FSB Public Consultation on Sound Practices for Responsible Adoption of Artificial Intelligence (AI)

Survey response 1

Respondent's Information

Name of jurisdiction:
United States of America

Name of jurisdiction: [Other]

Please provide your information:
Given name and last name: - Michael Bidun Email address: - mike@bidun.com Name of organisation: - Bidun Group LLC

Do you agree with your responses being made public on the FSB website?
Yes

Questions

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?
No Comment

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

Broadly yes. The sound practices are comprehensive and notably practical, and I support them as drafted. I offer one comment on an architectural pattern the report references but does not yet govern as such: dynamically spawned, ephemeral sub-agents.

Production agentic systems can operate as an orchestrator that dynamically spawns short-lived sub-agents to decompose and parallelise a task. In agentic software-development tools, where the pattern is most visible today, an orchestrator frequently spawns many sub-agents in a single task, created and destroyed in seconds to minutes, without human involvement in each instantiation. While this architecture is nascent in regulated financial production, the report's own use-case discussion (coding assistants; research agents and AI-powered assistants with access to system tools) describes the on-ramp. As agentic adoption scales, the governed unit will frequently not be "an agent" but "a delegation tree."

The report is not silent on sub-agents: Sound Practice 11 (Cyber and ICT risk management) applies the principle of least privilege to AI agents and their sub-agents. That is a sound instinct. But the report's governance architecture does not yet use the vocabulary its cyber practice already has. The documentation case study under Sound Practice 3 describes emerging practice well: agentic AI-specific identifiers, inventory attributes covering the tools an agent can access, agent-level risk tracking for reusable agents, and agent "certification" within defined boundaries. These practices presume agents are persistent, enumerable artifacts that can be individually registered and certified in advance. Ephemeral sub-agents satisfy neither condition: they cannot be individually onboarded, certified, or inventoried in any meaningful sense, and a registry of instances that lived for ninety seconds is an audit artifact, not a control.

The gap, precisely stated, is that sub-agents appear in the report's permissioning language but not in its identity, certification, threshold, or oversight architecture. The report recognises the ingredients of the problem, that multiple agents may interact in unanticipated ways, that an agent may take hundreds of intermediate steps and err at any of them, and that a low-materiality agent may still hold access to significant systems and data, but the sound practices do not yet say where identity, entitlements, thresholds, and oversight should attach when the agent population is dynamic.

My three recommendations are set out under Question 3. This comment addresses delegation within a single institution's boundary; delegation to or from third-party agents raises related questions that intersect Sound Practice 12 and merit separate treatment.

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 balance is broadly right, and the agentic treatment is more practical than most guidance to date. I recommend three additions, all addressing the delegation-tree pattern described under Question 2.

1. ANCHOR IDENTITY AND CERTIFICATION AT THE ORCHESTRATION BOUNDARY; TREAT SPAWNING AS A DELEGATED-IDENTITY EVENT.

The individual identifier and certification practices in Sound Practices 3 and 10 should attach to the orchestrating agent, the persistent, registrable artifact, with sub-agents operating as delegated identities under the parent's identifier, carrying scoped, time-boxed credentials. The governable, loggable event is the delegation (which parent spawned which child, with what scope, granted access to which tools), not the existence of the instance. This preserves the report's "synthetic employee" framing in workable form: the orchestrator is the employee of record; sub-agents are its supervised work, not additional employees.

2. REQUIRE RESTRICTIVE ENTITLEMENT INHERITANCE, AGGREGATE (NOT PER-INSTANCE) AUTHORISATION THRESHOLDS, AND INDEPENDENT AUTHORISATION.

Three failure modes deserve explicit language.

First, privilege escalation by spawning. Sound Practice 11 already directs least privilege to AI agents and their sub-agents; I recommend the final report state its delegation corollary explicitly: a sub-agent must not be able to acquire entitlements, data access, or tool access exceeding its parent's certified boundary, so that spawning can never function as an escalation path. Least privilege binds each instance; it does not bind the tree.

Second, and most consequential for the report's treatment of agents executing financial transactions and the approval thresholds it discusses: if approval or authorisation requirements bind per agent action, they are evadable by decomposition. To be clear, the concern is effect, not intent. Decomposing work across sub-agents is ordinarily a benign architectural choice, not an evasion attempt; but the sub-threshold-splitting effect is the same whether or not anyone intended it, and it is structurally identical to the transaction-structuring pattern that financial-crime controls treat as evasion precisely because effect, not intent, is what matters at the control layer. I recommend the final report state that approval thresholds and transaction controls for agentic AI bind on the aggregate effect of the delegation tree within a defined window, not on individual agent instances.

Third, and following directly from the second: this bears on the dual authorisation the report recommends for agent-executed transactions above a certain value. Where dual authorisation is satisfied by agents within the same delegation tree, the independence the control exists to create is absent by construction. A checking agent sharing its maker's model, context, and objective does not supply a second judgment, and where it reports a confidence score on work its own tree produced, the human relying on that score is not receiving independent assurance. The report rightly warns about rubber-stamping and automation bias, but its proposed remedies are aimed at human motivation, incentives and oversight-quality metrics, and no such remedy reaches a checker that has no motivation to align. I therefore recommend the final report state that authorising agents sit outside the delegation tree they authorise. Where the check is a threshold rather than a judgment, moreover, it need not be an AI at all: aggregate limits are deterministically verifiable, and a checker that computes rather than reasons cannot inherit the maker's context, drift with it, or be argued out of its answer. This distinction, between checks requiring judgment and checks requiring only verification, deserves a place in the report's treatment of AI-in-the-loop oversight, since it determines whether adding an AI to the oversight layer strengthens the control or merely reproduces the maker inside it.

3. DEFINE OVERSIGHT, EXPLAINABILITY, AND CONTAINMENT AT THE TREE LEVEL.

Sound Practice 8's direction toward tracing the "full reasoning path" of multi-step agentic systems should extend explicitly to the full delegation tree: parent-child lineage, scopes granted, tools invoked, and intermediate outputs, as the unit of explainability and audit. Containment should likewise be stated as a design requirement at the orchestration layer: systems should be architected so that terminating the orchestrator deterministically terminates or

safely quiesces its delegation tree, including queued and in-flight actions. This is a property that must be engineered and tested, not assumed, since asynchronous and already-dispatched actions can otherwise outlive their parent. This framing also strengthens the report's recognition that human monitoring does not scale and that monitoring may require augmentation with another AI agent: the supervisory layer should be understood as attaching to the orchestration boundary, watching trees rather than chasing instances.

ON PROPORTIONALITY: these recommendations are consistent with the report's proportionality principle and are, in practice, proportionality-friendly. Because the controls attach at the orchestration platform rather than per agent instance, the compliance burden scales with the capability deployed, not with the number of sub-agents spawned, avoiding a regime in which institutions face per-instance obligations that grow with ordinary use.

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

This question goes to the heart of my comment. The sound practices are flexible in principle, but their agentic provisions currently rest on an assumption that newer architectures are already breaking: that an agent is a persistent, enumerable artifact that can be individually identified, certified, inventoried, and thresholded.

The dynamically spawned, ephemeral sub-agent pattern described under Question 2 does not satisfy that assumption, and it is not speculative. It is standard practice today in agentic software-development tools, which the report itself identifies as a live financial-institution use case. Where the practices attach controls to the agent as the unit (identifiers, certification, per-action approval thresholds, per-instance oversight), they will not extend cleanly as agent populations become dynamic.

The fix is not to add practices for each new architecture as it appears, which would not be sustainable, but to attach controls at the boundary that remains stable while the population inside it churns: the orchestration layer. Identity anchored at the orchestrator, thresholds binding on the aggregate delegation tree, and oversight watching trees rather than instances are all architecture-agnostic. They hold whether an orchestrator spawns three sub-agents or three hundred, and they do not require the report to anticipate each new pattern in advance. That is what would make the practices genuinely durable rather than merely revisable.

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.

No Comment

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

No Comment

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

The definitions are clear, but the glossary has a gap that the report's own text exposes. Sound Practice 11 refers to "AI agents and their sub-agents," yet the glossary defines neither "sub-agent" nor any term for multi-agent structure. As supervisory expectations for agentic AI form, shared vocabulary for that structure will matter, and the report is currently using a term it has not defined.

I recommend the glossary add three terms:

ORCHESTRATOR: an agent that decomposes goals and delegates tasks to other agents.

SUB-AGENT: an agent instantiated by an orchestrator, typically ephemeral, operating under delegated authority.

DELEGATION CHAIN / TREE: the recorded lineage of such delegations, including the scope and tool access granted at each step.

These would give the sound practices, and supervisors reading them, a precise way to say where identity, entitlements, thresholds, and oversight attach when the agent population is dynamic rather than fixed.

8. Are there any comments you would like to make on this report? Please fill in.
One note on scope. My comments under Questions 2, 3, 4 and 7 all address a single architectural pattern: dynamically spawned, ephemeral sub-agents, and the delegation trees they form. I have deliberately confined them to delegation within a single institution's boundary. Delegation to or from third-party agents raises related questions that intersect Sound Practice 12 - whose entitlements govern, where liability attaches, and how a delegation record spanning two institutions is audited when neither controls the other's logs - and in my view merits separate treatment, either in the final report or in subsequent work. I would be glad to discuss any of the above, and can provide a formatted copy of these comments as a single letter on request. Thank you for the opportunity to comment, and for a report that is notably practical in its treatment of agentic AI.