FSB FINANCIAL STABILITY BOARD

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

Response to Consultation

VeriLinkOS

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?**

   I strongly agree with the FSB's assessment of benefits and risks. However, one risk dimension deserves more emphasis: the gap between documented policies and verifiable enforcement.

   Most financial institutions today rely on logs, dashboards, and attestations to demonstrate compliance. None of these constitute independent evidence. A regulator cannot verify a log without trusting the institution that produced it.

   The FSB's sound practices correctly identify the need for "tamper-evident records" (Sound Practice 3) and "audit trails of agent transactions" (Sound Practice 10). What is missing is a recognition that these records must be verifiable by third parties without access to the institution's internal systems. Cryptographic receipts – anchored to tamper-evident ledgers – address this gap. I recommend the FSB add "lack of independent verifiability" as a distinct risk, and reference cryptographic audit trails as an emerging mitigation.

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

   The sound practices are comprehensive and well-structured. However, they are principles-based. Financial institutions need concrete implementation guidance.

   For example, Sound Practice 10 requires "meaningful human oversight" and "tamper-evident records" but does not specify what constitutes evidence of meaningful oversight. In practice, this requires: (i) cryptographic binding of the human approval to the specific agent action; (ii) timestamped proof that the approval occurred before the action; (iii) auditable records of override rates to detect automation bias ("rubber-stamping").

   I have built a reference implementation that addresses these requirements. I would be pleased to share technical details with the FSB and its members.

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?**

The sound practices correctly identify that agentic AI requires additional attention. Section 1.3 (Agentic AI risks) is excellent, particularly the recognition of unauthorised actions, the impracticality of real-time human monitoring, and the difficulty of overriding or remediating autonomous actions.

However, the sound practices could go further in specifying what governance of agentic AI requires in practice. Three technical requirements are essential:

1. Real-time authority validation. Before an agent acts, the system must verify that the agent's authority remains valid – not just that permission was granted at deployment. Context changes. Permissions expire. The FSB should explicitly require that agentic systems validate authority at the moment of each action, not just at deployment.

2. Cryptographic binding of human oversight. When a human approves an agent action, the approval must be cryptographically bound to the specific action, timestamped, and made verifiable by third parties. This prevents disputes about what was approved and when.

3. Tamper-evident, third-party verifiable receipts. Financial institutions need evidence that survives system decommissioning and can be verified by regulators without access to internal logs. Cryptographic receipts anchored to public or permissioned ledgers address this.

I have implemented these three requirements in a production system currently under evaluation in pilot. I would welcome the opportunity to share this implementation with the FSB.

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

Yes. The principles-based approach is appropriate. The most important requirement for flexibility is that governance mechanisms must be independent of specific AI models or architectures.

Cryptographic receipts meet this test. They do not depend on the underlying AI model (traditional ML, GenAI, or agentic). They simply record what happened, who authorised it, under what policy, and what evidence was considered. This approach will remain valid regardless of how AI technology evolves.

I encourage the FSB to maintain this technology-neutral stance while continuing to highlight emerging risks from agentic AI.

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 existing case studies are helpful but focus on large, internationally active banks. I would be pleased to contribute a case study on cryptographic governance for agentic AI, applicable to banks, asset managers, insurers, and nonbank financial institutions.

The case study would demonstrate:

An agentic AI workflow for transaction monitoring (applicable to banks and payment providers)

Integration of human-in-the-loop approval for high-risk actions

Generation of tamper-evident, third-party verifiable receipts

Detection of automation bias through override rate analytics

I have attached a technical briefing for the FSB's reference. I am available to provide additional detail or a live demonstration.

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

The case studies provide useful illustrations but are largely qualitative. Financial institutions need technical specifications they can implement.

The most actionable insight would be a reference architecture for cryptographic audit trails of agentic AI actions. This would include:

API specifications for authority validation

Receipt schema for tamper-evident records

Verification protocols for third-party auditors

I have built such a reference architecture and would be pleased to share it with the FSB and its members as a basis for future guidance.

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

The glossary is clear and appropriately technology-neutral. I note that "agentic AI" is defined, which is helpful. However, the glossary does not define "tamper-evident record" or "cryptographic receipt". Given the importance of these concepts to Sound Practices 3 and 10, I recommend adding definitions in future iterations.

Suggested definitions:

"Tamper-evident record: A record whose alteration is cryptographically detectable, even by the party that created it."

"Cryptographic receipt: A digitally signed, verifiable record of an AI action that binds together the action, its authorising authority, applicable policy, and any human approval."

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

I commend the FSB for this comprehensive and timely consultation. The sound practices provide an excellent foundation. I urge the FSB to consider cryptographic audit trails as an emerging best practice for operationalising agentic AI governance, and I offer my technical implementation as a reference for future guidance.

I am available to provide a technical briefing to the FSB secretariat or its members at their convenience.