Agentic AI: Accountability and Reasonable Automation in Financial Institutions



15 November 2026

Introduction

Agentic artificial intelligence ("AI") offers financial institutions ("FIs") opportunities to deliver services faster, reduce operating costs and increase capacity. Systems that can select and carry out steps within a workflow can change how work is organised, allowing staff to focus on cases requiring their attention. The Monetary Authority of Singapore (the "Authority") recognises AI's potential to improve productivity, services and risk management.[1]

An FI is accountable for how it conducts its business, including its use of agentic AI. Greater autonomy changes what a system can do without individual human approval. It does not transfer the FI's accountability to the system. The practical question is whether the FI can justify the work it entrusts to AI and the arrangements it uses to control that work.

Some commentary nevertheless creates unnecessary doubt about this starting point. A recent article published on UNESCO's Global AI Ethics and Governance Observatory describes an "accountability gap" in agentic workflows and links accountability to the actor able to constrain a system when it acts. It also recognises that responsibility remains with people or legal entities.[2] In our view, that framing gives too much weight to control at the moment of execution. The FI remains accountable for how it selects, configures, deploys and oversees the system, whether or not a person approves each action.

This article explains an FI's continuing supervisory accountability, distinguishes it from contractual responsibility for failures, and sets out a practical framework for assessing whether automation is reasonable. It also offers starting steps for implementing the Authority's AI risk management guidelines proportionately.

---

[1] Monetary Authority of Singapore, MAS Sets Out Supervisory Expectations on Responsible AI Adoption by Financial Institutions (7 October 2026), paragraphs 2 and 6.

[2] Kasia Odrozek, Who answers for the machine? Closing the accountability gap in agentic workflows, Global AI Ethics and Governance Observatory (30 October 2026), opening paragraph and discussion of liability across value chains. The author's views are not necessarily those of UNESCO.

Empowering a System Requires Management Judgment

Delegating work to agentic AI requires an FI to make familiar decisions about authority, limits and supervision. For example, an FI decides what work to entrust to a relationship manager, assesses the person's capabilities, sets limits and determines how performance will be supervised. Giving the manager discretion does not relieve the FI of its obligations. Nor would saying "the employee decided" answer a supervisory concern about the FI's arrangements.

What may the system do? What must it refer for approval? What evidence supports giving it that scope? How will the FI detect unacceptable performance and intervene? The analogy concerns the allocation of authority and control, not the legal status of the software. It becomes imperfect where probabilistic systems act at machine speed across interconnected services. Oversight must reflect those characteristics.[3]

The scope of that delegation matters. A system permitted to recommend that a payment gateway be isolated has a different role from one permitted to disconnect it. Moving from recommendation to execution changes the consequences of an error and the safeguards required.

Supervisory Accountability and Provider Responsibility

The Authority expressly states that FIs remain accountable for their use of AI, including systems supplied by third parties.[4] Its Guidelines on Artificial Intelligence Risk Management (the "AI Guidelines") address the assurance an FI should obtain about a provider's system and its suitability for the FI's intended use.[5]

Suppose a provider designs a payment system and its approval controls, while an FI configures credentials and transaction limits. Their agreement allocates responsibilities for configuration, testing and updates. A model update changes how the system selects payment tools. An S$80,000 payment bypasses the intended S$20,000 approval threshold. Investigation finds that the new route fell outside the approval control and that the FI's credentials permitted its use.

Contractual responsibility may be divided. The analysis would consider the promised design, the FI's configuration obligations, responsibility for testing and notifying changes, and what caused the loss. Neither the update nor the configuration issue, taken alone, settles breach or liability. Any recovery depends on the agreement, the evidence and applicable contractual limitations.

Those contractual questions are separate from the FI's continuing supervisory accountability. The FI should still be able to explain its provider selection, assurance, configuration, change-control, monitoring and response arrangements. A provider's failure may be relevant to that assessment. It does not transfer the FI's accountability to the provider.[6]

Continuing accountability does not mean that every system failure establishes a regulatory breach. The assessment concerns the FI's conduct and controls in the circumstances, including the risks it could reasonably identify and manage.

---

[3] Authority, Guidelines on Artificial Intelligence Risk Management (7 October 2026), paragraphs 1.9–1.11; Authority, MAS Partners Industry to Develop Safeguards for AI Agents in Finance (3 July 2026), paragraph 2.

[4] Authority, MAS Sets Out Supervisory Expectations on Responsible AI Adoption by Financial Institutions (7 October 2026), paragraph 4, third bullet.

[5] Authority, Guidelines on Artificial Intelligence Risk Management (7 October 2026), paragraphs 5.10–5.11.

[6] AI Guidelines, paragraphs 5.10–5.11 and 5.23–5.25.

The Legal Effect of Automated Actions

Supervisory accountability differs from whether a transaction binds an FI. A human representative's actual authority may be express or implied.[7] Apparent authority depends on a representation attributable to the FI and the counterparty's reliance; the representative cannot create authority merely by asserting it.[8] These questions are distinct from whether the FI's supervision was appropriate.

The Electronic Transactions Act 2010 accommodates automated communications through its definitions of "originator" and "information system", without resolving every attribution or authority dispute.[9] Section 15 prevents a contract formed through automated message systems from being denied validity or enforceability solely because no natural person reviewed or intervened in each action or the resulting contract. It does not itself establish attribution, authority or contractual assent, or resolve questions of mistake.[10]

If a system sends an acceptance of amended settlement terms on an FI's behalf, whether the FI is bound still depends on the parties' agreement, communications and conduct, and the applicable contract rules.

Assessing Whether Automation Is Reasonable

An FI's legal and contractual obligations remain the starting point. Its risk appetite cannot enlarge its contractual right to place a hold on customer funds. A well-designed process does not supply a missing contractual basis for the hold.

Agentic AI can turn an incorrect output into an action in another system. The tools, credentials and permissions available to it determine what it can access and do. Errors can spread through successive steps or between agents, causing failures elsewhere in the workflow. Persistent memory can retain incorrect information or malicious instructions that influence later tasks. Malicious instructions embedded in external content can also redirect the system's behaviour through indirect prompt injection. Changes to models, tools, data or operating conditions can alter performance after deployment. The FI should therefore assess the complete workflow, including access permissions, interactions between components and the controls that prevent or contain errors.[11]

Six connected questions help turn those risks into a judgment about the proposed scope and controls. They bring together the AI Guidelines' risk and control expectations as a practical assessment, rather than a separate regulatory test or a scoring formula.

1. Authority granted: What actions may the system take, using which tools, credentials and accounts, and within what transaction or exposure limits? Permission to recommend a payment hold differs from permission to impose it. For an S$80,000 merchant hold, the FI needs evidence that the process selects the correct account, amount and duration.[12]

---

[7] Alphire Group Pte Ltd v Law Chau Loon [2020] SGCA 50, paragraph 7.

[8] Guy Neale and others v Ku De Ta SG Pte Ltd [2015] SGCA 28, paragraphs 93–98 and 103.

[9] Electronic Transactions Act 2010, section 2(1) (definitions of "information system" and "originator"). An intermediary is excluded from the definition of originator.

[10] Electronic Transactions Act 2010, section 15. This does not displace other requirements for a binding contract.

[11] Infocomm Media Development Authority, Model AI Governance Framework for Agentic AI (Version 1.5, 20 May 2026, updated 5 June 2026), sections 1.1.1–1.1.3, 1.2.1–1.2.3, 2.1.2 and 2.3.2–2.3.3; Cyber Security Agency of Singapore, Advisory on Cybersecurity Risks of OpenClaw (28 May 2026), introduction and "For organisations"; AI Guidelines, paragraphs 1.10–1.11 and 5.23–5.25.

2. Foreseeable severity of harm: What could a reasonably foreseeable error do to the FI or its customers? Thirty errors in 10,000 decisions cannot be judged better than 80 without examining their consequences. A S$2,000 hold preventing payroll may matter more than an S$80,000 hold causing little disruption. Average accuracy alone is insufficient.

3. Scale and speed: How many actions could be affected before a problem is contained? Assess concurrent activity and chains of dependent actions, including the harm that could accumulate before monitoring or intervention takes effect. Controls that contain one mistaken payment may be inadequate for a rapid sequence of payments.

4. Detectability and reversibility: What records, alerts and checks would reveal an error in time, and what can still be stopped, corrected or undone? Identify consequences that correction cannot reverse: releasing funds on Monday does not undo missed payroll on Friday. The less reversible the outcome, the stronger the case for safeguards before execution.

5. Reliability of safeguards: What evidence shows that access restrictions, approval gates and other controls work under normal and adverse conditions? Test whether they remain effective after relevant model, configuration or operating changes. Consider how controls could fail together and whether the remaining risk is acceptable; the existence of a documented limit is not evidence that it will hold.

6. Effectiveness of human intervention: Does the reviewer have the information, competence, authority and time to improve the outcome? Human review may identify an error, repeat it or delay intervention. The Authority indicates that higher-risk uses, particularly those affecting customers, typically warrant stronger human oversight and less autonomy.[13] Where action is too fast for timely review, the design needs effective approval gates, limits or interruption arrangements.

These factors should be considered together. Severe, irreversible consequences may justify a narrower scope even where errors are infrequent. Reliable limits and timely detection may support greater autonomy in a more contained use case. The AI Guidelines support such judgments through risk assessment, evaluation, testing, oversight and monitoring.[14] The FI should record why the combination is reasonable and revisit it when the evidence or operating conditions change.

Starting Steps for Implementing the AI Guidelines

Implementation should be proportionate to an FI's size, activities and risk profile, and to each use case's impact, complexity and reliance on AI.[15]

- **Assign ownership and resources:** Identify who is responsible for deployment, monitoring and incident management. Give those people the capabilities, resources and authority needed to perform their roles.[16]
- **Prepare and approve deployment:** Complete and document the applicable testing, validation, reviews

[12] AI Guidelines, paragraphs 3.4(b) and footnote 18, 4.7–4.8, 5.15 and 5.22(a). Specific transaction or exposure limits are implementation choices informed by the risk assessment.

[13] Authority, Response to Feedback Received on Guidelines on Artificial Intelligence Risk Management (7 October 2026) (the "Consultation Response"), paragraphs 10.48–10.50. See also AI Guidelines, paragraph 5.9.

[14] AI Guidelines, paragraphs 3.4(b), 4.10–4.12, 5.3, 5.9, 5.14–5.16 and 5.18–5.25; Consultation Response, paragraphs 10.29 and 10.48–10.49. The six questions are the author's synthesis; the Guidelines do not prescribe a six-factor test or require every transaction to be reversible.

[15] AI Guidelines, paragraphs 1.6, 2.1–2.3, 4.10–4.13 and 5.2; Consultation Response, paragraph 10.48.

[16] AI Guidelines, paragraphs 3.5(d) and (g), 5.1, 5.23(c) and 6.1–6.2.

and approvals before use, and implement any conditions attached to approval. For use cases assessed as high risk, establish and test fallback arrangements and consider suitable override or shutdown controls.[17]

- **Monitor performance and manage changes:** Monitor against agreed thresholds, investigate incidents and reassess use when findings or changed conditions warrant it. Subject significant changes to appropriate review and re-approval before implementation.[18]

FIs are expected to implement the applicable provisions of the AI Guidelines in accordance with the Authority's phased timetable: the governance, oversight and risk-management expectations in sections 3–4 apply from 7 October 2027, and the lifecycle and capability expectations in sections 5–6 are to be implemented by 7 October 2028. The Authority expects appropriate lifecycle controls for higher-risk uses as soon as possible.[19]

Conclusion

Agentic AI can transform how an FI conducts its business. The FI's task is to decide how much authority to delegate and support that decision with evidence that its safeguards are proportionate and effective. That justification should be revisited as performance and operating conditions change. When an automated action is challenged, the FI should be able to explain why its chosen process was reasonable. The system may act autonomously; the decision to give it that authority remains the FI's.

How HM Strategy Can Help

HM Strategy helps FIs implement agentic AI and the AI Guidelines proportionately to their business and use cases. We assist with use-case and risk assessments, operating limits, approval and escalation arrangements, third-party reviews, and policies for deployment, monitoring and incident response. We also support staff training and the documentation of management decisions, helping FIs explain how their implementation addresses the relevant risks and supports their existing obligations.

For further information, contact:

Chris Holland, Partner | chris.holland@hmstrategy.com

www.hmstrategy.com

Disclaimer: The material in this article represents general information only and should not be relied upon as legal advice. Holland & Marie Pte. Ltd. is not a law firm and may not act as an advocate or solicitor for purposes of the Legal Profession Act 1966 of Singapore.

[17] AI Guidelines, paragraphs 5.3, 5.14–5.15, 5.18–5.22 and 5.23(b).

[18] AI Guidelines, paragraphs 5.23–5.25. Paragraph 5.25(a) recognises that significant changes may occur without prior review and calls for compensating controls, such as enhanced monitoring, in relation to high materiality use cases.

[19] AI Guidelines, paragraphs 1.8 and 2.2–2.5; Consultation Response, paragraphs 13.5–13.6. FIs meeting paragraph 2.3's criteria should implement the basic governance policies and procedures in paragraph 2.5 by 7 October 2027. Paragraph 13.5 of the Consultation Response addresses earlier lifecycle controls for high-risk uses.