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

Response to Consultation

Fintech Association for Consumer Empowerment (FACE)

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?**

The benefits and risks covered in the report are broadly comprehensive. We would, however, suggest the following additions.

ADDITIONAL BENEFITS

1. Synthetic financial data

AI-generated synthetic financial data can help FinTechs develop and test financial products and processes, such as credit-scoring, fraud-detection, and risk-management models, while protecting customer privacy.

2. AI-powered KYC and customer onboarding

The report could discuss AI-powered KYC and customer onboarding beyond payment authentication. AI can automate identity verification, detect fraudulent documents, perform facial recognition, and assess customer risk, enabling faster and more secure onboarding. Covering this use case would highlight how AI reduces onboarding time, lowers costs, and advances financial inclusion.

3. Voice AI for customer support

AI-powered voice assistants can handle routine customer queries, provide account information, manage claims, and offer multilingual assistance around the clock. For example, Bank of Baroda, a major public sector bank in India, has deployed conversational AI to automate customer interactions, while platforms such as Skit.ai and Gnani.ai are used by banks including SBI, ICICI Bank, Axis Bank, and IDFC FIRST Bank to manage millions of customer conversations. Including this application would demonstrate how AI improves customer experience while reducing operational costs.

4. AI-powered collection reminders

AI can personalise repayment reminders based on customer behaviour, repayment history, and preferred communication channels, making collections more effective than generic

notifications. For instance, Indian startups such as Credgenics provide AI-powered debt-collection solutions.

5. AI-driven grievance handling

AI can automatically classify customer complaints, recommend solutions, and escalate complex cases to human agents where necessary. Agentic AI platforms, for example, can detect customer sentiment in real time, resolve routine issues independently, and transfer complex complaints to human representatives with the full conversation context.

6. Agentic AI for payment protocols in e-commerce

Agentic AI can autonomously compare prices, apply discounts, select the most suitable payment method, and complete transactions in line with user-defined preferences. For example, Google has developed the CASE (Conversational Agent for Scam Elucidation) framework for Google Pay in India, in which agentic AI autonomously gathers scam intelligence to strengthen payment security. This illustrates how autonomous agents can support next-generation payment ecosystems, and including such an emerging application would make the report more forward-looking.

ADDITIONAL RISKS

1. Legacy infrastructure risk

Many banks still rely on decades-old core banking systems that were not designed for AI, and integrating AI into these systems can be complex, costly, and prone to security or operational issues. For example, a 2026 Capgemini survey found that banks spend 43% of their IT budgets maintaining legacy systems, limiting their capacity to scale AI initiatives, and that over 80% of executives reported technology innovation programmes were not delivering the expected gains, partly owing to legacy constraints. While the report emphasises AI governance and risk management, it offers limited guidance on overcoming legacy core banking systems that can hinder secure and effective AI implementation.

2. Long-term data exposure risk

AI systems require large volumes of historical customer and transaction data for training and analytics, and the longer this data is retained, the greater the consequences of a breach. In 2026, for instance, an IT failure at Lloyds Banking Group exposed the personal information of nearly 500,000 customers, illustrating how large stores of financial data can amplify the impact of system failures. The report addresses data governance but does not specifically consider the risks of retaining large volumes of historical data for AI training over extended periods.

3. Privacy risks beyond data security

The report could also address privacy risks that go beyond data security. While it covers data security and governance, it provides limited treatment of AI-specific privacy threats, such as models memorising and revealing personal information, membership inference attacks, re-identification of anonymised data, and challenges arising from cross-border data

transfers. Addressing these would strengthen the report, particularly in the context of India's Digital Personal Data Protection Act (DPDP Act) and data localisation requirements.

4. Predatory inclusion and proxy discrimination

The report also does not discuss predatory inclusion and proxy discrimination. Although it addresses fairness and bias, it does not specifically consider how AI can use alternative data to discriminate against particular groups or to enable digital redlining indirectly. AI may also identify financially vulnerable customers who are more likely to accept high-cost credit, leading to predatory lending. Including this risk would give a more complete assessment of AI ethics by showing how AI can worsen financial exclusion and consumer harm even while appearing to expand access to credit.

5. Indirect prompt injection

The report could add indirect prompt injection as an attack vector against AI agents. This occurs when an attacker conceals malicious instructions within external data that the AI reads, fetches, or analyses, even though the user issuing the prompt is entirely legitimate. For example, an AI assistant summarising a user's unread emails may encounter an email containing hidden text, such as white text on a white background or concealed HTML metadata, instructing it to disregard prior instructions, search for financial documents, transmit them to an external site, and delete the email. Because the agent cannot reliably distinguish the content it is meant to read from instructions it is meant to follow, it may execute the hidden command as though it were a legitimate system order.

This risk is heightened for modern AI agents, which can browse the live web, interact with APIs, write and execute code, and control computer interfaces. As they autonomously ingest large volumes of unpredictable external data, they are especially exposed to indirect prompt injection: an agent that visits a compromised website, reads a poisoned document, or queries a tampered database can be hijacked to steal data, delete files, or send phishing messages through the institution's own infrastructure.

References

https://skit.ai/in/

https://www.credgenics.com/

https://www.bis.org/review/r250319j.htm

https://arxiv.org/abs/2509.12589

https://arxiv.org/abs/2508.19932

https://www.bankingdive.com/news/banks-struggle-scale-ai-legacy-tech-budgets-capgemini/814908/

https://www.meity.gov.in/static/uploads/2024/06/2bf1f0e9f04e6fb4f8fef35e82c42aa5.pdf

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

Under Sound Practice 2, the report recommends measures such as AI oversight committees, dedicated risk teams, multiple layers of review (the "three lines of defence"), and centralised AI monitoring systems. These structures may be realistic for a large bank with significant staff strength, but not for a start-up or mid-sized FinTech. A FinTech with 80 employees and a small engineering team may not have a separate risk function to manage AI-related risks, yet it could still be serving hundreds of thousands of customers through AI. A principle of proportionality should therefore be built into frameworks for AI governance in financial services. The FSB could draw on the approach of the Monetary Authority of Singapore (MAS), whose FEAT principles (Fairness, Ethics, Accountability, and Transparency), together with the Veritas framework, serve as a global benchmark for responsible AI in finance and provide practical testing methodologies for algorithmic fairness, ethics, accountability, and transparency.

Apart from this, we also propose the following points -

1. Ensuring the absence of customer harm from AI processes

Sound Practice 6 could be strengthened by explicitly requiring that the selection and deployment of AI use cases incorporate a structured assessment of potential customer harm, covering both direct and indirect harm as well as the risk of deceptive or misleading outcomes over the short and long term. Relevant international approaches offer useful reference points: the OECD AI Principles (2019) emphasise that AI systems should be robust, safe, and aligned with human-centred values, including the avoidance of unintended harm.

2. Data anonymisation and non-identifiability

While Sound Practice 7 appropriately emphasises data governance and security, further consideration could be given to strengthening requirements on data anonymisation and minimisation, so that datasets used for AI development are not reasonably identifiable, whether in isolation or in combination with other sources. Supervisory guidance could also encourage synthetic data generation as a complementary approach for AI model training and testing. Synthetic datasets that replicate the statistical properties of real-world data without containing identifiable personal information may reduce reliance on sensitive historical data, mitigating breach impact and re-identification risk while also reducing dependence on extensive consent frameworks for secondary data use.

References -

https://www.mas.gov.sg/-/media/MAS/News-and-Publications/Monographs-and-Information-Papers/FEAT-Principles-Final.pdf

https://www.oecd.org/en/topics/sub-issues/ai-principles.html

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 FSB applies the same core sound practices, namely governance, human oversight, lifecycle risk management, data governance, and cybersecurity, but does not specify whether agentic AI should face stricter approval requirements. There are two principal reasons why it might.

1. Higher autonomy directly increases systemic risk

Agentic AI operates with greater autonomy and decision-making capability than traditional AI. Whereas traditional systems typically provide recommendations, agentic systems can execute actions such as payments, trades, and credit decisions. This raises the risk of rapid, large-scale errors and makes pre-deployment scrutiny more important in preventing systemic disruption.

2. Error amplification and speed of harm

More complex AI systems act at machine speed and scale, so small model errors can propagate instantly across thousands or millions of transactions. Stricter approval would ensure that high-impact systems are validated more rigorously before deployment, reducing the risk of fast-moving and largely irreversible financial harm.

On this basis, the report could set out the risks and benefits that relate specifically to agentic AI.

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

Overall, the sound practices provide a strong foundational framework for AI governance in financial services, particularly in model risk management, data governance, transparency, and accountability. Their flexibility for newer and rapidly evolving forms of AI could, however, be strengthened by making certain principles more explicit and forward-looking.

1. Recognising potential AI harm explicitly

The framework could more clearly emphasise the need to anticipate potential harm from AI systems, rather than only mitigating risks after deployment. Such harm includes behavioural manipulation, biased decision-making through proxy variables, and the unintended exclusion of vulnerable customers. A stronger preventive orientation would improve adaptability to generative and agentic AI.

2. Built-in guardrails across the AI lifecycle

To remain future-ready, the sound practices should explicitly require guardrails at every stage of the AI lifecycle, including real-time monitoring and intervention mechanisms for high-risk outputs, particularly in increasingly autonomous systems.

3. Addressing AI-enabled exploitation risks

With the rise of generative and agentic AI, the framework should explicitly address the risk of AI-driven exploitation, such as personalised persuasion, behavioural targeting, and

dynamic pricing based on customer vulnerability. Strengthening guidance in this area would improve consumer protection in evolving digital finance ecosystems.

4. Ensuring adaptability for newer AI paradigms

While the principles are broadly technology-neutral, the framework would benefit from clearer expectations for continuous updating as new AI capabilities emerge. This is especially important for use cases such as autonomous agents in payments, lending, and advisory services, which may not fit neatly within traditional model risk governance.

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 bank- and platform-heavy and less detailed for non-bank financial companies (NBFCs), FinTechs, and market infrastructure players in developing markets. The following additional examples may be considered.

1. KreditBee (NBFC and digital lending)

KreditBee uses AI for automated credit underwriting, enhanced customer service, optimised loan collections, and instant digital loan approvals for young and first-time borrowers. This enables minutes-level approvals and disbursal and extends financial inclusion to thin-credit users.

2. Paytm (payment FinTech)

Paytm uses AI to analyse spending, manage balances, deliver personalised financial insights, prevent fraud, and assess credit risk for fast and secure lending. This supports the instant blocking of suspicious transactions and large-scale fraud prevention across millions of daily payments.

3. PhonePe (payment FinTech)

PhonePe applies AI to simplify digital transactions, automate merchant workflows, and enhance security, reducing payment fraud and enabling faster and safer UPI transaction approvals.

4. TransUnion CIBIL (credit bureau)

TransUnion CIBIL uses AI for fraud prevention, credit scoring and risk assessment, custom marketing insights and segmentation, and predictive analytics. This delivers better credit risk assessment than score-only systems and helps lenders identify low-risk borrowers with limited credit history.

References -

https://www.digitap.ai/enhancing-credit-underwriting-for-kreditbee.html; https://analyticsindiamag.com/ai-startups/how-bengaluru-based-kreditbee-leverages-ai-and-ml-to-democratise-credit/

https://www.electronicpaymentsinternational.com/news/paytm-redesigned-app/; https://www.aicerts.ai/news/ai-market-recognition-paytm-morgan-stanley-report/

https://www.millenniumpost.in/business/how-phonepes-ai-is-changing-the-way-700-million-indians-manage-money-658516

https://www.transunion.com/about-us/ai-machine-learning

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

   Yes. The case studies are informative and relevant to the report. We have added examples alongside our responses above, which the FSB may wish to develop into further case studies.

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

   Yes. The definitions are clear and aligned with industry practice, and we have identified no discrepancies.

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

   All comments are mentioned above.