



AI governance: The architecture of trust in an age of intelligent power

Inspired by Aristotle's ethics, Leonardo da Vinci's systems
thinking and modern AI risk science



August 2026

kpmg.com/in

KPMG. Make the Difference.

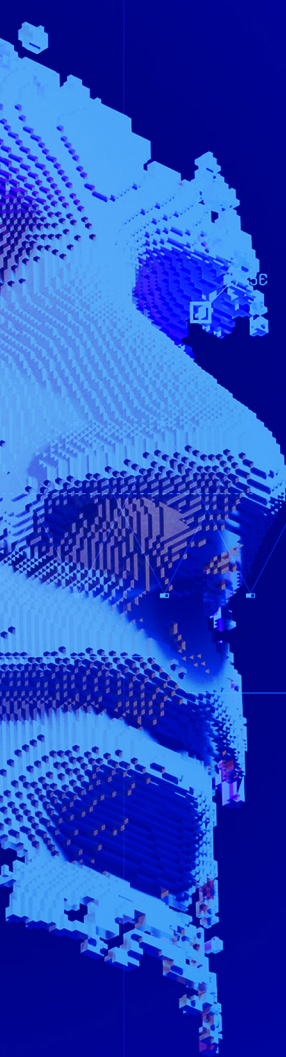
Why AI needs governance, not hesitation

“

Intelligence becomes wisdom only when power meets balanced judgment, and design is combined with responsibility.

”





Artificial Intelligence is not just another technology wave. It is a new form of delegated cognition. Like fire, electricity, finance, or the corporation itself, AI extends human capability, but also magnifies human weakness. Aristotle would ask whether AI serves human flourishing. Leonardo da Vinci would ask whether the design reveals harmony between purpose, form, and consequence. A modern AI scientist would add a third question: can the system be measured, monitored, challenged, and governed across its lifecycle?

Done well, AI can compress cycle times, improve decision quality, augment human creativity, detect risks earlier, personalise customer experiences, strengthen productivity, and transform operating models. KPMG International research across almost 1,400 AI executives describes AI as more than technology: it is an enterprise-wide transformation requiring strategy, culture, and operations to be addressed holistically. The research identifies three phases of value creation: enabling people, embedding AI in work, and evolving the enterprise and ecosystem.

Many organisations confuse AI adoption with AI value. Installing AI models, chatbots, analytics models, or AI agents is only the beginning. Value is created when AI changes the economics of work, improves the quality of decisions, and increases the speed at which the organisation adapts.

The value of AI is not only in the model. It is in the partnership between the model and the human. A global study by KPMG International and a highly reputed university analysed more than 1.4 million real AI prompts and found that effective AI use was distinguished not simply by technical skills or perfect prompts, but by how users challenged initial answers, asked more ambitious questions, and intentionally chose the right tools.

In Aristotelian terms, AI should strengthen practical wisdom. In Leonardo's terms, AI should be an instrument of disciplined imagination: powerful, elegant, useful, and proportionate to its purpose.

KPMG International Trusted AI approach provides a practical framework to secure an organisation from the unfolding risk surface. It translates Responsible AI into enterprise principles: values-driven, human-centric, and trustworthy, supported by pillars such as reliability, security, safety, privacy, sustainability, explainability, data integrity, transparency, fairness, and accountability. KPMG International Trusted AI framework states that trust should be embedded across the AI lifecycle, from design and data to deployment, monitoring and assurance. These principles guide how AI is built, bought, and used responsibly, emphasising that AI should enhance human work, not replace human judgment.

The risk: When intelligent systems escape ordinary controls

Three years ago, an organisation could describe its AI risk posture by describing its models. Were they accurate? Were they biased? Would they fabricate? Those questions still matter, but they now cover a shrinking fraction of the exposure.

As AI moves from experimentation to enterprise-scale deployment, risk emerges across three levels:

- 1 **AI the organisation cannot see**
- 2 **AI that acts or decides**
- 3 **AI under scrutiny**

1

AI the organisation cannot see

Shadow AI: Using unapproved AI tools to summarise, draft reports, write code, classify tickets, or prepare analysis. The productivity benefit may be immediate, but the organisation may lose control over data, intellectual property, confidentiality, compliance, and auditability.

Risk from Third-party and embedded AI: Many AI capabilities now arrive within SaaS platforms, service management tools, analytics platforms, workflow products etc. The complexity increases when third parties rely on AI tools in service delivery. The organisation may not control the model, data flows, security posture, or explainability, yet may remain accountable for the business outcome.

AI that acts or decides

Risk from autonomous AI: Traditional AI produces output. Autonomous AI agent can plan, route, approve, escalate, modify records, access systems, and trigger workflows. Risks such as unauthorised actions, inaccurate decision-making, sensitive data exposure, and limited human oversight are some examples.

Risk from AI influenced decisions: AI generated suggestions may cause biased outcomes, inaccurate decisions, lack of explainability, and contaminated human decisions.

2

AI under scrutiny

Regulatory Non-compliance: Data sovereignty/localisation, inconsistent privacy protections, and cross-border transfer restrictions are some of the challenges in leveraging AI in a global environment.

Other emerging risks include:

Agent identity and entitlement sprawl: As AI agents become digital workers, the key question is not only what the agent says, but what the agent can access and do. Agents may inherit permissions, invoke tools, update records, trigger approvals, and move data across applications.

Information integrity and synthetic-content contamination: AI can contaminate the enterprise knowledge fabric with false, duplicated, synthetic, or low-quality/unverified content. Over time, organisations may lose confidence in what is authentic, what is human-created, what is AI-generated, what is source-backed, and what has simply been repeated enough times to appear true.

3

NIST's work on synthetic content also focuses on provenance, watermarking, detection, and digital content transparency.

How to govern AI practically

AI governance framework must not be just an artifact sitting outside your business process but must be well entrenched in it. Following are some key standards/regulations to build upon an organisational AI Risk Management and Governance Framework.

- ISO/IEC 42001 adds management-system discipline by specifying requirements for establishing, implementing, maintaining, and continually improving an Artificial Intelligence Management System
- NIST's AI Risk Management Framework is useful because it is voluntary, rights-preserving, non-sector specific, and designed to help organisations manage AI risks while promoting trustworthy and responsible AI across design, development, deployment, and use
- The EU AI Act brings a risk-based regulatory lens, classifying AI by risk categories such as prohibited, high-risk, limited-risk, and minimal-risk systems. European Commission guidance on prohibited AI practices highlights harmful manipulation, social scoring, and certain biometric identification practices as unacceptable because of risks to fundamental rights.

Together, KPMG Trusted AI, NIST AI RMF, ISO/IEC 42001, and EU AI Act guardrails point to a practical governance architecture.

Path to AI governance

01

Build an AI inventory before building an AI strategy:

No organisation can govern what it cannot see. Every AI use case, including embedded AI in third-party tools, copilots, models, chatbots, agents, analytics engines, and workflow automation, should be recorded with owner, purpose, users, data categories, vendor, geography, criticality, and decision impact.

02

Classify risk by use case, not by technology label:

A generative AI tool used to draft a note is not the same as an AI model influencing employment, credit, clinical, fraud, safety, legal, or customer-impacting decisions. Risk category should drive review intensity.

03

Establish governance forums with clear cadence:

AI governance requires strategic oversight, management review, and design-level challenge. KPMG's Trusted AI Governance Model describes a governance through a Trusted AI Council, Trusted AI Guardians, and Trusted AI Design Authority, with accountabilities covering strategy, assurance, escalation, lifecycle review, impact assessments, and risk evaluation.

04

Create a mapped AI risk and control matrix:

A single AI control library should map regulatory expectations, privacy, cybersecurity, model risk management, third-party risk, and sector regulations as an integrated controls repository, more importantly, these controls to be embedded across the development and production lifecycle.

06

Govern third-party AI as enterprise AI:

Third-party AI should be governed with the same seriousness as internally built AI. Contracts and vendor controls should address data usage, training restrictions, residency, sub-processors, model changes, audit rights, explainability, incident notification, security, exit rights and traceability.

05

Embed Trusted AI controls into the lifecycle:

Controls should be embedded from ideation to retirement.

- Ideation: Define purpose, users, benefits, and decision rights.
- Data stage: Assess lawful basis, confidentiality, quality, lineage, representativeness, and retention.
- Model/agent build stage: Test accuracy, bias, robustness, explainability, prompt safety, adversarial risk, and drift.
- Deployment: Define human oversight, fallback, logging, disclosures, exception handling and incident response.
- Retirement: Perform impact assessment, plan alternatives and manage data.

07

Treat AI agents as governed identities:

Every AI agent should have a named business owner, clear purpose, scoped permissions, action boundaries, approval thresholds, logs, exception handling, and kill-switch mechanisms. Agents should not inherit broad permissions by default.

08

Measure value and trust together:

Governance fails when it is seen only as a brake. It should be a value accelerator. Every material AI use case should maintain a value ledger and a trust ledger. The value ledger should track productivity, cycle time, cost-to-serve, quality improvement, customer impact, and risk reduction. The trust ledger should track bias exceptions, privacy events, hallucination rates, complaints, drift, security issues, unapproved use, and control effectiveness.

The governed enterprise will win

The enterprise that wins with AI will not be the one that deploys the most models. It will be the one that learns fastest without losing accountability.

Aristotle would remind us that excellence is a habit. Leonardo Da Vinci would remind us that design is the union of art, science, and consequence. AI governance brings these principles together for secure innovation at scale.

KPMG in India contacts:

Akhilesh Tuteja

Head, Clients and Markets
E: atuteja@kpmg.com

Vikash Somani

Partner, Governance, Risk and Compliance Services
E: vsomani@kpmg.com

Kartik Ramakrishnan

Partner and GCC Sector Leader, Healthcare and Life sciences
E: kartikramakrishnan@kpmg.com

Sumit Kapoor

Partner, Governance, Risk and Compliance Services
E: sumitkapoor@kpmg.com

Annapurna Alladi

Partner, Cyber Risks and Controls - South
E: aalladi@kpmg.com

kpmg.com/in



Access our latest insights on
KPMG Insights Edge

Follow us on:

kpmg.com/in/socialmedia



The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavour to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

KPMG Assurance and Consulting Services LLP, Lodha Excelus, Apollo Mills Compound, NM Joshi Marg, Mahalaxmi, Mumbai - 400 011 Phone: +91 22 3989 6000, Fax: +91 22 3983 6000.

© 2026 KPMG Assurance and Consulting Services LLP, an Indian Limited Liability Partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization.

This document is meant for e-communications only.