



TOPICS:

Model Risk

SOURCE:

Infocomm Media Development Authority

IMDA: Updated Model AI Governance Framework for Agentic AI

- The Infocomm Media Development Authority (IMDA) of Singapore's Model Framework provides practical guidance for organisations deploying **AI agents that can autonomously plan, decide and take actions on behalf of humans**. The framework recognises that agentic AI offers significant productivity and innovation benefits but introduces new governance challenges because agents can access sensitive data, interact with external systems, execute transactions and collaborate with other agents.
- The framework outlines **four key governance dimensions**: 1) **Assess and bound risks upfront** – Organisations should evaluate agent use cases based on factors such as autonomy, access to sensitive data, ability to modify systems and potential business impact. Risks should be limited through least-privilege access, clear operational boundaries, and defined permissions; 2) **Make humans meaningfully accountable** – Responsibility for agent actions remains with organisations and their personnel.

Clear roles, governance structures and human approval checkpoints should be established, particularly for high-risk or irreversible actions; 3) **Implement technical controls and processes** – Organisations should deploy safeguards throughout the AI lifecycle, including access controls, guardrails, testing, monitoring, logging, threat modelling, and change management. System-level controls are preferred over prompt-based restrictions for critical tasks; 4) **Enable end-user responsibility** – Users should understand agent capabilities, limitations, and risks, supported by training that promotes effective oversight and preserves professional judgment.

- Overall, the framework promotes a **risk-based, human-centric approach** to deploying agentic AI responsibly while balancing innovation, accountability, and operational resilience.

