FSB FINANCIAL STABILITY BOARD

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

Response to Consultation

Cyber Risk Institute

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?**

   CRI agrees with the report's assessment of AI's core benefits, particularly its capacity to strengthen real-time cyber defense mechanisms and automate high-volume transaction monitoring. However, the FSB should consider highlighting three nuances related to emerging risk profiles from a governance and risk management perspective:

   • Timeline Compression in Risk Remediation. Advanced AI models significantly lower the barrier to entry for malicious actors by rapidly identifying system vulnerabilities. This drastically compresses the time available for institutions to identify and mitigate risks, meaning traditional, manual remediation and patching cycles can no longer keep pace. Effectively, this rapid acceleration forces risk managers to make critical risk mitigation decisions faster than legacy governance processes and committee structures were originally designed to support.

   • Systemic Operational Dependency and Concentration. Vulnerabilities can emerge when multiple sectors (e.g., banking, insurance, and asset management) rely on a limited number of similar core models or underlying technology infrastructures. From a governance and risk perspective, a severe operational disruption or systemic failure in a widely adopted model could instantly ripple across the entire financial ecosystem. Regulators should view concentration risk not merely as an individual firm vendor problem, but as a shared stability concern that requires coordinated, system-wide oversight.

   • Internal Governance Gaps & Shadow AI. Organizations may implement technical and governance controls intended to prevent the use of unapproved or external consumer-grade AI applications with proprietary corporate data. However, employees may still find ways to use such tools outside of established processes, creating “Shadow AI” workflows that reduce organizational visibility and oversight, and may increase the risk of unintended data leakage or exposure.

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

The sound practices are generally comprehensive and clear to enable financial institutions' adoption of AI. In particular, CRI appreciates the clear focus on board governance in Sound Practice 1 and the operational structure of the Three Lines of Defense (3LOD) model described in Sound Practice 2.

However, CRI suggests further clarification around Sound Practice 8 (Explainability and Transparency) to better balance supervisory expectations with operational realities. Because advanced models and autonomous tools are fundamentally complex and opaque, institutions may be unable to comprehensively demonstrate exactly how a model or system arrives at a specific decision or output. The sound practices should recognize these inherent technical limitations, shifting focus away from total engineering transparency.

Instead, the FSB should emphasize robust governance, testing, monitoring, and documentation as the primary, practical methods to satisfy transparency objectives. Supervisors should guide institutions to evaluate a system's performance boundaries, balancing explainability requirements against its predictive power and track record through governance-led compensating controls (e.g., rigorous input and output validation, operational benchmarking, staged deployment environments, and strict boundary parameter restrictions).

Furthermore, to successfully implement Sound Practice 3 (Incorporation of AI Risks into Risk Management Frameworks), institutions need a scalable mechanism to transition high-level principles into daily operational control objectives.

The financial sector has already recognized this operational gap. To ensure the FSB's guidance remains naturally aligned with current industry execution, we recommend referencing the FS AI RMF in the final report's appendix alongside the National Institute of Standards and Technology (NIST) AI Risk Management Framework . Developed specifically for the financial sector under the auspices of the U.S. Treasury and the Financial Services Sector Coordinating Council, the FS AI RMF was engineered to translate the abstract concepts of the NIST AI RMF into practical governance tiers. While the NIST framework outlines three high-level tiers (Functions, Categories, and Subcategories), the FS AI RMF delivers a practical fourth tier: concrete operational control objectives, sample control designs, and clear evidentiary expectations optimized for financial institutions and their regulators.

Importantly, because the FS AI RMF is already structurally mapped to the widely accepted NIST Cybersecurity Framework and CRI Profile, this linkage allows firms to scale up their existing technology assessment architectures rather than constructing a redundant, siloed control mechanism for AI risk governance and oversight.

CRI recognizes that this Report is focused on financial institutions' responsible adoption of AI rather than broader macroprudential or systemic risk policy. At the same time, Section 1.3 frames several of the risks it identifies, including third-party concentration, herding and correlated behavior, and procyclicality, in systemic terms, noting that correlated AI use could "exacerbate market stress, liquidity crunches, and asset price vulnerabilities" during periods of financial instability. Institution-level sound practices, while necessary, may not fully address risks of this nature on their own. CRI encourages the FSB to consider, as part of its

separate and ongoing work, whether further system-level analysis would be useful to complement the institution-level focus of this Report.

CRI appreciates the Report's acknowledgment that “adopting the sound practices does not absolve a financial institution from its obligations to meet local legal and regulatory requirements related to AI adoption.” In practice, institutions operating across borders may still face difficulty reconciling this Report's materiality-and-risk framework with regimes that apply a different risk-tiering approach, including the national AI guidance referenced elsewhere in the Report (e.g., MAS, HKMA, OSFI, JFSA, FINMA) as well as jurisdiction-specific AI legislation adopted or under consideration in other markets. This challenge is not unique to cross-border institutions. In the United States, financial institutions may also need to reconcile state-level AI regulation that is inconsistent from state to state. CRI encourages the FSB to consider, when finalizing this Report, whether additional clarity on navigating this coexistence would be helpful to institutions, while recognizing that reconciling specific regulatory conflicts is a matter for national authorities.

Finally, with respect to Sound Practice 3 (Incorporation of AI Risks into Risk Management Frameworks) and Sound Practice 11 (Cyber and ICT Risk Management), CRI recommends the FSB consider adding reference to AI Bill of Materials (AI-BOM) as part of financial institutions’ documentation and ICT risk management practices. Similar in concept to a software bill of materials, an AI-BOM would allow institutions to track the components, dependencies, and data sources underlying their AI use cases, enabling more real-time assessment of vulnerability posture when a component (e.g., a shared foundation model or open-source library) is found to be compromised or deprecated. This would complement, rather than duplicate, the AI inventories and third-party registers already contemplated under Sound Practice 3.

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 FSB consultation report effectively balances these considerations by tracking advanced agentic behaviors, such as data tampering and unauthorized autonomous task execution.

To strengthen Sound Practice 10 (Human Oversight), global policies should recognize that generative environments require a fundamental paradigm shift from historical, fixed software testing to continuous risk and performance monitoring. Generative systems can produce varied outputs over time, even under identical conditions. Risk and governance frameworks should consequently prioritize broader governance oversight, output consistency evaluations, ongoing input monitoring, and specialized subject-matter expert spot-checks rather than traditional systems-testing protocols.

Further, it is important to recognize the emerging considerations associated with treating AI agents as “synthetic employees,” including how institutions may consider responsibility when an agent operates outside its intended scope. Consideration should also be given to risks that may emerge through interactions among AI agents across institutions, shared infrastructure, data environments, or third-parties, which may not be fully captured through institution-level oversight. As agentic AI continues to evolve, the sound practices should

similarly evolve to reflect changing capabilities, expectations, and sector-level considerations or be broad enough to encompass these nuances.

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

   CRI applauds the FSB's explicit focus on proportionality. Allowing risk frameworks to scale dynamically based on an institution's size, transaction velocity, and the core materiality of the underlying use case ensures that smaller community institutions are not crushed by compliance costs tailored for large, complex financial institutions. However, it is important to ensure that proportionality is clearly understood across jurisdictions so that it can be consistently applied.

   To protect the document's longevity, the FSB should maintain a technology-neutral perspective. Regulatory descriptions should focus on fundamental system behaviors, functional boundaries, and risk exposures rather than specific technical concepts like Large Language Models (LLMs), which may be replaced as the technology landscape matures.

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.**

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

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

   The glossary's alignment with international bodies like NIST and ISO/IEC 42001 is commendable. We suggest adding two precise definitions to maximize industry clarity from a risk management standpoint:

   • AI Red-Teaming. A specialized, adversarial validation process where multidisciplinary teams deliberately simulate realistic risk scenarios throughout an AI system's lifecycle to discover vulnerabilities, logical weaknesses, or training biases. This should be clearly distinguished from traditional IT network penetration testing.

   • Confabulation. An essential refinement of the standard term "LLM hallucination". Confabulation occurs when a generative system constructs a response that is completely internally consistent and seemingly logical, yet remains factually incorrect or disconnected from real-world datasets.

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