Human Factors and Systems Interaction, Vol. 154, 2024, 77–84
https://doi.org/10.54941/ahfe1005355

Trusting AI: Factors Influencing Willingness of Accountability for AI-Generated Content in the Workplace

Ulrike Aumüller[1] and Eike Meyer[2]

[1]Forschungs- und Entwicklungszentrum FH Kiel GmbH, Kiel, Germany
[2]AI Applications R&D Centre at the University of Applied Sciences, Kiel, Germany

ABSTRACT

In the rapidly evolving landscape of Artificial Intelligence (AI) and business ethics, a critical area of focus has emerged: the willingness of leadership to assume responsibility for AI-generated content in decision-making processes. While the current public discourse predominantly addresses AI's impact on customer service, potential biases, and job displacement, etc., a less explored yet significant aspect is how AI reshapes tasks and roles within organizations, particularly in decision-making. AI's capability to analyse vast data sets expeditiously supports both operational and strategic decisions across various sectors. However, this support comes with ambivalent outcomes, ranging from enhanced efficiency to risk of taking decisions with negative business impact based on AI outputs with hidden biases. Such ambiguity can undermine trust in AI, especially when the rationale behind AI-generated recommendations is opaque. The central question of this paper will be the extent to which leaders are prepared to be accountable for decisions made based on AI insights. This encompasses scenarios where leaders themselves develop AI-driven decision foundations as well as when such bases are prepared by their teams. Additionally, the perspective of top management will be examined, specifically their expectations regarding the plausibility and responsibility attribution in AI-supported decision-making, even when the decision path is not entirely transparent. In understanding the adoption of AI in decision-making, key factors influencing trust and usage of algorithms emerge. Research suggests that trust extends beyond algorithm accuracy, significantly influenced by social validation such as prior adoption by others, which can reduce cognitive load and improve engagement. Furthermore, cultural and age differences play a crucial role; for example, older adults may exhibit higher trust in automation compared to younger individuals. Additionally, the expectation of near-perfection performance from automated systems can lead to scepticism, especially when an algorithm falters, impacting ongoing trust and usage. These elements are vital in evaluating the readiness to assume responsibility for AI-generated decisions in the workplace. This paper aims to identify and categorize those criteria that have a central effect on the willingness to assume responsibility for AI facilitated decisions and AI generated content in companies. Those categories may later serve as a framework to be considered by management when adopting a strategy concerning their policies for AI based decision-making processes. do not use titles for name, such as Ph.D., Prof., etc. **The recommended page size of full papers is 6 pages with a maximum of 10 pages,** including all paper components such as references, appendices, acknowledgements. When preparing your paper please use Harvard referencing style (unnumbered, alphabetized).

Keywords: GenAI, AI leadership, Accountability, AI adoption, AI readiness

INTRODUCTION

The advent of artificial intelligence (AI) in the workplace has heralded a new era in the dynamics of content creation and management, bringing to the forefront the critical issue of assuming responsibility for AI-generated content. This paper embarks on an exploration of the factors influencing employees' willingness to assume such responsibility. Drawing upon the framework of technology acceptance models (Davis, 1989), we delve into individual psychological attitudes towards AI, reflecting on the perceived ease of use and perceived usefulness as determinants (Venkatesh and Davis, 2000), and the influence of AI as an automated system and the additional dimensions this brings to the acceptance not only of the technology itself, but identifications with its' outputs as part of organizational processes and activities.

Furthermore, the study expands its scope by incorporating insights from organizational behaviour studies (Robbins & Judge, 2019), examining how organizational culture, ethical guidelines, and leadership styles collectively shape this willingness. Acknowledging the complexity of AI ethics and governance (Mittelstadt *et al.*, 2016), the research also considers broader ethical and legal frameworks impacting employee responsibility. Three focus research questions have been investigated as part of the project:

- Which traits in leaders enhance willingness to assume responsibility for AI based decisions?
- Which traits of AI systems are influencing the identification with AI based decisions?
- Which role does an organisation's culture play in the willingness to assume responsibility?

Through this multifaceted approach, the study aims to contribute to the literature on AI in the workplace by offering a nuanced understanding of the interplay between human and artificial agents in professional settings, thus providing valuable insights for organizations in the effective and responsible integration of AI technologies.

Which Traits in Leaders Enhance Willingness to Assume Responsibility for AI Based Decisions?

In the context of integrating Artificial Intelligence (AI) into decision-making processes, a pivotal question arises: What traits in leaders amplify their willingness to assume responsibility for AI-generated decisions? The exploration of this question leads into the complex interplay of human psychology, technology acceptance, and leadership qualities within an organizational setting.

Some hints to answering this question may be found in a study carried out by Alexander et al. The authors examined neurophysiological reactions and task execution as their participants were engaged in solving mazes (Alexander *et al.*, 2018). They found some patterns that may potentially be applied to the adoption of AI as well.

One of the traits they looked at was trust: The willingness of a person to trust and subsequently take responsibility for AI-based decisions is intricately

linked to their perception of the technology. The study illustrates some of the nuanced factors influencing this trust. Notably, the adoption of algorithms was significantly influenced by social validation – information about others' adoption showed to reduce the subjects' cognitive load and enhanced performance. This insight suggests that leaders who value social proof and collective experience might demonstrate greater openness to integrating AI into their decision-making process when the technology is already being used by other colleagues.

Furthermore, according to the study, leaders' trust in algorithms can also be influenced by dispositional trust, which may vary with culture, age, gender, and other personality traits. For instance, as per them, older adults showed higher trust in automation compared to younger individuals. This might in turn be reflected in their leadership styles and decision-making approaches, when AI is involved. Additionally, cultural factors like politeness norms might affect how leaders perceive and interact with AI systems (Nass *et al.*, 1999).

Alexander et al. found that those subjects in their study that were of the opinion that "others are generally trustworthy" showed a willingness to use an algorithm more than two times higher than other participants. Thus, more trusting leaders might also be more inclined to accept an AI influenced decision making process.

Interesting insights also emerge from the physiological responses to AI usage. Alexander and colleagues found that algorithm use was associated with a decrease in heart rate, indicating lower stress levels when relying on AI. Conversely, those solving tasks without AI assistance experienced increased heart rates, suggesting higher arousal or stress. For leaders, this could imply that reliance on AI might offer a more relaxed decision-making environment, potentially enabling clearer thinking and better judgment. Which in turn might make the use of AI more appealing to them.

Apart from a person's reactions and attitudes towards algorithms and AI, other – more general – factors might also play a role in the willingness to assume responsibility for actions and decisions prepared by them. For example, an individual's demand for decision autonomy and their desire to avoid responsibility should not be overlooked. Leaders' willingness to assume responsibility may, among other things, be influenced by their desire for autonomy and authority, with individual and contextual characteristics playing a significant role. Men, for example, seem to be more likely to demand both autonomy and authority simultaneously. Additionally, individuals with presumably more regard for others tend to be more cautious when making decisions on behalf of those (Ertac *et al.*, 2020).

Apart from a person's desire for authority and autonomy, their qualifications and their overall willingness to lead should also be considered when examining this research question. According to Li et. al., the willingness to lead and a person's individual qualifications come into play at an earlier stage of the person's work life: they tend to influence the selection of leaders. When the "willingness to lead" is considered as crucial in a selection process, than less women than men get chosen for those leadership positions, because the men are viewed as more willing to lead than their female counterparts (Li *et al.*, 2021). Considering that women seem to also react differently

to algorithms (Alexander *et al.*, 2018), the recruiting process and choice of characters in an organisation might also be factors influencing how willing their staff are towards opting for adaption of AI based decision-making.

Reading the above, it may appear that an individual's personality traits are stable and consistent and thus their approach to AI based decision-making would also be consistent. However, personality traits and the connected leadership decisions may vary depending on the environment: In situations of crises, the same personality traits are applied differently than in regular times and may lead to different leadership styles (Damti and Hochman, 2022). Thus, patterns in the willingness to assume responsibility for AI based decisions that are observed based on individuals' characteristics might have to be examined in different situations to ensure their viability throughout changing environments.

In summary, a leader's willingness to assume responsibility for AI-generated decisions is multifaceted, influenced by factors such as trust in technology, social validation, cultural and personal predispositions, and even physiological responses. Understanding these traits might be a tool to guide organizations in identifying and nurturing leaders who can effectively and responsibly integrate AI into their decision-making processes.

In the next section of the paper, the focus will shift from the individual, internal perspective to the technological facets of AI itself that may influence the identification with AI based decision-making.

Which Traits of AI Systems Are Influencing the Identification With AI Based Decisions?

Focusing on AI itself, there are various elements of this technology that impact the identification with AI-based decisions and content.

Transparency and Explainability: Understanding how the AI makes its decisions is vital. People are more likely to trust and identify with AI decisions if the process is clear and understandable (Ribeiro, Singh and Guestrin, 2016). This is a key reason for the rise of explanable AI in recent years which has the potential to greatly contribute to the acceptance of more autonomous AI systems (Arrieta *et al.*, 2020).

Accuracy and Reliability: Davenport and Ronanki (2018) highlight the importance of accuracy and reliability in AI systems, noting that consistent and correct outcomes are crucial for building trust and identification.

Fairness and Unbiased Decision-Making: Utilizing AI systems with low bias and that make fair decisions, especially regarding gender, race, and other personal attributes, is critical. People are more likely to identify with decisions perceived as equitable and impartial. The critical role of fairness and unbiased decision-making in AI is discussed by Selbst et al. (2019), who delve into the challenges of ensuring AI systems are free from biases.

User-Friendly Interface and Interaction: How users interact with the AI, including the intuitiveness of the interface and the clarity of information presentation, greatly affects their acceptance and identification with AI decisions (Bader and Kaiser, 2019).

Ethical Considerations and Alignment with User Values: Decisions that adhere to ethical standards and align with the users' values and societal norms are more likely to be embraced (Jobin, Ienca and Vayena, 2019). This includes respecting privacy and considering the broader impact of decisions on society (Manheim and Kaplan, 2019).

These are five key AI system traits that will determine willingness of accountability for AI content. The subsequent section will discuss the role of organizational culture in this context.

Which Role Does an Organisation's Culture Play in the Willingness to Assume Responsibility?

The role of an organization's culture and the willingness to assume responsibility for AI-generated content involves aspects of organizational behaviour, management theory, and ethics. Organizational culture, as defined by Schein (2010), encompasses the shared values, beliefs, and practices that dictate how members of an organization interact among themselves and with external entities. This will be a basis to discuss an organization's approach to AI and its outputs as part of its AI readiness (Jöhnk, Weißert and Wyrtki, 2021). Through this lens this paper looks at five aspects that can be considered key factors for organizations: ethical orientation, risk perception and management, leadership and governance, knowledge and understanding of AI and transparency and openness.

Ethical Orientation: The ethical orientation of an organization is a pivotal factor. Treviño et al. (1998) emphasize the importance of an ethical work climate in shaping members' behaviour and decision-making. In organizations where ethical standards and accountability are prioritized, there is likely to be a higher inclination to assume responsibility for AI-generated content. One basis of this would be integrating clear ethical guidelines for AI use into the organizational culture.

Risk Perception and Management: How an organization perceives and manages risk, as discussed by Slovic (2016), also plays a crucial role. Risk-averse cultures may exhibit a lower willingness to assume responsibility for content that only has limited transparency in the process of its creation. In contrast, risk-taking cultures may embrace AI and its content as part of innovative endeavours.

Leadership and Governance: The influence of leadership in shaping organizational culture is well-documented (Bass and Bass, 2009). Leaders who demonstrate accountability for AI-generated content can influence their organization's stance on responsibility. Furthermore, effective governance structures around AI usage, as proposed by Floridi and Cowls (2022), reinforce the significance of responsibility.

Knowledge and Understanding of AI: An organization's collective understanding and knowledge of technology, as highlighted by Hirschheim and Newman (1988) and Kim and Kankanhalli (2009), can impact the willingness to assume responsibility. Cultures that encourage continuous learning and possess a high level of AI literacy are likely more confident in their ability to oversee AI-generated content responsibly.

Transparency and Openness: Transparency and openness in organizational culture, as advocated by Bennis et al. (2008), foster a greater willingness to assume responsibility. This would likely extend to clear communication about AI systems' capabilities, the data they use, and their decision-making processes, as well as the foundations of assuming responsibility for any decisions taken on such basis. As such this transparency would facilitate accountability and responsibility in AI deployment.

In summary, it can be assumed that the different factors of an organization's culture significantly influence its approach to assuming responsibility for AI-generated content. This influence is manifested through the organization's ethical orientation, risk management strategies, leadership styles, level of AI understanding, and commitment to transparency. They are used as a basis for the AI management responsibility model presented in the next section.

AI ACCOUNTABILITY WILLINGNESS FRAMEWORK

It is clear that a narrow focus on AI adoption will likely miss important factors that subsequently hinder an organization to be managed effectively. To address this, the identified factors have been further synthesized into a framework that provides an overview of relevant criteria for the willingness of assuming responsibility for AI generated content. It encompasses the most relevant factors from all three dimensions that were observed earlier.

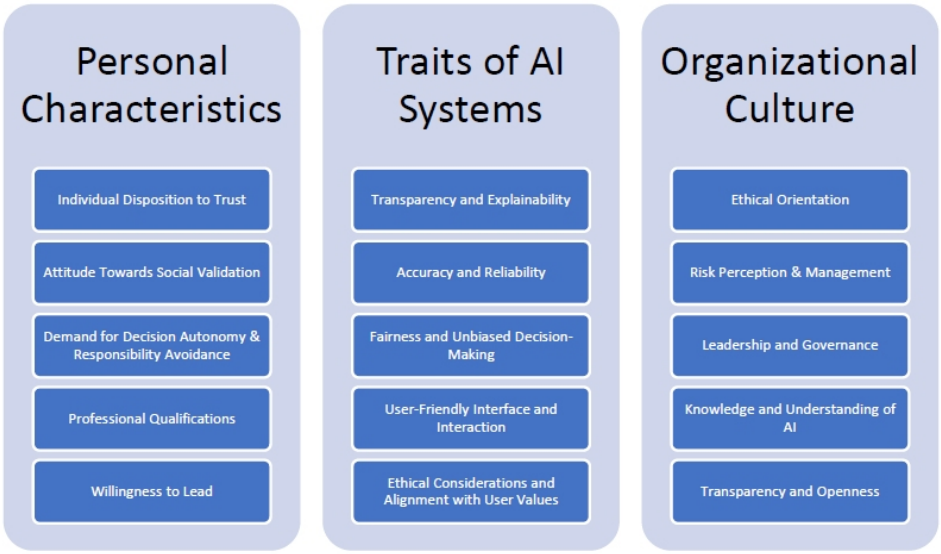

Figure 1: AI accountability willingness framework.

CONCLUSION

Drawing on fields such as organizational change, technology adoption as well as leadership, relevant factors influencing willingness for accountability for AI generated content were identified in this paper. It also shows overarching

categories that bring a general structure to this field. These categories and the corresponding factors may provide organizations with a structured approach to AI adoption with the focus on bearing AI accountability in mind, potentially allowing for a more efficient introduction of this key technology into a company.

The framework allows for further research to test and evaluate the identified criteria to further refine future approaches to accountability in this important field. As with any AI-related research, the highly dynamic developments of the topic will require constant re-evaluation of findings and their academic and practical implication and relevance.

REFERENCES

Alexander, V., Blinder, C. and Zak, P. J. (2018) 'Why trust an algorithm? Performance, cognition, and neurophysiology', *Computers in Human Behavior*, 89, pp. 279–288.

Arrieta, A. B. *et al.* (2020) 'Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI', *Information fusion*, 58, pp. 82–115.

Bader, V. and Kaiser, S. (2019) 'Algorithmic decision-making? The user interface and its role for human involvement in decisions supported by artificial intelligence', *Organization*, 26(5), pp. 655–672.

Bass, B. M. and Bass, R. (2009) *The Bass handbook of leadership: Theory, research, and managerial applications*. Simon and Schuster.

Bennis, W., Goleman, D. and O'Toole, J. (2008) *Transparency: How leaders create a culture of candor*. John Wiley & Sons.

Damti, S. and Hochman, G. (2022) 'Personality Characteristics as Predictors of the Leader's Ethical Leadership in Regular Times and in Times of Crisis', *Sustainability*, 14(16), p. 9800. Available at: https://doi.org/10.3390/su14169800.

Davenport, T. H. and Ronanki, R. (2018) 'Artificial intelligence for the real world', *Harvard business review*, 96(1), pp. 108–116.

Davis, F. D. (1989) 'Perceived usefulness, perceived ease of use, and user acceptance of information technology', *MIS quarterly*, pp. 319–340.

democracy', *Yale JL & Tech.*, 21, p. 106.

Ertac, S., Gumren, M. and Gurdal, M. Y. (2020) 'Demand for decision autonomy and the desire to avoid responsibility in risky environments: Experimental evidence', *Journal of Economic Psychology*, 77, p. 102200. Available at: https://doi.org/10.1016/j.joep.2019.102200.

Floridi, L. and Cowls, J. (2022) 'A unified framework of five principles for AI in society', *Machine learning and the city: Applications in architecture and urban design*, pp. 535–545.

Hirschheim, R. and Newman, M. (1988) 'Information systems and user resistance: theory and practice', *The computer journal*, 31(5), pp. 398–408.

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, pp. 5–20.

Jobin, A., Ienca, M. and Vayena, E. (2019) 'The global landscape of AI ethics guidelines', *Nature machine intelligence*, 1(9), pp. 389–399.

Kim, H.-W. and Kankanhalli, A. (2009) 'Investigating user resistance to information systems implementation: A status quo bias perspective', *MIS quarterly*, pp. 567–582.

Li, Y., Wang, X. and Zheng, J. (2023) *Qualifications, Willingness to Lead, and Gender Debiasing in Leader Selection*. preprint. SSRN. Available at: https://doi.org/10.2139/ssrn.4388577.

Mittelstadt, B. D. *et al.* (2016) 'The ethics of algorithms: Mapping the debate', *Big Data & Society*, 3(2), p. 205395171667967. Available at: https://doi.org/10.1177/2053951716679679.

Nass, C., Moon, Y. and Carney, P. (1999) 'Are People Polite to Computers? Responses to Computer-Based Interviewing Systems[1]', *Journal of Applied Social Psychology*, 29(5), pp. 1093–1109. Available at: https://doi.org/10.1111/j.1559-1816.1999.tb00142.x.

Ribeiro, M. T., Singh, S. and Guestrin, C. (2016) '"Why should i trust you?" Explaining the predictions of any classifier', in *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135–1144.

Schein, E. H. (2010) *Organizational culture and leadership*. John Wiley & Sons.

Selbst, A. D. *et al.* (2019) 'Fairness and abstraction in sociotechnical systems', in *Proceedings of the conference on fairness, accountability, and transparency*, pp. 59–68.

Slovic, P. (2016) *The perception of risk*. Routledge.

Treviño, L. K., Butterfield, K. D. and McCabe, D. L. (1998) 'The ethical context in organizations: Influences on employee attitudes and behaviors', *Business ethics quarterly*, 8(3), pp. 447–476.

Venkatesh, V. and Davis, F. D. (2000) 'A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies', *Management Science*, 46(2), pp. 186–204. Available at: https://doi.org/10.1287/mnsc.46.2.186.11926.