Response To the Financial Stability Board Consultation On Sound Practices For Responsible Adoption Of Artificial Intelligence

Andrew J Devin, Independent Executive & Practitioner (ORCID 0009-0008-6315-9517)

Submitted: June 19 2026

---

1. Respondent Introduction

This response is submitted by Andrew J Devin in a personal capacity.

I am an independent executive and practitioner working on AI governance, institutional defensibility and decision-level evidence for material AI-assisted reliance. My background includes senior operating, transformation and digital leadership experience across regulated financial services, technology, public-sector and enterprise environments. My current work focuses on the evidential conditions under which AI-assisted decisions can be verified, reconstructed and defended under supervisory, audit, legal or fiduciary scrutiny.

This response is intentionally narrow. It does not argue that the Financial Stability Board has missed the main AI risk categories. The consultation report already identifies the principal issues with appropriate breadth, including governance, model risk, data quality, human oversight, explainability, cyber risk, ICT risk, third-party dependencies and concentration risk.

The question addressed here is more specific: whether the final sound practices should specify the evidential condition under which material AI-assisted decisions can be treated as defensible.

This narrower question sits within a wider pattern of regulatory convergence. The EU AI Act addresses high-risk AI governance, oversight, transparency and logging, while DORA addresses digital operational resilience and ICT third-party risk in the financial sector. In the United States, revised model-risk guidance continues to emphasise risk-based governance for banking organisations. In Australia, operational-risk and service-provider management expectations have also strengthened. Other supervisory initiatives referenced in the FSB report, including work in the United Kingdom, Singapore, Hong Kong, Canada and by international standard-setting bodies, point in the same general direction. These regimes and initiatives are important, but they do not by themselves specify the decision-level evidential condition under which material AI-assisted reliance can be verified, reconstructed and defended. That is the narrower gap this response addresses.

---

2. Executive Position

The proposed sound practices are directionally strong and should be retained. They provide a coherent foundation for responsible AI adoption by financial institutions and are appropriately non-prescriptive, proportionate and capable of refinement as AI technologies evolve.

The proposed refinement is that responsible adoption should not stop at authorised use, governance frameworks, performance monitoring, human oversight or third-party assurance. Those controls are necessary. They are not always sufficient to show why reliance on a specific AI-assisted output, recommendation, monitored result or action was justified at the point of decision.

For material AI-assisted decisions, financial institutions should maintain governance, evidence and assurance arrangements sufficient to demonstrate that reliance on AI-assisted outputs,

recommendations, monitored results or actions was authorised, operated within approved parameters, subject to verification before reliance, reconstructable after the event and defensible under supervisory, audit, legal or fiduciary scrutiny. The full operative formulation, including the independence standard for verification, is set out in Annex A.1.

The trigger should not be AI use as such. The trigger should be material AI-assisted reliance.

Lower-risk productivity use, administrative support and incidental AI assistance should not carry the same evidential burden as decisions affecting credit, capital allocation, claims handling, market execution, compliance determinations, operational resilience, customer outcomes or cyber-risk response. The requirement should be proportionate to the materiality, risk, autonomy, complexity and explainability of the AI use case.

This response therefore recommends one targeted addition to the sound-practice architecture: a decision-level evidential proof condition for material AI-assisted reliance. That condition would make the practices more testable without requiring the FSB to prescribe a particular technical architecture, vendor model, assurance product or control design.

In practical terms, the final report should help boards, senior management, supervisors and assurance providers ask one question: can the institution evidence why reliance on the AI-assisted output or action was justified at the point of decision?

The responses below apply that question to the consultation questions and propose targeted wording for Sound Practices 3, 9, 10 and 12, together with limited glossary additions.

This response does not ask the FSB to prescribe a technical architecture, require manual review of every AI output, treat all productivity use as material reliance, impose a uniform burden across all institutions, require deterministic reproduction of model internals or treat vendor assurance as irrelevant. It asks only that, where AI contributes to a material decision, the institution should be able to evidence why reliance was justified at the point of decision.

Proposed Refinements At A Glance

Issue Identified	Requested Refinement	Where Addressed
Material AI-assisted reliance can occur without decision-level evidence	Recognise unverified reliance at the point of decision as a cross-cutting evidential risk	Q1
Documentation may evidence the use case without evidencing the reliance decision	Clarify that Sound Practice 3 should include, or link to, decision-level evidence for material AI-assisted reliance	Q2, Annex A.2
Reasoning traces may be mistaken for proof of the decision pathway	Clarify that reasoning traces and chain-of-thought logs should be corroborated by decision-level evidence	Q3, Annex A.3
Performance monitoring may be mistaken for output-level verification	Clarify that Sound Practice 9 should support verification of outputs or actions before material reliance	Q2, Q3, Annex A.4
Human oversight may be nominal rather than evidential	Clarify that Sound Practice 10 should preserve the basis of human review where human oversight is relied upon as verification	Q2, Annex A.5
AI monitoring another AI system may reproduce correlated failure modes	Clarify that AI-assisted oversight should itself be evidenced, tested and reconstructable	Q3, Annex A.6

Vendor assurance may support system governance without proving institutional reliance	Clarify that Sound Practice 12 should require institution-held decision-level evidence for material reliance	Q2, Annex A.7
Current case studies are adoption-focused	Add a proof-oriented non-bank case study showing reconstruction after challenge	Q5, Q6, Annex A.8, Annex C
The glossary describes AI systems but not reliance-proof concepts	Add four terms: material AI-assisted reliance, verification of an AI-assisted output before reliance, decision-level evidence and reconstructability	Q7, Annex B

Question 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?

This response broadly agrees with the benefits and risks of AI adoption described in the report.

The report appropriately recognises that AI can support efficiency, improve risk monitoring, enhance customer service, strengthen fraud detection, assist cyber defence and improve decision-making across front, middle and back-office functions (Section 1.2). It also rightly observes that not adopting AI can itself carry risk, for example in fraud monitoring (Introduction). Responsible adoption should not be framed only through risk avoidance.

The report also identifies the principal risk categories with appropriate breadth (Section 1.3): third-party dependencies and service-provider concentration, market correlations, cyber risks, model risk, data quality, governance challenges, explainability limitations, consumer and market-conduct risks and the implementation challenges associated with human oversight. That is the right risk foundation.

The substantive risk that would benefit from clearer treatment is not a separate operational risk category. It is a cross-cutting evidential risk: unverified reliance at the point of decision.

This risk arises where an AI-assisted output, recommendation or agentic action contributes to a material financial decision, but the institution cannot later evidence why reliance on that contribution was justified. The use case may have been approved, the system monitored, a human present and a vendor's assurance on file, yet the institution may still be unable to reconstruct what was checked, what was accepted, what was rejected and why reliance was reasonable on the evidence available at the time. The standard is not whether the decision proved correct in hindsight; it is whether reliance was justified at the point of decision. That is not only a documentation issue. It is a defensibility issue.

The report already points towards this risk. It recognises that even where information about an AI model or system is available, neither developers nor deployers may understand how it translates inputs into outputs (Section 3.5). It identifies as a misconception the assumption that human involvement ensures effective oversight (Section 1.3). Its third-party case study records accountability gaps, with institutions remaining accountable for outcomes even where the AI was not built in-house (Section 3.9). It also observes that explainability is not an end in itself, but enables accountability and conceptual soundness (Section 3.5). The same is true of evidence: it is not an end in itself, but what makes reliance defensible.

The remaining step is to connect these observations to the decision-level proof question. Where material AI-assisted reliance occurs, the supervisory question is not only whether governance, controls and oversight were in place, but whether the institution can evidence why reliance was justified in the

specific decision context. This applies across the named risk categories. Model performance, explainability, human oversight and third-party assurance each describe the environment around a decision; none of them, individually or together, evidences the decision itself. The response to Question 2 develops this application against the relevant sound practices.

One dimension of this risk is systemic, which is why it belongs in an FSB report rather than only in firm-level supervision. The report warns that reliance on common models, datasets or infrastructure can produce correlated behaviours across financial institutions (Section 1.3). The same correlation can apply to evidence. Where many institutions rely on the same third-party systems with the same evidential gap, a single model failure or vendor incident can leave a sector of institutions simultaneously unable to defend materially similar decisions. This is correlated indefensibility: correlated reliance without decision-level evidence. This is offered as an inference from the report's analysis, not as a finding already made in the report. It is the decision-evidence counterpart of correlated model, data or infrastructure dependency: the systemic issue is not only that institutions may rely on common systems, but that they may share the same inability to evidence material reliance when those systems are challenged.

The final report should therefore recognise unverified material AI-assisted reliance as a cross-cutting risk that sits beneath several named categories. It does not replace those categories. It explains why they become harder to supervise when material AI-assisted decisions cannot be verified, reconstructed and defended after the event.

The benefit of adding this risk is practical: it makes the sound practices testable. Supervisors, boards, accountable executives and assurance providers would be able to ask one specific question: can the institution evidence why reliance on the AI-assisted output was justified at the point of decision? If the answer is no, responsible adoption has not failed in every respect. Governance, monitoring and controls may all be present. But the material decision remains evidentially weak.

The corresponding sound-practice response is an evidential proof condition for material AI-assisted decisions, proportionate to materiality. The institution should be able to show that reliance was authorised, operated within approved parameters, subject to verification of the output or action before reliance through evidence, controls or methods sufficiently independent of the system being checked, reconstructable and defensible under supervisory, audit, legal or fiduciary scrutiny. The proposed addition does not constrain responsible innovation. It protects the benefits identified in the report by making material AI-assisted reliance capable of being defended when tested.

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

The proposed sound practices are directionally strong, but evidentially incomplete.

Taken together, they provide a coherent structure for responsible AI adoption, covering board and senior management oversight, organisation-wide governance, risk management frameworks, documentation, materiality assessment, performance management, human oversight, cyber and ICT risk management and third-party AI risk management. This response supports that architecture, and notes that the report itself states the practices are not exhaustive and will be refined as AI technologies evolve. The refinement proposed here is offered in that spirit, and it is narrower than the practices' current scope.

The proposed sound practices do not yet specify what evidence must exist before a material AI-assisted decision can be treated as verified, reconstructable and defensible.

That distinction matters because responsible adoption and defensible reliance are not the same thing. A financial institution may be able to show that an AI use case was approved, documented, monitored, subject to human oversight and supported by third-party assurance, and still be unable to show why reliance on a specific AI-assisted output was justified at the point of decision.

Sound Practice 3 (Section 2.3) is the natural place to address this gap. Its documentation requirements are necessary, but they operate principally at use-case, lifecycle and governance level. The inventory it describes records descriptions, risk ratings, dependencies, responsible personnel, approvals, assumptions, limitations, validation findings and third-party information: principally a record of the use case and its control environment, not a record of why a specific output was relied upon in a material decision. It does not require decision-level evidence. For a material decision, that means a record showing that the output was assessed before reliance through evidence, controls or methods sufficiently independent of the system being checked; that the basis for reliance was recorded; and that the decision can later be reconstructed.

Sound Practice 9 (Section 3.6) provides a strong performance-management foundation. However, performance assessment at model, system or use-case level should not be treated as a substitute for evidence that a material output was subject to verification before reliance. For material AI-assisted decisions, testing and monitoring should be connected to the reliance event, not only to the general performance of the use case.

Sound Practice 10 (Section 3.7) correctly emphasises meaningful human oversight rather than nominal compliance, and the report elsewhere identifies as a misconception the assumption that human involvement ensures effective oversight (Section 1.3). The practice already requires that overseers have sufficient ability, authority and incentive to intervene. The refinement is evidential: where human review is relied upon as verification for a material decision, the basis of that review should be preserved. A bare approval record, without the basis of review, should not by itself be treated as evidence that the output was tested before reliance.

Sound Practice 12 (Section 3.9) would benefit from one clarification, which the report's own case study supports: it records that institutions face accountability gaps, remaining accountable for AI outcomes even where the AI was not built in-house. Transparency tools, certifications, audit rights and contractual controls support system governance and vendor oversight. They do not establish why the deploying institution relied on a specific output in a specific material decision. For material reliance, vendor assurance should be supplemented by institution-held decision-level evidence.

The final report would therefore be clearer if it distinguished three related but separate layers: adoption governance, whether the institution may use AI; operational control, whether the system is used within approved parameters; and reliance proof, whether the institution can evidence why reliance on the AI-assisted output was justified at the point of decision.

The first two layers correspond to authority and operation. The reliance-proof layer is evidenced through verification, reconstructability and defensibility.

This does not require the FSB to prescribe a technical architecture, vendor model or control design. Nor does it require manual review of every output. The evidential expectation should remain proportionate where AI-assisted reliance occurs through high-volume automated processes. A bespoke narrative proof pack should not be required for every lower-consequence output. Where individual decisions are lower in consequence but occur at scale, the evidential burden may be satisfied at the level of the governing control system, including approved decision logic, model or rule thresholds, operating parameters, rule-

based or deterministic controls, independent data validation, exception handling, override pathways, monitoring records and reconstructable sampling or challenged-case evidence. The depth of sampled or challenged-case reconstruction should scale with the volume, correlation, concentration and cumulative reliance characteristics of the use case. This preserves proportionality while preventing scale from becoming an evidential escape route.

Materiality should therefore be assessed not only by reference to the severity of a single decision, but also by reference to volume, correlation, concentration and cumulative institutional reliance. A low-consequence individual output may not require the same evidence as a high-consequence individual decision. But high-volume automated reliance can still become material where repeated outputs affect financial exposure, customer outcomes, market behaviour, operational resilience, cyber response or prudential risk.

On implementation cost, the proposed refinement operates at the margin of existing requirements rather than as new infrastructure. The report already expects institutions to maintain inventories, documentation, logging, performance monitoring and oversight records; the evidential floor principally requires that, for material decisions or material high-volume reliance systems, those records be linked to the specific reliance decision, governing control system or reconstructable sample and preserved. The cost should also be assessed against the alternative. Where decision-level or control-system evidence is not preserved at the point of reliance, the institution bears the post-event cost of forensic reconstruction under supervisory, audit or legal challenge, at a time and on terms it does not control. Proportionate pre-positioning of reliance evidence is generally less costly than retrospective reconstruction, and the materiality trigger and control-based verification options described above confine the burden to decisions and systems where consequence attaches.

For material AI-assisted decisions, financial institutions should maintain governance, evidence and assurance arrangements sufficient to demonstrate that reliance on AI-assisted outputs, recommendations, monitored results or actions was authorised, operated within approved parameters, subject to verification of the output or action before reliance through evidence, controls or methods sufficiently independent of the system being checked, reconstructable after the event and defensible under supervisory, audit, legal or fiduciary scrutiny.

In short, the five capabilities are authority, operation, verification, reconstructability and defensibility.

This refinement would make the sound practices more complete without making them prescriptive, and it supports the report's own proportionality framing. Lower-risk productivity uses should not carry the same evidential burden as material decisions affecting credit, capital allocation, claims handling, market execution, compliance determinations, operational resilience, customer outcomes or cyber-risk response. Equally, high-volume reliance should not escape evidential expectations merely because each individual output appears low consequence when viewed in isolation.

Question 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 proposed sound practices broadly strike the right balance by remaining technology-neutral while recognising that GenAI and agentic AI may require additional controls (Introduction; Section 3.7). That balance should be retained.

The refinement proposed here is not a separate governance regime for GenAI or agentic AI. It is to clarify that these systems sharpen the evidential question already present across all material AI-assisted decisions: what evidence must exist to show that reliance on the output, recommendation, action or monitored result was justified at the point of decision?

GenAI and agentic AI make this question harder for three reasons.

First, GenAI can produce outputs that are fluent, plausible and useful without those outputs carrying their own verification. This is not simply a hallucination problem; it is a reliance problem. A generated explanation, recommendation, summary or rationale may assist a human reviewer, but it should not be treated as proof that the underlying decision basis was sound. Where a GenAI output is used as part of a material decision, the institution should preserve evidence showing how that output was checked, accepted, rejected or escalated.

Second, agentic AI changes the risk profile because the system may plan, retrieve information, call tools, sequence tasks, initiate actions or interact with other systems. The report already addresses much of the resulting record-keeping: Sound Practice 10 calls for documenting agents' reasoning at key decision points, the tools and data accessed, defined boundaries and final outcomes (Section 3.7), and Sound Practice 9 recognises that monitoring agentic AI may require assessing specific actions, not just outputs (Section 3.6). Those requirements are sound. The remaining gap is linkage: the report requires these records to exist, but does not require them to be connected to the reliance decision they supported, or to a checking method sufficiently independent of the system that produced them. Records without that linkage are runtime artefacts, not reliance evidence.

Third, the report recognises that human monitoring may become impractical at scale and that monitoring may need to be continuous or performed by AI, while institutions and individuals retain ultimate accountability (Sections 1.3 and 3.7). That is a practical and important acknowledgement, but it creates a further evidential question.

An AI monitoring layer should not be treated as verification of an output or action before reliance merely because it monitors another AI system.

The monitoring system may share training data, input data, retrieval sources, prompting assumptions, benchmark incentives, vendor dependencies or plausibility patterns with the system being monitored, and may therefore reproduce the same failure mode in a form that appears more controlled. The report's own analysis supports this concern at system level: Section 1.3 warns that reliance on common AI models, datasets or infrastructure can produce correlated behaviours across financial institutions. The same correlation logic applies within an institution, between the monitoring layer and the monitored system.

The final report should therefore distinguish technical separation from methodological independence. A second AI system, separate model, separate vendor or separate deployment should not automatically count as verification of the output or action before reliance. The report already states the relevant principle in its treatment of benchmarking: Sound Practice 9 notes that benchmarking is most helpful when the alternative output is produced by a model or system taking a different perspective, such as different data sources or methodologies (Section 3.6). This response asks that the same principle be applied to the verification of material AI-assisted reliance. For material decisions, verification of the output or action before reliance should be assessed by whether the checking method can test the output or action through evidence, controls or methods sufficiently independent of the system being checked.

The same discipline should apply to reasoning traces. The report refers to chain-of-thought logging as a way to help explain outputs (Section 2.3), and recognises in its explainability examples (Annex 2) that some LLM reasoning may not be accurate and that a model may describe one set of steps while employing another. Reasoning traces may assist investigation, monitoring and review, but they should not be treated as proof of the actual decision pathway unless corroborated by decision-level evidence.

The appropriate balance is therefore not between general AI governance and special rules for newer AI. It is between technology-neutral principles and evidence-sensitive controls. The same supervisory question should apply across AI forms; the evidence needed to answer it will vary with the system's materiality, autonomy, complexity, explainability and role in the decision. For material GenAI and agentic AI use cases, the institution should be able to demonstrate the five capabilities set out in response to Question 2, with the evidential weight calibrated to the system's autonomy and the monitoring arrangements applied.

This approach keeps the sound practices flexible and technology-neutral while making them more precise where GenAI and agentic AI create higher evidential risk. It does not require manual review of every output, nor does it prevent institutions from using automated or AI-assisted monitoring. It requires that, where material reliance occurs, the institution can produce evidence that the relevant output or action was checked in a way that supports defensible reliance.

GenAI and agentic AI do not merely require stronger monitoring. They require monitoring arrangements that can themselves be evidenced, tested and reconstructed. The issue is not whether AI can assist oversight, which it can. It is whether AI-assisted oversight produces evidence capable of supporting material reliance.

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

The proposed sound practices are broadly flexible enough to accommodate newer types of AI and responsible adoption over time. Their non-prescriptive form, proportionality framing and emphasis on continuous learning and adaptation under Sound Practice 4 (Section 2.4) are appropriate, and that flexibility should be retained. The final report should not prescribe a fixed technical architecture, vendor model, control design or assurance method.

The refinement proposed here is that flexibility should be anchored to a stable evidential outcome. If flexibility is defined only by governance process, oversight form, documentation type or control design, the framework leaves institutions uncertain about what must be shown when a material AI-assisted decision is later tested. As AI systems become more autonomous and more dependent on third-party models, the form of governance will change. The proof question should not. The stable question is the one proposed in response to Question 1: can the institution evidence why reliance on the AI-assisted output or action was justified at the point of decision? This question is technology-neutral. It applies to traditional AI, GenAI, agentic AI, composite pipelines, AI-assisted monitoring and systems not yet in common use, because it does not depend on the internal architecture of the model. It depends on the institutional act of reliance.

The report already contains the elements needed to support this approach. It states that the practices are not exhaustive and will be refined as AI technologies evolve (Introduction), and its own case study on agentic AI inventories (Section 2.3) anticipates that the attributes captured for such systems are expected to develop further as norms and expectations mature: an acknowledgement that mechanism-level

guidance dates, while the evidential question does not. The report also already governs by outcome where mechanism is unavailable: Sound Practice 7 asks institutions facing third-party data limitations to deliver the overall outcomes through alternative or compensating means (Section 3.4). This response asks that the same outcome-anchoring be applied to reliance evidence.

The final report should therefore clarify that future flexibility has two dimensions:

1. **Flexible implementation:** institutions may adopt different controls, tools, assurance methods and monitoring arrangements according to materiality, complexity, autonomy, explainability and risk; and

2. **Stable evidential outcome:** where material AI-assisted reliance occurs, the institution should be able to show that the output or action was subject to verification before reliance through evidence, controls or methods sufficiently independent of the system being checked, and that the reliance decision can be reconstructed and defended after the event, demonstrating the five capabilities set out in response to Question 2.

The control design may vary across time, technologies and institutions, proportionate to the factors the report identifies. The evidential outcome should remain consistent. A current use case may involve human review of a model output before a credit decision; a future use case may involve an agentic system retrieving information, calling tools and initiating action. The evidence retained will differ; the supervisory question should not. The same applies to oversight arrangements that become automated or AI-performed as adoption scales, for the reasons set out in response to Question 3.

The FSB could also consider, as its monitoring work develops, whether AI adoption indicators should be supplemented by proof-capacity indicators. Adoption indicators show where AI is used and vulnerability indicators show where risk may accumulate; proof-capacity indicators would show whether institutions can evidence, reconstruct and defend material AI-assisted reliance when tested. This would align future monitoring with the distinction proposed throughout this response.

This approach avoids two errors. The first is over-prescription: freezing today's control forms into guidance that newer AI architectures may quickly outgrow. The second is under-specification: allowing each new generation of AI systems to be governed through flexible process language while the proof condition remains undefined. The report need not specify how every institution must verify every AI-assisted output. It should specify what responsible adoption must be able to show when AI contributes to a material decision.

Flexibility should operate at the level of method, not at the level of evidential responsibility. The sound practices are sufficiently flexible in structure. They would be stronger if that flexibility were anchored to a stable proof condition for material AI-assisted reliance.

Question 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 and should be retained. They show how different types of financial institutions and public-private initiatives are applying AI in practice across credit risk, fraud detection, operational efficiency, AI inventories, capability-building, explainability, testing, resilience and third-party risk. They

also support the report's proportionality framing by showing that responsible adoption does not have a single form.

The final report would, however, benefit from one additional type of case study: a proof-oriented case study showing how a financial institution reconstructs a material AI-assisted decision after challenge. This would complement the adoption examples by showing not only how AI creates benefit, but how responsible adoption is evidenced when a material decision is tested.

In direct response to the request for additional case studies, particularly for nonbanks, two stylised non-bank examples - an insurer claims-triage and reserving decision and an asset-manager portfolio-risk decision - are provided in Annex C for the FSB's use or adaptation. They are stylised supervisory and audit scenarios rather than reports of actual cases, and they follow the template set out in response to Question 6. The case-study question they answer is: can the institution show what AI-assisted output or action was relied upon, what verification occurred, what evidence was available at the point of decision, who or what accepted the result and how the reliance decision was reconstructed after the event?

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

The case studies provide actionable insights on governance choices, documentation, testing, resilience and third-party risk. Their main limitation is not relevance but evidential completeness: they do not appear designed to illustrate the post-event reconstruction of a specific material AI-assisted reliance decision under challenge, understandably given the report's adoption focus.

Since the report states that its case studies are intended to illustrate how the sound practices may be applied in practice, a proof-oriented case study would simply extend that purpose to Sound Practices 3, 9 and 10 applied to a single decision.

The most actionable addition would follow this template:

1. the material decision context;
2. the AI-assisted output, recommendation, monitored result or action;
3. the approved-use boundary and operational controls;
4. the verification method applied before reliance;
5. the evidence preserved at the point of decision;
6. the human or automated review path;
7. the reconstruction of the reliance decision after challenge; and
8. the lessons for governance, controls and documentation.

This template maps directly to the five capabilities identified in response to Question 2: authority, operation, verification, reconstructability and defensibility. It asks whether the institution can show that the AI-assisted reliance decision was authorised, operated within approved parameters, subject to verification before reliance, capable of reconstruction and defensible when tested.

The case studies already show how AI can be adopted responsibly. One additional proof-oriented case study, preferably involving a non-bank financial institution, would show how material AI-assisted reliance can be defended. Annex C provides two stylised examples following this template.

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

The definitions in the glossary are clear and aligned with industry practice for describing AI systems, capabilities and risks. Entries such as artificial intelligence, GenAI, agentic AI, explainability, LLM hallucinations, model risk and supply chain support a common vocabulary for responsible AI adoption, and the treatment of explainability as a continuum rather than a binary property is a strength.

The alignment gap concerns recent developments rather than existing entries. As adoption shifts towards agentic systems and AI-assisted monitoring of decision pathways, developments the report itself describes, the glossary describes the systems but not the decision-level concepts needed to determine whether a financial institution can evidence why reliance on an AI-assisted output was justified.

This response recommends adding four terms: material AI-assisted reliance; verification of an AI-assisted output before reliance; decision-level evidence; and reconstructability. These additions are deliberately limited. The purpose is not to introduce a parallel vocabulary or require the FSB to adopt respondent-specific terminology, but to clarify the evidential condition that should apply where AI contributes to a material financial decision.

The most important addition is material AI-assisted reliance, because it defines the trigger: incidental productivity use should not carry the same evidential burden as material decisions of the kind identified in response to Question 2. The other three terms support that trigger. Proposed definitions are set out in Annex B.

Annex A: Proposed wording for incorporation into the sound practices and case-study section

The amendments below are offered as targeted refinements to the existing sound-practice architecture. They do not require the FSB to prescribe a specific technical architecture, vendor model, control design or assurance product. They clarify the evidential outcome that should apply where AI contributes to material financial decisions.

A.1 Headline Proof-Condition Paragraph

> *For material AI-assisted decisions, financial institutions should maintain governance, evidence and assurance arrangements sufficient to demonstrate that reliance on AI-assisted outputs, recommendations, monitored results or actions was authorised, operated within approved parameters, subject to verification of the output or action before reliance through evidence, controls or methods sufficiently independent of the system being checked, reconstructable after the event and defensible under supervisory, audit, legal or fiduciary scrutiny. The requirement should be proportionate to the materiality, risk, autonomy, complexity and explainability of the AI use case.*

A.2 Proposed Addition To Sound Practice 3 - Documentation And Decision-Level Evidence

> *For material AI-assisted reliance, documentation should not only identify the AI use case, responsible personnel, approvals, dependencies, assumptions, limitations and validation findings. It should also preserve, or be capable of linking to, decision-level evidence showing why reliance on a specific AI-assisted output, recommendation, monitored result or action was justified at the point of decision.*
>
> *Such evidence may include the material decision context, approved-use boundary, relevant input, retrieved context where applicable, output or action relied upon, verification of the output or action before reliance, exception or escalation path, human or automated review record, acceptance rationale and final decision or execution record.*

A.3 Proposed Addition To Sound Practice 3 - GenAI, Agentic AI And Reasoning Traces

> *Enhanced documentation and logging for GenAI and agentic AI may assist risk management, oversight, investigation and post-event review. However, reasoning traces, chain-of-thought logs or other explanatory records should not be treated as proof of the actual decision pathway unless corroborated by decision-level evidence.*
>
> *Where such records are used to support material reliance, financial institutions should preserve evidence showing how the output or action was checked, accepted, rejected or escalated before reliance.*

A.4 Proposed Addition To Sound Practice 9 - Performance Management And Verification Of Outputs Before Reliance

> *Performance assessment, testing and ongoing monitoring should support, but not substitute for, verification of AI-assisted outputs before material reliance through evidence, controls or methods sufficiently independent of the system being checked.*
>
> *Where benchmarking, dual-system comparison or AI-assisted monitoring is used to support verification, financial institutions should assess whether the checking method is sufficiently independent in data, method, perspective or control design to test the output being checked. Technical separation alone should not be treated as methodological independence.*

A.5 Proposed Addition To Sound Practice 10 - Human Oversight And Preserved Basis Of Review

> *Where human oversight is relied upon as verification for a material AI-assisted decision, the institution should preserve the basis of the review.*
>
> *This should include evidence of what the reviewer considered, what output or action was checked, what exception or escalation was considered and why the output or action was accepted, rejected or escalated. A bare human approval record, without the basis of review, should not by itself be treated as evidence that the output was tested before reliance.*

A.6 Proposed Addition To Sound Practice 10 - AI-Assisted Oversight And Monitoring

> Where AI is used to monitor, triage, assess or support oversight of another AI system in a material decision pathway, the monitoring arrangement should itself be evidenced, tested for performance, coverage and correlated failure modes, and reconstructable.
>
> An AI monitoring layer should not be treated as verification of an output or action before reliance merely because it monitors another AI system. Financial institutions should preserve evidence of the monitoring control's design, operation, coverage, exception handling and connection to the final reliance decision.

A.7 Proposed Addition To Sound Practice 12 - Third-Party AI Risk And Institution-Held Decision Evidence

> *Vendor transparency, model or system cards, certifications, standards, access rights, audit rights, contractual controls and third-party assurance support system governance and third-party oversight. They do not establish why a deploying institution relied on a specific AI-assisted output, recommendation, monitored result or action in a specific material decision.*
>
> *For material AI-assisted reliance, third-party transparency and assurance should be supplemented by institution-held decision-level evidence.*
>
> *Where vendor architecture, contractual limitations or operational practice prevents the institution from accessing decision-critical information needed to verify, reconstruct and defend material reliance, that limitation should be assessed as a constraint on the suitability of the AI system for the relevant material use case.*

A.8 Optional Insertion Into The Case-Study Section

The final report could also include a short proof-oriented case study illustrating how a financial institution reconstructs a material AI-assisted decision after challenge. The case study could show the material decision context, AI-assisted output or action, approved-use boundary, verification method, evidence preserved at the point of decision, human or automated review path, reconstruction after challenge and lessons for governance, controls and documentation.

Annex B: Proposed Glossary Additions

B.1 Material AI-Assisted Reliance

The use of an AI-assisted output, recommendation, classification, monitored result or action as part of the substantive basis for a decision capable of producing financial, legal, prudential, conduct, customer, market, cyber or operational consequence.

Note: the defining condition is reliance rather than AI use as such. AI use falls within this definition where the AI contribution materially informs, alters, ranks, recommends, validates, rejects, escalates or provides reasoning used in the final decision. Incidental productivity use falls outside the definition unless the AI-generated content is relied upon as part of the material decision basis without verification of the output before reliance through evidence, controls or methods sufficiently independent of the system being checked.

B.2 Verification Of An AI-Assisted Output Before Reliance

The assessment of an AI-assisted output, recommendation, monitored result or action before material reliance, using evidence, controls or methods sufficiently independent of the system being checked.

Note: this is distinct from the verification stage of the AI lifecycle described in Section 3.1, which tests whether a system meets specified requirements. The term defined here concerns the assessment of a specific output before reliance. Verification may be human, automated or combined, including deterministic guardrails, threshold controls, independent data validation, sampling, exception testing, dual-system comparison or escalation. Where verification is performed by another AI system, independence is not assumed from technical separation; the relevant question is whether the checking method is sufficiently independent in data, method, perspective or control design to test the output being checked.

For this purpose, sufficiently independent does not require a particular institutional structure or a separate AI system. It means that the checking method is capable of testing the output, recommendation, monitored result or action through a materially different evidential basis, data source, methodology, perspective, control design or escalation path. A second model, separate vendor, separate deployment or AI monitoring layer should not be treated as sufficiently independent merely because it is technically separate from the system being checked.

B.3 Decision-Level Evidence

Evidence preserved at the point of decision sufficient to show why reliance on an AI-assisted output, recommendation, monitored result or action was reasonable on the evidence available at the time.

Note: this may include the decision context, system used, approved-use boundary, relevant input and retrieved context, output or action relied upon, verification of the output or action before reliance, exception or escalation path, review record, acceptance rationale and final decision or execution record. It is distinct from general governance documentation, vendor assurance, model cards, inventories and audit logs unless those materials are linked to the specific reliance decision.

B.4 Reconstructability

The institutional capability to reconstruct and evidence the decision path for a material AI-assisted reliance decision after the event.

Note: for GenAI and agentic AI, reconstructability does not require deterministic reproduction of the model's internal generation path or every transient technical state. It requires preservation of sufficient decision-state evidence to show what was relied upon, what checks were applied, what was accepted or rejected, who or what approved or executed the action and why reliance was reasonable at the point of decision. Reconstructability should be proportionate to materiality, decision velocity, system autonomy, explainability, institutional scale and applicable privacy, confidentiality, security and cross-border data-flow obligations.

Annex C: Stylised Non-Bank Proof-Oriented Case Studies

The two case studies below are stylised. They are offered for the FSB's use or adaptation in the final report, in response to the request in Question 5 for additional case studies, particularly for nonbanks. Each follows the template set out in response to Question 6 and illustrates the application of Sound Practices 3, 9, 10 and 12 to a single material decision.

Non-Bank Case Study 1: Insurer Claims Triage, Settlement Recommendation And Reserving Impact

Context: A non-bank insurer deploys an AI system to assist claims triage and settlement handling. The system reviews claim forms, policy wording, supporting documents, historical claims patterns and fraud indicators. It recommends whether a claim should be approved, rejected, escalated or adjusted, and may also generate a proposed rationale and settlement range.

Potential benefit: The use case may deliver faster claims handling, more consistent triage, improved fraud detection, better allocation of specialist claims review and earlier identification of reserve pressure.

Material reliance point: The risk arises where the AI-assisted recommendation materially shapes the claims outcome or feeds into reserving, pricing or finance sign-off. At that point, the issue is not only whether the claims tool was approved, monitored and documented. It is whether the insurer can show why reliance on the recommendation was justified in the specific claim or reserving decision.

Verification gap: A gap arises if the insurer can produce model documentation, dashboards, claims notes and high-level monitoring reports, but cannot link the specific recommendation to the claim evidence, policy terms, verification performed, review basis and final decision. A generated rationale or claims summary should not be treated as evidence that the underlying decision was verified.

Where present, verification might include an independent policy-rule check, actuarial or claims-specialist review of the proposed settlement range against policy terms and claims history, validation of key data fields against source documents and documented escalation where the recommendation falls outside approved thresholds.

Reconstruction trigger: The need for reconstruction may arise through a customer complaint, ombudsman review, internal audit, external audit, actuarial review or supervisory inquiry. The insurer may be asked to show what the system recommended, what evidence was available, what checks were performed, who or what accepted the recommendation and why the final decision was reasonable at the point of decision.

Evidence the institution should be able to produce: The case study should show the claim type and materiality, the AI-assisted output or recommendation, the approved-use boundary, the relevant data and retrieved context, verification of the output before reliance, any exception or escalation path, the human or automated review record, the acceptance rationale and the final claims or reserving decision.

Sound-practice lesson: The claims file should not only show that an AI system was approved for claims triage. For material outcomes, it should show why reliance on the AI-assisted recommendation was justified. That lesson would apply directly to Sound Practices 3, 9, 10 and 12.

Non-Bank Case Study 2: Asset Manager Or Investment Firm Portfolio-Risk Monitoring And Rebalancing Recommendation

Context: An asset manager or investment firm deploys AI to monitor portfolio risk, liquidity, concentration, counterparty exposure, market signals, issuer news and client mandate constraints. The system identifies a material deterioration in a position or exposure and recommends rebalancing, hedging, escalation, temporary trading restrictions or no action.

Potential benefit: The use case may improve risk detection, accelerate escalation, support more consistent monitoring and help investment or risk committees identify emerging exposures before they crystallise.

Material reliance point: The risk arises where the AI-generated alert, recommendation or proposed action materially informs a portfolio decision, risk committee decision, trading restriction, client communication or escalation decision. At that point, system-level performance monitoring is not enough. The firm must be able to show why reliance on the particular recommendation was justified in the decision context.

Verification gap: A gap arises if the firm can show that the monitoring tool was approved, tested and subject to governance, but cannot reconstruct the data sources used, the relevant market context, the risk-limit check, the independent data validation, the human or automated review path and the basis on which the recommendation was accepted, modified or rejected.

Where present, verification might include independent market-data validation, risk-limit and mandate checks, comparison against a non-AI benchmark or risk model, human review by the portfolio or risk function and documented escalation where the proposed action exceeds approved autonomy or materiality thresholds.

Reconstruction trigger: The need for reconstruction may arise through client challenge, internal risk review, audit, supervisory inquiry, market-event review or litigation. The firm may be asked to show why it acted, or failed to act, on the AI-assisted recommendation.

Evidence the institution should be able to produce: The case study should show the portfolio, mandate or risk limit affected, the AI-generated alert or recommendation, the approved-use boundary and any prohibited autonomous action, the data sources and market inputs used, verification of the output before reliance, the decision-maker's acceptance, rejection or modification of the recommendation, the resulting action or non-action and how the decision state was reconstructed after challenge.

Sound-practice lesson: A record that the system was monitored or that the portfolio action was authorised is not enough. For material investment or risk decisions, the institution should preserve decision-level evidence connecting the AI-assisted recommendation, checks performed, authority exercised and final action taken. That lesson would apply directly to Sound Practices 3, 9, 10 and 12.

These examples are deliberately stylised. The purpose is not to prescribe sector-specific controls, but to illustrate a general case-study pattern that is particularly useful for nonbanks: where AI contributes to a material decision, the case study should show what was relied upon, what was checked, what evidence was preserved, who or what accepted the result and how the reliance decision was reconstructed after the event.

Sources And Selected Supporting Work

A. FSB Consultation And Related FSB Materials

Relevance note: These materials are included because they are the primary FSB documents to which this response is addressed or from which the response draws its framing. They establish the consultation context, the proposed sound-practice architecture, the FSB's prior analysis of AI-related financial-stability vulnerabilities and its existing work on AI adoption, third-party dependency and operational resilience.

1. Financial Stability Board, Sound Practices for Responsible Adoption of Artificial Intelligence: Consultation report, 10 June 2026.
   https://www.fsb.org/2026/06/sound-practices-for-responsible-adoption-of-artificial-intelligence-ai-consultation-report/
   PDF: https://www.fsb.org/uploads/P100626.pdf

2. Financial Stability Board, consultation webpage for Sound Practices for Responsible Adoption of Artificial Intelligence, including response deadline and submission process.
   https://www.fsb.org/2026/06/fsb-consults-on-sound-practices-for-the-responsible-adoption-of-artificial-intelligence-ai/

3. Financial Stability Board, The Financial Stability Implications of Artificial Intelligence, 14 November 2024.
   https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/
   PDF: https://www.fsb.org/uploads/P14112024.pdf

4. Financial Stability Board, Monitoring Adoption of Artificial Intelligence and Related Vulnerabilities in the Financial Sector, 10 October 2025.
   https://www.fsb.org/2025/10/monitoring-adoption-of-artificial-intelligence-and-related-vulnerabilities-in-the-financial-sector/
   PDF: https://www.fsb.org/uploads/P101025.pdf

5. Financial Stability Board, Enhancing Third-Party Risk Management and Oversight: A Toolkit for Financial Institutions and Financial Authorities, 4 December 2023.
   https://www.fsb.org/2023/12/final-report-on-enhancing-third-party-risk-management-and-oversight-a-toolkit-for-financial-institutions-and-financial-authorities/
   PDF: https://www.fsb.org/uploads/P041223-1.pdf

B. Selected Regulatory, Supervisory And Standard-Setting Materials

Relevance note: These materials are included to show the wider regulatory and supervisory convergence around AI governance, model risk, operational resilience, human oversight, third-party risk and supervisory accountability. They are not cited to suggest that those regimes already define the decision-level evidential proof condition proposed in this response. They are included to show that the proposed refinement sits within, and narrows, an existing international direction of travel.

6. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act).
   https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

7. Regulation (EU) 2022/2554 of the European Parliament and of the Council of 14 December 2022 on digital operational resilience for the financial sector (DORA).
   https://eur-lex.europa.eu/eli/reg/2022/2554/oj/eng

8. Board of Governors of the Federal Reserve System, SR 26-2: Revised Guidance on Model Risk Management, 17 April 2026.
https://www.federalreserve.gov/supervisionreg/srletters/SR2602.htm

9. Bank of England, Prudential Regulation Authority, SS1/23: Model Risk Management Principles for Banks, May 2023.
https://www.bankofengland.co.uk/prudential-regulation/publication/2023/may/model-risk-management-principles-for-banks-ss
PDF: https://www.bankofengland.co.uk/-/media/boe/files/prudential-regulation/supervisory-statement/2023/ss123.pdf

10. Bank of England, Financial Conduct Authority and HM Treasury, Joint statement on frontier AI models and cyber resilience, 15 May 2026.
https://www.bankofengland.co.uk/news/2026/may/boe-fca-and-hm-treasury-joint-statement-on-frontier-ai-models-and-cyber-resilience

11. Australian Prudential Regulation Authority, Prudential Standard CPS 230: Operational Risk Management.
https://www.apra.gov.au/sites/default/files/2023-07/Prudential%20Standard%20CPS%20230%20Operational%20Risk%20Management%20-%20clean.pdf

12. Monetary Authority of Singapore, Artificial Intelligence Model Risk Management, 5 December 2024.
https://www.mas.gov.sg/publications/monographs-or-information-paper/2024/artificial-intelligence-model-risk-management

13. Hong Kong Monetary Authority, Responsible Innovation with GenA.I. in the Banking Industry: Practical Insights from the GenA.I. Sandbox, 31 October 2025.
https://brdr.hkma.gov.hk/eng/doc-ldg/docId/20251031-6-EN
PDF: https://brdr.hkma.gov.hk/eng/doc-ldg/docId/getPdf/20251031-6-EN/20251031-6-EN.pdf

14. International Association of Insurance Supervisors, Application Paper on the Supervision of Artificial Intelligence, July 2025.
https://www.iais.org/2025/07/the-iais-publishes-application-paper-on-the-supervision-of-artificial-intelligence/
PDF: https://www.iais.org/uploads/2025/07/Application-Paper-on-the-supervision-of-artificial-intelligence.pdf

15. International Organization of Securities Commissions, Artificial Intelligence in Capital Markets: Use Cases, Risks and Challenges, Consultation Report, March 2025.
https://www.iosco.org/library/pubdocs/pdf/IOSCOPD788.pdf

C. Selected Supporting Work By The Respondent

Relevance note: These works are selected from a wider KEV research corpus on verification, governance and institutional defensibility in AI-assisted reliance. They are not relied upon as external authority. They are included to show that the evidential proof condition proposed in this response forms part of a developed legal-evidential analysis concerning verification absence, human oversight, institutional proof, liability exposure, insurance retreat and reconstruction under scrutiny.

16. Andrew J Devin, Epistemic Debt: The Economics Of Ungoverned AI.
https://ssrn.com/abstract=6135728

17. Andrew J Devin, From J = 0 To J = 1: Human Presence Is Not Proof Of Human Verification.
https://zenodo.org/records/20305912

18. Andrew J Devin, Permission Is Not Proof: The Evidential Burden Of AI-Assisted Execution Under V = 0, Legal-Evidential Note No. 1. https://zenodo.org/records/20648490

19. Andrew J Devin, Deployment Entanglement: Frontier AI Providers, Operational DNA Transfer And Liability In The Reliance Chain, Legal-Evidential Note No. 2. https://zenodo.org/records/20214194

20. Andrew J Devin, Human Oversight Is Not A Liability Firewall, Legal-Evidential Note No. 3. https://zenodo.org/records/20305986

21. Andrew J Devin, When AI-Assisted Decisions Reach Litigation, What Evidence Actually Exists? https://zenodo.org/records/19251905

22. Andrew J Devin, Uninsurable By Design - Risk Transfer And The Macroeconomic Consequences Of Verification Absence Under V = 0. https://zenodo.org/records/19339957

Acknowledgement

I am grateful to Michael Magoulias for strategic review and comment during the preparation of this response. I also benefited from informal comments on an earlier draft. The views and any errors remain my own.

Supplementary Note To FSB Consultation Response: Sovereign Interruptibility Of Frontier-Model Access

Andrew J Devin, Independent Executive & Practitioner (ORCID 0009-0008-6315-9517)

Submitted: June 19 2026

1. Purpose And Scope

This short note supplements my response to the Financial Stability Board consultation on *Sound Practices for Responsible Adoption of Artificial Intelligence*.

The scope of this note is limited to the financial-stability and evidence-continuity implications of sudden changes in frontier-model access. It does not address the legal, national-security or export-control merits of any government action, nor the conduct of any specific AI provider. Its purpose is to identify a risk condition that became concrete after publication of the consultation report and to explain how it reinforces the decision-level proof condition proposed in my main response.

The proposed risk condition is **sovereign interruptibility of frontier-model access**.

By this I mean the condition in which access to a frontier AI model, or the ability of particular users, institutions or jurisdictions to use that model, may be suspended, restricted, altered or withdrawn because of legal, regulatory, sovereign or national-security action outside the control of the deploying financial institution.

This note also responds to the FSB's own caveat in its 10 June 2026 consultation press release: "They are also not developed to address recent risks that have emerged related to frontier AI models, although some sound practices would help financial institutions respond to such risks." Sovereign interruptibility is one such risk condition. It sits inside third-party, ICT, operational resilience and governance concerns, but is not fully captured by ordinary vendor-risk language.

The issue is not whether financial institutions should avoid frontier models. The issue is whether material AI-assisted reliance remains authorised, operated within approved parameters, subject to verification before reliance, reconstructable and defensible when access to an upstream frontier model changes suddenly.

This issue also sits alongside recent regulatory attention to frontier AI as a cyber-resilience, ICT, third-party and supply-chain concern. The dimension addressed here is adjacent but distinct: whether financial institutions can preserve access-continuity and evidence-continuity for material AI-assisted reliance when an upstream frontier-model dependency changes suddenly.

2. Structural Risk Class And Illustrative Event

The risk condition addressed in this note is structural. A frontier-model dependency may become unavailable, restricted, altered, substituted or legally constrained for commercial, technical, legal, regulatory, sovereign or national-security reasons outside the deploying institution's control. The point is not whether any particular interruption persists. The point is that access can be withdrawn, differentiated or restored on legal, sovereign or status-based timelines that the institution neither controls nor predicts.

On 12 June 2026, after publication of the FSB consultation report, Anthropic stated that the United States government, citing national-security authorities, had issued an export-control directive requiring it to

suspend all access to Fable 5 and Mythos 5 by any foreign national, whether inside or outside the United States, including foreign-national employees. The provider stated that, to ensure compliance, it had to disable both models abruptly for all customers while working to restore access, and that access to all other models was unaffected.

This note does not rely on that event as evidence of misconduct by any party. It uses the provider's statement as a recent and verifiable illustration of a wider access-continuity and evidence-continuity risk. The relevant feature is that frontier-model access was, according to the statement, restricted on a nationality and status basis outside the control of downstream users.

The event illustrates that frontier-model dependency is not limited to ordinary vendor outage, cyber compromise, service concentration or model performance risk. A model may remain technically functional, and the provider may remain operational, while access is restricted or suspended because a state authority treats the model as strategic capability.

That creates a distinct financial-stability question: can a regulated financial institution preserve defensible reliance when upstream model access is suddenly interrupted, restricted, differentiated or restored for sovereign reasons?

---

3. Why This Is Distinct From Ordinary Third-party Risk

The FSB consultation report already addresses third-party dependencies, service-provider concentration, cyber risk, ICT risk, operational resilience, model risk, data quality, governance, explainability and human oversight. Those categories remain valid.

Sovereign interruptibility adds a narrower risk condition within those categories.

Ordinary third-party resilience planning commonly assumes some combination of notice, contractual recourse and a provider capable of performing continuity obligations. Sovereign interruption may break all three. Access may change suddenly, contractual remedies may not reach the state action that caused the interruption and the provider may itself be unable to perform continuity commitments because it is legally required to restrict or suspend access.

The interruption may also be jurisdictional or status-based. Access may differ by nationality, employee location, customer jurisdiction, model class, use case, licence status or legal authorisation. The operational question is therefore not only whether the provider is available, but whether the institution remains entitled and able to use the relevant model for the relevant material pathway.

This is distinct from sanctions risk, which generally applies to designated entities, persons, jurisdictions or transactions. Sovereign interruptibility applies to access to the model itself, regardless of the deploying institution's own sanctions or compliance status.

The evidential problem may persist after access is restored. The institution may later need to reconstruct decisions made before, during or after the interruption, including decisions made under fallback arrangements, partial access, changed controls or substituted models.

For non-material experimentation, this may be an operational issue. For material AI-assisted reliance, it may become a decision-evidence issue. If the records needed to explain the reliance decision were not preserved at the point of decision, access interruption can become evidence interruption.

4. Relevance To The Main Response

My main response proposes that material AI-assisted decisions should be governed by a decision-level evidential proof condition, with proposed wording set out in Annex A and related definitions in Annex B. In particular, it proposes that financial institutions should be able to demonstrate that material AI-assisted reliance was authorised, operated within approved parameters, subject to verification before reliance, reconstructable after the event and defensible under supervisory, audit, legal or fiduciary scrutiny.

Sovereign interruptibility reinforces that proposal.

It shows why the proof condition should not be limited to model governance, vendor approval, lifecycle documentation, performance monitoring or human oversight. Those controls are necessary, but they may not be sufficient if upstream model access changes suddenly after the institution has embedded the model in a material decision pathway.

The additional point is evidence continuity. Where a frontier model supports material reliance, the institution should be able to reconstruct the reliance decision even if the model is later unavailable, restricted, replaced, altered or subject to a changed access condition.

This does not require the FSB to prescribe a specific technical architecture, vendor model, control design or assurance product. Nor does it require manual review of every output. It requires that, for material AI-assisted reliance, the institution preserve sufficient decision-state evidence to show why reliance was justified at the point of decision and how the institution responded if the upstream model dependency changed.

5. Implications For The Proposed Sound Practices

This note does not require a new sound practice. It supports targeted refinement of existing practices.

The FSB has acknowledged that some sound practices would help financial institutions respond to recent frontier-model risks. The refinements below identify how four of those practices could be sharpened for sovereign interruptibility of frontier-model access.

Sound Practice 3 - Incorporation Of AI Risks Into Risk Management Framework

Under Sound Practice 3, financial institutions are expected to adopt risk management frameworks that effectively address AI risks and include processes for AI identification, documentation, materiality and risk assessment. For material AI-assisted reliance, those processes should also capture whether the use case depends on frontier models whose access may be affected by legal, regulatory, sovereign, national-security or provider-level restrictions.

Documentation should preserve, or be capable of linking to, decision-level evidence showing which model or system was used, what access condition applied, what output or action was relied upon, what verification occurred before reliance and what fallback arrangement was available if access changed.

Sound Practice 9 - Performance Management

Performance monitoring should not be limited to whether the model performs under normal access conditions. For material use cases, institutions should assess how reliance is controlled when access is suspended, restricted, degraded, substituted or restored.

Where a model is replaced, constrained or subject to fallback operation, the institution should preserve evidence showing whether the substituted pathway remained within approved parameters and whether outputs or actions remained subject to verification before reliance.

Sound Practice 10 - Human Oversight

Human oversight arrangements should specify who has authority to suspend reliance, approve fallback operation, escalate exceptions or require manual review when model access or model behaviour changes suddenly.

A human approval record should not by itself be treated as evidence of defensible reliance unless the basis of review is preserved.

Sound Practice 12 - Third-Party AI Risk Management

Third-party AI risk management should distinguish ordinary service interruption from sovereign interruptibility of frontier-model access.

For material reliance, due diligence and contractual controls should be supplemented by evidence-continuity requirements. Institutions should understand whether they can preserve, access, export or reconstruct decision-critical evidence if model access is suspended, altered or withdrawn.

Where provider architecture, legal restrictions, contractual terms or operational practices prevent the institution from retaining decision-critical evidence, that limitation should be assessed as a constraint on the suitability of the AI system for the relevant material use case.

---

6. Proposed wording for FSB consideration

The FSB may wish to add the following paragraph to the final report, either under Sound Practice 12 or in the discussion of third-party, ICT and operational resilience risks:

> *For material AI-assisted reliance, financial institutions should consider whether dependence on frontier AI models creates access-continuity and evidence-continuity risks, including circumstances in which model access, model behaviour or provider availability may change suddenly because of legal, regulatory, sovereign, national-security or provider-level action. Institutions should be able to demonstrate that material reliance remains authorised, operated within approved parameters, subject to verification before reliance, reconstructable and defensible where such changes occur. This may require preservation of decision-level evidence, model and version records, access-condition records, fallback arrangements, verification records, escalation decisions and final action records proportionate to the materiality and risk of the use case.*

The FSB may also wish to consider whether this condition warrants recognition in the glossary or explanatory text. One possible formulation is:

Sovereign Interruptibility Of Frontier-Model Access

The condition in which access to a frontier AI model, or the ability of particular users, institutions or jurisdictions to use that model, may be suspended, restricted, altered or withdrawn because of legal, regulatory, sovereign or national-security action outside the control of the deploying financial institution.

This term is offered not to expand the consultation beyond its responsible-adoption scope, but to clarify a specific form of third-party and operational dependency that may be material to financial stability.

---

7. Conclusion

The FSB's proposed sound practices remain directionally sound. The recent access-interruption event does not displace the consultation framework. It reinforces the need for an evidential floor.

The financial-stability issue is not whether any particular government action or provider response was correct. It is whether financial institutions can preserve defensible reliance when upstream frontier AI dependencies become interruptible by legal, sovereign or operational action.

For material AI-assisted reliance, responsible adoption should include not only AI governance, vendor assurance, monitoring and human oversight, but also evidence continuity under sudden changes in model access, behaviour or availability.

Material AI-assisted decisions require an evidential proof condition. Sovereign interruptibility of frontier-model access makes that condition more urgent.

Sources And Selected Supporting Work

A. FSB Consultation And Related Materials

Relevance note: These materials are included because this note supplements the FSB consultation response and addresses a risk condition relevant to the FSB's proposed sound-practice architecture, particularly third-party AI risk, ICT risk, operational resilience and frontier-model risks.

1. Financial Stability Board, Sound Practices for Responsible Adoption of Artificial Intelligence: Consultation report, 10 June 2026.
   https://www.fsb.org/2026/06/sound-practices-for-responsible-adoption-of-artificial-intelligence-ai-consultation-report/
   PDF: https://www.fsb.org/uploads/P100626.pdf

2. Financial Stability Board, FSB consults on sound practices for the responsible adoption of artificial intelligence, 10 June 2026.
   https://www.fsb.org/2026/06/fsb-consults-on-sound-practices-for-the-responsible-adoption-of-artificial-intelligence-ai/

3. Financial Stability Board, The Financial Stability Implications of Artificial Intelligence, 14 November 2024.
   https://www.fsb.org/2024/11/the-financial-stability-implications-of-artificial-intelligence/
   PDF: https://www.fsb.org/uploads/P14112024.pdf

4. Financial Stability Board, Enhancing Third-Party Risk Management and Oversight: A Toolkit for Financial Institutions and Financial Authorities, 4 December 2023.
   https://www.fsb.org/2023/12/final-report-on-enhancing-third-party-risk-management-and-oversight-a-toolkit-for-financial-institutions-and-financial-authorities/
   PDF: https://www.fsb.org/uploads/P041223-1.pdf

B. Recent Frontier-Model Access Event

Relevance note: This source is included as the factual trigger for the supplementary note. The note does not rely on the event as evidence of misconduct by any party. It uses the event as a recent and verifiable illustration of access-continuity and evidence-continuity in material AI-assisted reliance.

5. Anthropic, Statement on the US government directive to suspend access to Fable 5 and Mythos 5, 12 June 2026.
   https://www.anthropic.com/news/fable-mythos-access

C. Selected Regulatory Context

Relevance note: This material is included to show that recent supervisory attention to frontier AI already includes cyber-resilience, ICT and supply-chain concerns. The supplementary note addresses a narrower and complementary issue: evidence continuity where model access, model behaviour or provider availability changes suddenly.

6. Bank of England, Financial Conduct Authority and HM Treasury, The Bank, FCA and HM Treasury joint statement on Frontier AI models and cyber resilience, 15 May 2026.

https://www.bankofengland.co.uk/news/2026/may/boe-fca-and-hm-treasury-joint-statement-on-frontier-ai-models-and-cyber-resilience

Acknowledgement

I am grateful to Michael Magoulias for strategic review and comment during the preparation of this note. I also benefited from informal comments on the related consultation response. The views and any errors remain my own.