RBI FREE-AI committee report on Framework for Responsible and Ethical Enablement of Artificial Intelligence

KPMG. Make the Difference.

A brief Summary

AI is a powerful technology reshaping work, business, and society, especially in emerging economies like India. It offers solutions for financial inclusion and innovation in sectors like finance, but also poses risks like bias, lack of transparency, and data security concerns. Effective regulation is key to balancing its benefits and harms. In order to encourage the responsible and ethical adoption of AI in the financial sector, the FREE-AI Committee was constituted by the Reserve Bank of India.

The RBI conducted two surveys to understand current AI adoption and challenges in the financial sector. The Committee referenced these surveys and, in addition, undertook extensive stakeholder consultations to gain further insights. the Committee formulated 7 Sutras that represent the core principles to guide AI adoption in the financial sector including:

7 Sutras

- Trust is the Foundation
- People First
- Innovation over Restraint
- Fairness and Equity
- Accountability
- Understandable by Design
- Safety, Resilience and Sustainability

Using the above background, the Committee recommends an approach that fosters innovation and mitigates risks, treating these two seemingly competing objectives as complementary forces that must be pursued in tandem. This is achieved through a unified vision spread across 6 strategic Pillars that address the dimensions of innovation enablement as well as risk mitigation. Under innovation enablement, the focus is on Infrastructure, Policy and Capacity and for risk mitigation, the focus is on Governance, Protection and Assurance.

Risk Associated with AI Models

Model Risk Factors	Bias, Opacity, Data risk etc.
Operational Risk	Amplification of errors, Model Drift, System Degradation
Third Party Risk	Vendor Flaws, Hidden Dependencies.
Liability Challenges	Non-deterministic outputs lead to Legal and reputational risks
AI-Driven Collusion	Autonomous agents may fix prices.
Financial Stability Risk	Procyclicality, Herding effect, Model convergence Risk
Cyber-Security Risk	Poisoning, Adversarial inputs, Deepfakes
Data Privacy & Security	Over-collection of data, Localization laws violation.
Consumer Protection & Ethics	Algorithmic bias, Opaque decisions
AI Inertia	Risk of non-adoption, Missed opportunities in inclusion and innovation

The information presented in this document is derived from the official publication titled "Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)" released by the Reserve Bank of India (RBI) on 13 August 2025.

Free AI Framework

This is the FREE-AI vision: a financial ecosystem where the encouragement of innovation is in harmony with the mitigation of risk. It outlines pillars for driving AI innovation in financial services through infrastructure, policy and capacity building while ensuring governance, protection and assurance for responsible adoption. At its core lie the 7 Sutras, principles that emphasize trust, accountability, fairness and sustainability in AI deployment

Innovation Enablement

Infrastructure

1. Financial Sector Data Infrastructure
2. AI Innovation Sandbox
3. Incentives and Funding Support
4. Indigenous Financial Sector-Specific AI Models
5. Integrating AI with DPI

Policy

6. Adaptive and Enabling Policies
7. Enabling AI-Based Affirmative Action
8. AI Liability Framework
9. AI Institutional Framework

Capacity

10. Capacity Building within REs
11. Capacity Building for Regulators and Supervisors
12. Framework for Sharing Best Practices
13. Reinforce and Reward Responsible AI Innovation

7 Sutras

Trust is the Foundation

Trust is non-negotiable and should remain uncompromised

People First

AI should augment human decision-making but defer to human judgment and citizen interest

Innovation over Restraint

Foster responsible innovation with purpose

Fairness & Equity

AI outcomes should be fair and non-discriminatory

Accountability

Accountability rests with the entities deploying AI

Understandable by Design

Ensure explainability for trust

Safety, Resilience and Sustainability

AI systems should be secure, resilient and energy efficient

Risk Mitigation

Governance

14. Board Approved AI Policy
15. Data Lifecycle Governance
16. AI System Governance Framework
16. Product Approval Process

Protection

18. Consumer Protection
19. Cybersecurity Measures
20. Red-Teaming
21. Business Continuity Plan for AI Systems
22. AI Incident Reporting and Sectoral Risk Intelligence Framework

Assurance

23. AI Inventory within REs and Sector-wide Repository
24. AI Audit Framework
25. Disclosures by REs
26. AI Toolkit

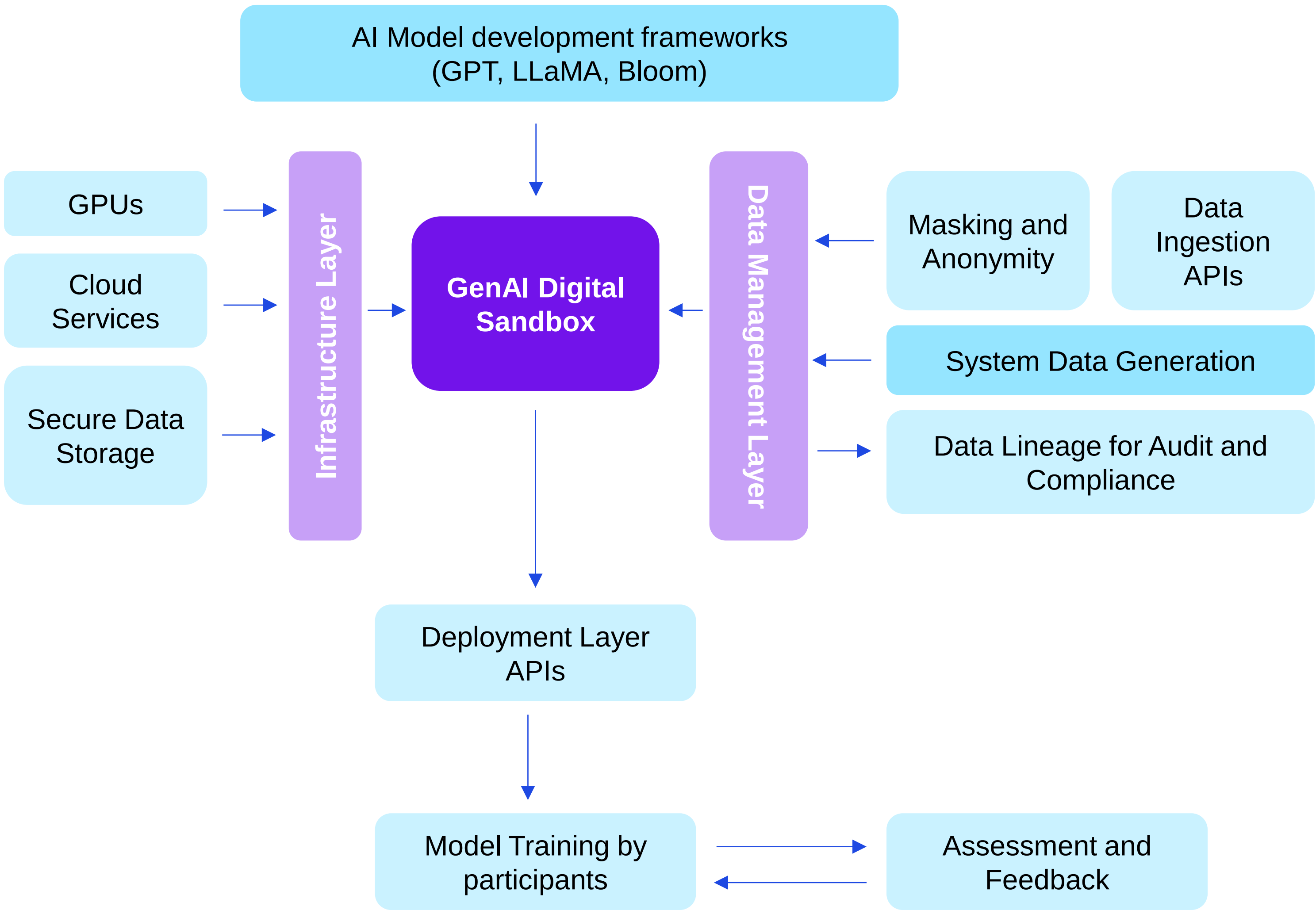
The information presented in this document is derived from the official publication titled "Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)" released by the Reserve Bank of India (RBI) on 13 August 2025.

Gen AI Sandbox Model

The GenAI Digital Sandbox is a secure, controlled testing environment designed to help financial institutions and fintechs experiment safely with Generative AI (GenAI) technologies. It allows them to develop, test, and validate AI models without risking customer data or violating compliance norms.

Inside the sandbox, developers can access synthetic or anonymized data, use popular GenAI frameworks like GPT or LLaMA, and deploy models in a simulated environment. It includes tools for real-time monitoring of risks like bias, hallucination, and inaccuracy, and ensures that models comply with standards like AML, KYC, and consumer protection.

The sandbox also supports explainability tools, helping institutions understand how AI decisions are made, which is crucial for transparency and trust. Overall, it's a safe space for innovation, enabling responsible AI development before real-world deployment.



The information presented in this document is derived from the official publication titled "Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)" released by the Reserve Bank of India (RBI) on 13 August 2025.

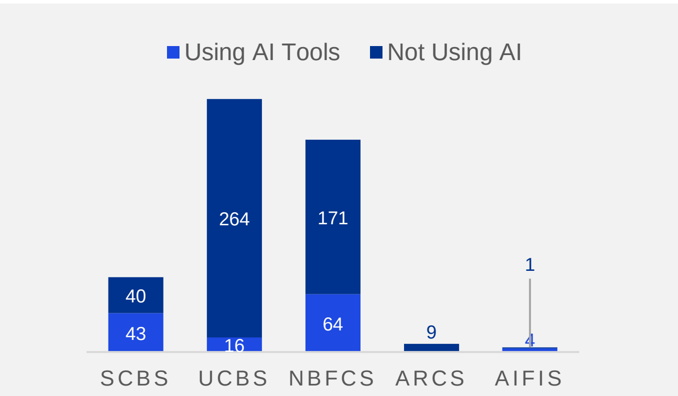
Insights from Surveys and Stakeholder Engagements

Survey Methodology

Two targeted surveys were carried out, covering SCBs, NBFCs, AIFI and FinTechs. Follow-up interactions were conducted with select CDOs/CTOs to understand the extent to which AI had been adopted in the Indian financial services industry and any associated challenges.

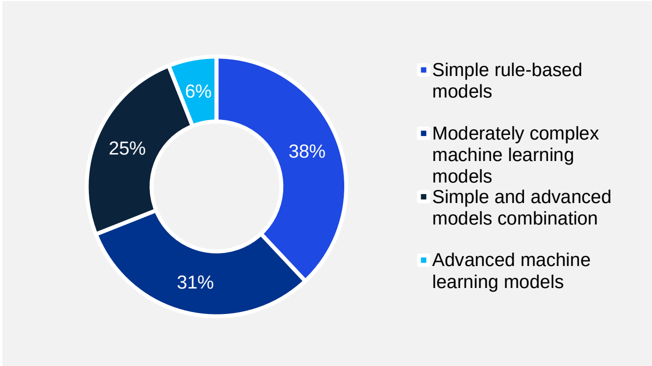
AI Adoption by Supervised Entities

AI adoption among supervised financial entities remains limited, primarily due to low uptake by smaller institutions such as Urban Co-operative Banks (UCBs) and Non-Banking Financial Companies (NBFCs). Within UCBs, especially Tier 1, there was no reported use of AI, and adoption among Tier 2 and Tier 3 remained minimal. Similarly, AI usage among NBFCs was relatively low, and Asset Reconstruction Companies (ARCs) showed no adoption at all. In contrast, larger public and private sector banks have begun exploring AI, though mostly through basic rule-based systems or early-stage models. The overall low adoption is attributed to challenges such as limited institutional capacity, lack of a strong business case, and high infrastructure costs. Where AI is being used, it is typically for improving process efficiency, enhancing customer interaction, and supporting decision-making, with common applications including predictive analytics, lead generation, and customer service chatbots.



Complexity of Model Deployed

Most respondents largely relied on simple rule based non learning AI models and moderately complex ML models, with limited adoption of advanced AI models. In interactions with these entities, it became clear that simpler models were preferred due to ease of implementation, compatibility with legacy systems, and greater control and explainability. There was a preference towards cloud-based deployments for lower cost, scalable solutions and expansion of digital services, with 35% respondents using the public cloud.



Use Cases of AI tools in Financial Institutions

Out of 583 AI applications either in production or under development, the most common use cases are in customer support, sales and marketing, credit underwriting, and cybersecurity. These areas are favored due to their low risk, structured workflows, and predictable outcomes, making them easier to implement.

Cybersecurity solutions are often third-party tools that integrate well with existing systems. In contrast, internal administrative tasks and coding assistants are more common among applications still under development.

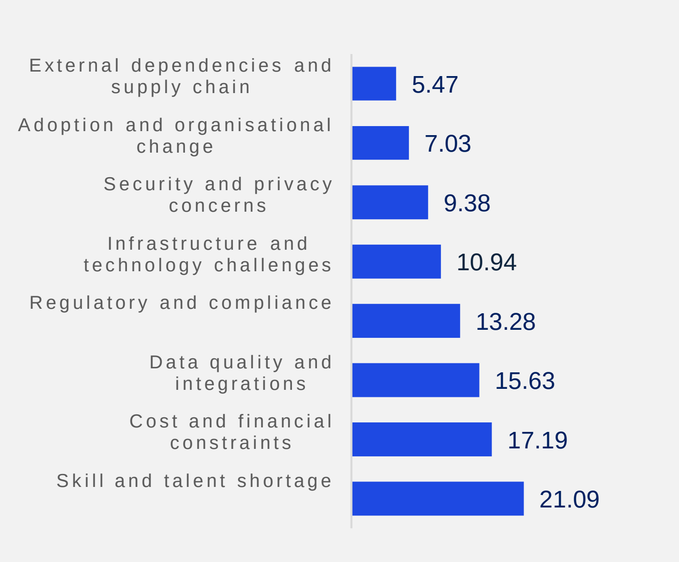
The FTD survey also revealed growing interest in Generative AI (Gen AI), with 67% of surveyed entities exploring at least one use case. However, most Gen AI applications are still experimental and focused on internal productivity tools like chatbots, with limited adoption in customer-facing financial services due to concerns around data sensitivity, explainability, and bias.

The information presented in this document is derived from the official publication titled "Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)" released by the Reserve Bank of India (RBI) on 13 August 2025.

Insights from Surveys and Stakeholder Engagements

Frictions in AI Adoption

The respondents cited several barriers to wider AI adoption that included the AI talent gap, high implementation costs, lack of high-quality data for model training, insufficient access to computing power, and legal uncertainty. Smaller entities, particularly those with resource constraints, highlighted a need for low-cost environments where they could securely experiment before deploying their use cases.

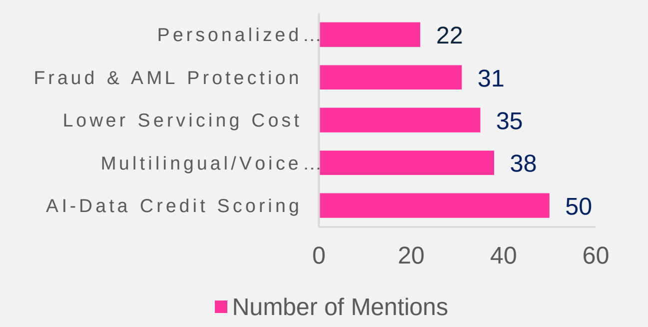


AI Inclusion-Oriented Use Cases:

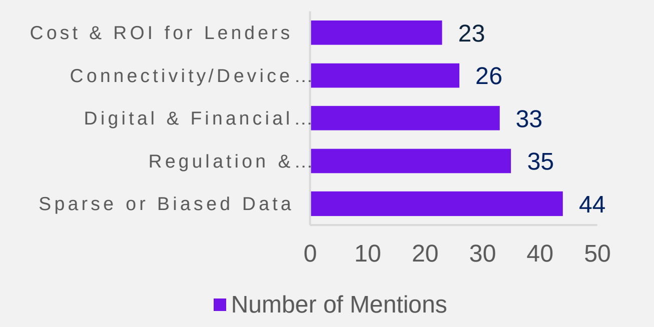
Inclusion-Oriented Use Cases: During the interactions held by FTD, entities suggested that AI has the potential to expand the reach of financial services to the underserved and unserved population through solutions like alternate credit scoring, multilingual chatbots, automated KYC, and agent banking. There were, however, bottlenecks such as sparse data, financial literacy gaps, cost and RoI.

(Analysis bases on 80 Unique Responses)

Potential Benefits (Number of Mentions)



Top 5 Roadblocks (Number of Mentions)



With the exception of large banks and NBFCs, most of the entities were focused on use cases that provide a short-term return on investment. Their apprehensions included the concern that their investments in AI could become obsolete in a short time, considering the pace of hardware evolution, model developments and training parameters. The respondents pointed out that AI applications are not plug-and-play, and require high-quality data, domain-specific customization, and skilled human capital to deliver the desired outcomes

Challenges, Gaps, and Expectations

Financial institutions expressed significant concerns about adopting advanced AI due to risks like data privacy, cybersecurity, governance gaps, and reputational damage. Many lacked formal governance structures, with only a third having board-level oversight and even fewer with incident response mechanisms. Data management practices were fragmented, with no dedicated policies for AI training data, and limited use of tools for bias detection or model monitoring. Most institutions did not track model drift or conduct real-time audits. Skill gaps were also evident, with few structured training programs and a reliance on self-learning. Despite these challenges, there was strong demand—85% of respondents—for a clear regulatory framework. Institutions sought guidance on data privacy, algorithmic transparency, bias mitigation, and cross-border data flows.

The information presented in this document is derived from the official publication titled “Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)” released by the Reserve Bank of India (RBI) on 13 August 2025.

RBI's Recommendations for AI in Financial Services

Innovation Enablement Framework

A. Infrastructure

1. **Financial Sector Data Infrastructure:** Create a shared, high-quality data platform integrated with India AI's AI Kosh
2. **AI Innovation Sandbox:** Establish a secure environment for REs, FinTechs, and TSPs to test AI models
3. **Incentives & Funding:** Provide financial support and shared compute infrastructure for smaller institutions
4. **Indigenous AI Models:** Develop sector-specific AI models (LLMs, SLMs) as public goods
5. **AI + DPI Integration:** Combine AI with Digital Public Infrastructure to enhance inclusion and service delivery

B. Policy

1. **Adaptive and Enabling Policies:** Regulators should periodically review and update policies to address AI-specific risks
2. **Enabling AI-Based Affirmative Action:** Encourage AI use for financial inclusion with proportionate compliance requirements
3. **AI Liability Framework:** Introduce a flexible liability framework that supports innovation while ensuring accountability
4. **AI Liability Framework:** Set up a Standing Committee and a dedicated institution for continuous AI oversight

C. Capacity

1. **REs' Capacity Building:** Train board members and staff on AI governance and risks
2. **Regulators' & Supervisor's Capacity Building:** Establish an AI institute for training regulators and supervisors
3. **Framework for sharing Best Practices:** Create a platform for sharing AI use cases and lessons learned
4. **Recognize and Reward AI innovation:** Launch awards and grants to promote responsible AI innovation

Risk Mitigation Framework

A. Governance

1. **Board-Approved AI Policy:** Every RE should have a formal AI policy covering governance, risk, and accountability
2. **Data Lifecycle Governance:** Implement strong internal controls for data used in AI systems
3. **AI System Governance:** Ensure robust model lifecycle management, especially for autonomous AI
4. **Product Approval Process:** Include AI-specific risk checks in product approval processes

B. Protection

1. **Consumer Protection:** Ensure transparency, fairness, and grievance redressal in AI interactions
2. **Cybersecurity Measures:** Strengthen defenses against AI-specific threats and use AI for threat detection
3. **Red Teaming:** Conduct regular adversarial testing of AI systems to uncover vulnerabilities
4. **Business Continuity Plan:** Include AI-specific contingencies in BCPs to handle model failures
5. **Incident Reporting & Sectoral Risk Intelligence:** Establish a sector-wide AI incident reporting framework with a non-punitive approach

C. Assurance

1. **AI Inventory & Repository:** Maintain internal AI inventories and contribute to a sector-wide repository
2. **AI Audit Framework:** Implement risk-based internal and third-party audits of AI systems
3. **Public Disclosures:** Include AI governance and usage disclosures in annual reports
4. **AI Compliance Toolkit:** Develop standardized tools to help REs assess and demonstrate responsible AI use

The information presented in this document is derived from the official publication titled "Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI Committee Report)" released by the Reserve Bank of India (RBI) on 13 August 2025.

KPMG in India contacts:

Vishnu Pillai
Partner and Head TE-FS & OMP Kochi
T: +91 9819576349
E: vishnupillai@kpmg.com

Nitin Saxena
Partner
T: +91 9004489361
E: nitinsaxena1@kpmg.com

Ripunjay Gauttam
Director
T: +91 9892171473
E: ripunjaygauttam@kpmg.com

kpmg.com/in

Access our latest insights on
KPMG Insights Edge

Follow us on:

kpmg.com/in/socialmedia