FSB FINANCIAL STABILITY BOARD

Sound Practices for Responsible Adoption of Artificial Intelligence (AI): Consultation report

Response to Consultation

Thomas W (independent response)

1. **Do you agree with the benefits and risks of AI adoption by financial institutions described in this report? Are there any substantive benefits or risks not covered?**

   Yes. The report captures the principal benefits and risks of AI adoption by financial institutions, including efficiency, risk management, customer-service benefits, model risk, data risk, third-party concentration, explainability, cyber risk, and the specific concerns raised by generative and agentic AI.

   One additional risk deserves slightly more explicit treatment: whether records of agentic AI activity can be verified and relied upon outside the trust boundary of the system that produced them. This becomes important where supervisors, auditors, insurers, counterparties, or institutions relying on third-party platforms must reconstruct consequential agent actions after the fact.

2. **Are the sound practices sufficiently comprehensive and clear to enable financial institutions' responsible AI adoption?**

   The sound practices are comprehensive with respect to evidence production and oversight inside the adopting institution. Sound Practice 3 calls for enhanced documentation and logging for agentic use cases. Sound Practice 8 notes that logs of agent activity can support real-time tracking or after-the-fact evaluation. Sound Practice 10 addresses audit trails of agent transactions, payment-system restrictions, and monitoring of agent processes.

   A structural clarification would make the practices stronger. For agentic AI, three functions should be kept distinct: producing a record, independently verifying the record, and deciding whether the verified record can be relied upon for a defined purpose.

   A log entry may show that an event occurred, but it does not by itself show that the event was authorized. A credential may show access, but not delegated authority for a particular protected action. Enhanced logging should therefore bind consequential agent actions to the authority claimed for them, including source, scope, expiry, and revocation status at execution time.

3. **Do the sound practices strike an appropriate balance between managing risks relating to all forms of AI, and addressing some of the risks relating to emerging and new complex forms of AI, such as GenAI and agentic AI?**

   Yes, broadly. The practices appropriately cover AI generally, while recognizing that GenAI and agentic AI create additional risks.

   The distinction proposed above is particularly important for agentic AI because agents may execute actions, use tools, access systems, or interact with third parties. In those settings, governance should not stop at whether an agent was approved or logged. It should also ask whether a particular action was within authority, within scope, and supported by evidence that can be tested outside the producing system's boundary.

4. **Are the sound practices sufficiently flexible to accommodate and address newer types of AI and responsible adoption over time?**

   Yes. The practices are appropriately principles-based and flexible over time.

   The distinction between record production, independent verification, and reliance is also technology-neutral. It should remain relevant as AI systems evolve, because there will always be parties outside the producing system's trust boundary who must decide what weight to give a record. Naming those functions would future-proof the practices without prescribing a specific technical architecture.

5. **Do the case studies in this report sufficiently highlight how different types of financial institutions can benefit from responsible AI adoption? Are there additional case studies for inclusion in the report? If so, please provide such case studies, particularly for nonbanks.**

   The case studies are useful. One additional case study that may be valuable, especially for nonbanks, would involve a regulated buyer, insurer, asset manager, or other nonbank financial institution relying on an external agentic AI platform.

   Such a case could show how the institution obtains evidence of consequential agent actions from a third-party provider, how that evidence is verified outside the provider's own trust boundary, and how the relying institution decides whether the evidence is sufficient for audit, insurance, procurement, or supervisory purposes.

6. **Do the case studies in this report provide actionable insights for financial institutions in their responsible AI adoption?**

   Yes. The case studies provide useful operational insight. They would become more actionable if they more clearly distinguished between ordinary activity logs and records capable of independent verification.

   For agentic AI, a useful case study should show not only that an agent's actions were logged, but also whether the evidence identifies the issuing component, protects integrity, records the claimed authority, captures scope and expiry, and shows whether the authority was revoked or superseded at execution time.

7. Are the definitions in the glossary clear and aligned with industry sound practices, including recent developments in AI?

Two glossary additions would help.

First, distinguish *access* from *authority*. Access means a system can perform an action. Authority means the action is legitimately permitted under the institutional, contractual, or delegated conditions in force at execution time.

Second, consider defining *execution evidence* or *evidence of authority*: a record binding a specific consequential action to the authority claimed for it, with integrity protection and issuer identification sufficient to support verification by a party that did not produce the record.

8. Are there any comments you would like to make on this report? Please fill in.

The final report could strengthen coordination and information-sharing by adding short clarifications to Sound Practices 3, 10, and 12.

For Sound Practice 3, enhanced documentation and logging for consequential agent actions should distinguish ordinary activity records from records capable of independent verification.

For Sound Practice 10, where oversight shifts from human-in-the-loop toward human-on-the-loop or human-in-command, verifiable post-execution evidence of authority should be recognized as a compensating control.

For Sound Practice 12, institutions relying on external agentic platforms should be able to obtain action-level evidence in a form inspectable and verifiable outside the provider’s own trust boundary.

These distinctions are elaborated in ProofRail, a public reference project exploring portable evidence, verification, and reliance for protected agentic AI actions, offered as one open articulation of the concepts rather than as a proposed standard or product.