

Bank Governance Oversight and Controls for Agentic Outsourcing

The Immediate Audit

Phil Hatch | Akholi

October 23, 2026

Every one of the study's 150 engagements contained a reported departure from intended governance, oversight, or control functions. Bank boards should commission an independent audit of delivery, contain material exposure, and require evidence that corrections operate.

The Board Should Commission the Audit Now

Bank boards should authorize an immediate, independent audit of their agentic outsourcing engagements. In the study of 150 engagements serving Tier 1 financial institutions, every engagement had at least one respondent reporting that intended governance, oversight, and control functions were not implemented as designed. The finding concerns the bank's ability to govern a service delivered by agents and people across two firms. A provider assurance statement alone cannot establish whether the bank's controls operate. Figure 1 presents the finding that requires investigation.¹

¹ Phil Hatch, *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*, version 1.8.2 (Akholi, forthcoming manuscript dated November 1, 2026), Introduction, "Executive Summary" and "Evidentiary Basis and Its Limits," <https://doi.org/10.2139/ssrn.7164559>.

Figure 1. The finding that requires a board decision

100%

At least one respondent reported that intended governance, oversight, or control functions were not implemented as designed.

150 of 150 sampled engagements

Commission the independent operating-control audit now.

Reported implementation gaps; not independently established legal breaches or a population estimate.

Source: Phil Hatch, *Agentic AI Outsourcing at Tier 1 Financial Institutions*, version 1.8.2, Introduction, principal findings, and Part 3, “Reported Governance Weaknesses Across All 150 Engagements.”

The board should approve the mandate, resources, independent reporting line, and suspension authority within seven days. Management should preserve records and contain material exposure while the review proceeds. The proposed program requires an inventory within two weeks, supported priority assessments by week eight, and verified initial corrections with a funded plan for remaining work by week twelve. These targets should reflect the portfolio’s exposure, size, and available resources.²

The bank retains governing responsibility, although its legal expression differs by jurisdiction. FINMA’s Swiss supervisory observations address ownership, testing, controls, and contractual responsibility in AI outsourcing. EIOPA’s DORA interpretation confirms that in-scope EU financial entities retain responsibility when ICT services are supplied by third parties. The US agencies’ nonbinding interagency guidance likewise preserves bank responsibility and separates board oversight from management implementation. The audit should therefore examine retained bank duties and provider arrangements under each entity’s applicable requirements.³

Accordingly, directors should approve the audit’s scope and decision criteria while management establishes operating limits and corrects deficiencies. The chief compliance

² Hatch, *Agentic AI Outsourcing*, Part 6, “The Executive Action Plan,” table 6.2 and “Proposed Implementation Schedule.”

³ Swiss Financial Market Supervisory Authority FINMA, *Governance and Risk Management When Using Artificial Intelligence*, FINMA Guidance 08/2024, December 18, 2024, sec. 2.1, 4, https://www.finma.ch/~media/finma/dokumente/dokumentencenter/myfinma/4dokumentation/finma-aufsichtsmittelungen/20241218-finma-aufsichtsmittelung-08-2024.pdf?sc_lang=en; European Insurance and Occupational Pensions Authority, “3501 - DORA 286,” final Q&A, “EIOPA Answer,” accessed October 23, 2026, https://www.eiopa.europa.eu/qa-regulation/questions-and-answers-database/3501-dora-286_en; Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, and Office of the Comptroller of the Currency, *Interagency Guidance on Third-Party Relationships: Risk Management*, June 7, 2023, 1, 20–22, <https://www.federalreserve.gov/supervisionreg/srletters/sr2304a1.pdf>.

officer should connect findings to obligations, the chief audit executive should protect assurance objectivity, and the executive sponsor should resolve dependencies between bank functions and providers. All engagements belong in the inventory; severity and exposure determine the order and depth of examination. A reported implementation gap warrants investigation, but does not by itself establish a legal breach.

Approval Records Do Not Establish Substantive Review

The study identifies a specific reason approval records can mislead directors: the review step may survive after substantive examination has disappeared. An individual contributor in client servicing and communications reported, “when queue is long the checker just click through,” attributing the practice to high message volume and small teams on both sides. A technology engineering contributor described approvals returned in under a minute for a change touching twenty files. The records indicated approval without establishing the scrutiny that made it protective.⁴

This reported behavior constitutes compliance theater in the ordinary sense: an approval represents scrutiny without demonstrating that examination occurred. The phrase is an interpretation of the accounts, not a measured study category. The audit should therefore determine whether reviewers had the evidence, time, knowledge, and authority to challenge material changes. Counting approvals or inspecting policy documents cannot answer that question.

Severity nevertheless varied. Written governance weaknesses appeared on all 150 engagements, including some with comparatively favorable overall ratings. The appropriate response is to preserve functioning controls and prioritize departures capable of unacceptable customer, financial, regulatory, or operational harm. A favorable aggregate assessment cannot excuse one consequential exception, and universal inclusion in review does not require uniform suspension.⁵

The UK FCA’s October 2026 observations make the same distinction between vendor assurance and tested protection. They identify unchallenged vendor assurances as an AI oversight weakness and discuss inventories, model-change notice, contingency arrangements, and exit options. For the bank’s audit, those UK examples support examining whether promised protections remain usable when delivery changes.⁶

⁴ Hatch, *Agentic AI Outsourcing*, Part 2, secs. 4.2–4.3, “Human Review Capacity Against Agent Output Volume” and “Routine Approval of Agent Output at Human Review Points.”

⁵ Hatch, *Agentic AI Outsourcing*, Part 3, “Reported Governance Weaknesses Across All 150 Engagements.”

⁶ Financial Conduct Authority, “AI in Financial Services: Good Practice and Areas for Improvement,” published October 19, 2026, updated October 20, 2026, sec. 4.2, “Third-party Relationships,” <https://www.fca.org.uk/publications/good-poor-practice/ai-financial-services>.

Five Explanations Require Operating Tests

The reported gap between intended controls and operating practice has five connected explanations: unsuitable control design, excessive review volume, inadequate auditability, delayed decisions, and inherited processes. They are audit hypotheses rather than proven causes on every engagement. Testing them together matters because a reviewer with adequate skills may still lack records or intervention authority. Figure 2 converts the explanations into proposed tests supporting a decision on continued operation.

Figure 2. Five explanations, five tests of operating control

Explanation to test	Operating proof required
Control function not fitted to agents	Demonstrate who authorizes, prevents, reverses, and escalates material agent actions.
Output exceeds review capacity	Show substantive review at peak demand and a controlled response when capacity is exceeded.
Material work cannot be reconstructed	Trace a selected outcome across both firms using usable, accessible records.
Decisions lag agent action	Demonstrate containment before exposure exceeds the approved tolerance.
Pre-AI processes remain in place	Test delegated decisions, change acceptance, bank response times, and escalation against actual cadence.

Proposed audit application, not a validated diagnostic instrument.

Source: Author's proposed audit application of the study's Part 2 indicators C148, C149, C152, C153, and C154 and its account of inherited governance cadence; not a validated diagnostic instrument.

The Control Function Must Fit Agent Actions

The first test is whether control responsibilities fit the actions agents can take. Respondents described governance designed for manual checking and decisions that neither firm's representatives had authority to make. Auditors should identify actions that can create a commitment, change a record, release information, or affect a customer, then establish who can authorize, prevent, reverse, or escalate them. Assigning an owner without matching authority leaves the exposure unresolved.⁷

⁷ Hatch, *Agentic AI Outsourcing*, Part 2, sec. 5.3, "Joint Governance Model Designed for Work Performed by People."

Existing model classifications cannot complete this allocation. Revised US model risk guidance, SR 26-2, excludes generative and agentic AI from its scope, while directing banks to use risk management and governance practices to guide controls for excluded tools and systems. The relevant question is therefore which functions govern the agent's actions and consequences, including those outside traditional model coverage.⁸

Review Volume Must Stay Within Demonstrated Capacity

The second test is whether substantive review remains possible at the arriving workload. Respondents described output exceeding joint teams' checking capacity and analysts carrying review duties alongside delivery. The audit should establish how arrival rates, review effort, qualified staffing, backlog, and defect detection interact at normal and peak demand. Those operating measures are needed because the study accounts describe perceived capacity, not a measured staffing ratio.⁹

When demand exceeds demonstrated capacity, the operating design must provide a controlled response: pausing work, using a bounded queue or qualified fallback staff, or narrowing authorization. Mayer Brown's agentic governance analysis identifies alert fatigue and automation bias as threats to effective human oversight, supporting an examination of whether assigned review remains meaningful under pressure.¹⁰

An automated checker can support that response only if its limitations are understood. Management should identify defects that could pass through both agent and checker, and demonstrate how exceptions reach an accountable person. Auditors should test the combined arrangement under relevant defects and volumes. Evidence of detected and resolved errors, alongside completed work, establishes more than a larger queue of approvals.

Auditability Requires a Usable History Across Both Firms

The third test concerns the history available to the bank. Respondents described missing intermediate records, proprietary traces that the other firm could not query, and a release reconstructed from personal notes because the system retained only the final state. These accounts identify why an apparently acceptable output may be insufficient evidence of controlled delivery.¹¹

⁸ Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, and Office of the Comptroller of the Currency, "Supervisory Guidance on Model Risk Management," April 17, 2026, sec. II and n. 3, <https://www.federalreserve.gov/frs/guidance/supervisory-guidance-on-model-risk-management.htm>; see also SR 26-2, <https://www.federalreserve.gov/supervisionreg/srletters/sr2602.htm>.

⁹ Hatch, *Agentic AI Outsourcing*, Part 2, secs. 4.2–4.3, "Human Review Capacity Against Agent Output Volume" and "Routine Approval of Agent Output at Human Review Points."

¹⁰ Arsen Kourinian, "Governance of Agentic Artificial Intelligence Systems," Mayer Brown, February 24, 2026, "Mitigation Measures," <https://www.mayerbrown.com/en/insights/publications/2026/02/governance-of-agentic-artificial-intelligence-systems>.

¹¹ Hatch, *Agentic AI Outsourcing*, Part 2, sec. 5.1, "Agent Actions Visible Only Inside Each Firm's Own Tools."

For selected outcomes, reviewers should trace source inputs, relevant agent and model configurations, authorized tool actions, material human alterations, approvals, and bank reporting. The objective is an observable action history and decision basis suited to the control, rather than access to a model's private internal reasoning. Records must be obtainable within applicable rights, protected against inappropriate access, and usable during an incident. An unspecified promise that the provider can assemble them later may not meet the bank's intervention needs.

Intervention Must Occur Before Exposure Exceeds Tolerance

The fourth test connects evidence to the time available for intervention. An individual contributor in corporate and back-office business process outsourcing reported: "When the agent picks wrong supplier record, nobody is switching it off. I raise, then two firms discuss about it and by then hundred more lines are done." The account shows why auditors should examine the entire interval from detection through notification, decision, restriction, and recovery, rather than the speed of one escalation step.¹²

That interval should be compared with the time in which the activity could exceed the bank's exposure tolerance. Where intervention cannot reliably occur in time, management should reduce autonomy, constrain throughput or permissions, introduce preventive limits, or use an alternative delivery route. EY's agentic governance analysis supports placing decision rights, monitoring, and human escalation within this operating design.¹³

Inherited Processes Must Be Redesigned Where Their Cadence No Longer Fits

The fifth test is whether inherited processes support those intervention requirements. A portfolio director described a monthly governance forum attempting to govern continuously changing agent work. Periodic strategic oversight can remain useful while operational decisions require a different cadence. The audit should distinguish delegated decisions, changes requiring prior acceptance, immediate exceptions, and matters reserved for periodic review. It must also examine bank access provisioning, data handovers, and response times that determine what the provider can deliver.¹⁴

McKinsey's agentic security playbook recommends extending established policies and third-party risk practices to agent capabilities, assigning ownership through the life cycle, and maintaining visibility into dependencies and access. Applied to outsourcing, that means repairing the existing control functions around the service and its supporting providers. A new

¹² Hatch, *Agentic AI Outsourcing*, Part 2, sec. 5.2, "Joint Decision Speed Mismatched to Agent Speed."

¹³ David Guarrera, "Why Agentic AI Governance Needs Real-Time Trust Layers," EY, June 8, 2026, discussion of the trust-layer playbook, https://www.ey.com/en_us/insights/ai/agentic-ai-governance-and-real-time-trust.

¹⁴ Hatch, *Agentic AI Outsourcing*, Part 2, "The Shift From Managing People to Managing Agents," opening portfolio-director account.

committee cannot compensate for unresolved authority, evidence, capacity, or intervention time.¹⁵

The Audit Must Follow the Service Across Both Firms

Together, the five tests require a single record connecting each material obligation to operating control. The proposed record should identify the control objective, responsible bank and provider roles, performance evidence, exception route, and authority to restrict operation. It gives the audit a consistent basis for determining whether responsibility, action, and evidence address the same exposure.

The audit should then follow delivery from input to bank acceptance, including work outside the approved process map. Reconcile procurement and outsourcing registers with technology, model, payment, and access records to identify pilots, transitions, production, and material dependencies. Protected staff interviews should accompany records where informal approvals, manual corrections, or parallel spreadsheets may sustain reported performance. Conflicting accounts and missing evidence should remain explicit findings until resolved.¹⁶

Operating tests should independently repeat selected reviews and reconcile performance reports to underlying work. They should cover peak-load defect detection, empowered escalation, suspension, and fallback. Legal and compliance specialists should assess departures against applicable obligations, identify reporting or notification requirements and relevant clocks, and retain the facts supporting their judgments.

Where outcomes may be defective, correction must address delivered work as well as the control that failed. Management should establish the potentially affected population, preserve investigative evidence, and assign responsibility for correction, communication, or redress. Repairing the future approval workflow does not establish that earlier outputs were sound. The board should receive the impact assessment and outstanding customer obligations with the control finding.

Independent assurance must remain separate from management's correction duties. The US interagency guidance addresses review of control design and operation, staffing and expertise, and conflicts. The bank should establish suitable assurance independence, manage adviser conflicts, and define provider access arrangements before fieldwork, including the response to

¹⁵ Benjamin Klein et al., "Deploying Agentic AI with Safety and Security: A Playbook for Technology Leaders," McKinsey, October 16, 2025, "Guiding Principles for Agentic AI Security," <https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/deploying-agentic-ai-with-safety-and-security-a-playbook-for-technology-leaders>.

¹⁶ Hatch, *Agentic AI Outsourcing*, Part 6, "The Executive Action Plan," table 6.2 and "Proposed Implementation Schedule."

unavailable evidence. Figure 3 sets out the proposed timetable for translating findings into correction and continuation decisions.¹⁷

The Initial Twelve Weeks Must Produce Decisions and Evidence

Figure 3. Proposed initial audit and correction program

When	Management action	Evidence the board should receive
Week 1	Authorize and contain	Mandate, independent assurance line, preserved records, suspension authority.
Week 2	Reconcile the portfolio	Inventory covering pilots, transitions, production, and material dependencies.
Weeks 3–4	Prioritize exposure	Risk-ranked reviews, protected reporting, evidence requests, interim restrictions.
Weeks 5–8	Test operating controls	Supported priority assessments, impact findings, owners, and correction plans.
Weeks 9–12	Verify initial corrections	Independent closure evidence, decisions on critical exposure, funded remaining work.

Proposed targets. Unresolved findings remain open and visible; the timetable does not promise full closure.

Source: Adapted from the study's Part 6, table 6.2 and "Proposed Implementation Schedule." Dates are proposed planning targets; timing depends on exposure, portfolio size, evidence access, and resources.

The first twelve weeks should produce supported decisions and initial corrections while leaving unresolved work visible. Prioritize engagements by customer impact, financial exposure, regulatory relevance, autonomy, volume, and known control gaps, beginning where failure could cause the greatest harm before intervention.¹⁸

By week eight, management and the independent audit lead should present priority assessments that specify each critical finding's exposure, owner, containment, correction target, impact assessment, and closure evidence. This links urgency to a stated consequence instead of an aggregate score. Escalate critical matters when identified; do not defer them to the next scheduled report.

By week twelve, independently verified initial corrections should accompany a funded plan for open findings and remaining engagements. Closure must demonstrate the corrected control

¹⁷ Board of Governors, FDIC, and OCC, *Interagency Guidance on Third-Party Relationships*, 1, 20–22.

¹⁸ Hatch, *Agentic AI Outsourcing*, Part 6, "The Executive Action Plan," table 6.2 and "Proposed Implementation Schedule."

performing under relevant conditions, with retained evidence and a named assurance owner. Revised procedures, completed training, or provider attestation alone cannot establish effectiveness. Persistent critical exposure requires an explicit recommendation for restrictions, suspension, redesign, transfer, or exit.¹⁹

Continued Operation Must Rest on Demonstrated Control

The resulting board report should connect assurance findings to the conditions of continued delivery. For each material service, show unresolved critical exposure and containment, failed reconstruction or review tests, overdue decisions, observed capacity limits, and verified closures. Management should state whether operation remains within approved conditions and who accepted residual exposure. An absence of reported incidents is not evidence that the control operated.

The continuation decision should specify permitted activity, autonomy and throughput limits, supporting evidence, unresolved exposure, the accountable approver, and the next review trigger. Where material actions cannot be reconstructed or constrained before unacceptable harm, a defensible alternative control or restriction is required. Where correction is demonstrated, appropriately bounded delivery can proceed. Provider confidence and sunk investment cannot replace that evidence.

The board should commission the audit now. Its purpose is to establish whether authorized actions, substantive review, usable records, and timely intervention function together across both firms. Continued or expanded operation should depend on that proof, with proportionate restrictions retained until material weaknesses are corrected or otherwise resolved.

Evidence and Source Note

The study records 850 delivery-team respondents' experiences during January–June 2026 in a nonrepresentative, single cross-section. Banks and vendors were not approached for corroboration; the findings establish neither causation nor population prevalence. This paper uses the author-supplied version 1.8.2, dated November 1, 2026, as forthcoming material, with public sources checked through October 23, 2026. External authorities provide context, not independent validation. The diagnostic, control record, and continuation tests are proposed applications requiring institution-specific testing.²⁰

¹⁹ Hatch, *Agentic AI Outsourcing*, Part 6, “The Executive Action Plan,” table 6.2 and “Proposed Implementation Schedule.”

²⁰ Hatch, *Agentic AI Outsourcing*, Introduction, “Executive Summary” and “Evidentiary Basis and Its Limits”; Hatch, *Agentic AI Outsourcing*, Part 3, “Reported Governance Weaknesses Across All 150 Engagements.”

Bibliography

- Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, and Office of the Comptroller of the Currency. *Interagency Guidance on Third-Party Relationships: Risk Management*. June 7, 2023. <https://www.federalreserve.gov/supervisionreg/srletters/sr2304a1.pdf>.
- Board of Governors of the Federal Reserve System, Federal Deposit Insurance Corporation, and Office of the Comptroller of the Currency. "Supervisory Guidance on Model Risk Management." April 17, 2026. <https://www.federalreserve.gov/frs/guidance/supervisory-guidance-on-model-risk-management.htm>.
- European Insurance and Occupational Pensions Authority. "3501 - DORA 286." Final Q&A. Accessed October 23, 2026. https://www.eiopa.europa.eu/qa-regulation/questions-and-answers-database/3501-dora-286_en.
- Financial Conduct Authority. "AI in Financial Services: Good Practice and Areas for Improvement." Published October 19, 2026. Updated October 20, 2026. Accessed October 23, 2026. <https://www.fca.org.uk/publications/good-poor-practice/ai-financial-services>.
- Guarrera, David. "Why Agentic AI Governance Needs Real-Time Trust Layers." EY. June 8, 2026. https://www.ey.com/en_us/insights/ai/agentic-ai-governance-and-real-time-trust.
- Hatch, Phil. *Agentic AI Outsourcing at Tier 1 Financial Institutions: Full Study Report*. Version 1.8.2. Akholi. Forthcoming manuscript dated November 1, 2026. Author-supplied manuscript consulted October 23, 2026. <https://doi.org/10.2139/ssrn.7164559>.
- Klein, Benjamin, Charlie Lewis, and Rich Isenberg. "Deploying Agentic AI with Safety and Security: A Playbook for Technology Leaders." McKinsey. October 16, 2025. <https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/deploying-agentic-ai-with-safety-and-security-a-playbook-for-technology-leaders>.
- Kourinian, Arsen. "Governance of Agentic Artificial Intelligence Systems." Mayer Brown. February 24, 2026. <https://www.mayerbrown.com/en/insights/publications/2026/02/governance-of-agentic-artificial-intelligence-systems>.
- Swiss Financial Market Supervisory Authority FINMA. *Governance and Risk Management When Using Artificial Intelligence*. FINMA Guidance 08/2024. December 18, 2024. https://www.finma.ch/~media/finma/dokumente/dokumentencenter/myfinma/4dokumentation/finma-aufsichtsmittelungen/20241218-finma-aufsichtsmittelung-08-2024.pdf?sc_lang=en.