

BOARD BRIEF

Agentic AI: The Board Has Approved the Investment. Has It Authorised the Action?

APAARNA S KUMMAR | Founder – Nexora Tech | Former CIO – SBI & HSBC | Board Advisor

A Governance Brief for Independent Directors, Audit & Risk Committees, and Technology Oversight Functions

Executive Summary

Boards that approved AI investments in 2023 and 2024 largely understood what they were approving: a model, a tool, a capability. The system would analyse, recommend, or generate. A human would decide. **That understanding is now structurally outdated.**

> Agentic AI does not recommend. It acts. It executes payments, sends communications, modifies records, accesses external systems, and triggers downstream workflows — autonomously, at speed, and often without a human decision point in the loop.

The question a board asked about its large language model deployment — *"Is the model accurate?"* — is no longer the right question. The question now is: **"Who authorised the agent to act, and under what conditions?"**

This is not a technology question. It is a governance question that sits squarely within the board's fiduciary responsibility.

The strategic risk exposure is material. Agentic systems operating across enterprise functions — finance, procurement, customer engagement, compliance — create new categories of operational, regulatory, and reputational risk that existing board governance frameworks were not designed to address. The agent that executes a payment without a human approval threshold is not a technology failure. It is an **authority design failure**. The agent that leaks proprietary data through an indirect prompt injection attack is not a cybersecurity failure alone. It is a **governance control failure**.

Regulators are beginning to reflect this. The EU AI Act's human oversight language anticipates override and interrupt capability in high-risk contexts. NIST's Generative AI Profile explicitly addresses prompt injection and data poisoning as AI value-chain vulnerabilities. India's Digital Personal Data Protection Act creates liability frameworks that agentic deployments will increasingly be subject to. The governance gap between where agentic AI is operating and where board oversight has been designed to reach is widening — and the consequences of that gap are no longer theoretical.

The capital implications are equally significant. Agentic AI investments are being positioned internally as productivity multipliers and cycle-time reducers. Those productivity narratives are reaching boards through management presentations that emphasise benefit metrics without paired risk metrics. A board that approves capital for agentic deployment based on productivity projections alone — without a corresponding agent authority policy, a permissioning framework, and a human override protocol — has not made a capital allocation decision. **It has decided to accept the liability.**

> This Board Brief defines the governance architecture boards must put in place **before** the next agentic AI deployment is approved, not after the first incident requires explanation.

Context & Urgency Framing

For most of the past three years, enterprise AI governance has operated within a relatively comprehensible framework: a model is trained or procured, deployed for a defined task, produces an output, and a human acts on that output. The governance questions were consequential — bias, accuracy, explainability, data provenance — but **the human remained the actor. The AI remained the instrument.**

Agentic AI fundamentally disrupts this architecture. An AI agent is not an instrument that informs human action. It is a system that takes action itself — browsing the web, calling APIs, executing transactions, modifying databases, sending emails, and coordinating with other agents — in pursuit of a goal it has been given. The human defines the goal. **The agent defines the path.**

Traditional AI

Model produces output → Human decides → Human acts

Agentic AI

Human defines goal → Agent decides path → Agent acts autonomously

This shift from AI-as-tool to AI-as-actor is moving faster than governance frameworks are adapting to it. The UK AI Security Institute has reported rapid capability gains in frontier model performance across tested domains, with autonomy skills in software tasks advancing at a pace that makes last year's risk assessments obsolete. What required human intervention twelve months ago may not require it today. **Boards that set agent governance parameters in 2024 and have not revisited them are governing a system different from the one they approved.**

The enterprise deployment trajectory reinforces this urgency. Agentic AI is no longer a proof-of-concept consideration. It is operational in customer service functions, financial processing workflows, procurement pipelines, and compliance monitoring systems across regulated industries. In each of these contexts, the agent's authority to act — what it can do, what it cannot do, and what requires a human decision — is largely a management-level configuration decision. In most organisations, that configuration has not been reviewed at the board level, validated by the risk committee, or stress-tested by internal audit.

That is the governance gap this Brief addresses.

Risk Taxonomy — Strategic & Operational Risks

Autonomy Creep

The most consequential strategic risk of agentic AI is **autonomy creep** — the gradual expansion of agent authority as frontier model capabilities improve, without a corresponding board decision to authorise that expansion. An agent deployed today for a narrow task may be technically capable of a significantly broader range of actions within 12 months. If the agent's permission boundaries are not explicitly defined and periodically reviewed, autonomy expands by default. Default expansions of authority in enterprise systems are not innovation. **They are governance failures in slow motion.**

Supply-Chain Dependency

A second strategic risk is **supply-chain dependency**. Agentic deployments are typically built on foundation model APIs, external tool integrations, and third-party data connectors. Each integration point is a dependency, and the aggregate creates a systemic exposure to vendor concentration, API deprecation, and third-party security posture — none of which are typically visible in the board-level reporting that accompanies AI investment approvals.

Operational Risks

Agentic systems introduce operational risks that differ categorically from traditional software risk. An agent executing a flawed instruction set does not produce an incorrect report. It executes incorrect actions — potentially at scale across multiple systems — before the error is detected. The error correction challenge in an agentic environment is not retrospective analysis. **It is an operational reversal, which may not always be possible.**

> The segregation of duties principle — foundational to financial controls and audit assurance — is structurally challenged by agents that can initiate, approve, and record a transaction within a single automated workflow. Boards with Audit Committee responsibility for financial controls cannot treat this as a technology configuration matter. **It is a control architecture matter.**

Risk Taxonomy — Regulatory, Cyber & Reputational Risks

Regulatory Risks

The EU AI Act's provisions for high-risk AI systems include explicit requirements for human oversight capability — the ability to override, interrupt, or correct AI system outputs. India's DPDPA creates liability exposure for personal data processed or transmitted by automated systems. The agent's autonomy does not create a regulatory exemption. It creates a **regulatory attribution question: who is responsible for what the agent did?**

Cyber & Data Risks

Indirect prompt injection is among the most operationally significant risks in agentic deployments and among the least understood at the board level. Unlike direct prompt injection, indirect prompt injection embeds malicious instructions in content the agent retrieves during task execution: a webpage, an email, a document, a database record. NIST has explicitly flagged prompt injection as a key vulnerability in AI value-chain security. OWASP's framework for agentic risk identifies it as a primary attack vector. **Most enterprise security architectures were not designed to detect or contain it.**

Reputational Risks

An agent operating in a customer-facing context — service, communications, complaint handling — carries reputational risk disproportionate to its operational footprint. A single agent interaction that produces a harmful, discriminatory, or legally problematic output is not experienced by the customer as a technology error. **It is experienced as an organisational decision.** Courts in multiple jurisdictions have already found organisations liable for chatbot outputs. Agentic systems operating with greater autonomy carry proportionally greater exposure.

Risk Taxonomy — Capital Allocation Risks

> The productivity narrative that accompanies agentic AI investment proposals is typically well-constructed. The risk narrative is typically absent.

A board approving capital for agentic deployment without a paired risk quantification — incident probability, containment cost, regulatory penalty exposure, reputational remediation — is approving a one-sided business case. **The capital allocation decision is incomplete.**

What Boards Typically See

Productivity projections, cycle-time reductions, cost savings, competitive positioning — benefit metrics without paired risk metrics.

What Boards Must Also Require

Incident probability, containment cost, regulatory penalty exposure, reputational remediation — a complete, two-sided business case.



Regulatory Landscape — India

Digital Personal Data Protection Act 2023 (DPDPA)

The DPDPA establishes obligations for data fiduciaries that agentic systems will increasingly implicate. An agent that processes personal data autonomously — without explicit consent mapping, purpose limitation controls, and data minimisation architecture — is a compliance liability that the board cannot outsource to the technology function. The DPDPA's accountability framework places responsibility at the **organisational level, not the system level.**

RBI Technology Risk & IT Governance

RBI's guidance on technology risk and IT governance, including its frameworks for model risk management in financial institutions, establishes expectations for validation, monitoring, and oversight of automated decision systems. Agentic deployments in lending, payment processing, and customer onboarding functions in regulated financial institutions are subject to these expectations, and **boards of regulated entities are responsible for oversight of compliance.**

SEBI Cybersecurity Frameworks

SEBI's cybersecurity frameworks for market infrastructure institutions and registered intermediaries establish requirements for system resilience and incident response that agentic deployments must meet. The Cyber Resilience Framework for the securities sector reflects a regulatory direction of travel toward **explicit board accountability for technology risk** — a direction that agentic AI will accelerate.

Regulatory Landscape — Global Parallels

EU AI Act

Establishes a tiered risk classification for AI systems, with high-risk applications subject to mandatory human oversight, conformity assessments, and post-market monitoring obligations. Agentic systems operating in hiring, credit, critical infrastructure, or law enforcement contexts are likely to be classified as high-risk. Organisations with EU market exposure — directly or through group structures — **cannot treat these obligations as non-applicable to their India-domiciled AI deployments.**

DORA

The EU's Digital Operational Resilience Act applies to ICT risk management across financial entities and their critical third-party providers. Agentic systems built on third-party foundation model APIs may trigger ICT concentration risk assessment requirements. **Boards of financial entities with EU exposure should explicitly assess the applicability of DORA to their agentic AI vendor relationships.**

NIST Generative AI Risk Management Profile

Provides a framework for identifying, assessing, and managing risks specific to generative and agentic AI systems. It is not a regulatory instrument, but **regulators and auditors increasingly reference it as a benchmark for governance adequacy.**

Strategic & Capital Implications

The productivity case for agentic AI is real. Cycle-time reductions in document processing, code generation, customer query resolution, and compliance monitoring are measurable and significant. The competitive case is also real: organisations that deploy agentic capabilities effectively will operate with structural cost and speed advantages over those that do not.

> The governance question is not whether to capture these advantages. It is whether the board has the oversight architecture in place to capture them without accumulating unpriced risk.

Upside: Enterprise Value

Agentic AI deployments that are well-governed — with clear authority frameworks, robust monitoring, and demonstrable human oversight capability — are increasingly a signal of operational maturity that institutional investors and rating agencies are beginning to assess.

Downside: Governance Failure

A material incident — a payment error at volume, a data exfiltration event, a regulatory enforcement action — in an agentic system whose governance the board cannot demonstrate will be assessed as a **governance failure, not a technology failure.** The valuation consequence is institutional.

Investor expectations are evolving in the same direction. ESG frameworks are beginning to incorporate AI governance as a component of the 'G' in governance assessment. Boards that cannot demonstrate structured AI oversight — including for agentic deployments — face increasing scrutiny from institutional investors, PE/VC governance diligence, and proxy advisory assessment.

Operating Model Considerations — Governance Structures

The most consequential governance gap in most organisations' agentic AI operating models is the absence of an **Agent Authority Policy** — a board-approved framework that defines which actions agents are permitted to take autonomously, which require human approval, and which trigger escalation. This policy cannot be drafted by the technology function and filed as a management document. It requires board-level deliberation and approval because it defines the boundary of autonomous action within the enterprise.

Committee Responsibilities

Risk Committee

Has primary responsibility for approving the agent authority policy and defining the risk appetite for autonomous system action. It should require **quarterly reporting** on agent actions executed, override rates, near-miss events, and autonomy boundary breaches.

Audit Committee

Responsible for assurance over agent controls: permissioning architecture, segregation-of-duties compliance in agent-mediated workflows, logging and auditability of agent actions, and incident response adequacy. This is **not a post-deployment audit task.** It requires a controls design review before agentic systems are deployed in regulated functions.

Technology or Innovation Committee

Where it exists, is the appropriate venue for competitive capability assessment: evaluating the productivity and innovation case for agentic deployment against the systemic risk and dependency risks it introduces.

Operating Model Considerations — Talent Implications

> Agentic AI governance requires a capability that most organisations do not currently have at the board or senior management level: the ability to assess agent architecture, evaluate permission boundary design, and interpret agent monitoring telemetry.

Building this capability — through Board education, management development, and selective external advisory — is **not optional. It is a governance infrastructure investment.**



Board Education

Structured governance briefings — not technology briefings — to equip directors with the conceptual framework to ask the right questions about agentic deployments.



Management Development

Building internal capability to assess agent architecture, evaluate permission boundary design, and interpret agent monitoring telemetry at the senior management level.



Selective External Advisory

Engaging specialist external advisors to supplement internal capability gaps, particularly for controls design review and regulatory alignment assessments.

Maturity Framework: Agentic AI Governance

1

Level 1 — Unaware

The board has not distinguished between AI tools and agentic AI systems. Agent deployments are managed entirely within the technology function. There is no agent authority policy, no board-level risk appetite definition for autonomous action, and no audit assurance over agent controls.

2

Level 2 — Reactive

The board receives information about agentic deployments after incidents or regulatory enquiries. There is awareness of the risk category, but no structured governance response. Management is developing policies, but they have not been reviewed or approved by the board.

3

Level 3 — Structured

An agent authority policy exists and has been reviewed by the Risk Committee. Permissioning frameworks are documented. Audit Committee has received a controls assessment for at least one agentic deployment. Monitoring telemetry is defined, though reporting to board committees is irregular.

4

Level 4 — Integrated

The board has approved an enterprise-wide agent authority policy. Quarterly reporting on agent actions, override rates, and near-miss events reaches the Risk Committee. Audit Committee receives annual assurance on agent controls. Capital allocation decisions for agentic AI include paired risk assessments. Board members have received structured education on agentic risk.

5

Level 5 — Embedded Governance

Agentic AI governance is a standing agenda item for relevant board committees. The agent authority policy is reviewed and updated annually or following material capability changes in deployed systems. Board reporting integrates productivity metrics and risk metrics for agentic deployments. The organisation can demonstrate to regulators, auditors, and investors that agentic AI is governed to the same standard as other material operational risks.

Board-Level Questions to Ask Management — Part 1

The following questions are designed for Independent Directors, Audit & Risk Committees, and Technology Oversight Functions to pose to management regarding agentic AI deployments.

1

Complete Inventory

What is the complete inventory of agentic AI systems currently operating in this organisation — and which functions do they touch?

2

Autonomous Actions & Thresholds

For each agentic deployment, what actions can the agent take autonomously, and what approval thresholds govern actions above a defined consequence threshold?

3

Agent Authority Policy

Does the organisation have a board-approved Agent Authority Policy? If not, what is the timeline for developing one?

4

Least-Privilege Design

Are agent permissions designed on a least-privilege basis — and has that design been independently validated?

1

Kill-Switch Protocols

Do we have documented kill-switch or safe-stop protocols for every agentic deployment — and have they been tested?

2

Prompt Injection Defence

How are we defending against indirect prompt injection in tool-using agentic systems, and what is the detection and response capability?

3

Segregation of Duties

Has the internal audit assessed segregation of duties compliance in workflows where an agent can initiate, process, and record a transaction?

Board-Level Questions to Ask Management — Part 2

01

Monitoring Telemetry & Reporting

What monitoring telemetry exists for agent actions, and what reporting does the Risk Committee receive — including override rates, near-miss events, and autonomy boundary breaches?

02

Third-Party & API Risk

How are third-party foundation model providers and API integrations assessed for concentration risk and security posture in the context of our agentic deployments?

03

DPDPA Obligations

Has the organisation assessed its DPDPA obligations in relation to personal data processed or transmitted by agentic systems?

04

Escalation & Incident Response

What is the escalation and incident response protocol if an agentic system takes an action that causes material financial, regulatory, or reputational harm?

05

Productivity vs. Risk Case

How does management distinguish between the productivity case and the risk case when presenting capital proposals for agentic AI, and are risk metrics included in post-deployment performance reporting?

Case Illustrations

Case A: The Autonomous Procurement Agent

A manufacturing organisation deploys an agentic AI system to manage routine procurement workflows — supplier queries, purchase order generation, and invoice matching. The agent is configured with approval authority up to a defined threshold. Within eight months, operations management has informally raised the threshold twice to reduce processing delays. No board or Risk Committee review was triggered.

A prompt-injection attack embedded in a supplier invoice causes the agent to redirect a series of payments to a fraudulent account. The investigation reveals that the agent's permission boundaries had not been reviewed since deployment, and no monitoring telemetry was in place to detect anomalous payment patterns.

The financial loss is recoverable. The control's failure is not. **The board cannot demonstrate that it had a governance framework for autonomous payment systems.**

Case B: The Governed Agent Deployment

A financial services organisation prepares to deploy an agentic AI system for customer service query resolution. Before deployment, the Risk Committee reviews and approves an agent authority policy defining permissible actions, escalation thresholds, and human override protocols. The Audit Committee receives a controls assessment covering permissioning architecture, logging, and incident response.

Quarterly reporting to the Risk Committee includes agent actions executed, override rates, and customer escalation events. When a regulatory examination reviews the deployment eighteen months later, the organisation provides a governance paper trail demonstrating board-level oversight from design through operation. **No enforcement action follows.**

The contrast with a peer institution — whose similar deployment attracted regulatory scrutiny for the absence of human oversight documentation — is instructive.

12–24 Month Action Roadmap

1 Months 1–3: Inventory and Awareness

Commission a complete inventory of agentic AI systems currently operating across the enterprise. Present the inventory to the Risk Committee with an initial assessment of authority boundaries and control gaps. Initiate board education on agentic AI risk — not a technology briefing, but a **governance briefing.**

2 Months 3–6: Policy and Framework Development

Develop a draft Agent Authority Policy for Risk Committee review. Define approval thresholds, human override requirements, and escalation protocols for each category of agent action. Assess existing audit frameworks for adequacy against agentic deployment controls requirements.

3 Months 6–12: Controls Implementation and Assurance

Implement permissioning architecture on a least-privilege basis for all operational agentic deployments. Establish monitoring telemetry for agent actions. Commission Internal Audit to perform a controls assessment for the highest-risk agent deployments. Define the reporting framework for quarterly Risk Committee updates.

4 Months 12–18: Regulatory Alignment

Assess DPDPA compliance obligations for all agentic deployments involving personal data. Evaluate DORA applicability for entities with EU exposure. Review the agent authority policy against the EU AI Act high-risk classification criteria for relevant deployments.

5 Months 18–24: Governance Maturation

Conduct annual review of agent authority policy, incorporating any material changes in deployed system capabilities. Integrate agentic AI risk metrics into enterprise risk reporting. Assess governance maturity against the framework in this Brief and identify remaining gaps. Present a maturity assessment to the Audit Committee.

Strategic Conclusion

> The history of enterprise technology governance contains a consistent pattern: the governance frameworks that would have prevented the most consequential failures were available before those failures occurred. They were not absent. They were not applied.

Agentic AI is at that inflexion point now. The governance architecture required to oversee autonomous systems — agent authority policies, least-privilege permissioning, human override protocols, monitoring telemetry, audit assurance — is not experimental. It is defined. OWASP has mapped the risk taxonomy. NIST has profiled the vulnerability landscape. The EU AI Act has articulated the human oversight standard. **The question facing boards today is not whether the governance framework exists. It is whether the board has applied it.**

The organisations that will derive the greatest long-term value from agentic AI are not the ones that deploy fastest. They are the ones that govern best — whose boards have defined what agents are authorised to do, whose risk committees monitor what agents actually do, and whose audit functions can demonstrate that those two things remain aligned.

That governance quality is not a constraint on competitive advantage. In an environment where regulators, investors, and courts are increasingly assessing organisational accountability for autonomous system behaviour, **it is a competitive advantage.**

> The board that asks **"Who authorised the agent to act?"** — before the first incident — is exercising the governance leadership this moment demands.

One-Page Whitepaper Executive Summary

AGENTIC AI: THE BOARD HAS APPROVED THE INVESTMENT. HAS IT AUTHORISED THE ACTION?

The Governance Shift

- Agentic AI does not recommend — it acts. It executes payments, modifies records, and triggers workflows autonomously.
- The governance question has changed from "Is the model accurate?" to "Who authorised the agent to act?"
- Most boards have approved agentic AI investments without approving an agent authority framework.

The Risk Categories

- **Strategic:** Autonomy creep — agent capability expanding beyond original board authorisation
- **Operational:** Segregation of duties failures in agent-mediated financial workflows
- **Regulatory:** DPDPA liability for personal data in autonomous workflows; EU AI Act human oversight requirements
- **Cyber:** Indirect prompt injection enabling data exfiltration or malicious action via retrieved content
- **Reputational:** Courts have found organisations liable for chatbot outputs; agentic exposure is proportionally greater
- **Capital:** Investment proposals presenting productivity metrics without paired risk quantification

The Regulatory Direction

- **EU AI Act:** Human override capability mandated for high-risk AI systems
- **NIST GenAI Profile:** Prompt injection identified as primary AI value-chain vulnerability
- **India DPDPA:** Automated data processing carries organisational accountability
- **DORA:** ICT concentration risk assessment applies to the foundation model API dependencies

The Governance Architecture Required

- Board-approved Agent Authority Policy defining permissible autonomous actions and thresholds
- Least-privilege permissioning for all agentic deployments, independently validated
- Kill-switch and safe-stop protocols, tested and documented
- Quarterly Risk Committee reporting: agent actions, override rates, near-miss events
- Audit Committee assurance over agent controls before deployment in regulated functions

The Maturity Question: Can your board demonstrate — to regulators, auditors, and investors — that your agentic AI deployments are governed to the same standard as other material operational risks?



Thank You

We extend our sincere gratitude for your valuable time and attention in reviewing this Board Brief on Agentic AI. We hope this document has provided profound insights into the critical governance considerations essential for navigating this transformative technological landscape, particularly from an Indian regulatory perspective.

This comprehensive brief was prepared by **Apaarna S Kummar**, Founder of Nexora Tech and former CIO at SBI and HSBC, bringing extensive expertise in technology strategy and financial sector governance.

> At Nexora Tech, our philosophy is encapsulated in our tagline: **Clarity. Architecture. Governance. Impact.**

Connect with Nexora Tech

- Email: **sales@nexoratechsolutions.com**
- Whatsapp: +91-8799918172

We look forward to partnering with you in building robust and responsible AI governance frameworks.