworks, particularly in the context of Generative AI. Examining how employees attribute accountability to different actors, such as themselves, the AI system, or its creators, offers valuable insights into addressing the "many hands" problem. By exploring these dynamics, we can advance frameworks that promote equitable responsibility-sharing and minimise bias in human-AI collaborations.

This paper aims to contribute to this growing discourse by investigating the perception of accountability among employees when interacting with GAI systems. Drawing upon attribution theory and concepts like self-serving bias, this study seeks to unravel the factors shaping accountability attributions and their implications for the ethical integration of GAI in professional and organisational contexts.

2. BACKGROUND

Generative AI and Its Role in Transformative Work
Generative AI (GAI) has emerged as a powerful subset of artificial intelligence, capable of creating human-like content such as text, music, and visuals while synthesizing data for analytical insights. Unlike traditional AI, GAI leverages advanced pattern recognition to generate unique and original outputs, demonstrating transformative potential across diverse domains (Dasborough, 2023). The adoption of GAI in professional environments has been associated with increased efficiency, accessibility, and cost-effectiveness, driving significant advancements in digital evolution and business processes (Goldstein et al., 2023). For instance, in Human Resource Management, GAI enhances tasks such as job description creation and intelligent search, empowering professionals to innovate and shift their focus toward high-value, creative work (Saengrith et al., 2022).

At the individual level, GAI fosters operational excellence by augmenting decision-making and problem-solving capabilities, marking a paradigm shift in workplace dynamics. Its integration not only optimizes repetitive tasks but also redefines roles, enabling employees to engage in more strategic and meaningful activities (Chen et al., 2023). This dual capacity to generate and analyze content makes GAI an indispensable tool in modern organizational transformation.

to the "problem of many hands," where multiple actors contribute to the system, creating difficulty in attributing responsibility (Cooper et al., 2022).

This ambiguity is particularly evident in disagreements over who should be held accountable for AI outcomes. While some argue that algorithm developers bear primary responsibility (Diakopoulos, 2016), others contend that designers and managers also share accountability for societal impacts (Shin & Park, 2019). The legal framework adds another layer of complexity, as algorithms lack legal personhood, limiting their accountability under current laws (Deibel, 2021). Despite these challenges, studies in human-computer interaction suggest that people often perceive AI as a social actor, applying expectations and accountability norms typically reserved for humans (Moon, 2003).

This multifaceted debate highlights the urgency of clarifying accountability mechanisms for GAI systems. As organizations increasingly rely on these technologies, the need for transparent frameworks that delineate responsibilities among stakeholders is paramount.

Challenge	Description	Example
Diffused Responsibility	Shared accountability across multiple actors makes pinpointing liability difficult	Autonomous vehicle accidents
Lack of Legal Clarity	Absence of specific laws for AI systems	Unregulated AI tools in healthcare
Ethical Concerns	Potential for bias and unethical decisions	Algorithmic discrimination in hiring

Table 2: Challenges in AI Accountability and Legal Context

3. LITERATURE REVIEW

Accountability remains a pivotal focus in AI governance, emphasizing transparency, fairness, and adherence to ethical standards. Sinclair (1995) and Theodorou & Dignum (2020) highlight the need for structured accountability frameworks to navigate AI's sociotechnical complexity. Current policies, especially within the European Union, advocate for robust accountability mechanisms that include compliance, reporting, oversight, and enforcement to mitigate potential risks. These mechanisms aim to enhance trust and ensure AI systems align with societal values. However, challenges persist in defining accountability across diverse contexts, illustrating gaps in harmonizing its conceptualization across sociotechnical and legal frameworks.

The rapid evolution of AI technologies has exposed inadequacies in current legal systems, particularly regarding liability for AI's autonomous actions. Baldwin (2019) and Sukhodolov et al. (2020) explore scenarios such as accidents

Domain	Key Tool	Benefits
Human Resources (HR)	ChatGPT	Automated resume screening
Education	Duolingo AI	Personalized learning pathways
Content Creation	DALL-E	Efficient visual content generation

Table 3: Generative AI: Transformative Use Cases and Benefits

AI Accountability and the Complex Web of Responsibility
The integration of GAI introduces challenges in accountability, particularly in addressing ethical concerns and ambiguities surrounding responsible actors. Accountability in AI systems is conceptualized as the obligation of involved parties—developers, users, and decision-makers—to explain, justify, and take responsibility for outcomes influenced by AI, as defined by Bovens (2007). However, achieving accountability becomes convoluted due

caused by self-driving cars, where culpability remains unclear. The legal debate extends to whether AI entities could be granted legal personhood or subjected to modified criminal liability principles. Existing frameworks often fail to account for the distributed responsibilities among developers, users, and operators, highlighting a pressing need for updated regulations. These gaps raise significant concerns about the effectiveness of current laws in addressing the nuanced challenges posed by AI, particularly in its potential to autonomously commit harmful actions.

Ethical governance in AI intertwines deeply with its legal dimensions. Researchers, including Al-Qusi (2018), emphasize preserving human rights and prioritizing societal welfare as foundational to AI system development. The IEEE SA (2016) further advocates embedding ethical principles into AI design, ensuring safety and justice remain uncompromised. The discussion extends to AI's potential to operate autonomously, necessitating a proactive approach to regulation that incorporates both technological innovation and ethical accountability. A sociotechnical perspective, integrating diverse stakeholders' input, is essential for crafting governance mechanisms that are equitable and effective.

While existing studies provide valuable insights, significant gaps remain in integrating accountability frameworks with evolving AI capabilities. Current research often overlooks the practical implementation of these frameworks across diverse sociotechnical systems. Additionally, limited consensus exists on the scope and nature of AI's legal personality, leaving ambiguities in its application. Future research must address these inconsistencies by proposing adaptive, multidisciplinary solutions that align legal, ethical, and technical dimensions.

This review underscores the critical interplay between accountability, legal adaptability, and ethical standards in AI governance. The subsequent discussion explores strategies for bridging identified gaps, focusing on developing cohesive frameworks that balance innovation with societal values.

4. METHODOLOGY

The methodology for this research review paper involved a systematic approach to collecting and analyzing relevant literature on the themes of accountability, legal frameworks, and ethical implications in Artificial Intelligence (AI). The primary focus was to identify studies addressing governance mechanisms, liability issues, and the ethical design of AI systems. A thematic organization was employed to synthesize insights and highlight interconnections between the identified themes.

The following databases were extensively used:

Databases: IEEE Xplore, SpringerLink, ScienceDirect, JSTOR, and Google Scholar.

Search Keywords: "AI accountability frameworks," "legal implications of AI," "ethical AI governance," "AI liability," "sociotechnical systems in AI," and "autonomous systems regulations."

Selection Criteria:

Peer-reviewed articles published between 2015 and 2024.

Studies focusing on European AI governance and broader global perspectives.

Articles addressing practical case studies, such as self-driving cars and autonomous AI systems.

Literature emphasizing multidisciplinary approaches involving legal, ethical, and technical dimensions.

This systematic approach was chosen to ensure the inclusion of high-quality, relevant studies that comprehensively address the multifaceted nature of AI governance. Peer-reviewed sources were prioritized to uphold academic rigor. The combination of thematic and chronological organization allowed for a clear analysis of evolving discussions and emerging gaps. By integrating diverse perspectives, this methodology ensured a robust foundation for identifying inconsistencies and proposing actionable solutions in the subsequent discussion.

5. DISCUSSION

This study advances our understanding of how accountability attributions function in the context of Generative AI (GAI) systems, particularly when deployed in professional environments. Rooted in attribution theory and the self-serving bias, our findings provide nuanced insights into how employees navigate accountability in success and failure scenarios, shedding light on the evolving dynamics of human-AI collaboration.

Framework	Key Features	Strengths	Limitations
Responsible AI Design	Ensures ethical principles in AI development	Proactive risk mitigation	Implementation challenges
Explainable AI (XAI)	Focus on transparency and interpretability	Enhances user trust	Limited applicability for complex models
Legal Personhood Model	Treats AI as a legal entity	Clarifies liability	Contradicts traditional legal systems

Table 1: AI Accountability Frameworks and Their Key Features

Principal Findings

Our results reveal an unexpected departure from traditional self-serving bias. Contrary to established research, employees attributed accountability primarily to themselves, regardless of whether the outcome of using GAI was success or failure. While self-attribution during success aligns with the desire to preserve self-worth, the attribution of accountability during failures, often linked to self-critical reflection, indicates a deeper sense of professional responsibility. This finding underscores the unique implications of GAI in fostering heightened personal accountability among users, a behavior less commonly observed with non-GAI systems.

Interestingly, participants demonstrated an acute awareness of shared accountability, recognizing the interconnected roles of developers, organizations, and the AI systems themselves. These findings support the "many hands" problem, emphasizing the collective responsibility inherent in GAI system development and use. Notably, higher levels of shared accountability were observed in success scenarios, reflecting users' efforts to balance personal achievements with a collaborative acknowledgment of external contributions.

Contributions to Research

This study contributes to accountability literature in several ways. First, it provides empirical evidence that GAI systems elicit accountability dynamics distinct from other AI technologies. Unlike traditional AI, GAI's opaque decision-making processes amplify users' perceptions of control and accountability, suggesting the need for more tailored frameworks to address the unique characteristics of GAI systems.

Second, while existing studies often examine accountability in isolation, focusing on individual actors, this research highlights the interplay between multiple stakeholders. By demonstrating that employees simultaneously attribute accountability to themselves, developers, and organizations, this study enriches our understanding of multi-actor accountability attribution. It also provides a foundation for future research on balancing personal and collective accountability in AI governance.

Lastly, this study offers a novel perspective by examining accountability in both success and failure contexts, revealing distinct attribution patterns. This dual-outcome approach provides a more comprehensive understanding of accountability behaviors, enabling future researchers to explore how different outcomes influence long-term accountability perceptions.

Practical Implications

Our findings have several implications for practice. For employees, the tendency to self-attribute accountability highlights the importance of fostering critical engagement with GAI systems. Organizations should encourage employees to view GAI as a tool for augmentation rather than automation, promoting informed decision-making and reducing over-reliance on AI outputs.

For developers, the shared attribution of accountability underscores the need to design systems that align with users' expectations and facilitate transparency. Developers should proactively address concerns about GAI's opaque nature, ensuring users feel supported and informed in their interactions with these systems.

Organizations must also recognize their role in shaping accountability perceptions. By providing training, establishing clear guidelines, and fostering a culture of shared responsibility, organizations can enhance trust in GAI systems and improve their integration into workplace practices.

Limitations and Future Research

This study is not without limitations. First, the reliance on a single experimental context limits the generalizability of our findings. Future studies should replicate this research across diverse industries, cultural settings, and GAI applications to validate our conclusions. Additionally, while our methodology focused on employees' self-reports, incorporating observational or longitudinal approaches could provide richer insights into accountability dynamics over time.

Second, the complexity of accountability attributions in GAI systems warrants further exploration. Future research could investigate how factors such as organizational culture, task complexity, and user expertise influence attribution patterns. Incorporating moderators like self-efficacy or cognitive load could also deepen our understanding of accountability behaviors in different contexts.

Finally, the role of GAI in reshaping team dynamics remains an open question. Future studies could explore how shared accountability among team members affects collaboration and decision-making, particularly in high-stakes environments. Addressing these questions will be crucial for developing holistic accountability frameworks that support ethical and effective GAI integration.

REFERENCES

[1] L. McDonald, "AI systems and liability: An assessment of the applicability of strict liability & a case for limited legal personhood for AI," *University of St Andrews Law Journal*, vol. 3, no. 1, 2023, pp. 5–21, doi:10.15664/stalj.v3i1.2645.

[2] N. S. Uzougbo, C. G. Ikegwu, and A. O. Adewusi, "Legal accountability and ethical considerations of AI in financial services," *GSC Advanced Research and Reviews*, vol. 19, no. 2, 2024, pp. 130–142, doi:10.30574/gscarr.2024.19.2.0171.

[3] G. Du, S. Lins, I. Blohm, and A. Sunyaev, "My fault, not AI's fault. Self-serving bias impacts employees' attribution of AI accountability," *ICIS 2024 Proceedings*, no. 27, 2024.

[4] G. N. Reeke Jr. and G. M. Edelman, "Real brains and artificial intelligence," *Daedalus*, vol. 117, 1988, pp. 143–173, [Online]. Available: <https://www.jstor.org/stable/20025142>.

[5] H. Felzmann, E. F. Villaronga, C. Lutz, and A. Tamo-Larrieux, "Transparency you can trust: Transparency requirements for artificial intelligence between legal norms and contextual concerns," *Manuscript submitted for publication*, 2024.

[6] C. Novelli, M. Taddeo, and L. Floridi, "Accountability in artificial intelligence: What it is and how it works," *Journal of Legal Studies, University of Bologna*, vol. 27, 2023, pp. 401–418.

[7] Y. Ellamey and A. Elwakad, "The criminal responsibility of artificial intelligence systems: A prospective analytical study," *Corporate Law & Governance Review*, vol. 5, no. 1, 2023, pp. 92–100, doi:10.22495/clgrv5i1p8.

[8] J.-C. Bélisle-Pipon, E. Monteferrante, M.-C. Roy, and V. Couture, "Artificial intelligence ethics has a black box problem," *AI & Society*, vol. 38, no. 6, 2022, pp. 1507–1522.

[9] M. Busuioc, "Accountable artificial intelligence," *Public Administration Review*, vol. 81, no. 5, 2021, pp. 825–836, doi:10.1111/puar.13437.

[10] N. Diakopoulos, "Accountability in algorithmic decision making," *Communications of the ACM*, vol. 59, no. 2, 2016, pp. 56–62.

[11] K. E. Martin, "Ethical implications and accountability of algorithms," *Journal of Business Ethics*, vol. 160, no. 4, 2018, pp. 833–850.

[12] B. D. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, 2016, pp. 1–21, doi:10.1177/2053951716679679.