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

Response to Consultation

Digital China Information Service Group Co., Ltd (DCITS), AI Innovation Center

1. **Do you agree with the benefits and risks of AI adoption by financial institutions described in this report? Are there any substantive benefits or risks not covered?**

   We agree with the benefits and risks described, and we find Section 1.3's treatment of agentic AI risks the most operationally precise published to date by an international body.

   On benefits, we would add one observation supporting the report's global relevance: the sound practices converge closely with recent supervisory guidance in our home jurisdiction, which likewise requires process-level logging of AI reasoning paths retained through the life of the business relationship, mandatory human review checkpoints for decisions with material customer impact, and explicit labelling of AI-generated content to customers. In our experience this convergence materially lowers the cost, for internationally active institutions and their vendors, of building one control architecture that satisfies multiple regimes — a benefit the final report might note when discussing cross-border coordination.

   On risks, we identify one substantive risk not yet covered: risks arising at the decommissioning of AI agents. When an agent is retired, orphaned credentials, residual tool and data authorisations, open inventory entries, and stale memory or retrieval stores (which remain available for poisoning) persist unless explicitly closed out. These are the mirror image of the onboarding risks the report addresses thoroughly. We elaborate a corresponding control suggestion in our response to Question 2.

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

   The sound practices are comprehensive and clear, with one gap and two enhancement suggestions.

   The gap: agent decommissioning.

   Section 3.1's lifecycle rightly includes a Retirement stage, and Sound Practice 10's agentic measures thoroughly address an agent's entry into service (identifiers, certification, boundary-setting, checkpoints). The exit path, however, is not yet addressed: when an agent is decommissioned, its credentials must be revoked, its tool and data authorisations

recalled, its entries in the inventory closed out, and its accumulated memory and retrieval stores disposed of or archived under data-governance rules. This is the natural completion of the report's own "synthetic employees" framing — onboarding implies offboarding — and omitting it leaves the risks noted in our response to Question 1. We suggest adding decommissioning controls to the Sound Practice 10 agentic list, cross-referenced to the Retirement stage.

Enhancement one: recognize versioned, human-readable procedural assets as a compensating control under Sound Practice 8.

Structuring the judgment applied by an AI system as versioned, human-readable procedural assets — explicit instructions, decision rules, and workflows distilled from expert practice and historical outcomes, sitting between the underlying model and the use case — has proven effective in our production deployments. The primary audit object shifts from model weights, which second- and third-line functions cannot read, to procedural documents, which they can read, challenge, approve per version, and roll back. This directly supports Sound Practice 3's documentation requirements (prompt versioning, version control enabling rollback); in our client deployments compliance staff review and sign off on these assets version by version. A further benefit is reuse: several hundred such assets accumulated in a governed library are shared across agents in different business domains, so a newly certified agent inherits already-reviewed judgment rather than restating it — compounding both efficiency and consistency of control.

Enhancement two: extend the portability principle to procedural assets.

Sound Practice 12 observes that "ensuring data access and portability can enable substitutability." Where skills and workflows are stored in an open, human-readable form decoupled from any particular runtime or model, institutions can migrate them between providers and models — directly mitigating the vendor lock-in and concentration concerns the report raises.

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?**

Yes, the balance is appropriate. We offer two suggestions to strengthen the agentic side of that balance, drawn from production deployment experience.

First, recognize the platform layer as the natural enforcement point for agentic safeguards. The agentic provisions in Sound Practices 3, 10, and 11 — unique agent identifiers, agent certification, tool-access monitoring for scope creep, least-privilege enforcement including sub-agents, dynamic identity and access management, graduated kill switches, and process-level logging of intermediate steps — share a structural property: they are impractical to implement use case by use case, and only become reliable when enforced by a shared runtime layer through which all agents, tools, and data connections must pass. In our deployments, least-privilege permissioning, sandbox isolation, controls on high-risk operations, and audit-trail generation are implemented once, as properties of the runtime, and inherited by every agent loaded onto it; scope-creep detection presupposes a mandatory gateway mediating every tool invocation; and intervention mechanisms are only

testable when they are runtime primitives rather than per-application features. We suggest the final report make this explicit — for example, by noting in Section 3.7 or 3.8 that financial institutions may implement agentic safeguards through a common agent runtime or orchestration layer, and that the presence and configuration of such a layer could provide useful evidence that controls are systematic rather than ad hoc. This would also serve proportionality: smaller institutions that cannot build these controls in-house can obtain them through a platform, provided third-party accountability under Sound Practice 12 is maintained.

Second, action-level evaluation of agentic AI is feasible today and could be referenced more concretely. Section 3.6 correctly observes that assessing agentic AI must consider not only whether the agent achieved its objective but whether its actions were appropriate, and notes this is a live research area. We respectfully note that action-level evaluation is already implementable: benchmark methodologies that test whether an agent's action sequences comply with stated business policies under realistic, multi-turn conditions (for example, the publicly available tau-bench family of benchmarks developed by Sierra) can be adapted to institution-specific policies. We have adapted such published methodologies to banking scenarios: client-facing agents are evaluated pre-deployment against scripted multi-turn dialogues spanning standardized client-situation types, scored against the institution's own conduct rules with zero-tolerance criteria for defined red-line behaviours, and re-evaluated on every version change of the agent or its underlying skills. We suggest the final report reference policy-compliance evaluation of action sequences as an available technique, to prevent institutions from concluding that agentic performance assessment must await further research.

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

Yes. The proportionality principle and the explicitly non-exhaustive framing provide the flexibility needed. The agent decommissioning gap noted in our responses to Questions 1 and 2 is, in our view, the main lifecycle stage where flexibility currently lacks an anchor to attach to: once decommissioning controls are added to the framework, future agent forms (multi-agent systems, long-memory agents) will have a natural place to attach their exit-stage requirements.

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 well chosen. We offer two anonymized case studies from production or pilot deployments in China, and would be pleased to provide further detail bilaterally.

Case study A: Governing reusable procedural assets across a multi-domain agent portfolio.

A technology provider operating a shared agent platform for banks maintains a governed library of several hundred versioned procedural assets (skills) distilled from expert practice, subject to admission review, controlled change management, and rollback. Agents across distinct business domains — wealth-management marketing, corporate banking, consumer-

protection compliance review, advisor training — load these assets through a unified runtime that enforces least privilege, sandbox isolation, controls on high-risk operations, and full audit trails. When a new scenario is introduced, its agent reuses already-reviewed assets rather than restating judgment, so second-line review effort concentrates on the genuinely new increment. Risk tiering of agents follows the logic of Sound Practice 5: tier is assigned by data and system exposure rather than task prominence. This case illustrates how platform-level asset governance lets smaller institutions obtain controls they could not economically build alone.

Case study B: Pre-deployment behavioural evaluation of a client-facing advisory agent.

A bank deploying an agent in its private-banking advisory business — simulating clients to train relationship managers in handling objections — required pre-deployment evidence of behavioural compliance. The agent portfolio uses a three-role architecture (client simulator, scorer, reviewer); evaluation runs scripted multi-turn dialogues across thirteen standardized client-situation types, scores against the bank's conduct rules with a zero-miss requirement for defined red-line behaviours, and is re-run on each version change of the agent or its skills. Human oversight is configured as human-in-the-loop for anything client-facing. [Detailed evaluation metrics available bilaterally.]

On nonbanks specifically: our automotive-finance deployment experience suggests the sound practices transfer well to nonbank lenders, with Sound Practice 5's materiality assessment requiring the most local adaptation. We would be glad to develop this into a fuller nonbank case study bilaterally. We also note that consumer-protection review of marketing and collections materials — closely mirroring the compliance-review use case in Section 1.2 — has proven a strong entry scenario, because its output is itself a control.

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

Yes — the case studies are among the report's most valuable content. We particularly commend the agentic fraud-detection case study in Section 1.2 and suggest the final report generalize its design lesson explicitly: the agent processes a high-frequency signal stream that no human could review, but produces low-frequency, high-leverage intermediate assets (detection rules) — and human approval is placed on the assets, not the signals. Stating this placement pattern as a general principle ("place human approval at the low-frequency, high-leverage intermediate layer") would reconcile the report's two findings that per-decision human review does not scale (Section 1.3) and that oversight must remain meaningful (Sound Practice 10), and would give institutions a concrete design rule rather than a menu alone.

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

The definitions are clear. We suggest two additions:

Agent certification: the documentation case study in Section 2.3 introduces this term without a glossary entry. A definition along the lines of "a documented review and approval process under which an AI agent is authorized to operate within defined boundaries, with its certified

status, permitted scope, and validity recorded in the institution's AI inventory" would aid consistent adoption.

Procedural asset (or equivalent term): a versioned, human-readable artifact — instructions, decision rules, or workflow definitions — that governs an AI system's behaviour in a use case and serves as a unit of documentation, review, approval, rollback, and reuse. Naming this object would give Sound Practices 3 and 8 a common vocabulary for the layer most institutions actually govern day to day.

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

About the respondent.

DCITS (Digital China Information Service Group Co., Ltd) is a financial technology company with forty years of experience building core banking and enterprise IT systems for financial institutions in China. Through its AI Innovation Center, DCITS operates an enterprise-grade agent runtime platform for financial institutions — providing a unified runtime, least-privilege permissioning, sandbox isolation, controls on high-risk operations, and end-to-end audit trails for AI agents — together with a governed library of reusable, versioned procedural assets ("skills") shared across agents. Deployments and pilots span wealth management, corporate banking, consumer-protection compliance review, advisor training, and software engineering scenarios at banks and an automotive finance company. We respond as a third-party technology provider — the category addressed in Sound Practice 12 — and as a practitioner implementing the governance mechanisms described in Sound Practices 3, 8, 9, and 10 in production banking environments. The views expressed are solely those of DCITS as a commercial technology provider; they do not represent the position of any financial authority or jurisdiction. Our comments address engineering and implementation practice only.

Summary of our four suggestions, in order of the questions above:

(i) add agent decommissioning controls to Sound Practice 10, cross-referenced to the Retirement stage (Questions 1, 2, 4); (ii) recognize versioned, human-readable procedural assets as a compensating control under Sound Practice 8, and extend Sound Practice 12's portability principle to them (Question 2); (iii) note that agentic safeguards may be implemented through a common agent runtime layer, whose presence and configuration could evidence systematic control (Question 3); and (iv) reference policy-compliance evaluation of action sequences as an available technique for agentic performance assessment (Question 3).

We support the FSB's intention to finalize the report this year and would welcome the opportunity to contribute implementation experience from the Chinese banking technology market to future workshops or follow-up work. We consent to publication of this response on the FSB website.