Technical Note

Singapore's Agentic AI Governance and AI Risk Management



Issued

7 October 2026

Issuing agencies

Infocomm Media Development Authority
Monetary Authority of Singapore

Primary applicability

Technology, Risk Management, Data Governance, Compliance and Business.

© Management Solutions 2026. All rights reserved

1. Executive summary

2. Model AI governance framework for agentic AI. Assess and bound risks upfront

3. Model AI governance framework for agentic AI. Make humans meaningfully accountable

4. Model AI governance framework for agentic AI. Implement technical controls and processes

5. Model AI governance framework for agentic AI. Enable end-user responsibility

6. Guidelines on Artificial Intelligence Risk Management. AI Oversight, Capability & Capacity

7. Guidelines on Artificial Intelligence Risk Management. Key AI Risk Management Systems, Policies and Procedures

8. Guidelines on Artificial Intelligence Risk Management. AI Life Cycle Controls

A.1 Why Management Solutions

A.2 Abbreviations

© Management Solutions 2026. All rights reserved

Singapore's AI governance framework and risk management guidelines promote responsible AI adoption by addressing the risks of autonomous AI systems and strengthening AI risk management in financial institutions

Context

- In recent years, **MAS has promoted the responsible use of AI in the financial sector**, notably through the **2018 FEAT principles** (Fairness, Ethics, Accountability and Transparency) followed by further initiatives addressing emerging risks, including those associated with **generative AI**.
- In November 2025, **MAS launched a public consultation** on draft Guidelines on AI Risk Management for Financial Institutions, which closed in January 2026. Subsequently, in October 2026, MAS published the **final Guidelines**, setting out supervisory expectations for AI risk management.
- Building on Singapore's broader AI governance initiatives, in May 2026, IMDA updated its **Model AI Governance Framework for Agentic AI**, which provides **voluntary** guidance to support the responsible development and deployment of increasingly autonomous AI systems.

Next steps

The **MAS Guidelines** establish the following implementation timeline**:**

- **7 October 2027**. End of the transition period: financial institutions are expected to meet the expectations in **Sections 3–4**.
- **7 October 2028**. Financial institutions are expected to meet the expectations in **Sections 5–6**.
- **During 2027.** MAS plans a further consultation on **agentic AI**.

Key aspects of the publications

Model AI Governance Framework for Agentic AI

Assess and bound risks upfront	• Recommends upfront impact analysis, least-privilege permissions, bounded workflows, and emergency stop controls.
Make humans meaningfully accountable	• Allocates lifecycle ownership across internal teams and vendors, requiring informed approvals for high-stakes actions.
Implement technical controls and processes	• Enforces runtime guardrails, rigorous multi-agent testing, immutable audit trails, and gradual rollouts.
Enable end-user responsibility	• Proposes AI disclosures, user training, feedback loops, and core skill preservation for operational continuity.

Guidelines on Artificial Intelligence Risk Management

AI Oversight, Capability & Capacity	• Establishes board and senior management accountability for AI governance while requiring adequate technical infrastructure, continuous training, and skilled resources for operational resilience
Key AI Risk Management Systems, Policies and Procedures	• Requires maintaining an updated AI inventory and assessing risk materiality to apply proportionate controls.
AI Life Cycle Controls	• Sets risk-based safeguards across data management, transparency, fairness, human oversight, third-party risks, testing, monitoring and change management.

© Management Solutions 2026. All rights reserved

The approval decision helps define exactly what the agent may access, decide and execute

Control focus	What to implement	Banking example	Evidence
Assess suitability	• **Assess domain** criticality, sensitive data, external access, action scope, reversibility, autonomy, task/system complexity and vendor opacity. **Evaluate residual risk** after controls.	• For a **KYC assistant**, distinguish **drafting analysis** from **updating customer risk** classifications or approving onboarding.	• Use-case assessment, risk rating, residual-risk decision.
Model threats	• **Map workflows** and how **untrusted inputs** reach memory, tools and other agents. Assess **prompt injection**, **memory poisoning, tool misuse and privilege compromise**. Update the **threat model**.	• Trace whether **instructions hidden** in uploaded customer documents could cause an agent to **transmit data externally**.	• Data/Action flow map and threat-control mapping.
Bound autonomy	• Define **permitted workflows and SOPs**. Prefer **deterministic boundaries** at system/tool level. Contain impact through **isolation** and mechanisms to take agents **offline**.	• Permit **document extraction** and **memo drafting**. Technically block onboarding decisions and production database writes.	• Workflow specification, allowed actions and stop procedures.
Control identity and access	• Issue **unique centrally managed** identities with **accountable** ownership. Scope permissions by **task/session**, avoid **transferable authority** and record delegation. Keep delegated rights within the **authorising** human's rights.	• Give a **document-review agent** read access only to assigned files. Prevent a spawned sub-agent from inheriting broader access.	• Agent register, permissions matrix and delegation records.
Contain ecosystem risks	• Assess **visibility** and **control** over vendor agents and external tools. **Limit connections** where assurance is insufficient and re-evaluate risk as complexity or permissions change.	• Initially restrict a **vendor agent** to its approved platform. Review cross-platform access before enabling it.	• Vendor assessment, integration boundaries and review triggers.

© Management Solutions 2026. All rights reserved

Human accountability can be supported by named owners, informed approvals and evidence that oversight works

Control focus	What to implement	Banking example	Evidence
Assign lifecycle ownership	• Allocate responsibilities for **goals** and **policies**, product design, development, testing, monitoring, cybersecurity and responsible use. **Adapt governance** as capabilities evolve.	• Name a **business owner** for a compliance agent and **designate who approves** scope changes, monitors performance and handles incidents.	• Responsibility matrix and escalation contacts.
Clarify supplier obligations	• Set **contractual responsibilities** for security, performance and data protection. Seek **transparency** on agent capabilities and **data handling**, plus scoped authentication and observable tool access.	• Ask a vendor to **demonstrate** per-agent access logs and data-use controls. **Reduce** sensitive-data access if visibility remains insufficient.	• Contract provisions, vendor disclosures and assurance review.
Design approval boundaries	• **Consider requiring approval** before high-stakes, irreversible or atypical actions. Keep requests contextual and digestible. **Allow plan** edits and additional justification where risk warrants it.	• Before an **external customer communication** or **payment**, show the proposed action, recipient, source evidence and risk context.	• Checkpoint map and approval records.
Verify oversight quality	• **Train reviewers** with domain expertise. Audit override rates, review times and unusual approval patterns. **Low override rates or short reviews** can signal bias and warrant investigation.	• **Sample approved KYC memos** against underlying documents. Investigate **unusually rapid** approvals rather than assuming they prove strong performance.	• Review samples, oversight metrics and corrective actions.
Support and protect reviewers	• **Complement approvals** with real-time monitoring and escalation. **Deny action** by default if approval infrastructure **fails** or an action has **no established approval policy**.	• **Hold a sensitive action** when the approver is unavailable. Escalate **attempted out-of-scope tool** use to an accountable supervisor.	• Fail-closed configuration, alert rules and escalation tests.

© Management Solutions 2026. All rights reserved

Controls can help constrain execution, demonstrate reliability and remain effective after deployment

Control focus	What to implement	Banking example	Evidence
Enforce technical boundaries	• Use **deterministic access/workflow controls**, validated tool inputs, runtime limits, trusted MCP servers and sandboxed execution. Constrain shared memory and structure agent-to-agent communications.	• **Reject** malformed tool calls, **block** unapproved endpoints and stop repeated retries. Keep customer data outside tools that do not require it.	• Control specifications, configurations and enforcement tests.
Test complete workflows	• **Test task execution**, policy adherence and tool use across entire workflows. **Test agents individually and together**, repeatedly, with varied inputs in realistic but safely constrained environments.	• **Test** hallucinated data, indirect prompt injection, access leakage, tool failure and propagation of errors between agents.	• Evaluation datasets, scenario coverage, results and defects.
Roll out gradually	• **Limit** initial users, available tools and exposed systems. **Use deployment experience** to refine controls before expanding capabilities.	• Start with **trained compliance staff** and **read-only document** access. **Enable additional integrations** only **after** reviewing **pilot** findings.	• Pilot scope, rollout plan and expansion decision.
Monitor and intervene	• **Log and trace** actions and interactions. **Prioritise** high-risk activity, preserve immutable trails and set risk-proportionate alerts, human escalation, termination and fallback mechanisms.	• **Alert** on unauthorised writes, anomalous tool sequences and repeated failures. **Halt the workflow** and route the task to a **human**.	• Trace logs, alert thresholds, incident/fallback runbook.
Control changes	• **Version the system** and **define** technical, environmental, performance and regulatory **review triggers**. Scale review to risk and continue testing after deployment.	• **Re-evaluate** a model update, new tool or autonomy increase for cascading effects. **Retest affected workflows** before broader release.	• Version history, impact assessment and regression results.

© Management Solutions 2026. All rights reserved

Clear boundaries and appropriate skills enable users to review agents and maintain continuity when they fail

Control focus	What to implement	Banking example	Evidence
Explain the interaction	• Disclose **AI identity in the interface** at the point of **interaction**. Explain **authorised actions**, limitations, user responsibilities and human contact points.	• A customer-service interface **identifies the agent as AI**, states what it **can resolve** and provides a route to a human for **disputed decisions**.	• Interface notices, capability factsheet and contact routes.
Explain data use	• Describe collection, storage and use of data **in line with privacy policies**. Obtain **explicit consent** where necessary and explain **relevant data-use restrictions**.	• Tell staff which **customer documents** an agent **may process** and whether information persists between sessions. Explain applicable **restrictions**.	• Data-use notices, usage rules and consent design.
Train users for their role	• For staff integrating or overseeing agents**, teach suitable/restricted** use cases, prompting, action scope, common failure modes and effective oversight. **Provide refreshers** as features evolve.	• Train analysts to **detect** unsupported source-of-wealth claims, outdated policy references and looping workflows, and to **check source evidence**.	• Role-based training materials, completion and refreshers.
Enable feedback and control	• Provide ways to **flag wrong actions and overrides**. Support **human escalation**. Where appropriate, allow **stricter user approval** thresholds within organisation-defined limits.	• Give reviewers a way to **reject** a draft, **correct** the result and **report** the issue to the product owner for investigation and improvement.	• Feedback channel, issue records and improvement actions.
Preserve core skills	• Identify **foundational capabilities** at risk of **erosion**. Maintain **training and work exposure** so employees can perform critical processes **manually** when agents fail or are unavailable.	• **Retain analyst practice** in manual document review and **rehearse a workflow** that operates without the agent during an outage.	• Core-skill plan and manual continuity procedures.

© Management Solutions 2026. All rights reserved

The Guidelines establish a comprehensive AI governance framework, defining clear oversight and accountability responsibilities while requiring adequate expertise, resources and technology infrastructure to ensure effective AI risk management

AI Oversight	
General Oversight	• **Ensure effective oversight of AI-related risks**, integrating them into existing risk management frameworks and fostering an appropriate risk culture. • Establish **consistent governance standards**, clear accountability and cross-functional coordination to manage AI risks across the institution. • Ensure **AI use remains aligned with regulatory requirements and supervisory expectations**.
Board responsibilities	• **Approve and periodically review AI governance arrangements**, including strategic direction, risk management frameworks, policies and procedures. • Ensure **AI risks are incorporated into the risk appetite framework**, with clearly defined board and senior management responsibilities. • Maintain **adequate AI knowledge and oversight capabilities**, regularly reviewing governance arrangements to address emerging risks and business developments.
Senior Management responsibilities	• **Implement and regularly review AI risk management frameworks**, ensuring appropriate controls throughout the AI life cycle. • Establish **clear responsibilities across business and control functions**, including AI identification, inventory management and risk assessments. • Define **escalation and reporting procedures** for material AI risks and incidents, ensuring timely board updates and adequate resources.
Local Management responsibilities	• Ensure **group AI frameworks address Singapore-specific requirements**, with adequate local oversight, access to relevant information and escalation mechanisms. • Maintain **accountability to MAS** for effective AI risk management, even when relying on group-level arrangements.

AI Capability & Capacity	
AI Risk Management Capabilities	• Ensure **adequate staff expertise and proper conduct** through recruitment, training and support for responsible AI use and effective risk management. • Allocate **sufficient human, technological and financial resources**, proportionate to the risk profile of AI use cases. • **Periodically review staff capabilities and training programmes**, ensuring they remain effective and address risks from emerging AI technologies.
Technology Infrastructure for AI	• Ensure **adequate technology infrastructure**, including computing capacity, network resources, system memory and secure data pipelines. • Maintain **AI performance, scalability, resilience and cybersecurity**, considering applicable technology risk management requirements and industry frameworks.

© Management Solutions 2026. All rights reserved

The Guidelines establish consistent processes for AI identification, inventory management and risk materiality assessment, enabling financial institutions to apply proportionate controls and maintain effective oversight of AI-related risks.

AI Identification

- **AI identification processes. Establish consistent criteria, policies and procedures** to identify AI use across business functions, including AI embedded in services provided by **material third-party providers**.
- **Roles and responsibilities**. **Designate a control function** to oversee AI identification, resolve classification decisions and ensure processes are **documented, regularly reviewed and updated.**
- **Unidentified AI risks**. Identify and **mitigate risks** arising from limitations in detecting AI use, including **shadow AI** and **embedded third-party AI**, ensuring **residual risks** remain within the institution's **risk appetite.**

AI Inventory

- **Inventory maintenance**. **Maintain an accurate and regularly updated inventory** of AI use cases, systems and models, reflecting **new, modified and decommissioned** AI applications.
- **Integration with existing inventories**. **Incorporate AI into existing inventories** or establish a **dedicated register**, ensuring appropriate links with other relevant inventories to facilitate oversight.
- **Key inventory attributes. Capture relevant information** on AI purpose, approved scope, model type, data, dependencies**, life cycle status, risk materiality**, ownership and supporting documentation.
- **Regular inventory reviews. Periodically reassess inventory design** and information requirements to reflect **emerging AI technologies**, additional safeguards and **third-party AI considerations**.
- **Ownership and independent oversight**. Define responsibilities for inventory management and **designate a control function** to establish procedures and independently oversee its **accuracy and completeness**.

AI Risk Materiality Assessment

- **Assessment methodology**. **Establish a consistent framework** to evaluate the **risk materiality** of each AI use case, ensuring that **higher-risk applications** are subject to **stronger controls and scrutiny.**
- **Inherent and residual risks**. Assess risks before and after applying controls, ensuring residual risk remains within the institution's **risk appetite prior to deployment** and reviewing assessments regularly.
- **Risk dimensions**. Evaluate AI risk materiality considering at least three dimensions: **impact,** including potential consequences for the institution and stakeholders; **complexity,** reflecting technological characteristics and explainability; and **reliance,** considering AI autonomy and human oversight.
- **Assessment responsibilities**. **Assign clear roles** and **designate a control function** to establish the assessment framework, **approve or arbitrate risk ratings** and ensure consistent, documented and regularly reviewed assessments.

© Management Solutions 2026. All rights reserved

The Guidelines require financial institutions to implement proportionate controls throughout the AI life cycle, with clear responsibilities, periodic reviews and appropriate safeguards, including contingency arrangements for high-risk AI use cases

Data Management	• **Ensure data suitability, representativeness, and quality** for the intended AI use, implementing checks to identify **errors, anomalies, bias, and data drift.** • Establish **data classification, security, and privacy controls**, including access restrictions, protection of sensitive data, and **regulatory compliance.** • Maintain **data traceability and auditability**, documenting sources, processing activities, quality assessments, and remediation measures throughout the **AI life cycle.**
Transparency and Explainability	• Establish **transparency and explainability controls** to enable relevant stakeholders to understand AI use, its limitations and the factors influencing AI-generated outcomes. • Apply **risk-based transparency requirements**, particularly for AI use cases affecting customer decisions, considering appropriate **disclosures, explanations and mechanisms for customer redress.**
Fairness	• Define **fair outcomes** and implement controls to **mitigate harmful bias**, ensuring that AI systems do not produce unjustified discriminatory outcomes. • Conduct **fairness assessments** where relevant, using appropriate metrics and considering potential adverse impacts on **different population groups.** • Document **assessment results and mitigation measures**, including identified limitations and actions taken to address unfair outcomes.
Human Oversight	• Establish human oversight mechanisms **proportionate to AI risks**, defining responsibilities, decision-making authority and **escalation procedures** throughout the AI life cycle. • Ensure personnel have the necessary **expertise and authority** to review, challenge and **intervene in AI-generated decisions**, supported by appropriate system design. • Document and **periodically review human interventions**, assessing oversight effectiveness and addressing risks such as **automation bias and decision fatigue**.
Third-Party AI Management	• Apply **risk-based controls** across onboarding, development, and deployment, including testing, documentation, and managing **limited visibility into external systems.** • Maintain **accountability for third-party risks**, assessing provider transparency, **supply chain vulnerabilities, concentration, and system updates.** • Establish **contractual safeguards and contingency arrangements**, considering provider obligations, **audit rights, disruptions, capabilities**, and solution complexity.

© Management Solutions 2026. All rights reserved

The Guidelines establish controls throughout the AI life cycle, covering system evaluation, cybersecurity, pre-deployment reviews, ongoing monitoring, change management and decommissioning, with enhanced safeguards for high-risk AI systems.

Selection and Evaluation	• Select **AI systems, models and features** considering performance, complexity, fairness and explainability, documenting decisions and involving experts where relevant. • Assess **benefits and risks of complex AI**, including hallucinations, limited transparency and security vulnerabilities. • Conduct **risk-based testing** using measurable thresholds, representative datasets, stress and adversarial testing, and overfitting mitigation. • Implement **controls and guardrails** to address AI limitations and test failure modes, particularly for Generative AI and AI agents.
Technology, Cybersecurity and Auditability	• Ensure **secure AI environments** through network segmentation, input validation, authentication, encryption and data protection controls. • Apply **access and third-party security controls**, including role-based access, segregation of duties and oversight of external components. • Maintain **auditable AI documentation** covering data, model selection, training, testing, fairness, explainability and limitations.
Pre-Deployment Reviews	• Conduct **independent pre-deployment reviews** assessing AI suitability, robustness, performance and compliance with established controls. • Require **formal independent validation for high-risk AI**, covering design, data, testing, fairness and limitations; apply documented reviews to lower-risk AI. • **Report findings and remediation actions** to the relevant approval body and ensure recommendations are addressed. • **Perform technology and cybersecurity reviews**, including architecture, access controls, resilience, penetration testing and AI red teaming.
Post-Deployment Monitoring and Review	• Establish **ongoing AI monitoring** with performance thresholds, drift detection and appropriate incident management and escalation procedures. • Define **monitoring responsibilities**, maintaining records of incidents, remediation and AI activities, supported by staff training. • Conduct periodic reviews and revalidations, including independent assessments for high-risk AI and reviews triggered by significant changes.
Change Management and Decommissioning	• Implement **change management controls**, including impact assessments, approvals, version control and safeguards for automatic updates. • Establish **AI decommissioning procedures**, addressing system dependencies, data retention, secure removal and stakeholder notifications.

© Management Solutions 2026. All rights reserved

Management Solutions is experienced in reviewing and developing AI systems across all industries, while ensuring regulatory compliance and meeting supervisors' expectations.

1. **Specialized team.** MS has a team of **1,000+ Data Scientists** who combine **technical and quantitative skills with strong regulatory knowledge and certifications** in leading cloud providers (AWS, Azure and Google).

2. **AI models and regulatory practice.** MS has led the **development of numerous AI models** (supervised learning, unsupervised learning, NLP techniques, deep learning models...) with application in **multiple use cases**: fraud detection, risk classification, energy prediction, AML, XAI, and reputational risk or model risk measurement, among others. At the same time, MS has been involved in the implementation of various regulatory requirements across different industries (financial, telco, insurance...).

3. **Experience with regulators and supervisors.** MS is a "**highly qualified external service provider" to the European Central Bank**, with which it has signed 7 framework agreements in the last 6 years, and to national authorities. For the interpretability of advanced models, **MS works under the recommendations of the EBA in its "Report on Big Data and Advanced Analytics"**, according to its 7 elements of confidence for model development and interpretability.

4. **Interpretable models.** MS has **extensive experience in the development of interpretable models** and the application of interpretability techniques in the industries in which it operates: banking, insurance, energy, telecommunications and other industries.

5. **R&D area.** MS allocates **10% of its capacity to R&D,** allowing it to be at the forefront of Artificial Intelligence. **Co-founding of the iDANAE chair** (intelligence, data, analysis and strategy) **with the UPM** (Universidad Politécnica de Madrid), focused on the development of components that form part of the value cycle of the most important assets of today's society, such as information and knowledge.

6. **Close relationship with the RAC** (Royal Academy of Sciences) and active participation in several **research projects with AI applications in areas such as sustainability** (quantification of climate risk) **and efficient training of neural networks** (training optimization and interpretability in transfer learning architectures).

7. **In-house development of proprietary tools ModelCraft™,** with advanced AI/XAI techniques covering **multiple areas of advanced modeling, including dashboards and proprietary interpretability modules**, as well as management and **definition of architectures and cloud services**; **Gamma™,** a **model governance and MRM** tool, incorporating inventory, workflow management, document repository and MRM reporting; **and Hatari™,** a **reputational risk quantification** tool based on information from media and social networks, using **innovative artificial intelligence and NLP techniques.**

© Management Solutions 2026. All rights reserved

A.2. Abbreviations

Abbreviation	Meaning
AI	Artificial Intelligence
AML	Anti-Money Laundering
AWS	Amazon Web Services
EBA	European Banking Authority
FEAT	Fairness, Ethics, Accountability and Transparency
IMDA	Infocomm Media Development Authority
KYC	Know Your Customer
MAS	Monetary Authority of Singapore
MCP	Model Context Protocol
MRM	Model Risk Management
MS	Management Solutions
NLP	Natural Language Processing
R&D	Research and Development
RAC	Royal Academy of Sciences
SOP	Standard Operating Procedure
UPM	Technical University of Madrid
XAI	Explainable Artificial Intelligence

© Management Solutions 2026. All rights reserved



Javier Calvo
Partner at Management Solutions
javier.calvo.martin@managementsolutions.com

Marta Hierro
Partner at Management Solutions
marta.hierro@msspain.com

© Management Solutions, 2026

All rights reserved. Cannot be reproduced, distributed, publicly disclosed or transformed, whether totally or partially, free of charge or at no cost, in any way or by any means, without the express written authorization of Management Solutions.

The information contained in this publication is merely to be used as a guideline, is provided for general information purposes and is not intended to be used in lieu of consulting with our professionals. Management Solutions is not liable for any use that third parties may make of this information. The use of this material is not permitted without the express authorization of Management Solutions.

For more information please visit

www.managementsolutions.com

Or follow us at: