MSELS Management Systems Evolution and Leadership Studies

VOL. 2 | NO. 2 | 2025

The Evolution of Ethical Leadership Systems in AI-Driven Organizations: A Qualitative Exploration

Niloufar Kamali[1], Dariush Farahmand[1]*, Ehsan Mirzaei[2]

1. Department of Business Management, Ferdowsi University of Mashhad, Mashhad, Iran
2. Department of Industrial Management, University of Tabriz, Tabriz, Iran

Abstract

This study aimed to explore how ethical leadership systems evolve in AI-driven organizations and to develop a qualitative model explaining the leadership capabilities, governance mechanisms, and organizational conditions required for responsible AI-enabled management. This qualitative study was conducted using a conventional content-analysis approach. Participants were 23 senior managers, AI project leaders, human-resource specialists, compliance officers, and digital-transformation experts working in AI-oriented organizations in Tehran. Participants were selected through purposive sampling, followed by theoretical sampling until conceptual saturation was achieved. Data were collected through semi-structured interviews focused on ethical decision-making, algorithmic accountability, leadership responsibility, transparency, employee trust, and organizational governance in AI-based work systems. Interviews lasted between 45 and 75 minutes, were audio-recorded with informed consent, transcribed verbatim, and analyzed using open coding, axial categorization, and theme development. NVivo software was used to organize transcripts, compare codes, retrieve quotations, and develop thematic networks. Trustworthiness was enhanced through member checking, peer review, prolonged engagement with the data, and an audit trail. Analysis produced five main categories: algorithmic moral awareness, distributed accountability, transparency-centered communication, human-centered governance, and ethical learning capacity. Participants described ethical leadership in AI-driven organizations as a shift from individual moral role modeling toward a systemic leadership architecture in which leaders must align technological design, managerial judgment, employee participation, and institutional control mechanisms. Interviewees emphasized that AI increases decision speed and scale but also creates opacity, responsibility diffusion, data-related harms, and employee anxiety when ethical governance is weak. Ethical leadership in AI-driven organizations is not limited to personal integrity or compliance; it requires an integrated system of values, governance routines, technical literacy, participatory decision-making, and continuous ethical learning. The proposed model highlights that responsible AI adoption depends on leaders who can translate ethical principles into practical organizational mechanisms.

Keywords: Ethical leadership; artificial intelligence; AI governance; algorithmic accountability; qualitative research; responsible AI; organizational ethics.

✉ *CORRESPONDING AUTHOR'S EMAIL

dariush.farahmand.pub@gmail.com

❝ HOW TO CITE

Kamali, N., Farahmand, D., & Mirzaei, E. (2025). The Evolution of Ethical Leadership Systems in AI-Driven Organizations: A Qualitative Exploration. *Management Systems Evolution and Leadership Studies,* 2(2), 29-38.

1. Introduction

Artificial intelligence has become a strategic infrastructure of contemporary organizations rather than a peripheral technological tool. Organizations increasingly use AI systems for recruitment, performance evaluation, customer analytics, financial forecasting, production planning, risk assessment, knowledge management, and strategic decision support. This transformation changes not only the technical architecture of organizations but also the moral structure of leadership. In traditional organizations, ethical leadership has generally been understood through leaders' personal integrity, fairness, care, communication, and role modeling. Brown, Treviño, and Harrison (2005) conceptualized ethical leadership as the demonstration of normatively appropriate conduct through personal actions, interpersonal relationships, communication, reinforcement, and decision-making. This definition remains foundational, yet AI-driven organizations create ethical conditions that are more

complex than those assumed in classical ethical leadership models because decisions are increasingly shaped by data pipelines, predictive algorithms, automated recommendations, opaque models, and platform-based control systems.

The central problem is that AI redistributes decision agency across humans and machines. In AI-mediated environments, managerial judgment is no longer exercised only through direct human deliberation. Instead, decisions may emerge from interactions among executive intention, data quality, algorithmic classification, automated scoring, software architecture, vendor design, and employee interpretation. Shrestha, Ben-Menahem, and von Krogh (2019) argue that organizational decision-making in the age of AI differs across decision speed, interpretability, replicability, alternative-set size, and search-space specificity. Their work shows that organizations must decide how to combine human and AI-based decisions through delegation, sequential human–AI decision-making, or aggregated decision structures. This implies that ethical leadership must evolve from a person-centered influence process into a system-level capability for governing hybrid human–machine judgment.

Ethical leadership scholarship has long emphasized that leaders influence followers through social learning, moral attention, reward systems, and ethical climate. Brown and Treviño (2006) described ethical leaders as both moral persons and moral managers, meaning that leaders must possess ethical character and actively manage ethical conduct through standards, communication, and accountability. Empirical studies also show that ethical leadership is associated with ethical climate and lower employee misconduct. Mayer, Kuenzi, and Greenbaum (2010), for example, found that ethical climate mediates the relationship between ethical leadership and employee misconduct. In AI-driven organizations, however, ethical climate is not formed only through interpersonal leadership behaviors; it is also embedded in technical systems, data practices, model governance, explainability routines, procurement decisions, audit processes, and organizational responses to algorithmic error.

AI ethics literature has produced a broad set of principles such as transparency, fairness, privacy, accountability, non-maleficence, human agency, and beneficence. Jobin, Ienca, and Vayena (2019) reviewed global AI ethics guidelines and found convergence around major principles such as transparency, justice and fairness, non-maleficence, responsibility, and privacy, while also identifying significant divergence in how these principles are interpreted and implemented. Floridi and Cowls (2019) similarly proposed a unified framework of five principles for AI in society, including beneficence, non-maleficence, autonomy, justice, and explicability. These contributions demonstrate that the ethical vocabulary of AI is now relatively mature, but the translation of principles into organizational practice remains difficult.

This gap between principle and practice is especially important for leadership studies. Mittelstadt (2019) warned that principles alone cannot guarantee ethical AI because AI development lacks the professional norms, fiduciary obligations, implementation methods, and robust accountability mechanisms that support principled practice in fields such as medicine. Morley, Floridi, Kinsey, and Elhalal (2020) likewise argued that organizations need practical tools, methods, and processes to translate AI ethics from abstract statements into operational action across the machine-learning lifecycle. Therefore, ethical leadership in AI-driven organizations cannot be reduced to the publication of codes of ethics, value statements, or compliance checklists. Leaders must create the conditions under which ethical principles become embedded in data collection, model development, deployment, monitoring, stakeholder communication, and organizational learning.

A further challenge concerns opacity. AI systems, particularly machine-learning models, may generate recommendations whose internal logic is difficult for users, managers, or affected employees to understand.

Lebovitz, Lifshitz-Assaf, and Levina (2022) showed that professionals using AI for critical judgments had to develop practices for dealing with opacity and for determining whether and how to engage with AI outputs. In organizational settings, opacity can produce ethical ambiguity: managers may rely on algorithmic recommendations without fully understanding their assumptions, employees may experience decisions as impersonal or unjust, and responsibility may become diffused between users, developers, vendors, and senior leaders. Ethical leadership must therefore involve the capacity to interrogate AI systems, demand explanations, challenge automated recommendations, and maintain meaningful human responsibility.

The emergence of formal AI governance frameworks further confirms that responsible AI is now a management-system problem. The NIST AI Risk Management Framework organizes AI risk-management activities around govern, map, measure, and manage functions, with governance operating as a cross-cutting function across the AI lifecycle. ISO/IEC 42001:2023 has also been introduced as an AI management-system standard that provides structured guidance for managing AI-related risks, opportunities, ethical considerations, transparency, and continuous learning. At the regulatory level, the European Union AI Act entered into force on August 1, 2024, and establishes risk-based obligations for AI developers and deployers. The OECD AI Principles, first adopted in 2019 and updated in 2024, similarly promote trustworthy AI that respects human rights and democratic values. These developments show that ethical AI is increasingly institutionalized through governance systems, not merely individual moral judgment.

Despite these developments, the literature has not sufficiently explained how ethical leadership itself evolves inside AI-driven organizations. Existing ethical leadership theories explain how leaders shape ethical climates, but they often assume that leaders and followers are the main agents of organizational morality. AI ethics literature identifies principles and governance mechanisms, but it does not always explain how leaders interpret, negotiate, and institutionalize these mechanisms in everyday organizational life. Management literature explains automation and augmentation, but it often pays limited attention to ethical meaning-making, trust, and moral responsibility. Raisch and Krakowski (2021) argue that automation and augmentation are interdependent in management, meaning that AI does not simply replace or support human work in isolation; rather, human and machine capabilities reshape each other over time. This insight is crucial for ethical leadership because moral agency also becomes relational, distributed, and evolving.

AI-driven organizations require leaders who can integrate moral awareness, technological understanding, governance discipline, and participatory communication. Ethical leadership must shift from episodic ethical intervention to continuous ethical system design. Leaders must ask not only whether an action is fair, but whether the data infrastructure is fair; not only whether a manager behaves transparently, but whether AI-supported decisions are explainable; not only whether employees trust leaders, but whether they trust the human–machine decision environment. This requires a new conceptualization of ethical leadership as a systemic capability that includes values, routines, structures, tools, and learning processes.

The present study addresses this gap by qualitatively exploring how ethical leadership systems evolve in AI-driven organizations. Focusing on managers, AI project leaders, HR specialists, compliance officers, and digital-transformation experts in Tehran, the study investigates how organizational actors understand ethical leadership in AI-enabled contexts, what challenges they experience, and what leadership capabilities and governance mechanisms they consider necessary for responsible AI adoption. The aim of this study was to develop a qualitative model of ethical leadership systems in AI-driven organizations by identifying the core categories,

mechanisms, and organizational conditions through which leaders translate AI ethics into everyday management practice.

2. Methods and Materials

This study used a qualitative exploratory design based on conventional content analysis. A qualitative approach was appropriate because the purpose of the study was not to test predetermined hypotheses but to explore how organizational actors interpret, experience, and construct ethical leadership in AI-driven organizations. The study focused on organizations in Tehran that had implemented AI-enabled tools in areas such as human-resource analytics, customer relationship management, fraud detection, digital marketing, operational forecasting, process automation, and managerial dashboards.

Participants were selected using purposive sampling to ensure that they had direct experience with AI-related decision-making, ethical governance, digital transformation, or leadership responsibilities. Theoretical sampling was then used to refine emerging categories by adding participants whose experience could clarify underdeveloped concepts. The final sample consisted of 23 participants, at which point theoretical saturation was achieved. Saturation was considered achieved when the final three interviews produced no substantially new codes and mainly confirmed the existing categories.

The participants included 7 senior executives, 5 AI or digital-transformation project managers, 4 human-resource managers, 3 compliance and risk-management specialists, 2 data scientists, and 2 organizational-development consultants. Fourteen participants were male and nine were female. Participants' ages ranged from 31 to 57 years. In terms of education, 4 participants held bachelor's degrees, 13 held master's degrees, and 6 held doctoral degrees. Their professional experience ranged from 7 to 29 years, while their direct experience with AI-based organizational systems ranged from 1 to 9 years. All participants worked in Tehran-based private, semi-public, or knowledge-based organizations.

Data were collected through semi-structured interviews. Semi-structured interviews were suitable because they allowed the researcher to ask comparable core questions while also giving participants enough flexibility to describe their personal experiences, organizational contexts, and ethical concerns. The interview protocol included questions such as: "How has AI changed the meaning of ethical leadership in your organization?" "What ethical challenges have emerged after using AI-based systems?" "Who is responsible when an AI-supported decision harms an employee, customer, or stakeholder?" "How do leaders explain AI-based decisions to employees?" "What governance mechanisms are needed to make AI use more ethical?" and "What leadership capabilities are necessary for responsible AI adoption?"

Interviews were conducted in a private setting or through secure online meetings, depending on participant availability. Each interview lasted between 45 and 75 minutes. All interviews were audio-recorded after informed consent was obtained and were transcribed verbatim. Participants were informed that participation was voluntary, that they could withdraw at any time, and that identifying information would be removed from the transcripts. Codes such as P1, P2, and P3 were used to preserve confidentiality.

Data were analyzed using a three-stage qualitative coding process. In the first stage, open coding was conducted through repeated reading of interview transcripts. Meaning units related to ethical leadership, AI governance, transparency, accountability, responsibility, employee trust, algorithmic bias, human oversight, and organizational

learning were identified and labeled. In the second stage, related codes were compared and grouped into subcategories. In the third stage, subcategories were integrated into broader main categories representing the evolution of ethical leadership systems in AI-driven organizations.

NVivo software was used to organize transcripts, assign codes, retrieve quotations, compare code frequencies, and develop thematic relationships. Data analysis was iterative rather than linear. Early codes were revised as new interviews were added, and emerging categories were compared against the full data set. The credibility of the analysis was strengthened through member checking with six participants, peer review by two qualitative-research experts, and comparison of coding decisions across several transcripts. Dependability was supported through an audit trail containing interview notes, coding memos, category-development records, and analytic decisions. Confirmability was enhanced by reflexive memo writing, which helped the researcher distinguish participant meanings from prior assumptions.

3. Findings and Results

The demographic characteristics of the participants showed that the study included a diverse group of organizational actors involved in AI-related leadership and governance. Of the 23 participants, 14 were male and 9 were female. Seven participants were senior executives, 5 were AI or digital-transformation project managers, 4 were human-resource managers, 3 were compliance or risk-management specialists, 2 were data scientists, and 2 were organizational-development consultants. Regarding education, 4 participants held bachelor's degrees, 13 held master's degrees, and 6 held doctoral degrees. In terms of work experience, 6 participants had 7–10 years of professional experience, 9 had 11–20 years, and 8 had more than 20 years. Direct experience with AI-based organizational systems was 1–3 years for 8 participants, 4–6 years for 10 participants, and more than 6 years for 5 participants.

The analysis produced five main categories and fifteen subcategories. The main categories were: algorithmic moral awareness, distributed accountability, transparency-centered communication, human-centered governance, and ethical learning capacity. Together, these categories formed a model of ethical leadership systems in AI-driven organizations. The model suggests that ethical leadership evolves from personal morality toward a systemic capability in which leaders must recognize AI-related moral risks, define responsibility across human and technical actors, explain AI-supported decisions, protect human dignity, and institutionalize continuous ethical learning.

The first category was algorithmic moral awareness. Participants stated that leaders in AI-driven organizations must become aware of moral risks that are not always visible at the surface of managerial decisions. These risks include biased training data, discriminatory outputs, hidden assumptions in predictive models, privacy violations, surveillance effects, overreliance on automated recommendations, and unintended harm to employees or customers. Participants emphasized that ethical leadership now requires the ability to ask moral questions about technical processes. One participant explained, "Before AI, we asked whether the manager was fair. Now we also have to ask whether the data is fair, whether the model is fair, and whether the people affected by the system even know how the decision was made" (P8). Another participant stated, "The ethical problem is not always in the final decision. Sometimes it is hidden in the data that entered the system six months earlier" (P14). This category

included three subcategories: recognition of hidden algorithmic harms, sensitivity to data-based injustice, and ethical questioning of automation.

The second category was distributed accountability. Participants repeatedly described AI-supported decisions as ethically complex because responsibility is distributed among senior managers, technical teams, system vendors, data owners, users, and governance committees. They argued that one of the most dangerous effects of AI is responsibility diffusion, in which managers attribute outcomes to "the system" and technical experts attribute decisions to "business requirements." A senior executive noted, "When a wrong decision is made by a human manager, we know who must answer. But when the AI dashboard recommends a decision, everyone says the system suggested it. This is where accountability disappears" (P3). Another participant stated, "Ethical leadership means not hiding behind the algorithm. Someone must be clearly responsible for approving, monitoring, and correcting AI decisions" (P17). This category included responsibility mapping, human approval authority, and post-decision answerability. Participants suggested that AI-driven organizations need explicit accountability maps showing who is responsible at each stage of the AI lifecycle, from data collection to model deployment and monitoring.

The third category was transparency-centered communication. Participants believed that AI reduces trust when employees or customers experience decisions as technically obscure, unexplained, or imposed. They did not define transparency as full technical disclosure; rather, they described it as meaningful explanation appropriate to the audience. For employees, transparency meant knowing when AI is used, what data are considered, how decisions are reviewed, and how objections can be raised. For managers, transparency meant having enough interpretive access to challenge model outputs. For customers, transparency meant receiving understandable explanations of automated decisions that affect access, service, pricing, or evaluation. One HR manager stated, "Employees do not expect to read the algorithm, but they expect to know why the system evaluated them in a certain way and whether a human has reviewed the result" (P11). Another participant said, "A black-box system creates a black-box culture. People stop asking questions because they think the answer is technical and unavailable" (P19). This category included explainable communication, disclosure of AI use, and grievance channels.

The fourth category was human-centered governance. Participants argued that ethical leadership in AI-driven organizations requires governance structures that protect human dignity, agency, participation, and contextual judgment. They criticized approaches in which AI adoption is treated only as an efficiency project. According to participants, ethical AI governance must include cross-functional committees, human oversight, periodic audits, privacy safeguards, bias testing, vendor assessment, and employee participation. One participant explained, "The leader's job is not to slow technology down, but to make sure technology does not remove human judgment from places where human judgment is essential" (P6). Another stated, "If AI is used for performance evaluation, promotion, or dismissal, there must be human review. Otherwise, the organization becomes fast but unfair" (P21). This category included human-in-the-loop mechanisms, participatory governance, and dignity-preserving automation. Participants emphasized that AI should augment human judgment rather than replace moral deliberation in high-stakes organizational decisions.

The fifth category was ethical learning capacity. Participants described ethical leadership as a dynamic capability because AI systems, organizational uses, employee expectations, and regulatory requirements change over time. They argued that ethical leadership cannot rely on a one-time policy or initial training session. Instead,

organizations must learn continuously from errors, complaints, audits, new regulations, employee feedback, and unexpected system behavior. A compliance specialist stated, “An AI ethics policy is only the starting point. The real issue is whether the organization learns after the system creates a problem” (P4). Another participant added, “We need leaders who can say: the model worked technically, but socially it created fear, so we must redesign the process” (P16). This category included continuous ethical training, feedback-based redesign, and adaptive governance. Participants believed that leaders must create psychological safety so that employees can report AI-related concerns without fear of being seen as anti-innovation.

4. Discussion and Conclusion

The findings of this study indicate that ethical leadership systems in AI-driven organizations evolve through five interconnected dimensions: algorithmic moral awareness, distributed accountability, transparency-centered communication, human-centered governance, and ethical learning capacity. The first finding, algorithmic moral awareness, shows that leaders must expand their ethical attention from interpersonal conduct to sociotechnical systems. This aligns with Brown et al.’s (2005) view that ethical leaders communicate and reinforce appropriate conduct, but extends it by showing that “appropriate conduct” in AI-driven organizations includes the ethical design, use, and monitoring of data-based systems. The finding is also consistent with Jobin et al. (2019), who identified transparency, fairness, responsibility, privacy, and non-maleficence as recurring principles in AI ethics guidelines. However, the present study adds that these principles become meaningful only when leaders develop the practical awareness to detect where harm may be hidden inside data practices, model assumptions, and automated workflows.

The second finding, distributed accountability, suggests that AI transforms responsibility from a linear managerial relation into a networked governance problem. Participants emphasized that responsibility may become fragmented among managers, developers, vendors, data teams, and users. This supports Mittelstadt’s (2019) argument that AI ethics cannot rely only on principles because AI development lacks mature accountability mechanisms comparable to those in established professions. It also aligns with Raji et al. (2020), who proposed internal algorithmic auditing as a way to close accountability gaps across the AI development lifecycle. The present study contributes to this literature by showing that accountability is not only a technical or legal requirement but also a leadership practice. Ethical leaders must prevent the organization from using algorithmic complexity as a shield against responsibility.

The third finding, transparency-centered communication, shows that ethical leadership in AI-driven organizations depends on meaningful explanation, not merely technical openness. Participants emphasized that employees and customers need understandable information about when AI is used, what data are considered, how outputs are reviewed, and how decisions can be challenged. This finding is consistent with Floridi and Cowls’s (2019) principle of explicability and with Shneiderman’s (2020) human-centered AI framework, which emphasizes reliable, safe, and trustworthy systems. It also supports Lebovitz et al.’s (2022) work on professional engagement with opaque AI systems. When AI outputs cannot be interrogated, professionals may either over-trust the system or disengage from it. In both cases, ethical leadership weakens because leaders lose the ability to mediate between technical output and human judgment.

The fourth finding, human-centered governance, indicates that ethical leadership must be institutionalized through structures that preserve human agency and dignity. Participants did not reject AI adoption; rather, they argued that AI should be governed through human oversight, participatory mechanisms, privacy safeguards, bias assessment, and review procedures in high-stakes decisions. This finding aligns with Raisch and Krakowski's (2021) argument that automation and augmentation are interdependent in management. AI may automate some tasks, but its organizational value depends on how it augments human judgment and how humans supervise its consequences. The finding is also consistent with the NIST AI Risk Management Framework, which places governance as a cross-cutting function across AI risk mapping, measurement, and management.

The fifth finding, ethical learning capacity, shows that ethical leadership in AI-driven organizations is not static. Participants described AI ethics as a continuous learning process because systems evolve, data drift, organizational uses change, and new harms may appear after deployment. This supports Morley et al.'s (2020) argument that AI ethics must move from abstract principles to practical tools and lifecycle methods. It also aligns with ISO/IEC 42001:2023, which frames AI management as a structured system for managing risks and opportunities, including transparency, ethical considerations, and continuous improvement. The present study adds that continuous improvement must include moral and social learning, not only technical optimization. Leaders must learn from complaints, failures, audit findings, employee concerns, and stakeholder feedback.

Taken together, the findings suggest that ethical leadership in AI-driven organizations should be conceptualized as a leadership system rather than an individual trait or style. Classical ethical leadership theory remains important because moral role modeling, fairness, communication, and accountability continue to matter. However, AI changes the object of ethical leadership. Leaders must now govern decision infrastructures, not only human behavior. They must shape ethical climates that include human–machine interaction, data governance, model oversight, explainability, and lifecycle accountability. This extends Mayer et al.'s (2010) finding that ethical leadership shapes ethical climate by showing that ethical climate in AI-driven organizations is partly encoded into technical and procedural systems.

The findings also show that ethical leadership is central to organizational trust in AI. Employees are more likely to accept AI-enabled decisions when leaders explain the purpose of AI use, clarify human oversight, create appeal mechanisms, and acknowledge limitations. Conversely, when leaders present AI as neutral, objective, or unquestionable, they may intensify distrust. This point is important because AI systems can appear authoritative even when they reproduce bias, incomplete data, or flawed assumptions. Ethical leadership therefore requires epistemic humility: leaders must recognize that AI outputs are probabilistic, context-dependent, and fallible. This supports Dignum's (2019) argument that responsible AI concerns not only system outputs but also why AI is developed, how it is designed, who is involved, and how it is governed.

The study further suggests that ethical leadership in AI-driven organizations requires a balance between innovation and restraint. Participants did not view ethics as an obstacle to AI adoption; rather, they viewed ethical leadership as the condition for sustainable AI adoption. This is consistent with the OECD AI Principles, which promote innovative and trustworthy AI that respects human rights and democratic values. It is also consistent with the EU AI Act's risk-based approach, which seeks to support AI development while imposing obligations according to risk. In organizational practice, this means that leaders must avoid both technological enthusiasm without governance and excessive risk avoidance without innovation. Ethical leadership requires disciplined

experimentation: organizations should innovate, but they should also test, monitor, explain, and correct AI systems.

The proposed model has several theoretical implications. First, it expands ethical leadership theory by showing that moral influence operates through sociotechnical governance systems. Second, it connects AI ethics with leadership studies by explaining how principles such as transparency, fairness, accountability, and human agency become organizational leadership practices. Third, it contributes to AI management literature by showing that automation and augmentation are not only strategic or operational choices but also ethical leadership challenges. Finally, it highlights the importance of qualitative inquiry in understanding AI-driven transformation because organizational actors experience AI not only as a technology but also as a change in power, responsibility, trust, and identity.

Ethical Considerations

All procedures performed in this study were under the ethical standards.

Acknowledgments

Authors thank all who helped us through this study.

Conflict of Interest

The authors report no conflict of interest.

Funding

According to the authors, this article has no financial support.

References

Bedi, A., Alpaslan, C. M., & Green, S. (2016). A meta-analytic review of ethical leadership outcomes and moderators. *Journal of Business Ethics, 139*(3), 517–536. doi:10.1007/s10551-015-2625-1

Brown, M. E., & Treviño, L. K. (2006). Ethical leadership: A review and future directions. *The Leadership Quarterly, 17*(6), 595–616. doi:10.1016/j.leaqua.2006.10.004

Brown, M. E., Treviño, L. K., & Harrison, D. A. (2005). Ethical leadership: A social learning perspective for construct development and testing. *Organizational Behavior and Human Decision Processes, 97*(2), 117–134. doi:10.1016/j.obhdp.2005.03.002

Charmaz, K. (2014). *Constructing grounded theory* (2nd ed.). Sage.

Dignum, V. (2019). *Responsible artificial intelligence: How to develop and use AI in a responsible way*. Springer. doi:10.1007/978-3-030-30371-6

European Commission. (2024). *AI Act enters into force*. European Commission.

Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review, 1*(1). doi:10.1162/99608f92.8cd550d1

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? *Field Methods, 18*(1), 59–82. doi:10.1177/1525822X05279903

International Organization for Standardization. (2023). *ISO/IEC 42001:2023: Information technology—Artificial intelligence—Management system*. ISO.

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

Lebovitz, S., Lifshitz-Assaf, H., & Levina, N. (2022). To engage or not to engage with AI for critical judgments: How professionals deal with opacity when using AI for medical diagnosis. *Organization Science, 33*(1), 126–148. doi:10.1287/orsc.2021.1549

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage.

Mayer, D. M., Kuenzi, M., & Greenbaum, R. L. (2010). Examining the link between ethical leadership and employee misconduct: The mediating role of ethical climate. *Journal of Business Ethics, 95*(1), 7–16. doi:10.1007/s10551-011-0794-0

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

Morley, J., Floridi, L., Kinsey, L., & Elhalal, A. (2020). From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Science and Engineering Ethics, 26*(4), 2141–2168. doi:10.1007/s11948-019-00165-5

National Institute of Standards and Technology. (2023). *Artificial intelligence risk management framework: AI RMF 1.0*. U.S. Department of Commerce. doi:10.6028/NIST.AI.100-1

OECD. (2024). *OECD AI principles*. OECD.AI Policy Observatory.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192–210. doi:10.5465/amr.2018.0072

Raji, I. D., Smart, A., White, R. N., Mitchell, M., Gebru, T., Hutchinson, B., Smith-Loud, J., Theron, D., & Barnes, P. (2020). Closing the AI accountability gap: Defining an end-to-end framework for internal algorithmic auditing. *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 33–44. doi:10.1145/3351095.3372873

Saldaña, J. (2021). *The coding manual for qualitative researchers* (4th ed.). Sage.

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe and trustworthy. *International Journal of Human–Computer Interaction, 36*(6), 495–504. doi:10.1080/10447318.2020.1741118

Shrestha, Y. R., Ben-Menahem, S. M., & von Krogh, G. (2019). Organizational decision-making structures in the age of artificial intelligence. *California Management Review, 61*(4), 66–83. doi:10.1177/0008125619862257

Treviño, L. K., Brown, M., & Hartman, L. P. (2003). A qualitative investigation of perceived executive ethical leadership. *Journal of Business Ethics, 45*(1–2), 5–37. doi:10.1023/A:1024127808455