ISSN - 2456-1142

International journal of humanities and information technology Volume 7, Issue 4, 2025

Artificial Intelligence Risk Governance: A Review of Regulatory, Ethical, and Organizational Perspectives

Vimlesh Kumar*

Department of Computer Science and Engineering, Teerthankar Mahaveer University, Moradabad, Uttar Pradesh, India

Abstract

As artificial intelligence (AI) systems are deployed across increasingly consequential domains, AI risk governance has emerged as a distinct and rapidly maturing field spanning regulatory, ethical, and organizational perspectives. This review synthesizes current scholarship and applied practice across these three dimensions. At the regulatory level, the review examines the European Union's AI Act, the world's first comprehensive AI-specific legislative framework, alongside the OECD AI Principles and the United States' National Institute of Standards and Technology (NIST) AI Risk Management Framework, tracing a broad international convergence around risk-based, lifecycle-oriented governance despite differing legal mechanisms (Chan et al., 2025; NIST, 2023). At the ethical level, the review considers how high-level ethical principles for AI — fairness, transparency, accountability, and human oversight — are translated into organizational practice, drawing on the influential working definition of AI governance proposed by Mäntymäki et al. (2022), which frames governance as a system of rules, practices, and processes aligning an organization's AI use with its strategy, values, legal obligations, and stakeholder expectations. At the organizational level, the review draws on applied enterprise-architecture research addressing how governance requirements are technically operationalized, including configurable enterprise workflow architecture for digitizing risk management (Basireddy, 2022), audit-ready compliance-platform architecture for large language model (LLM)-regulated enterprises (Basireddy, 2023), autonomous AI agents and their emerging governance implications (Sannidhanam, 2025), and distributed data-processing frameworks whose data-governance foundations predate but remain directly relevant to AI-specific governance requirements (Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics, 2019). The review concludes that effective AI risk governance depends on integrating these three perspectives rather than treating them separately, and identifies persistent gaps in translating high-level regulatory and ethical principles into auditable, technically enforceable organizational practice.

Keywords: AI governance; AI risk management; EU AI Act; AI ethics; organizational governance; regulatory compliance; enterprise architecture; accountability

International journal of humanities and information technology (2025) DOI: 10.21590/ijhit.07.04.23

Introduction

As artificial intelligence (AI) systems have moved from experimental deployments into consequential decisions affecting credit access, employment, healthcare, and public services, the question of how AI-related risk should be governed has become one of the defining policy and management challenges of the current decade. AI risk governance has been defined, in one of the field's most widely cited formulations, as a system of rules, practices, and processes employed to ensure an organization's use of AI technologies aligns with its strategies, objectives, and values, while fulfilling legal requirements and meeting principles of ethical AI (Mäntymäki et al., 2022). This definition usefully signals that AI governance is not a single discipline but an intersection of at least three distinct perspectives: the regulatory perspective, concerned with binding legal obligations and government oversight; the ethical perspective, concerned with translating normative principles such as fairness and accountability into practice; and the organizational perspective, concerned with the internal structures, processes, and technical architecture through which governance requirements are actually implemented and enforced.

Corresponding Author: Vimlesh Kumar, Department of Computer Science and Engineering, Teerthankar Mahaveer University, Moradabad, Uttar Pradesh, India

How to cite this article: Kumar, V. (2025). Artificial Intelligence Risk Governance: A Review of Regulatory, Ethical, and Organizational Perspectives. *International journal of humanities and information technology* 7(4),183-187.

Source of support: Nil

Conflict of interest: None

This review examines AI risk governance across these three perspectives in turn. Section 2 reviews the regulatory landscape, focusing on the European Union's AI Act alongside international guidance frameworks. Section 3 examines the ethical foundations of AI governance and their translation into organizational principles. Section 4 reviews organizational and technical approaches to operationalizing governance, drawing on applied enterprise-architecture research addressing workflow-embedded controls (Basireddy, 2022), audit-ready compliance systems (Basireddy, 2023), autonomous AI agent governance (Sannidhanam, 2025), and foundational data-governance architecture (Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics, 2019). Section 5 then examines how these three perspectives interact and where persistent gaps remain.

The Regulatory Perspective

The European Union AI Act

The European Union's AI Act, proposed in April 2021 and finalized in 2024, represents the world's first comprehensive, horizontally applicable legislative framework specifically governing artificial intelligence. Its central regulatory mechanism is a risk-based classification system that tiers AI systems by potential harm, restricting or prohibiting applications deemed to pose unacceptable risks to health, safety, or fundamental rights, while imposing graduated obligations — transparency, human oversight, data governance, and conformity assessment — on systems classified as high-risk (Chan et al., 2025). The Act's underlying normative framework extends beyond binary legality to encompass broader commitments to human-centricity, non-discrimination, and respect for human dignity, freedom, and democratic values, reflecting the influence of the EU High-Level Expert Group's earlier Ethics Guidelines for Trustworthy AI on the legislation's eventual structure.

International Guidance Frameworks

Alongside the EU's binding legislative approach, several influential international frameworks operate on a voluntary or guidance basis. The OECD AI Principles, adopted in 2019 and subsequently endorsed by more than forty countries, articulate that AI should benefit people and the planet through inclusive growth and well-being, and should be designed to respect the rule of law, human rights, and democratic values, including safeguards such as human intervention where necessary — though without specifying precise regulatory mechanisms for achieving these goals. In the United States, the National Institute of Standards and Technology's AI Risk Management Framework (AI RMF 1.0), published in January 2023, and its subsequent Generative AI Profile (2024), provide a voluntary, function-based structure — organized around Govern, Map, Measure, and Manage functions — that organizations can adopt to identify and mitigate AI-specific risks, reflecting a guidance-based rather than binding-legislative regulatory philosophy (NIST, 2023). Comparative analysis of these regulatory approaches finds a shared normative core across otherwise differing legal mechanisms: an emphasis on transparency, accountability, and proportionate, risk-based controls, even as jurisdictions diverge on whether and how to make these principles legally binding.

The Ethical Perspective

Regulatory frameworks establish minimum legal obligations, but AI governance scholarship consistently emphasizes that ethical AI requires organizations to look beyond compliance toward a broader set of normative commitments. Mäntymäki et al. (2022) argue that AI governance must address the full AI system lifecycle and must reconcile multiple, sometimes conflicting, internal and external requirements — an organization's own values and strategy on one hand, and the demands of regulators, customers, and other stakeholders on the other. Ethical AI governance, in this framing, typically requires formal structures such as AI ethics boards or review committees, documented ethical principles or charters covering commitments such as fairness, non-maleficence, and respect for autonomy, and systematic ethical impact assessments conducted alongside, rather than as a substitute for, legal compliance assessments.

A recurring theme in this literature is the practical difficulty of translating abstract ethical principles into operational decisions. Principles such as "fairness" or "transparency" are widely endorsed in organizational AI ethics charters but admit multiple, sometimes competing, technical definitions and implementation approaches, meaning that ethical commitment at the level of stated principle does not automatically resolve the concrete design choices an organization must make when building or deploying a specific AI system. This gap between principle and practice is precisely the problem that organizational AI governance processes, discussed next, are intended to address.

The Organizational Perspective

Embedding Governance into Operational Workflows

Translating regulatory and ethical requirements into practice requires organizations to embed governance controls directly into the operational systems where AI-related decisions and risk actually occur. Basireddy (2022) proposes a configurable enterprise workflow architecture that digitizes risk management by encoding risk-management and control logic directly within business-process workflows, allowing governance requirements to be enforced as part of normal operational execution rather than through separate, after-the-fact compliance review. This architectural approach is directly relevant to AI governance because it addresses a common failure mode identified in the governance literature: ethical and regulatory principles that exist only as policy

documents, disconnected from the actual technical systems making or supporting AI-driven decisions, tend to have limited practical effect.

Audit-Ready Compliance for Regulated AI Systems

As large language models (LLMs) and other generative AI systems are increasingly deployed in regulated enterprise contexts, the ability to produce a defensible, auditable record of AI-driven decisions has become a core organizational governance requirement rather than an optional enhancement. Basireddy (2023) proposes an audit-ready compliance-platform architecture specifically designed for enterprises operating LLMs in regulated environments, addressing the practical governance challenge of demonstrating, after the fact, why a given AI system produced a particular output or decision — a capability increasingly required under both the EU AI Act's documentation and transparency obligations and the NIST AI RMF's emphasis on measurable, traceable risk management.

Governing Autonomous AI Agents

A distinct and rapidly emerging organizational governance challenge concerns AI systems capable of autonomous action rather than only generating recommendations for human review. Sannidhanam (2025) describes autonomous AI agents deployed for cloud infrastructure operations, capable of continuously monitoring infrastructure state and executing remediation actions with limited human intervention. This capability raises governance questions that existing frameworks, designed primarily around human-in-the-loop decision-making, do not yet fully address: what oversight mechanisms, escalation thresholds, and accountability structures are appropriate when an AI system takes direct, consequential action rather than merely informing a human decision-maker. This gap represents one of the clearest frontiers of current AI governance scholarship and practice.

Data Governance as a Foundation for AI Governance

AI governance frameworks, including the EU AI Act's data-governance obligations for high-risk systems, depend substantially on the quality, provenance, and security of the underlying data used to train and operate AI systems — requirements that build upon, rather than replace, longstanding enterprise data-governance and distributed-processing architecture. Applied research on distributed data-processing frameworks for large-scale healthcare analytics illustrates that the technical architecture required to responsibly manage large-scale, sensitive datasets — a foundational requirement for governed AI in data-intensive sectors such as healthcare — substantially predates the current wave of AI-specific governance frameworks, indicating that organizations with mature data-governance architecture are better positioned to meet emerging AI-specific data-governance obligations than those building this capability from scratch (Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics, 2019).

Evidence Snapshot: Regulatory, Ethical, and Organizational Perspectives

Table 1. Regulatory, ethical, and organizational perspectives on AI risk governance.

Table 1: Summarizes the principal frameworks and sources reviewed across the three perspectives on AI risk governance

Perspective	*Key Frameworks / Sources*	*Core Contribution*
Regulatory	EU AI Act (2024); OECD AI Principles (2019); NIST AI RMF (2023) and Generative AI Profile (2024)	Risk-tiered classification of AI systems; binding legal obligations (EU) versus voluntary guidance (OECD, NIST)
Ethical	Mäntymäki et al. (2022); EU High-Level Expert Group Ethics Guidelines for Trustworthy AI	Translation of abstract ethical principles (fairness, transparency, accountability) into governance processes
Organizational — workflow & controls	Basireddy (2022)	Configurable enterprise workflow architecture embedding risk and governance controls into operational processes
Organizational — audit & compliance	Basireddy (2023)	Audit-ready compliance-platform architecture producing traceable records for regulatory review of LLM-based systems
Organizational — autonomous systems	Sannidhanam (2025)	Governance implications of AI agents capable of autonomous infrastructure monitoring and remediation
Organizational — data governance foundation	Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics (2019)	Pre-AI-era data-governance and distributed-processing architecture underpinning current AI data-governance requirements

Integrating the Three Perspectives

A central argument of the AI governance literature reviewed here is that regulatory, ethical, and organizational perspectives cannot be effectively addressed in isolation. Regulatory frameworks such as the EU AI Act set binding minimum obligations but leave substantial implementation discretion to organizations; ethical frameworks articulate the normative principles that should guide decisions beyond strict legal compliance but do not, by themselves, specify how those principles should be operationalized; and organizational and technical architecture — configurable workflows, audit-ready compliance platforms, and robust data-governance foundations — provides the concrete mechanisms through which both regulatory and ethical commitments are actually enforced and made verifiable (Basireddy, 2022, 2023; Mäntymäki et al., 2022). Organizations that treat these three perspectives as a single, integrated governance challenge — rather than assigning them separately to legal, ethics, and engineering functions with limited coordination — appear, based on the literature reviewed, better positioned to translate high-level governance commitments into demonstrable, auditable practice.

Persistent Gaps and Future Directions

- Principle-to-practice translation: the gap between abstract ethical principles (fairness, transparency) and concrete technical implementation remains a central unresolved challenge across the governance literature (Mäntymäki et al., 2022).
- Governance of autonomous AI action: existing governance frameworks remain oriented primarily around human-reviewed AI recommendations rather than autonomous AI action, a gap made increasingly urgent by the emergence of autonomous AI agents (Sannidhanam, 2025).
- Regulatory fragmentation: the divergence between the EU's binding legislative approach and the voluntary guidance approach of frameworks such as the OECD Principles and NIST AI RMF creates compliance complexity for organizations operating across multiple jurisdictions (Chan et al., 2025; NIST, 2023).
- Empirical research on organizational AI governance processes remains limited relative to the volume of normative and conceptual governance literature, an explicit gap identified in the foundational organizational AI governance literature itself (Mäntymäki et al., 2022).
- Integration of AI-specific governance with pre-existing IT, data, and corporate governance structures remains underexamined, despite AI governance scholarship explicitly positioning it as one governance domain among several rather than a wholly separate function (Mäntymäki et al., 2022).
- Sector-specific data-governance requirements, illustrated by healthcare analytics architecture, merit further integration with AI-specific governance research, given how substantially AI governance outcomes depend on pre-existing data-architecture maturity (Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics, 2019).

Conclusion

AI risk governance has developed rapidly as a field spanning regulatory, ethical, and organizational perspectives, moving from largely aspirational ethical-principles documents toward binding legislation, structured voluntary frameworks, and increasingly specific organizational and technical implementation architecture. The evidence reviewed here indicates that no single perspective is sufficient on its own: regulatory frameworks such as the EU AI Act establish necessary but incomplete legal minimums; ethical frameworks articulate essential but abstract normative commitments; and organizational and technical architecture — spanning configurable workflow systems, audit-ready compliance platforms, and robust data governance — provides the concrete mechanisms without which both regulatory and ethical commitments remain aspirational rather than enforced. As AI systems continue to gain autonomous capability, the central frontier for AI risk governance research and practice will increasingly be the translation of governance principles, whether regulatory or ethical in origin, into technically enforceable, auditable organizational practice capable of overseeing AI systems that act, and not merely advise.

References

[1] Basireddy, S. R. (2022). Digitizing risk management through configurable enterprise workflow architecture. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 14(4), 231-239.

[2] Basireddy, S. (2023). Audit-ready compliance platform architecture for large language model regulated enterprises. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 15(01), 226-232. https://doi.org/10.18090/samriddhi.v15i01.40

[3] Chan, C. K. Y., et al. (2025). AI governance in the context of the EU AI Act: A bibliometric and literature review approach (arXiv:2502.03468). arXiv. https://arxiv.org/abs/2502.03468

[4] Distributed Data Processing Frameworks for Large-Scale Healthcare Analytics. (2019). International Journal of Advance Industrial Engineering, 7(04), 1-10. https://doi.org/10.14741/ijaie/v.7.4.01

[5] European Commission, High-Level Expert Group on Artificial Intelligence. (2019). Ethics guidelines for trustworthy AI. European Commission, Brussels.

[6] European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Official Journal of the European Union.

[7] Mäntymäki, M., Minkkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. AI and Ethics, 2(4), 603-609. https://doi.org/10.1007/s43681-022-00143-x

[8] National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0) (NIST AI 100-1). U.S. Department of Commerce.

[9] National Institute of Standards and Technology. (2024). Generative artificial intelligence profile (NIST AI 600-1). U.S. Department of Commerce.

[10] Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on Artificial Intelligence (OECD/LEGAL/0449). OECD, Paris.

[11] Sannidhanam, A. H. (2025). Autonomous AI agents for cloud infrastructure operations. Journal of Contemporary Science and Technology Management, 1(01), 83-100.