

REVIEW

Open access

Organizational AI Governance: Integrating Ethical, Strategic, and Regulatory Dimensions

Beatriz Romero¹, Ines Castillo^{1*}, Arturo Medina², Lucia Vega¹

Abstract

The governance of artificial intelligence has emerged as a critical organizational capability as firms increasingly deploy AI systems that shape strategic outcomes, raise ethical concerns, and face expanding regulatory requirements. This article reviews the literature on organizational AI governance across management, information systems, and interdisciplinary scholarship. The analysis reveals that effective AI governance requires integration of three interconnected dimensions: ethical governance addressing bias, fairness, transparency, and accountability; strategic governance encompassing competitive positioning, oversight structures, and control mechanisms; and regulatory governance responding to evolving compliance mandates. The review identifies tensions between innovation imperatives and responsible governance, highlighting the need for organizations to balance rapid AI deployment with robust control systems. Key governance mechanisms include transparency frameworks, explainability tools, human oversight protocols, and risk management processes. The analysis concludes that AI governance must evolve from fragmented technical and compliance approaches toward integrated organizational systems that embed accountability across leadership, management, and operational levels.

Keywords Strategic oversight, Artificial intelligence governance, Ethical AI, Regulatory compliance, Organizational accountability

*Correspondence:

Ines Castillo
ines.castillo@gmail.com

¹ Department of Digital Enterprise and Strategic Innovation, University of Zaragoza, Zaragoza, Spain

² Department of Business Analytics and Intelligence Systems, University of Malaga, Malaga, Spain

Introduction

Artificial intelligence has transitioned from experimental technology to a strategic imperative across organizational contexts [1], fundamentally altering how firms create value, make decisions, and interact with stakeholders [2, 3]. As AI systems increasingly determine hiring outcomes, credit allocations, customer treatments, and operational processes [4], organizations confront unprecedented challenges in governing technologies that exhibit autonomy, opacity, and consequential decision-making capabilities [5, 6]. The governance of artificial intelligence encompasses the frameworks, structures, and processes by which organizations direct, control, and hold AI systems accountable throughout their lifecycle [1, 7, 8].

The urgency of AI governance derives from three converging pressures. First, ethical concerns surrounding AI systems have intensified as documented cases of algorithmic bias, discrimination, and harmful outcomes accumulate across sectors [9–11]. Organizations deploying AI face reputational, legal, and operational risks when systems produce unfair or inaccurate outputs [12], which can affect individuals and communities [13, 14]. Second, strategic considerations demand that organizations establish control mechanisms that balance AI-enabled innovation with risk management [1], ensuring that AI investments deliver competitive advantage without exposing firms to unacceptable vulnerabilities [7, 15]. Third, regulatory developments across jurisdictions, including the European Union's AI Act and sector-specific requirements

[9], impose mandatory compliance obligations that transform voluntary ethical commitments into enforceable legal duties [10, 16].

Despite growing recognition of AI governance's importance, organizations struggle to translate principles into practice. Research documents persistent gaps between ethical commitments and operational implementation [5], with firms adopting fragmented approaches that fail to address the systemic nature of AI risks [13, 17]. The literature reveals that effective governance requires moving beyond technical fixes toward integrated organizational systems [7] that embed accountability across leadership, management, and operational levels [18, 19]. However, the conceptualization of AI governance remains dispersed across disciplinary silos [2], with management scholars emphasizing strategic control, information systems researchers focusing on technical mechanisms, and ethics scholars prioritizing normative principles [3, 20].

This review synthesizes the interdisciplinary literature on organizational AI governance to address three objectives. First, the analysis examines how AI governance is conceptualized across management, information systems, and organizational scholarship, identifying core dimensions and frameworks. Second, the review explores how ethical, strategic, and regulatory dimensions intersect, sometimes conflict, and require integration for effective governance. Third, the analysis identifies major patterns, contradictions, and unresolved questions in the literature, providing directions for future research and managerial practice. The review contributes to understanding how organizations can develop governance capabilities that enable responsible and competitive AI deployment.

Literature Synthesis on Governance of Artificial Intelligence in Organizations

Ethical dimensions of organizational AI governance

The ethical governance of artificial intelligence constitutes a foundational domain in which organizations translate normative principles into operational practices. Research documents the proliferation of ethical guidelines, frameworks, and principles developed by industry consortia, professional associations, and multi-stakeholder

initiatives [9, 12, 21–24]. However, scholars consistently identify implementation gaps between espoused commitments and actual organizational practices [5, 13, 25]. This principle-practice divide reflects challenges in operationalizing abstract concepts such as fairness, accountability, and transparency within technical development processes and organizational decision-making structures [6, 14, 26].

Ethical AI governance encompasses mechanisms for identifying, assessing, and mitigating potential harms arising from the design, development, and deployment of AI systems. Organizations increasingly adopt ethics-based auditing approaches that evaluate systems against predetermined criteria, though questions remain about audit effectiveness, independence, and transparency [10, 11, 27]. Algorithmic impact assessments have emerged as governance tools that require organizations to systematically evaluate potential consequences before deployment, yet implementation varies widely across contexts [4, 18, 28].

A critical tension in ethical governance concerns the relationship between voluntary initiatives and mandatory requirements. While organizations initially approached AI ethics through voluntary commitments, regulatory developments increasingly mandate specific governance practices [9, 16, 29]. This shift raises questions about whether organizations develop genuine ethical capabilities or merely comply with minimum requirements, and whether compliance-oriented approaches adequately address systemic ethical risks [5, 13, 20].

The literature reveals that ethical governance cannot be separated from organizational culture, leadership commitment, and accountability structures. Research demonstrates that effective ethical AI governance requires embedding ethical considerations within organizational processes, including procurement, development, testing, and monitoring [6, 14, 21]. Organizations that establish dedicated ethics functions, review boards, and oversight committees demonstrate greater capacity to identify and address ethical concerns, though such structures vary considerably in authority, resources, and effectiveness [11, 19, 27].

Strategic oversight and organizational accountability

Strategic governance of artificial intelligence addresses how organizations direct, control, and derive value from AI investments while managing associated risks. Research conceptualizes AI governance as an organizational capability that enables firms to balance innovation imperatives with control requirements [1, 2, 8]. This perspective emphasizes that governance structures must support both competitive advantage and responsible deployment, recognizing that inadequate governance creates vulnerabilities that undermine strategic objectives [7, 15, 22].

Accountability constitutes a central concern in strategic AI governance, as organizations must establish clear responsibility for AI system outcomes. The distribution of accountability across technical teams, business units, and executive leadership remains contested, with research documenting the diffusion of responsibility that obscures accountability lines [3, 18, 23]. Effective governance requires specifying roles, responsibilities, and decision rights for AI systems throughout their lifecycle, from initial conception through deployment and monitoring [1, 7, 24].

Strategic oversight mechanisms include governance boards, review committees, and approval processes that evaluate AI initiatives before deployment. Organizations increasingly establish AI governance committees comprising representatives from legal, compliance, technical, and business functions to review high-risk applications [11, 19, 27]. These structures enable cross-functional deliberation but face challenges related to speed, expertise, and decision authority [8, 17, 25]. Research indicates that effective oversight requires both technical expertise to evaluate system capabilities and organizational authority to halt or modify deployments when risks outweigh benefits [2, 6, 28].

The relationship between strategic governance and innovation is marked by persistent tensions. Organizations face pressure to deploy AI rapidly to gain a competitive advantage, yet governance processes that introduce review requirements may slow development cycles [1, 12, 29]. Research explores how organizations can design governance systems that enable responsible innovation without unduly constraining experimentation and value creation [5, 13, 22]. Approaches include risk-based governance that applies more intensive oversight to higher-risk applications while maintaining flexibility for lower-risk initiatives [4, 9, 16].

Regulatory compliance and governance mechanisms

Regulatory developments increasingly shape organizational AI governance as jurisdictions enact binding requirements for AI systems. The European Union's AI Act establishes a comprehensive regulatory framework that classifies AI systems by risk level and imposes corresponding governance obligations [9, 10, 16]. This regulatory approach mandates transparency requirements, human oversight, risk management systems, and conformity assessments for high-risk applications [11, 20, 29]. Organizations subject to such regulations must develop governance capabilities that demonstrate compliance and enable regulatory engagement [4, 17, 26].

Regulatory governance encompasses mechanisms for interpreting and implementing legal requirements within organizational contexts. Research examines how organizations translate regulatory mandates into operational practices, including documentation, testing, monitoring, and reporting processes [2, 14, 24]. Compliance-oriented governance emphasizes traceability, in which organizations maintain records of system development, testing outcomes, and risk assessments to demonstrate regulatory compliance [7, 18, 27].

Sector-specific regulations impose additional governance requirements, particularly in financial services, healthcare, and employment contexts where AI systems have significant individual and societal impacts [3, 12, 21]. Organizations operating across jurisdictions face complexity from divergent regulatory approaches, requiring governance systems that accommodate multiple compliance regimes while maintaining operational efficiency [8, 15, 25]. **Table 1** differentiates the ethical, strategic, and regulatory governance logics of organizational AI, clarifying how each domain operates through distinct objectives, authority bases, and failure risks.

Table 1. Comparative governance logics in organizational AI governance

Governance dimension	Primary governance objective	Core question addressed	Primary authority basis
Ethical governance	Ensure AI use is normatively	Should the organization	Normative principles,

	acceptable and socially responsible	deploy and use this AI in this way?	stakeholder expectation, fairness, and rights consideration
Strategic governance	Align AI systems with organizational objectives while controlling risk and preserving decision quality	How should AI be governed to create value without undermining organizational control?	Executive authority, strategic priorities, internal decision rights, and performance logic
	Achieve conformity with binding legal and sectoral requirements	What must the organization demonstrate and document to satisfy external obligations?	Law, regulation, sector standards, and supervisory expectation
	Reconcile moral responsibility with business objectives	How can firms innovate without normalizing harmful or unfair outcomes?	Shared governance legitimacy across value and strateg
Ethical-strategic intersection			

Strategic-regulatory intersection	Align competitive deployment with legal accountability	How can firms scale AI under increasing regulatory scrutiny?	Internal strategic control plus external legal obligation
Ethical-regulatory intersection	Convert normative concerns into enforceable organizational practice	When does ethical expectation become a mandatory governance duty?	Convergence of public values and legal codification
Integrated governance capability	Coordinate ethical, strategic, and regulatory demands across the AI lifecycle	How can organizations govern AI as a coherent capability rather than a fragmented function?	Combined internal and external legitimacy

A critical question concerns whether regulatory compliance produces effective governance or merely symbolic adherence. Research suggests that compliance-oriented approaches may focus on meeting minimum requirements rather than developing robust governance capabilities [5, 13, 28]. Organizations may adopt governance practices that satisfy regulatory checklists without addressing underlying ethical and strategic concerns [1, 6, 22]. This risk highlights the importance of integrating regulatory requirements with broader ethical and strategic governance frameworks rather than treating compliance as a separate activity [9, 16, 23].

Human oversight, explainability, and trust

Human oversight constitutes a critical governance mechanism that addresses limitations of fully automated decision-making. Research examines how organizations design oversight systems that enable humans to effectively monitor, intervene in, and assume responsibility for AI system outputs [14, 19, 26]. Effective oversight requires consideration of human capabilities, including the ability to detect errors, understand system limitations, and exercise judgment when systems produce unexpected outcomes [17, 20, 29].

Explainability emerges as a foundational requirement for human oversight, as individuals must understand system reasoning to assess output validity and intervene appropriately [2, 7, 24]. Technical explainability methods provide insights into model behavior, but organizational governance must ensure that explanations are accessible, actionable, and integrated into decision-making processes [10, 15, 27]. Research distinguishes between technical interpretability and meaningful transparency, emphasizing that effective governance requires both appropriate technical explanations and organizational processes that enable human understanding and judgment [4, 12, 21].

Trust in AI systems depends on organizational governance practices that demonstrate reliability, accountability, and responsiveness to concerns. Organizations build trust by being transparent about system capabilities, limitations, and governance processes [6, 13, 18]. Research indicates that trust requires not only technical reliability but also organizational accountability structures that enable redress when systems cause harm [8, 14, 25].

The relationship between human oversight and automation presents governance challenges. Organizations must determine the appropriate allocation of decision-making authority between humans and AI systems, considering factors such as task complexity, risk level, and human capacity for effective oversight [3, 11, 28]. Governance frameworks must specify conditions under which human override is required, escalation procedures for uncertain cases, and accountability for decisions made under human oversight [1, 9, 22]. **Figure 1** presents the integrated organizational AI governance architecture linking ethical, strategic, and regulatory governance domains through cross-cutting mechanisms of transparency, explainability, accountability, risk management, and human oversight.



Figure 1. Integrated organizational AI governance architecture. The figure conceptualizes organizational AI governance as a multidimensional capability centered on AI system deployment and shaped by three interacting governance domains: ethical, strategic, and regulatory. These domains are operationalized through cross-cutting mechanisms of transparency, explainability, accountability, risk management, and human oversight. Governance is enacted by leadership, managers, compliance functions, and operational users, while feedback from outcomes, incidents, audits, and monitoring supports continuous refinement of governance structures and practices.

The Integration Challenge: Reconciling Competing Governance Demands

Tensions between innovation and responsible governance

Organizations confront fundamental tensions between the imperative for rapid AI innovation and the requirements of responsible governance. The competitive dynamics of AI adoption create pressure for speed that potentially conflicts with governance processes designed to ensure thorough risk assessment and ethical review [1, 4, 12]. Research documents that organizations face difficult trade-offs between capturing first-mover advantages and implementing comprehensive governance frameworks that may slow deployment timelines [2, 7, 13].

The nature of AI development compounds these tensions. Machine learning systems evolve through exposure to new data, creating governance challenges that differ from traditional information technology systems with stable functionality [3, 5, 8]. Organizations must govern not only

initial system development but also ongoing model behavior that may drift or adapt in ways that introduce new risks over time [6, 9, 17]. This dynamic quality requires governance approaches that extend beyond point-in-time assessments to encompass continuous monitoring, evaluation, and intervention capabilities [10, 14, 22].

Innovation governance tensions manifest differently across organizational contexts. Large established firms with substantial compliance infrastructure may experience governance processes that slow AI deployment relative to competitors. At the same time, smaller organizations may lack governance resources altogether, potentially exposing them to undetected risks [1, 11, 15]. Research suggests that governance systems must be calibrated to organizational capabilities, risk exposure, and strategic context rather than applying uniform approaches across diverse settings [2, 16, 24].

Ethical-regulatory interface: Beyond compliance

The relationship between ethical governance and regulatory compliance presents both opportunities and challenges for organizations. Regulatory developments increasingly codify ethical principles into enforceable requirements, potentially addressing gaps between voluntary commitments and actual practices [9, 10, 20]. However, compliance-oriented approaches risk reducing ethical governance to checklist exercises that satisfy legal requirements without developing genuine organizational ethical capabilities [5, 12, 25].

Research examines how organizations can integrate ethical and regulatory governance to achieve outcomes beyond minimum compliance. Studies suggest that organizations treating regulatory requirements as foundations rather than ceilings develop more robust governance systems that address both compliance obligations and broader ethical responsibilities [4, 16, 27]. This approach requires organizational cultures that value ethical considerations beyond legal risk management, supported by leadership commitment and accountability structures [6, 13, 21].

The ethical-regulatory interface also raises questions about the legitimacy of governance. Ethical frameworks developed through multi-stakeholder processes may command broader legitimacy than regulatory requirements imposed through political processes, yet ethical frameworks lack enforcement mechanisms that give regulations binding

force [9, 14, 28]. Organizations navigate this complexity by developing governance systems that reference both ethical principles and regulatory requirements, creating integrated frameworks that address diverse stakeholder expectations [2, 7, 18].

Strategic-governance alignment: Structures for accountability

Strategic governance requires alignment between organizational structures, processes, and accountability mechanisms that enable effective AI oversight. Research indicates that governance structures vary considerably across organizations, ranging from centralized AI governance committees to distributed approaches that embed governance within business units [1, 8, 19]. The appropriate structure depends on organizational factors, including AI maturity, risk exposure, regulatory requirements, and strategic objectives [3, 11, 23].

Centralized governance structures offer advantages in consistency, expertise concentration, and coordinated oversight across AI initiatives. Organizations establishing centralized AI governance functions can develop specialized expertise, maintain comprehensive inventories of AI systems, and ensure consistent application of governance requirements [10, 17, 26]. However, centralized structures may create bottlenecks, distance governance from operational realities, and reduce business unit ownership of responsible AI practices [4, 15, 24].

Distributed governance approaches embed accountability within business units and functional areas, enabling faster decision-making and closer alignment between governance and operational contexts. Research indicates that distributed approaches can increase business-unit engagement with responsible AI practices by integrating governance into existing workflows rather than imposing it externally [2, 6, 22]. However, distributed structures risk inconsistency, duplication of effort, and gaps in oversight where no single entity maintains visibility across organizational AI activities [5, 12, 29].

Effective strategic governance combines structural elements with process and accountability mechanisms. Research indicates that organizations benefit from establishing clear decision rights, escalation pathways for issues requiring executive attention, and accountability assignments that specify who is responsible for AI system outcomes [3, 9, 18]. Governance processes must be

embedded within AI development lifecycles, with checkpoints for ethical review, risk assessment, and compliance verification at appropriate stages [7, 14, 25].

Governance Mechanisms: From Principles to Practice

Transparency and explainability as operational requirements

Transparency mechanisms translate abstract governance principles into operational requirements that enable oversight, accountability, and trust. Research distinguishes between transparency regarding the existence and capabilities of AI systems, transparency regarding system operation and decision-making, and transparency regarding governance processes and accountability structures [14, 15, 24]. Organizations must address each dimension to develop comprehensive transparency approaches that meet stakeholder expectations and regulatory requirements [2, 17, 27].

Explainability constitutes a specific transparency requirement enabling understanding of how AI systems produce outputs. Technical explainability methods provide insights into model behavior, but organizational governance must ensure that explanations are appropriate to stakeholder needs and decision contexts [4, 10, 20]. Research indicates that effective explainability requires consideration of who requires explanation, for what purpose, and with what level of technical detail, as different stakeholders have varying needs for understanding system operation [6, 12, 22].

Organizations operationalize transparency through documentation requirements, disclosure practices, and communication protocols. Documentation standards specify information that must be maintained about AI system development, testing, and performance, enabling both internal oversight and external accountability [7, 16, 28]. Disclosure practices determine what information organizations share with affected individuals, regulators, and the public about AI system deployment and governance [1, 8, 21].

Risk management frameworks for AI systems

Risk management provides an organizing logic for AI governance, enabling organizations to prioritize governance efforts based on potential harms and likelihood of occurrence. Regulatory frameworks increasingly adopt risk-based approaches that impose more intensive governance requirements on higher-risk applications, while maintaining flexibility for lower-risk deployments [9, 10, 16]. Organizations similarly develop internal risk frameworks that classify AI systems by risk level and apply governance mechanisms proportionally [4, 11, 25].

Effective risk governance requires capabilities for identifying, assessing, and mitigating AI-specific risks that differ from traditional information technology risks. AI risks include technical failures such as inaccurate outputs or system instability, ethical risks such as bias or discrimination, operational risks such as integration failures, and reputational risks from harmful outcomes [3, 5, 13]. Organizations must develop expertise across these risk categories and integrate risk assessment into AI development processes [6, 14, 19].

Risk governance also encompasses monitoring and response capabilities. Organizations must maintain visibility into AI system performance and outcomes to detect emerging risks, with escalation processes for addressing identified issues [2, 15, 26]. Incident response procedures enable organizations to respond to harmful outcomes, including system modification, outcome reversal, and stakeholder communication [1, 7, 18]. Research indicates that organizations with mature risk governance capabilities treat incidents as learning opportunities that inform governance refinement rather than merely as failures to be contained [4, 10, 22].

Accountability structures and organizational roles

Accountability in AI governance requires specification of organizational roles and responsibilities across the AI lifecycle. Research identifies that diffusion of accountability remains a persistent challenge, as AI systems emerge from complex interactions among technical developers, business stakeholders, and operational users, with no single individual or function possessing comprehensive visibility or authority [3, 8, 17]. Organizations address this challenge by explicitly assigning accountability roles and decision rights [2, 12, 23].

Accountability structures typically designate roles including AI system owners responsible for overall system performance and outcomes, technical stewards responsible for development and maintenance, compliance reviewers responsible for governance verification, and operational users responsible for appropriate system application [1, 7, 24]. Role clarity enables accountability assignment when issues arise, though organizations must also establish processes for cross-functional coordination and issue resolution [4, 9, 20].

Executive leadership accountability represents a critical governance dimension. Research indicates that board and executive engagement with AI governance signals organizational commitment, allocates resources, and establishes accountability expectations throughout the organization [5, 11, 19]. Organizations with executive-level AI governance committees demonstrate greater capacity to address cross-cutting issues and balance governance requirements with strategic objectives [6, 13, 27].

Toward Integrated Organizational AI Governance

The governance capability framework

Synthesizing the literature reveals that organizational AI governance is a multidimensional capability that requires integration across ethical, strategic, and regulatory domains. Governance capability encompasses not only structures and processes but also organizational knowledge, cultural commitments, and adaptive capacity that enable organizations to respond to evolving AI risks and requirements [1, 2, 4]. Organizations developing governance capabilities move beyond compliance-oriented approaches toward integrated systems that embed governance across the AI lifecycle [3, 7, 15].

The capability framework emphasizes that governance cannot be reduced to technical tools or compliance checklists but requires organizational transformation. Effective governance integrates ethical considerations into development processes, aligns strategic objectives with control requirements, and embeds regulatory compliance within operational practices [5, 9, 14]. This integration demands that organizations develop cross-functional governance competencies, including technical expertise in

AI systems, ethical reasoning, risk assessment, and regulatory knowledge [6, 10, 18].

Governance capability development occurs through organizational learning processes. Organizations build governance capabilities through experience with AI deployment, incident response, and governance refinement [8, 12, 21]. Research suggests that organizations that treat governance as ongoing learning rather than static compliance develop more adaptive and effective governance systems capable of addressing emerging challenges [4, 13, 24].

Maturity pathways and organizational development

The literature suggests that organizations progress through governance maturity stages as they develop AI governance capabilities. Early-stage organizations typically adopt reactive approaches, addressing governance only when issues arise or when regulatory requirements demand attention [1, 5, 16]. Mid-stage organizations establish formal governance structures and processes, developing policies, committees, and review requirements that systematize governance activities [2, 7, 19]. Mature-stage organizations integrate governance into strategic decision-making, embedding governance considerations within innovation processes and organizational culture [3, 9, 22].

Governance maturity pathways vary by organizational context, with factors including industry sector, regulatory exposure, AI maturity, and organizational culture influencing development trajectories [4, 10, 25]. Organizations in highly regulated sectors may develop governance capabilities earlier due to compliance requirements, while organizations in less regulated contexts may face different pressures for governance development [6, 11, 28]. Understanding these contextual factors enables organizations to benchmark governance capabilities against relevant peers and identify appropriate development priorities [8, 14, 26]. **Table 2** compares alternative organizational design choices for AI governance, highlighting the structural trade-offs through which firms balance consistency, speed, expertise, accountability, and adaptive learning.

Table 2. Organizational design choices for AI governance: structural options, trade-offs, and capability consequences

Design dimension	Option A	Option B	Considerations				att resp
Governance structure	Centralized governance	Distributed governance	Centralized governance may be easier to implement but distributed governance may be more effective in the long run.	Compliance orientation	Minimum legal compliance	Capability-building compliance	immediate ease of implementation
Oversight timing	Ex ante review	Continuous lifecycle oversight	Ex ante review may be more effective in preventing risks but continuous oversight may be more effective in managing risks.	Human oversight model	Human-on-the-loop	Human-in-the-loop	Efficient low latency control less error
Governance trigger logic	Uniform controls for all AI systems	Risk-tiered governance	Uniform controls may be easier to implement but risk-tiered governance may be more effective in managing risks.	Transparency orientation	Technical interpretability focus	Stakeholder-oriented transparency focus	Better for debugging and evaluation
Accountability allocation	Collective/shared accountability	Named accountable owners	Collective/shared accountability may be easier to implement but named accountable owners may be more effective in managing risks.	Learning orientation	Incident containment	Incident-based governance learning	Fast turn of mind exercises
Ethics placement	Standalone ethics function	Embedded ethics in workflows	Standalone ethics function may be easier to implement but embedded ethics in workflows may be more effective in managing risks.	Strategic posture	Governance as a constraint on	Governance as an	May show

	innovation	enabler of responsible innovation	experience by researchers
--	------------	-----------------------------------	---------------------------

Future Research Priorities

The literature identifies several priority areas for future research on organizational AI governance. First, research should examine how organizations effectively integrate ethical, strategic, and regulatory governance dimensions, moving beyond siloed approaches toward comprehensive governance frameworks [1, 2, 5]. This includes investigation of organizational structures, processes, and capabilities that enable integration across governance domains [3, 7, 12].

Second, research should explore the effectiveness of different governance mechanisms in achieving desired outcomes. Studies should examine how transparency, explainability, human oversight, and accountability structures influence AI system outcomes, organizational performance, and stakeholder trust [6, 14, 22]. Longitudinal research tracking governance implementation and outcomes would provide insights into what works under what conditions [4, 9, 18].

Third, research should investigate organizational learning processes in AI governance. Studies should examine how organizations develop governance capabilities through experience, how incidents influence governance refinement, and how governance knowledge transfers across organizational contexts [8, 13, 23]. Understanding learning processes would inform the development of governance capabilities and the identification of best practices [10, 15, 27].

Fourth, research should examine the global governance landscape, including how organizations navigate divergent regulatory requirements across jurisdictions and how governance practices vary across cultural and institutional contexts [11, 16, 20]. Comparative research would illuminate how governance approaches adapt to different environments and identify opportunities for governance harmonization [2, 17, 24].

Conclusion

The governance of artificial intelligence represents a critical organizational capability that enables responsible and competitive AI deployment. This review has synthesized the interdisciplinary literature to reveal that effective AI governance requires integration across three interconnected dimensions. Ethical governance addresses normative concerns, including bias, fairness, transparency, and accountability, requiring organizations to translate principles into operational practices. Strategic governance encompasses oversight structures, control mechanisms, and accountability arrangements that align AI deployment with organizational objectives while managing risks. Regulatory governance responds to evolving legal requirements, mandating compliance frameworks that demonstrate conformity with jurisdictional obligations.

The review has identified persistent tensions that organizations must navigate as they develop governance capabilities. The imperative for rapid AI innovation conflicts with governance processes designed to ensure thorough risk assessment and ethical review, requiring organizations to design governance systems that enable responsible innovation rather than constrain it. The relationship between ethical commitments and regulatory compliance presents both opportunities and challenges, with organizations needing to integrate voluntary ethical frameworks with mandatory legal requirements. Strategic alignment requires governance structures and accountability mechanisms appropriate to the organizational context, with trade-offs between centralized and distributed approaches.

Governance mechanisms operationalizing principles include transparency frameworks that enable visibility into AI system operation and governance processes, explainability tools that support understanding and oversight, risk management systems that prioritize governance efforts based on potential harms, and accountability structures that assign roles and responsibilities across the organization. These mechanisms must be embedded within organizational processes, supported by leadership commitment, and integrated across governance domains to achieve effective outcomes.

The analysis contributes to scholarship by synthesizing dispersed literature across management, information systems, and interdisciplinary research, revealing common themes and persistent challenges. For practice, the review

provides frameworks for organizations developing governance capabilities, identifying governance dimensions, mechanisms, and integration requirements that enable responsible AI deployment. Organizations that treat AI governance as an integrated organizational capability rather than a fragmented compliance activity position themselves to realize AI benefits while managing associated risks.

Future research priorities include investigating governance integration across ethical, strategic, and regulatory dimensions; examining the effectiveness of governance mechanisms; studying organizational learning processes; and exploring global governance variation. Addressing these priorities will advance both scholarly understanding and organizational practice, contributing to the development of governance capabilities that enable trustworthy and sustainable AI deployment.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 05 Jul 2025 Revised: 15 Aug 2025 Accepted: 15 Sep 2025
Published online: 18 September 2025

Rights and permissions

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- . Schneider J, Abraham R, Meske C, vom Brocke J. Artificial intelligence governance for businesses. *Inf Syst Manag.* 2023;40(3):229-49.
- Chhillar D, Aguilera RV. An eye for artificial intelligence: insights into the governance of artificial intelligence and vision for future research. *Bus Soc.* 2022;61(5):1197-241.
- Mäntymäki M, Minkinen M, Birkstedt T, Viljanen M. Defining organizational AI governance. *AI Ethics.* 2022;2(4):603-9.
- Batool A, Zowghi D, Bano M. AI governance: a systematic literature review. *AI Ethics.* 2025;5(3):3265-79.
- Camacho Ibáñez J, Villas Olmeda M. Operationalising AI ethics: how are companies bridging the gap between practice and principles? An exploratory study. *AI Soc.* 2022;37(4):1663-87.
- Eitel-Porter R. Beyond the promise: implementing ethical AI. *AI Ethics.* 2021;1(1):73-80.
- Benjamins R. A choices framework for the responsible use of AI. *AI Ethics.* 2021;1(1):49-53.
- Shneiderman B. Bridging the gap between ethics and practice: guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Trans Interact Intell Syst.* 2020;10(4):26.
- Rai N. Why ethical audit matters in artificial intelligence? *AI Ethics.* 2022;2(1):209-18.
- Mökander J, Floridi L. Operationalising AI governance through ethics-based auditing: an industry case study. *AI Ethics.* 2023;3(2):451-68.
- Hadley E, Blatecky A, Comfort M. Investigating algorithm review boards for organizational responsible artificial intelligence governance. *AI Ethics.* 2025;5(3):2485-95.

Camilleri MA. Artificial intelligence governance: ethical considerations and implications for social responsibility. *Expert Syst.* 2024;41(7):e13406.

Larsson S. On the governance of artificial intelligence through ethics guidelines. *Asian J Law Soc.* 2020;7(3):437-51.

Rességuier A, Rodrigues R. AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data Soc.* 2020;7(2):2053951720942541.

Morley J, Floridi L, Kinsey L, Elhalal A. From what to how: an initial review of publicly available AI ethics tools, methods and research to translate principles into practices. *Sci Eng Ethics.* 2020;26(4):2141-68.

Hagendorff T. The ethics of AI ethics: an evaluation of guidelines. *Minds Mach.* 2020;30(1):99-120.

Taeihagh A. Governance of artificial intelligence. *Policy Soc.* 2021;40(2):137-57.

Roberts H, Cows J, Hine E, Morley J, Wang V, Taddeo M, et al. Governing artificial intelligence in China and the European Union: comparing aims and promoting ethical outcomes. *Inf Soc.* 2023;39(2):79-97.

Tallberg J, Erman E, Furendal M, Geith J, Klamberg M, Lundgren M. The global governance of artificial intelligence: next steps for empirical and normative research. *Int Stud Rev.* 2023;25(3):viad040.

Attard-Frost B, Brandusescu A, Lyons K. The governance of artificial intelligence in Canada: findings and opportunities from a review of 84 AI governance initiatives. *Gov Inf Q.* 2024;41(2):101929.

Corrêa NK, Galvão C, Santos JW, Del Pino C, Pontes Pinto E, Barbosa C, et al. Worldwide AI ethics: a review of 200

guidelines and recommendations for AI governance. *Patterns.* 2023;4(10):100857.

de Bruijn H, Warnier M, Janssen M. The perils and pitfalls of explainable AI: strategies for explaining algorithmic decision-making. *Gov Inf Q.* 2022;39(2):101666.

Hauer MP, Krafft TD, Zweig K. Overview of transparency and inspectability mechanisms to achieve accountability of artificial intelligence systems. *Data Policy.* 2023;5:e30.

Kingsman N, Kazim E, Chaudhry A, Hilliard A, Koshiyama A, Polle R, et al. Public sector AI transparency standard: UK Government seeks to lead by example. *Discov Artif Intell.* 2022;2(1):2.

Prunkl CEA, Ashurst C, Anderljung M, Webb H, Leike J, Dafoe A. Institutionalizing ethics in AI through broader impact requirements. *Nat Mach Intell.* 2021;3(2):104-10.

Laux J. Institutionalised distrust and human oversight of artificial intelligence: towards a democratic design of AI governance under the European Union AI Act. *AI Soc.* 2024;39(6):2853-66.

Langer M, Baum K, Schlicker N. Effective human oversight of AI-based systems: a signal detection perspective on the detection of inaccurate and unfair outputs. *Minds Mach.* 2025;35(1):1-30.

Sadek M, Kallina E, Bohné T, Mougenot C. Challenges of responsible AI in practice: scoping review and recommended actions. *AI Soc.* 2025;40(1):199-215.

Bogucka EP, Constantinides M, Šćepanović S, Quercia D. AI design: a responsible artificial intelligence framework for prefilling impact assessment reports. *IEEE Internet Comput.* 2024;28(5):37-45.