FSB FINANCIAL STABILITY BOARD

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

Response to Consultation

Design Pty Ltd

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?**

   We broadly agree. The report's risk analysis correctly identifies the precipitating threats: the convincing impersonation of authorised humans, and agents taking unauthorised actions whose remediation can be difficult or impossible. We would name one gap rather than a missing risk: the report treats human approval as a control without addressing how an approval is created, evidenced and later verified. That gap is the subject of our responses to questions 2, 3 and 5.

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

   They are close, with one substantive gap in Sound Practice 10. SP10 correctly names the control, "human approval or dual authorisation for transactions above a certain value" and "human intermediation for execution", but is silent on the instrument: the means by which a human's approval is created, evidenced and later verified. "Human approval" is not self-defining. It can mean a click on a screen, an entry in a log, or a cryptographically signed record bound to a specific person and a specific act. These are not equivalent, and the difference is the whole of the security property. In an environment where the precipitating threat is the convincing impersonation of authorised humans, a human-approval requirement that does not specify how the approval is made unforgeable risks being satisfied in name while failing in substance.

   A second distinction belongs in SP10: verifying who a person is does not establish that they deliberately authorised a particular act. The two are routinely conflated, and the conflation is now being tested at national scale. The United Kingdom made director identity verification mandatory in November 2025; within months, registrations fell by roughly a third and the documented fraud response was the use of paid proxy directors: verified identities standing in for absent intent. The deepfake cases that motivate this consultation are failures of the same kind. In the Arup matter (2024), every visible and audible identity signal on the call was authentic to the victim and entirely fabricated; the firm's own account is that no system was breached and "people were deceived." Where seeing and hearing an executive no longer constitutes proof of their intent, the authorising act must be carried by something the

impersonation cannot reproduce: a deliberate, physical act of authorisation that produces its own verifiable record.

We therefore suggest two additions to SP10's considerations for agentic AI: first, that human-approval controls should evidence intent, not merely confirm identity; second, that for the highest-consequence agent-executed actions, institutions specify how human approval is evidenced, favouring a verifiable, non-repudiable record produced by an instrument under the sole control of the authorising person, with scoped permissions and a revocation path.

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?**

   For agentic AI the balance is close but incomplete at the authorisation layer. The current generation of agent-commerce standards illustrates the gap. Protocols such as Google's Agent Payments Protocol, Visa's Trusted Agent Protocol and Mastercard's Agent Pay all represent the human's authorisation as a signed mandate, and all anchor that signature in whatever secure element happens to sit in the user's phone. The standards specify the receipt; they do not specify the pen. Where the consequence is high and the adversary is sophisticated, a software signing root co-located with the device most likely to be compromised, lost, or socially engineered is the weakest available foundation for the strongest available claim.

   Naming the instrument would be consistent with existing practice, not ahead of it. The European Union's eIDAS framework already reserves its strongest legal signature, the Qualified Electronic Signature, to signatures created by a certified Qualified Signature Creation Device, requiring that signing keys be "reliably protected by the legitimate signatory against use by others." The same logic survives in the notarial requirements that Germany and Switzerland retain for property and corporate transfers, and in Japan's deliberate retention of the registered seal (jitsuin) for high-consequence acts even after digitising routine procedures. Across very different legal cultures, routine authorisation dematerialises while consequential authorisation re-anchors in a verified, physical act.

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

   Yes, provided the practices are framed at the level of required properties rather than technologies. The instrument-layer consideration we propose under question 2 is technology-neutral: it asks that above-threshold human approval produce a verifiable, non-repudiable record created by an instrument under the sole control of the authorising person, and leaves the implementing technology open. Framed that way, the control holds regardless of how capable the executing AI becomes.

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.**

We offer one in brief, from a nonbank, as an emerging design pattern rather than a finished product. A bearer-held physical authorisation instrument, a secure element carried by the authorising individual, distinct from any general-purpose device, can serve as the signing root for above-threshold approvals. The human is presented with the proposed action; a single deliberate physical act produces a cryptographically signed record bound to that action; the record is publicly verifiable; the instrument's authority is scoped to defined actions and can be revoked. This pattern satisfies SP10's stated controls (human approval above a threshold, human intermediation before execution, per-agent and per-act audit trail) while closing the identity-versus-intent gap, because the authorising act cannot be performed by an impersonation of the person. We are developing such an instrument in Australia (theseal.io); we raise it here not to promote it but because the FSB's framework would benefit from contemplating the instrument layer explicitly, and from encouraging interoperability between any such instrument and the mandate-based protocols now being standardised.

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

Yes, within their scope. The addition we suggest is at the layer beneath them: for any case study involving agent-executed transactions, stating how human approval is evidenced would make the insight actionable at the control level rather than the policy level.

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

We suggest the glossary distinguish identity verification (establishing who a person is) from intent authorisation (evidencing that a person deliberately approved a specific act), since the sound practices depend on the second and the industry routinely conflates the two. We also suggest the glossary or commentary note the existing precedent for anchoring high-consequence human assent in a controlled physical instrument (eIDAS Qualified Electronic Signatures created by Qualified Signature Creation Devices; notarial and registered-seal regimes), so that the sound practices connect to established law rather than standing alone.

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

Our response reduces to four recommendations. First, in SP10's additional considerations for agentic AI, distinguish identity verification from intent authorisation, and state that above-threshold human-approval controls should evidence a deliberate human act, not merely confirm identity. Second, add an instrument-layer consideration: for the highest-consequence agent-executed actions, institutions should specify how human approval is created and verified, favouring a non-repudiable record produced by an instrument under the sole control of the authorising person, with scoped permissions and revocation. Third, note in the glossary or commentary the existing precedent (eIDAS QES/QSCD; notarial and registered-seal regimes) for anchoring high-consequence human assent in a controlled physical instrument. Fourth, encourage interoperability between human-authorisation instruments and the emerging signed-mandate protocols, so that the control SP10 requires has a specified, verifiable foundation rather than an assumed one.