



March 2026

LAUNCH OF NEW MODEL AI GOVERNANCE FRAMEWORK FOR AGENTIC AI

On 22 January 2026, the Ministry for Digital Development and Information (MDDI) announced the launch of the new Model AI Governance Framework ("**MGF**") for Agentic AI developed by the Infocomm Media Development Authority (IMDA).

The new MGF for Agentic AI builds on existing frameworks and resources and is targeted at organisations looking to deploy agentic AI, whether by developing AI agents in-house or using third-party agentic solutions.

BACKGROUND

Since the launch of the National AI Strategy in 2019, the government has published a range of frameworks and tools to guide the responsible development and deployment of AI. These include:

- + **Model AI Governance Framework for Traditional AI ("MGF 2020")**¹: First published in 2019 and updated in 2020, the MGF 2020 promotes the responsible use of traditional AI.

- + **Model AI Governance Framework for Generative AI ("MGF for GenAI")**²: Launched in 2024, the MGF for GenAI expands on the existing MGF 2020 to address concerns on generative AI and to foster a broader trusted ecosystem.
- + **AI Verify**³: Launched in 2022, AI Verify is an AI Governance Testing Framework and Toolkit developed by the IMDA and the Personal Data Protection Commission (PDPC) to help companies assess the responsible implementation of their AI system against internationally recognised AI governance principles. The testing framework has been updated to include considerations for generative AI applications.

More recently, on 12 February 2026, Prime Minister Lawrence Wong announced in his 2026 Budget Speech⁴ that the government will establish a new National AI Council to provide strategic direction and to drive Singapore's AI agenda. The government will also launch a new set of national AI Missions to transform key sectors of the economy—namely advanced manufacturing, connectivity, finance, and healthcare—using AI.

¹ The MGF 2020 can be found at <https://www.pdpc.gov.sg/-/media/files/pdpc/pdf-files/resource-for-organisation/ai/sgmodelaigovframework2.pdf>

² The MGF for GenAI can be found at <https://aiverifyfoundation.sg/wp-content/uploads/2024/06/Model-AI-Governance-Framework-for-Generative-AI-19-June-2024.pdf>

³ More information about AI Verify is available at <https://aiverifyfoundation.sg/>

⁴ The 2026 Budget Speech can be found at <https://www.singaporebudget.gov.sg/budget-speech/budget-statement/c-harness-ai-as-a-strategic-advantage#Driving-Transformation-through-AI-Missions>



More information on these initiatives will be announced in due course.

AGENTIC AI - WHAT IT IS AND ITS RISKS

Agentic AI systems are AI-powered agents that can plan and take action across multiple steps to achieve specified objectives. These agents are designed to operate with some independence, making decisions and carrying out tasks to meet user-defined goals.⁵

At their core, agents are software systems built on large language models (LLMs). The risks that arise are familiar ones, including traditional software vulnerabilities and LLM-specific risks like hallucination, bias, data leakage, and adversarial prompt injections.

However, agentic AI introduces new components, such as planning and reasoning, tools, and protocols. The once familiar risks can manifest in different ways:

- + **Planning and reasoning:** An agent can hallucinate and make a wrong plan to complete a task.
- + **Tools:** An agent can hallucinate by calling non-existent tools or calling tools with the wrong input, or calling tools in a biased manner.
- + **Protocols:** New protocols to handle agent communication can be poorly deployed.

Agentic AI systems often involve multiple agents working together. When multiple agents interact, new risks arise, such as:

- + **Escalation of errors:** A mistake by one agent can quickly escalate as its outputs are passed on to other agents.
- + **Unpredictable outcomes:** Agents may compete or coordinate in unintended ways, due to the interaction of complex optimisation algorithms.

MGF FOR AGENTIC AI

The MGF for Agentic AI aims to address these new risks arising from agentic AI so that organisations can develop and use agentic AI responsibly. The framework is structured around four key dimensions.

1. *Assess and bound risks upfront*⁶

- + Organisations should carry out risk assessments and consider agentic-specific factors such as the agents' access to sensitive data and level of autonomy.
- + After identifying the risks, organisations should minimise their impact through early design choices such as setting limits on agent autonomy, tool usage, and data access.

2. *Make humans meaningfully accountable for AI agents' actions*⁷

- + Responsibilities should be clearly allocated within and outside the organisation, by establishing chains of accountability across the agent value chain and lifecycle.
- + Organisations should design effective human oversight to guard against automation bias. For example, significant checkpoints should trigger human approvals, and the effectiveness of such approvals should be regularly audited.

3. *Implement technical controls and processes at each stage of the implementation lifecycle*⁸

- + **During design and development:** Design and implement technical controls to mitigate identified risks in new agentic components such as planning, tools, and protocols.
- + **Pre-deployment:** Test agents for baseline safety and security, to check that agents work as expected and controls are effective. Testing should be carried out across agents' entire workflow and at the multi-agent system level.

⁵ See part 1.1 of the MGF for Agentic AI.

⁶ See part 2.1 of the MGF for Agentic AI.

⁷ See part 2.2 of the MGF for Agentic AI.

⁸ See part 2.3 of the MGF for Agentic AI.



- + **During and post-deployment:** Gradually roll out agents to control risk exposure, limiting access to certain users or features initially. The rollout should be complemented by continuous monitoring and testing, with reporting and failsafe mechanisms in place for agent failures or unexpected behaviours.

4. *Enable end-user responsibility*⁹

- + **Transparency:** Inform users of the agent's capabilities, such as the scope of agent's access to user's data and actions the agent can take, and the contact points whom users can escalate to if the agent malfunctions.
- + **User education:** Train users on how to use agents responsibly and oversee agents effectively, while ensuring that users retain foundational skills.

IMPACT ON STAKEHOLDERS

For Organisations

- + **Prepare for compliance:** Organisations deploying agentic AI should familiarise themselves with the framework and review their governance processes, risk management practices, and technical controls to ensure that they are in line with the framework.
- + **Monitor developments:** The MGF for Agentic AI is subject to public consultation and further refinements. Organisations should stay up to date on any future iterations of the framework.

For End-Users

- + **Greater transparency and protection:** Users of agentic AI systems will benefit from having a more robust system of accountability.

NEXT STEPS

As the MGF for Agentic AI is a living document, IMDA invites interested parties to provide feedback to refine it and welcomes the submission of case studies that demonstrate how agentic AI can be responsibly deployed.

Further resources on Agentic AI and the framework can be found in Annex A of the MGF for Agentic AI.

CONTACT US



SHEETAL SANDHU

Partner at Virtus Law

+65 6661 6523
sheetal.sandhu
@stephensonharwood.com



EUNICE YAO

Partner at Virtus Law

+65 6661 6851
eunice.yao
@stephensonharwood.com



SEE JENG SOO

*Registered Foreign Lawyer
at Virtus Law*

+65 6661 6526
seejeng.soo
@stephensonharwood.com



SHU WEI LEE

Associate at Virtus Law

+65 6661 6897
shuwei.lee
@stephensonharwood.com



MING JING AIK

Associate at Virtus Law

+65 6602 6601
mingjing.aik
@stephensonharwood.com

The Singapore law aspects of this article were written by members of Virtus Law (a member of the Stephenson Harwood (Singapore) Alliance).

⁹ See part 2.4 of the MGF for Agentic AI.