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

Response to Consultation

Japanese Bankers Association (JBA)

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?

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We broadly agree with the benefits and risks associated with AI adoption by financial institutions described in the report. The report appropriately identifies key benefits, including greater operational efficiency, enhanced risk management, improved fraud detection and better customer service, as well as significant risks relating to governance, data quality, explainability, cyber security, third-party dependencies, and risks associated with the use of GenAI and agentic AI.

・However, the report could more explicitly recognise not only the risks arising from the use of AI, but also the risks associated with non-adoption or insufficient adoption of AI. AI can strengthen first-line activities such as customer support, business execution, financial crime prevention, cyber defence, document processing and risk identification; second-line activities such as compliance monitoring, alert triage, surveillance, policy review and risk analytics; and third-line activities such as evidence review, anomaly detection, population analysis and audit scoping. As financial institutions face increasing volumes of data, growing regulatory complexity, and evolving cyber and financial crime threats, failure to adopt AI appropriately may constrain their ability to maintain effective operational, risk management and assurance capabilities. The report could usefully acknowledge that, in some circumstances, overly cautious interpretations of AI-related risks may also impair operational effectiveness, risk management and assurance capabilities.

・The report could also highlight several additional risks, including risks associated with retrieval-augmented generation (RAG), insufficient auditability, over-reliance on AI outputs, diluted accountability, shadow AI, AI-enabled circumvention of controls, and the degradation of human expertise and professional judgement. AI may not only introduce new risks, but also amplify existing model, operational, conduct, cyber, third-party and reputational risks.

In particular, there is a risk that users may place undue reliance on AI-generated outputs without sufficient professional scepticism, while accountability for AI-supported decisions may become unclear.

・These examples should not be regarded as exhaustive. The cybersecurity threats posed by AI systems and the use of AI are not monolithic, and may differ depending on the AI architecture, deployment model, degree of autonomy, external connectivity, data environment and use case. In addition to currently recognised offensive and defensive risks, new attack vectors, vulnerabilities and misuse patterns may emerge as AI technologies, agentic capabilities, tool integrations and business applications continue to evolve. Accordingly, financial institutions should periodically review their AI risk taxonomies, control frameworks, monitoring arrangements and threat assessments in light of developments in technology, use cases, external dependencies and the threat environment.

・For large financial groups operating across multiple entities, jurisdictions, business lines, languages and data environments, these risks reinforce the importance of maintaining AI inventories, conducting risk assessments, clearly assigning accountability, and retaining evidence of approvals, changes and ongoing monitoring.

・They also underscore the need for dynamic and forward-looking AI risk management arrangements that can be updated as new forms of AI-related risk, including cyber-related risk, emerge over time.

・From a cyber security perspective, AI-related risks should not be treated as a single category. It would be beneficial to distinguish among: (i) attacks or misuse targeting AI systems themselves; (ii) ICT and operational risks arising from the use of AI; and (iii) the use of AI by threat actors. These categories involve different threat actors, attack paths and control requirements. In particular, the integration of agentic AI and external tools may introduce risks relating to excessive permissions, autonomous tool execution, prompt injection, RAG data contamination, and model or data poisoning, which require robust governance and control measures.

・The report should also more clearly recognise ecosystem-level risks arising from foundation model providers, cloud service providers, AI tooling providers and external APIs. Such risks may originate outside the direct control of financial institutions and cannot always be adequately mitigated through internal controls or vendor management alone. While financial institutions would implement appropriate governance and oversight, effective risk management also depends on provider transparency and accountability.

・This is particularly relevant for third-party AI services, including AI capabilities embedded in cloud services, SaaS platforms, vendor products and external APIs. In practice, financial institutions may face challenges where AI features are introduced or materially changed with limited prior notice, or where it is difficult to obtain sufficient contractual protections, assurance information, audit support or change notifications.

・Accordingly, greater emphasis should be placed on transparency regarding model changes, assurance artefacts, incident reporting, audit support and supply-chain visibility.

The report would also benefit from more explicitly recognising the respective responsibilities of financial institutions and AI service providers in managing AI-related risks.

・The FSB, other standard-setting bodies and relevant authorities could also help address such ecosystem-level risks by encouraging greater cooperation and transparency from AI service providers and the broader technology sector.

・Finally, risks associated with frontier AI services should not be considered solely through the lens of model performance and safety. Availability and access continuity risks are also becoming increasingly important. Model suspension, performance degradation, rate limiting, model replacement, jurisdictional restrictions, the withholding of outputs resulting from safeguards by providers, or fallback to lower-capability models may adversely affect legitimate business operations, cyber defence, financial crime controls, compliance monitoring and audit activities. The report could therefore encourage institutions to consider access continuity, contingency arrangements, alternative operating procedures and AI dependency planning as part of their broader business continuity and operational resilience frameworks.

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

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We consider the 12 sound practices to be broadly comprehensive and clearly articulated as a high-level framework for the responsible adoption of AI by financial institutions. In particular, it is valuable that the report distinguishes between enterprise-wide AI governance and use case-level AI lifecycle management, and addresses topics such as board and senior management oversight, materiality and risk assessment, human oversight, cyber and ICT risks, and third-party AI risk. The principles-based and risk-based nature of the framework appropriately allows institutions to tailor their approach according to their size, business model, and the materiality and complexity of their AI use cases.

・However, the final report would be more effective if it provided additional practical guidance to support consistent implementation across institutions. While the sound practices establish appropriate objectives, financial institutions would benefit from examples of the concrete control elements, artefacts, evidence and governance processes that may support implementation. Such examples include minimum expectations for AI inventories, AI risk taxonomies, materiality assessment criteria, explainability and transparency standards, performance evaluation and monitoring, AI incident management, third-party AI assessments, approval processes for the transition from proof of concept to production, change management, and documentation or evidence requirements. Without such guidance, institutions may implement the principles inconsistently, resulting in varying approaches to control, accountability, evidence retention and management reporting.

・The concept of proportionality could also be made more operational. Applying the same level of control to all AI use cases may unduly constrain lower-risk applications and may inadvertently encourage the use of unsanctioned or shadow AI solutions. For example, less stringent controls may be appropriate for assistive use cases such as first-line summarisation, search, drafting and research support; second-line alert triage, regulatory document analysis and policy mapping; and third-line audit scoping, population analysis and evidence review. By contrast, use cases involving customer interaction, decision support, external actions, autonomous execution, or material business and risk decisions may warrant enhanced controls, including stronger requirements for explainability, human approval, monitoring, logging, escalation and termination mechanisms. Providing examples of control expectations aligned to different risk levels and use-case categories would promote a more consistent application of proportionality.

・AI governance should also be integrated into existing enterprise risk management frameworks rather than implemented as a standalone discipline. AI-related risks interact with model risk, operational risk, conduct risk, cyber risk, data risk and third-party risk. Accordingly, AI inventories, risk assessments, risk and control self-assessments (RCSAs), issue management, key risk indicators, incident management, model risk management, cyber security oversight, data governance and internal audit processes should operate as part of an integrated risk management framework. Embedding AI governance within existing structures would reduce duplication, improve consistency and ensure that AI risks are incorporated into established risk management and assurance cycles.

・The report would further benefit from additional guidance on roles and responsibilities under the Three Lines Model. Clear allocation of accountability is essential to effective governance and auditability. The first line would be responsible for AI use cases, including business objectives, data usage, controls, operational management and residual risk. The second line would be responsible for establishing AI risk management standards, providing specialist review and challenge, monitoring compliance and risk trends, and reporting to senior management. The third line would be responsible for independently assessing the effectiveness of AI governance, inventories, risk assessments, evidence retention and second-line oversight. Examples of reporting expectations, governance forums, control objectives, evidence requirements and maturity considerations would make the sound practices more actionable and facilitate more consistent implementation across institutions.

・In addition, the report should strengthen its treatment of agentic AI and third-party AI. In these areas, effective risk management may depend not only on controls implemented by financial institutions, but also on transparency and accountability on the part of AI service providers. Human oversight remains important, but its form and intensity may need to be tailored to the nature, scale and risk profile of the relevant use case, and may in some cases need to be complemented by appropriate technical and operational controls. In some higher-risk contexts, additional controls may also be useful. These may include least-privilege access, allow-listing, predefined execution constraints, immutable logging, anomaly detection and emergency shutdown mechanisms, as well as automated monitoring or escalation mechanisms capable of responding to large volumes of anomalous activity at machine speed, including, where appropriate and proportionate, AI-enabled or agentic defensive controls operating under pre-agreed, human-defined policies and supported by appropriate governance arrangements and periodic review. The report should therefore

place greater emphasis on provider-side transparency, assurance artefacts, change notifications, incident reporting, audit support and supply-chain visibility, while recognising the practical limits of governance implemented by deployers where critical information remains under the control of AI providers. It would also be useful to promote greater clarity and consistency in expectations for AI-related cyber incident reporting across jurisdictions, particularly where such reporting depends on timely information from third-party AI providers.

・Finally, where frontier AI models or external AI services are embedded within critical business processes, the sound practices should more explicitly address operational resilience and dependency management. Reliance on external AI services introduces risks associated with service availability, access continuity, provider safeguards, model replacement, performance degradation and jurisdictional restrictions. Financial institutions would therefore assess AI dependency risks and establish appropriate resilience measures, including fallback models, manual workarounds, degraded-mode operating procedures, exit strategies, and monitoring of provider availability and latency. Incorporating these considerations would strengthen the practical applicability of the sound practices as AI adoption becomes increasingly embedded in critical business functions.

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?

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We consider that the sound practices strike an appropriate overall balance between risk management principles applicable to AI in general and risk responses tailored to emerging and more complex forms of AI, such as GenAI and agentic AI. We support the report's technology-neutral and risk-based approach, as well as its recognition of risks associated with hallucinations, prompt injection, autonomous execution, API connectivity and the challenges of maintaining effective human oversight. Given that traditional AI, machine learning, GenAI and agentic AI share common risks—including data quality, explainability, performance degradation, erroneous outputs, third-party dependencies and cyber risk—it is appropriate for the report to establish principles that apply across all forms of AI while also addressing risks specific to emerging AI technologies.

・At the same time, the final report should clarify that GenAI and agentic AI should not be treated uniformly as high-risk technologies. Risks should be assessed based on the nature of the use case rather than the technology itself. For example, assistive uses such as first-line summarisation, search support, drafting assistance and analytical support; second-line alert analysis, policy mapping and surveillance support; and third-line audit scoping, evidence review and population analysis present materially different risk profiles from customer-facing responses, automated decision support, external system actions or autonomous execution, even where the same underlying AI technology is used. The report

would therefore benefit from clarifying that the depth of controls should vary according to the pattern of use, degree of autonomy, potential impact and extent of human involvement.

・The report could also enhance its articulation of risk-based proportionality for emerging AI. AI use cases involving AML/CFT controls, regulatory reporting, customer protection, critical operations, external actions, fund transfers, customer-facing communications or material decision support should not necessarily be governed in the same manner as lower-risk use cases focused primarily on internal productivity and operational efficiency. Rather, the level of oversight and control should be calibrated according to factors such as the potential impact on customers, markets or regulatory reporting; the degree of involvement in decision-making; the level of autonomy; the financial impact; criticality of the underlying business process; dependency on external providers; operational resilience considerations; and explainability requirements. Clarifying these factors would support more consistent and proportionate implementation across institutions.

・The report should also provide additional guidance on the governance of agentic AI. While autonomy is an important consideration, many of the most significant risks arise from excessive functionality, excessive permissions and excessive autonomy. Financial institutions would therefore define, justify and periodically review which tools and functions an AI agent may use, what data and systems it may access, and the extent to which it may plan, decide, execute, retry actions, delegate tasks or interact with other agents and external services. For higher-risk or material use cases, additional controls may be necessary, including prior review, performance and safety testing, ongoing monitoring, human approval requirements, activity logging, anomaly detection and effective kill-switch mechanisms.

・The report should further clarify that human oversight, while important, may not take the same form in all circumstances. In high-speed, high-volume or highly autonomous environments, human oversight may not be capable of providing effective real-time control. It may therefore, where appropriate, may need to be complemented by technical controls such as least-privilege access, allow-listing, predefined technical constraints, immutable logging, anomaly detection and termination mechanisms.

・Finally, from a cyber security perspective, risks associated with agentic AI should be addressed not only within the context of human oversight, but also explicitly within cyber and ICT risk management. Agentic AI introduces distinct risks arising from tool invocation, API connectivity, interactions with external systems and potentially non-transparent execution paths. These characteristics may challenge traditional cyber security controls and create new attack surfaces.

・The cybersecurity threats posed by AI systems and the use of AI are not monolithic, and the examples of risks and controls should not be regarded as exhaustive. New attack vectors, vulnerabilities and misuse patterns may emerge as AI technologies, agentic capabilities and use cases continue to evolve.

・Accordingly, the report should explicitly address controls relating to agent identification and authentication, permission management, monitoring of tool usage and API activity, behavioural logging, anomaly detection, preservation of audit trails, impacts on connected systems and effective termination mechanisms. Such controls should be reviewed

periodically in light of developments in technology, use cases and the threat environment. Recognising these issues as cyber and ICT risks, as well as governance concerns, would strengthen the overall risk management framework for agentic AI.

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

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We consider the sound practices to be sufficiently flexible to accommodate newer forms of AI and the continued evolution of responsible AI adoption. Their emphasis on proportionality, risk-based implementation, technology neutrality, lifecycle management and organisational adaptability provides an appropriate foundation for responding to future developments in AI. Given the pace of technological change, a principles-based approach is preferable to one that relies on fixed technology definitions or static control frameworks that may quickly become outdated.

・At the same time, the final report would benefit from articulating that financial institutions should not be expected to redesign their entire governance framework each time a new AI technology emerges. A more practical approach would be to build on existing governance, inventory management, risk assessment, approval, monitoring and evidence management processes, while applying incremental enhancements where changes in functionality, autonomy, business impact, third-party dependencies, data usage or criticality alter the risk profile of an AI use case. Such an approach would promote both responsible innovation and effective risk management.

・The report could also more explicitly encourage a staged adoption approach. Responsible AI adoption is often most effectively achieved through progressive phases such as proof of concept, pilot, limited deployment and scaled deployment. A proof of concept may be used to assess technical feasibility and initial risks; a pilot may help evaluate usefulness, operational impacts and control effectiveness within a limited business environment; limited deployment may support monitoring within a defined scope; and scaled deployment can be supported by appropriate controls, training, monitoring, incident response and management reporting. This approach enables institutions to expand AI use safely and under appropriate conditions without becoming unduly hesitant to adopt AI.

・To enhance long-term adaptability, the report would benefit from providing illustrative reassessment triggers. Such triggers may vary across the Three Lines Model reflecting differences in use cases and risk profiles. For the first line, reassessment may be appropriate when AI use expands to customer-facing activities, production data environments, external actions, autonomous execution, new business processes or broader user populations. For the second line, triggers may include changes affecting compliance monitoring, alert handling, risk analytics, regulatory obligations or policy mapping. For the third line, reassessment may be warranted where AI increasingly influences audit scoping,

population analysis, evidence review or audit conclusions. Illustrative examples would help institutions apply reassessment requirements in a manner proportionate to their actual risk profile.

・The report would also benefit from emphasising a functional approach to AI governance rather than one based solely on technology classifications. In practice, risk profiles are more likely to change because of changes in functionality, autonomy, decision-making influence, data usage, external connectivity, third-party dependencies or business criticality than because of changes in technology labels alone. A function-based approach would therefore provide greater resilience as AI capabilities continue to evolve.

・Furthermore, financial institutions are increasingly adopting AI through vendor products, cloud services, SaaS platforms and AI capabilities embedded within existing business applications rather than solely through internally developed systems. As a result, risk profiles may change due to model updates, new features, changes in data handling practices, expanded API connectivity or provider-side specification changes. The report should therefore encourage periodic review of AI governance arrangements, maintenance of AI inventories, monitoring of technology developments, reassessment following significant changes to external services, and continued investment in specialist capabilities and AI literacy.

・Finally, for frontier AI services, reassessment triggers should not be limited to model updates or new functionality. Financial institutions may also be materially affected by changes in access conditions, jurisdictional restrictions, user eligibility requirements, rate limits, model replacement, provider-side safeguard mechanisms, output restrictions, additional review processes or unexpected service suspension. Although such developments may not directly alter model performance, they may nevertheless affect business continuity, operational resilience, auditability, compliance activities, cyber defence and financial crime controls. The sound practices would therefore be strengthened by explicitly encouraging institutions to monitor such developments and incorporate them into ongoing governance, dependency management and resilience planning.

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.

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・While the case studies included in the report provide useful illustrations of the benefits of responsible AI adoption, the report would benefit from a broader range of examples covering different types of financial institutions, particularly nonbank institutions. Financial institutions differ significantly in terms of business models, customer interactions, data characteristics, regulatory considerations, outsourcing arrangements, cloud dependencies and operational

resilience requirements. Additional sector-specific case studies would therefore enhance the practical usefulness of the report and demonstrate how the sound practices can be applied across a wider range of operating environments.

・Examples could include investment research, portfolio analysis and compliance reviews by asset managers; suitability assessments, transaction surveillance and market conduct monitoring by securities firms; fiduciary administration, asset servicing and document-intensive operations by trust banks and trust companies; real-time fraud detection, payment monitoring and customer support by payment service providers; and anomaly detection, operational monitoring and resilience management by financial market infrastructures such as clearing and settlement systems.

・The report would also benefit from including case studies that illustrate the practical progression of AI adoption. In many institutions, AI adoption begins with lower-risk internal use cases before expanding into more material business activities. Examples may include first-line use cases such as document summarisation, policy and procedure search, research support and approval memorandum drafting; second-line use cases such as compliance monitoring, alert analysis, regulatory document analysis and policy mapping; and third-line use cases such as audit scoping, population analysis, evidence review and large-scale document assessment. Such examples would help institutions understand where AI adoption can begin and how governance and controls may evolve as use cases become more material.

・In addition, the report could provide greater value by including case studies involving AI embedded within vendor products, cloud services, frontier AI models, agentic AI solutions and agent-to-tool integrations. As financial institutions increasingly deploy AI through third-party services rather than solely through internally developed systems, these deployment models are becoming increasingly relevant.

・For such case studies, it would be particularly useful to present not only the benefits achieved, but also the conditions that may enable responsible adoption. Examples would include risk assessments, control design, human review arrangements, independence safeguards, permission management, monitoring activities, logging and audit trails, incident response processes and residual risk assessments. For third-party AI and agentic AI use cases, examples could also consider provider transparency, assurance artefacts, supply-chain dependencies, cyber risks, business continuity considerations and exit strategies. Including both benefits and control considerations would enhance the feasibility of the case studies and help institutions understand how responsible AI adoption can be implemented in practice across different sectors and use cases.

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

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・The case studies included in the report provide useful practical insights for financial institutions pursuing responsible AI adoption. In particular, it is valuable that the examples address not only the benefits of AI, but also governance-related considerations such as AI inventories, risk assessment, third-party risk management, cyber and ICT risk management, and human oversight.

・However, to enhance the practicability and usability of the case studies across institutions, the final report should present them in a more structured and consistent manner. Financial institutions need to understand not only whether an AI use case delivered benefits, but also how it was governed, assessed and controlled throughout its lifecycle. A common framework could therefore include information such as the type of financial institution, business objective, AI type, proof-of-concept approach, materiality and risk rating, key risks identified, controls implemented, evidence retained, approval processes, benefit metrics, risk and control indicators, user training, operational monitoring, incident response arrangements, residual risk acceptance and lessons learned. Including these elements would help institutions link the case studies to their own AI inventories, risk assessments, proof-of-concept governance, production approvals, monitoring processes, management reporting and internal audit activities.

・The report would also benefit from greater clarity regarding the relationship between AI use cases, AI systems and AI models. In practice, ambiguity regarding these concepts may lead to inconsistent approaches to inventory management, risk assessment, approval, change management, monitoring and audit. This is particularly relevant where a single AI model supports multiple business processes, or where models and functions change while serving the same business purpose. Case studies could therefore explain what constitutes the AI use cases subject to governance, which AI systems support that use case, which underlying models are involved and how they are embedded within business processes. Similarly, a clearer distinction between assistive use and autonomous use would help institutions apply controls more consistently and proportionately.

・The case studies should also demonstrate that human oversight involves more than a formal approval step. Effective oversight requires that individuals are able to understand the purpose and limitations of AI outputs, identify anomalies or inappropriate outcomes, and intervene, correct, escalate or terminate activities where necessary. The appropriate form of oversight may vary depending on the context. Examples may include prior approval, post-use review, sample testing, threshold-based escalation and human final judgement. In addition, the level of oversight required may differ depending on whether AI is used primarily to support human judgement or whether it materially influences business, compliance or audit conclusions. Illustrating these distinctions would improve both supervisory predictability and practical implementation.

・Finally, case studies involving agentic AI and third-party AI would be particularly valuable if they addressed not only benefits, but also governance and accountability arrangements. Financial institutions are increasingly reliant on external models, cloud-based AI services, AI tools, APIs and agent-to-tool integration. In such cases, transparency and assurance provided by AI providers may be as important as controls implemented by deployers. Accordingly, case studies should consider topics such as provider transparency, assurance artefacts, change notifications, incident reporting, audit support and supply-chain

information, including how AI-related cyber incident reporting may be coordinated across different jurisdictional requirements.

・For agentic AI, case studies should also illustrate how institutions manage risks associated with excessive functionality, permissions and autonomy, including controls such as least-privilege access, allow-listing, logging and audit trails, anomaly detection and kill-switch mechanisms. Including these elements would enhance practicability of the case studies and help institutions understand how responsible AI adoption can be implemented and governed in real-world environments.

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

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We consider the glossary to be broadly clear and aligned with current industry practice and recent developments in AI. Definitions relating to AI, GenAI, agentic AI, foundation models, large language models (LLMs), hallucinations, retrieval-augmented generation (RAG), model drift and data drift are useful in establishing a common vocabulary across financial institutions, supervisors and standard-setting bodies.

・However, from an AI governance perspective, the glossary would benefit from greater clarity regarding the hierarchy of, and relationship among, AI models, AI systems, AI use cases, business processes, and their resulting customer, market or operational impacts. In practice, risks rarely arise from AI models in isolation, but from how a model is embedded within an AI system, how that system is deployed within a specific use case, and how the use case affects business processes, customers, markets or operations. Clarifying these relationships would support greater consistency in AI inventory management, risk assessment, approval, monitoring and audit activities across institutions.

・In particular, additional definitions or explanatory guidance relating to AI governance concepts would improve practical implementation. Definitions such as AI use case, material AI use case, high-risk AI use case, AI system, proof of concept, production use and third-party AI service would help establish more consistent approaches to inventory registration, risk classification, approval, change management, reassessment, monitoring and auditability. Such clarification would be particularly important where a single AI model supports multiple business processes or where multiple models are used to support a single business objective. Greater clarity regarding the unit of governance to which controls should be applied would improve consistency across institutions.

・The report should also ensure that the definitions support responsible AI adoption and do not unintentionally encourage overly conservative interpretations. Terms such as agentic AI, materiality, risk and human oversight may directly influence the level of controls institutions may consider necessary. If these concepts are interpreted too broadly or

ambiguously, disproportionately burdensome controls may be applied to lower-risk use cases such as summarisation, search, drafting support, research support, alert triage or evidence review, potentially discouraging appropriate AI adoption. Conversely, more robust controls may be appropriate where AI is involved in autonomous execution, external actions, customer interactions or activities that materially influence business, compliance or audit conclusions. Additional explanatory guidance would therefore support a more consistent application of proportionality in practice.

・The definition of human oversight would particularly benefit from further clarification. Human oversight should not be understood solely as prior approval of AI outputs. Effective oversight requires that individuals are able to understand the purpose and limitations of AI outputs, identify anomalies or inappropriate outcomes, and intervene, correct, escalate or terminate activities, where necessary. Depending on the nature of the use case and associated risks, different forms of oversight may be appropriate, including prior approval, post-use review, sample testing, threshold-based escalation and human final judgement.

・It would also be useful to recognise that the appropriate form and intensity of human oversight may vary according to the pattern of use, level of autonomy and a potential impact of the AI application. For example, first-line assistive uses that support human productivity may require different oversight arrangements from autonomous or externally facing activities. Similarly, second-line and third-line use cases that provide analytical support may warrant different controls from uses that materially influence compliance conclusions, risk assessments, audit findings or assurance outcomes. Clarifying these distinctions would make the glossary more operationally useful and support more consistent implementation of the sound practices across a wide range of AI use cases.

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

Note: The following comments are intended to align with the purpose of the report and to complement practical considerations and examples that may help institutions apply the sound practices in a manner proportionate to their business models, risk profiles, and stages of AI adoption. They are not intended to advocate for the uniform introduction of binding requirements or overly prescriptive supervisory expectations.

・We support the report's high-level, non-prescriptive, risk-based and proportionate approach. Responsible AI adoption should not be viewed as avoiding the use of AI; rather, it should enable financial institutions to test, learn from and progressively scale AI under appropriate governance, integration with existing risk management frameworks, and sufficient traceability and auditability of evidence. AI has the potential to generate benefits across all three lines of defence. In the first line, these benefits may include business execution, customer engagement, analytics, financial crime prevention and cyber defence. In the second line, AI may enhance compliance, risk management and monitoring. In the third-line, AI may support internal audit through evidence analysis and anomaly. Sound AI governance should therefore support safe and sustainable adoption rather than constrain innovation.

・As AI adoption expands across the organisation, it is important to clarify how governance responsibilities, evidential requirements, reassessment activities and escalation processes

should operate across the Three Lines Model. For example, the first line may be responsible for maintaining AI inventories, capturing use cases, promoting proof-of-concept activities and supporting business adoption. The second line may be responsible for conducting independent risk assessments, designing controls and providing monitoring and challenge. The third line may be responsible for assessing evidential traceability, explainability and auditability. Linking these responsibilities to control objectives, evidence expectations and maturity considerations would help institutions implement the sound practices more consistently.

・At the same time, the report would benefit from more explicitly recognising that systemic and ecosystem-level AI risks cannot be adequately managed through institution-level controls alone. Many of the most significant risks associated with third-party AI, frontier AI and agentic AI arise outside the direct control of financial institutions, including those related to model providers, cloud service providers, AI tooling providers and broader supply-chain dependencies. While robust vendor management and compensating controls remain critical, they may not be sufficient in isolation. The report would therefore benefit from a clearer articulation of provider accountability for transparency, security assurance, incident reporting, material change notification, audit support and supply-chain integrity, as well as greater clarity and consistency in expectations for AI-related cyber incident reporting across jurisdictions. Where relevant, further dialogue among relevant stakeholders, including those responsible for cyber security or other related areas, may also be useful.

・Relatedly, the report should recognise that governance implemented by deployers cannot always compensate for a lack of transparency on the part of providers. Financial institutions may have limited visibility into training data, model architectures, system prompts, fine-tuning processes, tool configurations or downstream dependencies. Without sufficient transparency and assurance artefacts on the side of providers, governance implemented by deployers may be insufficient to reduce residual risk to an acceptable level. Recognising the shared responsibilities of deployers and providers would strengthen the practical effectiveness of the framework.

・The report should also clarify that human oversight, while important, should not be regarded as a universal control. In high-speed, high-volume or highly autonomous environments, human oversight may become overloaded or reduced to a largely formal exercise. Effective governance therefore requires human oversight to be complemented by technical controls such as least-privilege access, allow-listing, automated policy enforcement, AI-enabled monitoring, automated control testing, tamper-resistant logging, anomaly detection, and escalation or kill-switch mechanisms.

・The treatment of agentic AI could also be strengthened. Agent-to-tool and agent-to-agent integrations, including architectures like the Model Context Protocol may create distinct attack surfaces and introduce risks such as tool poisoning, schema poisoning, excessive permissions and non-transparent agent activity. Accordingly, tool definitions, schemas, resource content, prompts and AI-generated outputs should be treated as untrusted inputs and protected through appropriate authentication, authorisation, integrity verification and logging controls. The report would also benefit from explicitly recognising that many agentic AI risks arise from excessive functionality, excessive permissions and excessive autonomy. In addition, the term “synthetic employees” should be used solely as governance analogy

for identity and role-management purposes and should not be interpreted as implying legal or operational equivalence to human personnel.

・Finally, as financial institutions increasingly rely on frontier AI services, the report should address not only model risk and third-party risk, but also access continuity, operational predictability of safeguards, fallback arrangements and operational resilience on the side of providers. Model suspension, performance degradation, rate limits, model replacement, jurisdictional restrictions, changes in access conditions, safeguard interventions, withholding of outputs, additional review delays or fallback to lower-capability models may materially affect legitimate banking activities, including financial crime controls, cyber defence, compliance monitoring, internal audit activities and incident response capabilities. This is not intended to require disclosure of sensitive safeguard mechanisms or to weaken safeguards implemented by providers. Rather, the objective is to support safety while ensuring, where practicable, entity-level notification mechanisms, logging, escalation channels, contractual clarity, fallback arrangements, manual workarounds, degraded-mode operating procedures and AI dependency planning as part of broader business continuity and operational resilience frameworks.

・More generally, illustrative and non-prescriptive practical guidance on factors that may distinguish lower- and higher-risk use cases, examples of reassessment triggers following changes to models or use cases, and considerations for adoption of third-party AI would improve predictability and support greater international consistency, while preserving institutions' flexibility to define risk tiers and associated controls based on their own business models, risk profiles and use cases.

・The continued publication of practical examples, FAQs and supervisory observations would further assist financial institutions in implementing the sound practices in a consistent and proportionate manner. It would also be helpful if such materials included illustrative examples of how global financial groups may balance group-wide consistency with local accountability, including governance approaches for data, technology and third-party dependencies across entities and jurisdictions.