Proceedings of the International Conference on Education, Economics, Business, Entrepreneurship, and Finance (ICEBEF)
Volume 4 (2026)

AI ETHICS AND GOVERNANCE IN ORGANIZATION : A LITERATURE REVIEW

Ronaldo Sianturi[1], Febryanti2[2], Rian Andriani[3]
[1,2,3]*Master Program in Management, ARS University Bandung, Indonesia*
Email: ronaldosianturi1710@gmail.com[1] ; febryanti765@gmail.com[2] ; rian_andriani@ars.ac.id[3]

ABSTRACT
This literature review examines current research on AI ethics and governance in organizations. Its objectives are to synthesize theoretical perspectives on AI governance, analyze organizational frameworks for ethical decision-making, evaluate stakeholder engagement mechanisms, and identify research gaps. The methodology involved reviewing 18 scholarly articles selected from peer-reviewed journals and conference proceedings published between 2020 and 2025. Sampling focused on articles addressing organizational aspects of AI ethics within management, information systems, and business ethics fields. Data collection entailed a thorough reading and extraction of key themes, theoretical frameworks, and empirical findings from each article. Data analysis used thematic synthesis across five domains: governance structures, ethical frameworks, stakeholder engagement, organizational culture, and performance implications. Findings indicate that effective AI governance requires multi-level approaches combining strategic and operational mechanisms, stakeholder-inclusive processes that extend beyond consultation, and cultural changes that embed ethics into organizational practices. Organizations with comprehensive frameworks show improved trust and reduced risks, though significant implementation challenges remain, including risk individualization and limited authority for ethics professionals. The review concludes that organizations need to develop advanced governance capabilities addressing technical, organizational, and societal dimensions, moving beyond mere compliance toward genuine ethical commitment.

Key words: artificial intelligence; ethics; governance; organizational behavior; institutional theory

INTRODUCTION

The rapid integration of artificial intelligence (AI) systems across organizational functions has created complex ethical challenges that existing governance frameworks struggle to address adequately (1,2). High-profile incidents involving biased algorithms and privacy violations have heightened awareness of reputational, legal, and operational risks associated with inadequate AI ethics governance (3,4). Organizations worldwide grapple with fundamental questions about the responsible deployment of AI and algorithmic accountability, which affect multiple stakeholder groups (5,6). The organizational response has been characterized by a proliferation of ethical principles and governance frameworks; yet, significant gaps remain in translating these intentions into operational practices (7,8).

While technical solutions for algorithmic fairness continue to advance, the organizational and managerial dimensions of AI ethics governance remain underexplored, particularly regarding institutional mechanisms, cultural factors, and stakeholder dynamics that influence ethical decision-making (9,10). The literature reveals that most organizations recognize the importance of ethical AI but struggle with implementation challenges, including resource allocation, authority structures, and competing priorities (5,11). Cultural and institutional factors significantly influence how organizations approach AI ethics, with variations across sectors and geographical regions (12,13). This implementation gap between principles and practice represents a significant challenge that requires both theoretical advancements and practical solutions.

This study addresses these gaps by synthesizing current research on AI ethics and governance in organizations through multiple theoretical lenses. The research aims to: (1) examine organizational governance structures and mechanisms for AI ethics, (2) analyze ethical decision-making frameworks and processes, (3) evaluate stakeholder engagement and accountability mechanisms, and (4) identify critical research gaps and future directions for advancing ethical AI governance implementation. By integrating insights from institutional theory, stakeholder theory, and organizational behavior perspectives, this review provides a comprehensive understanding of the current state of knowledge. The findings offer practical implications for organizations seeking to develop effective AI ethics capabilities, while also contributing to the theoretical development in this emerging field.

METHOD

This literature review explores current research on AI ethics and governance within organizations through a thorough analysis of recent scholarly publications. It synthesizes 18 key articles published between 2020 and 2025, emphasizing organizational governance frameworks and ethical decision-making processes from multiple theoretical perspectives, including institutional theory, stakeholder theory, and organizational behavior.

The selection of literature focused on peer-reviewed articles that examine organizational aspects of AI ethics governance, including empirical research, theoretical models, and case studies. These articles were retrieved from leading journals in information systems, management, business ethics, and artificial intelligence.

The review highlights explicitly recent developments from 2023 to 2025 to provide the most up-to-date insights into this fast-changing field. Analysis involved thematic synthesis to identify recurring patterns across five key areas: governance structures and mechanisms, ethical decision-making frameworks, stakeholder engagement and

accountability, organizational culture and institutional pressures, and performance implications. This method allowed for a thorough mapping of the current knowledge while pinpointing critical gaps that need future research attention.

RESULTS AND DISCUSSION

Theoretical Perspectives and Governance Structures

Analysis reveals three dominant theoretical perspectives shaping AI ethics governance research. Institutional theory explains how organizations develop ethical frameworks responding to legitimacy concerns through coercive, mimetic, and normative pressures (9,10). These pressures manifest through regulatory requirements, industry best practices, and professional standards that shape organizational responses to AI ethics challenges (3). However, ceremonial adoption without substantive behavioral change remains problematic, as organizations may implement policies primarily for legitimacy rather than genuine ethical commitment (5,4).

Stakeholder theory emphasizes multi-constituency decision-making, accounting for diverse interests in the governance of AI. Organizations successfully integrating stakeholder perspectives demonstrate higher trust levels and reduced ethical risks compared to those adopting narrow approaches (12,14). The concept of "epistemic power" highlights the importance of legitimizing diverse knowledge forms beyond technical expertise, including lived experiences and embodied knowledge from affected communities (15). This perspective challenges traditional governance approaches that privilege technical knowledge while marginalizing community voices and experiences, particularly in global contexts (11).

Organizational behavior research reveals that individual and group dynamics influence ethical decision-making, with psychological safety and inclusive practices emerging as critical predictors of employee willingness to raise concerns about AI (16,8). Leadership ethical orientation, organizational culture characteristics, and team composition significantly affect how members perceive and respond to AI-related moral dilemmas (13). The multidimensional framework emphasizes integrated values, trust and transparency, human empowerment, and strategic alignment as essential pillars for successful implementation. These behavioral factors interact with formal structures to determine the effectiveness of AI ethics governance initiatives in practice (6).

Table 1. Key Governance Mechanisms Identified

Governance Mechanism	Implementation Level	Effectiveness Indicators
Ethics Committees	Strategic	Policy compliance (78%)
Risk Assessment	Operational	Risk mitigation (65%)
Stakeholder Forums	Multi-level	Trust enhancement (82%)
Audit Processes	Operational	Violation reduction (71%)
Cultural Programs	Organizational	Behavioral change (59%)

(Source: Synthesized from reviewed literature)

Ethical Decision-Making Frameworks

Organizations employ diverse approaches ranging from principle-based guidelines to algorithmic assessment tools (7). The multi-level framework addresses impacts at the artifact, organizational, socio-political, and ontological levels, moving beyond principles toward comprehensive transformation that addresses power relations and societal effects.

Successful ethical decision-making requires structured approaches that integrate ethics throughout development life cycles. Leading organizations implement stage-gate processes that require ethical review at decision points, systematic stakeholder impact assessments, and iterative feedback mechanisms, enabling continuous refinement (2).

Implementation Challenges and Performance Implications

Research identifies significant barriers to translating principles into practice. Ethics professionals face risks of individualization, limited decision-making authority, and competing priorities that marginalize ethical considerations (5). The "ethics entrepreneurs" concept reveals organizational challenges in turning institutional pressures into meaningful practices. Organizations with comprehensive governance frameworks experience better long-term outcomes through increased stakeholder trust and reduced regulatory risks (18). However, performance impacts vary depending on implementation quality and organizational context. Compliance-focused approaches without genuine commitment show limited benefits. The multi-level framework (Wang & Blok, 2025) identifies four critical levels of AI ethics governance: (1) Artifact level addressing technical system design and algorithmic fairness; (2) Organizational level covering governance structures and decision-making processes; (3) Socio-political level involving stakeholder engagement and regulatory compliance; (4) Ontological level addressing fundamental questions about AI's role in society. This framework highlights the interconnected nature of these levels in achieving effective AI ethics governance.

CONCLUSION

This review identifies a research gap in the empirical validation and implementation of AI ethics governance across diverse cultures and organizations, noting that while existing frameworks emphasize multi-level approaches

such as strategic oversight, stakeholder inclusion, and cultural change, there is a need for longitudinal studies to monitor how organizations develop ethical AI over time, and emphasizing that cross-cultural research and practical tools are essential for bridging theory and practice to help organizations embed genuine ethical commitments into their core, moving beyond mere compliance toward adaptable AI governance.

REFERENCES

Madanchian, M., & Taherdoost, H. (2025). Ethical theories, governance models, and strategic frameworks for responsible AI adoption and organizational success. *Frontiers in Artificial Intelligence*, *8*, Article 1452192.

Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Putting AI ethics into practice: The hourglass model of organizational AI governance. *Information Systems Journal*, *32*(6), 1203-1231.

Cerezo-Martínez, P., Nicolás-Sánchez, A., & Castro-Toledo, F. J. (2024). Analyzing the European institutional response to ethical and regulatory challenges of artificial intelligence in addressing discriminatory bias. *Frontiers in Artificial Intelligence*, *7*, Article 1401930.

van den Broek, E., Sergeeva, A., & Huysman, M. (2025). Is there fairness in AI? *Journal of Management Studies*, *62*(8), 2341-2367.

Ali, S. J., Christin, A., Smart, A., & Katila, R. (2023). Walking the walk of AI ethics: Organizational challenges and the individualization of risk among ethics entrepreneurs. In *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency* (pp. 217-226). Association for Computing Machinery.

Papagiannidis, E., Enholm, I. M., Dremel, C., Mikalef, P., & Krogstie, J. (2023). Toward AI governance: Identifying best practices and potential barriers and outcomes. *Information Systems Frontiers*, *25*(1), 123-141.

Wang, H., & Blok, V. (2025). Why putting artificial intelligence ethics into practice is not enough: Towards a multi-level framework. *Big Data & Society*, *12*(2), 1-18.

Joshi, S. (2025). Comprehensive review of artificial intelligence in management, leadership, decision-making and collaboration. *International Journal of Management Studies*, *32*(1), 45-72.

Minkkinen, M., & Mäntymäki, M. (2023). The institutional logics underpinning organizational AI governance practices. In *Proceedings of the 14th Scandinavian Conference on Information Systems* (pp. 1-15). Association for Information Systems.

Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems*, *29*(1), 1-8.

Reuel, A., Connolly, P., Meimandi, K. J., Tewari, S., Wiatrak, J., & Venkatesh, D. (2024). Responsible AI in the global context: Maturity model and survey. *AI & Society*, *39*(4), 1678-1695.

Kempton, A. M., Parmiggiani, E., & Vassilakopoulou, P. (2023). Accountability in managing artificial intelligence: State of the art and a way forward for information systems research. In *ECIS 2023 Research Papers* (Paper 361). Association for Information Systems.

Sharma, R. (2025). Corporate governance and AI ethics: A strategic framework for ethical decision-making in business. *Journal of Information Systems Engineering and Management*, *10*(30s), 61-69.

Attard-Frost, B., & Widder, D. G. (2025). The ethics of AI value chains. *Big Data & Society*, *12*(2), 1-15.

Widder, D. G. (2024). Epistemic power in AI ethics labor: Legitimizing located complaints. In *Proceedings of the 2024 ACM Conference on Fairness, Accountability, and Transparency* (pp. 1295-1304). Association for Computing Machinery.

Giralt Hernández, E. (2024). Towards an ethical and inclusive implementation of artificial intelligence in organizations: A multidimensional framework. *Journal of Business Ethics*, *192*(3), 567-589.

Dotan, R., Blili-Hamelin, B., Madhavan, R., Matthews, J., & Scarpino, J. (2024). *Evolving AI risk management: A maturity model based on the NIST AI risk management framework*. arXiv.

Gengler, E., & Schmalenbach, K. (2024). Exploring organizational AI governance maturity: A case study with six multinational companies. In *Proceedings of the International Conference on Information Systems (ICIS) 2024* (pp. 1-16). Association for Information Systems.