



RBI's FREE-AI committee report in the financial sector

Framework for Responsible and Ethical
Enablement of Artificial Intelligence

September 2025

kpmg.com/in

KPMG. Make the Difference.



Table of Contents



Overview of the report	03
Benefits and risks of implementing AI in the financial sector	05
Overview of AI governance and ecosystem trends	07
FREE-AI framework	11
Way forward	17

01

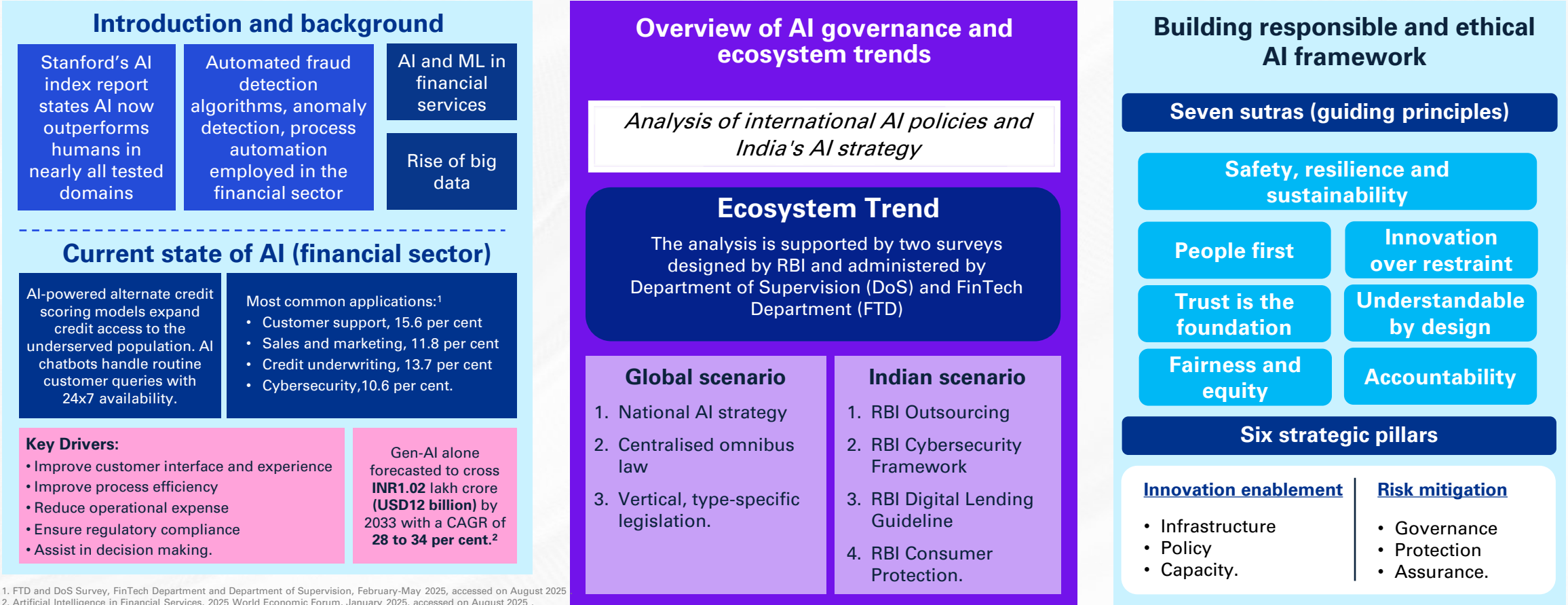
Overview of the report

Framework for responsible and ethical enablement of artificial intelligence



Overview

The Reserve Bank of India (RBI) has issued a report titled 'Framework for Responsible and Ethical Enablement of Artificial Intelligence in the Financial Sector' (hereinafter also referred to as '*RBI FREE AI Report August 25*'). This report sets out a framework to guide the use of AI in the financial sector, aiming to harness its potential while safeguarding against associated risks.



02

Benefits and risks of implementing AI in the financial sector



Benefits and risks



Benefits

AI for Financial Inclusion

- Alternative data like Mobile Usage and GST filings can be used to assess credit worthiness
- Chatbots can guide underserved populations with financial advice and grievance support
- Voice-enabled banking in local languages empowers illiterate users to access Financial services.

Leveraging AI in Digital Public Infrastructure

Boost access for smaller players: By building upon the strong base of India's Aadhaar and UPI infrastructure, AI can be implemented in KYC, UPI, account aggregators and public-good AI models, which in turn can boost access for smaller players.

Financial Sector Specific Models

- Developing indigenous models tailored to specific financial sector needs could ensure relevance, data sovereignty, and better support for local innovation
- Indigenous models or fine-tuned Small Language Models (SLMs) can offer cost-effective, multilingual, and secure solutions.

Autonomous AI Systems:

- Autonomous agents can break down complex goals and collaborate using protocols like MCP and A2A.
- This shift from task to decision automation could transform India's financial sector
- AI agents could help SME borrowers compare and secure loans in real time through intelligent interactions.

Synergies with other emerging technologies

- AI and quantum computing can enhance each other- AI can optimise quantum algorithms, while quantum can speed up AI training
- Though early-stage, these synergies hint at powerful next-gen AI systems for finance.



Risks

Model risk factors

- AI model risk stems from deviations from expected outcomes, due to bias/poor data/opaque design
- Data/design/calibration/implementation flaws, may cause cascading failures and erode trust
- While AI can help manage model risk, it also introduces 'model-on-model' risks and challenges like hallucinations and low explainability.

Operational Risks-systems under stress

- Automating mission-critical processes reduces human error but can amplify faults at scale
- AI systems like fraud detection may misfire due to model drift or data issues, causing financial/reputational harm.

Third-party risks

- Limited visibility into vendor and subcontractor controls hampers due diligence and risk management
- Concentration risk also arises from reliance on a few dominant providers, increasing systemic vulnerability.

Liability considerations in probabilistic and non-deterministic systems

- Bias from poor training data or opaque models complicates assigning responsibility among developers, deployers, or data providers
- Non-deterministic outcomes in credit or investment decisions make liability even harder to determine.

Cybersecurity and privacy of data

- AI systems often over-collect data, violating principles like data minimisation and purpose limitation
- Global AI infrastructure may conflict with data localisation laws and expose sensitive info via mosaic attacks
- Decryption during processing can create vulnerabilities like memory scraping or privileged access threats.

Risks to consumers and ethical concerns

- Algorithmic bias, data privacy violations, and opaque decision-making can harm vulnerable groups
- Manipulation, lack of informed consent, and autonomous decisions raise ethical and liability concerns.

Potential impact on financial stability

- May amplify procyclicality, herding effects, and model convergence risk, increasing market volatility
- Opacity and lack of stress-testing → unpredictability during crises, risking misclassification and credit tightening
- Events like the 2010 flash crash show how mismanaged automated trading can threaten financial markets.

Risk of AI-driven collusion

- Poses a theoretical risk to fair competition, especially in dynamic pricing and high-frequency trading
- Autonomous AI agents may sustain supra-competitive prices, breaching market conduct rules.

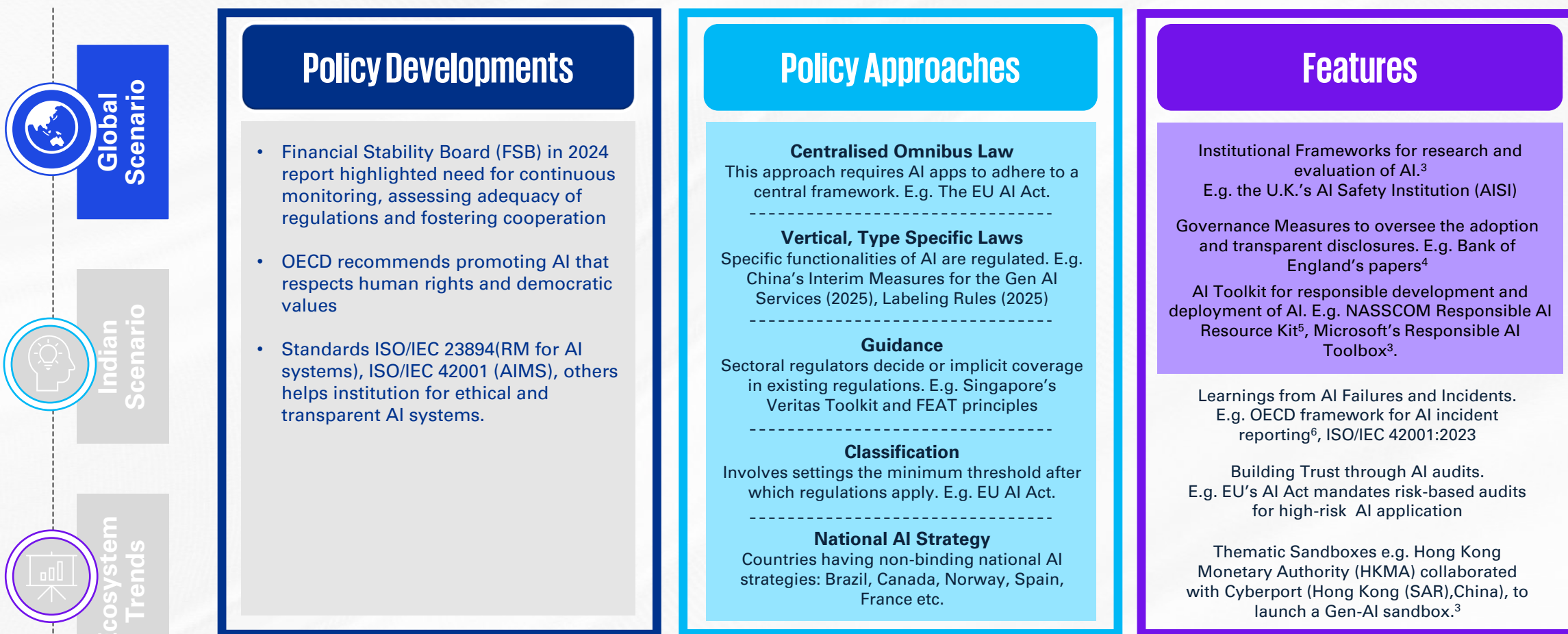
03

Overview of governance and ecosystem trends

Opportunities and challenges



Overview: Global scenario



3. FREE-AI Committee Report, RBI, 13 August 2025 accessed on August 2025

4. Bank of England and FCA – Discussion Paper on Artificial Intelligence and Machine Learning in UK Financial Services (Oct 2022) accessed on August 2025

5. NASSCOM Responsible AI Resource Kit, NASSCOM, 11 October 2022 accessed on August 2025.

6. Towards a common reporting framework for AI incidents, OECH, February 2025 accessed on August 2025.

Overview: Indian scenario



Existing Regulations and Policies³

Regulation/ Policy	AI Risks touchpoints under existing policies	Suggested AI specific enhancements
Outsourcing of Financial Services	(i) Accountability for outsourced activities (ii) Risk management and governance of third-party services.	(i) Clauses addressing AI specific risks may be included in the Outsourcing Agreement (ii) Clauses disclosing the use of AI by third-party vendors may be incorporated into the outsourcing agreement.
Cyber Security Framework in Banks (2016)	(i) Data confidentiality, integrity, and availability (ii) Incident reporting and response mechanisms.	(i) AI specific threats such as model poisoning in risk assessments under Cyber Security Policy (ii) May establish protocols for monitoring and mitigating AI related cybersecurity incidents.
Digital Lending (Sept 2, 2022)	(i) Data privacy and consent in digital lending (ii) Accountability of REs for third-party digital lending apps.	(i) Transparency in AI driven credit assessments, including disclosure of use of AI (ii) Include implementation of fairness audits to detect and mitigate algorithmic biases.
Customer Service in Banks (2015)	(i) Customer rights and grievance redressal mechanisms (ii) Board level oversight of customer service.	(i) Customer awareness when interacting with AI systems (ii) Establishment of processes for customers to contest AI driven decisions.
Fraud Risk Management (2024)	(i) Framework for early warning signals and fraud detection (ii) Risk management policies approved by the Board.	(i) AI driven fraud detection mechanisms can be encouraged (ii) May suggest to regularly test AI models for accuracy and bias in fraud detection.
IT, Governance, Risk, Controls and Assurance Practices (2023)	(i) Broad IT governance (ii) Oversight of information systems and related risks.	(i) May include Access Control measures for the autonomous AI.
Outsourcing of IT Services (2023)	(i) Risk assessments and due diligence of IT service providers (ii) Data protection and incident reporting obligations.	(i) May be required for the service providers to disclose use of AI in service delivery under (ii) May amend to include AI specific risk assessments.

Policy developments

- India aspires to be a global leader in ethical, responsible and innovation-driven AI
- Following a pro-innovation approach to promote beneficial AI use cases and limit harm to users
- IT Act [2000], Intermediary Rules and Guidelines and Bhartiya Nyaya Sanhita [2023] sufficiently address current risks

Policy approaches

- NITI Aayog's National Strategy for AI
 - Guidelines for leveraging AI across sectors
 - Set of Principles for responsible AI.
- India AI Mission for research and innovation backed by **INR10,372 Cr** in 2024 Union Budget.
- SEBI's paper on guidelines for responsible usage of AI/ML in Indian Security Markets⁷.

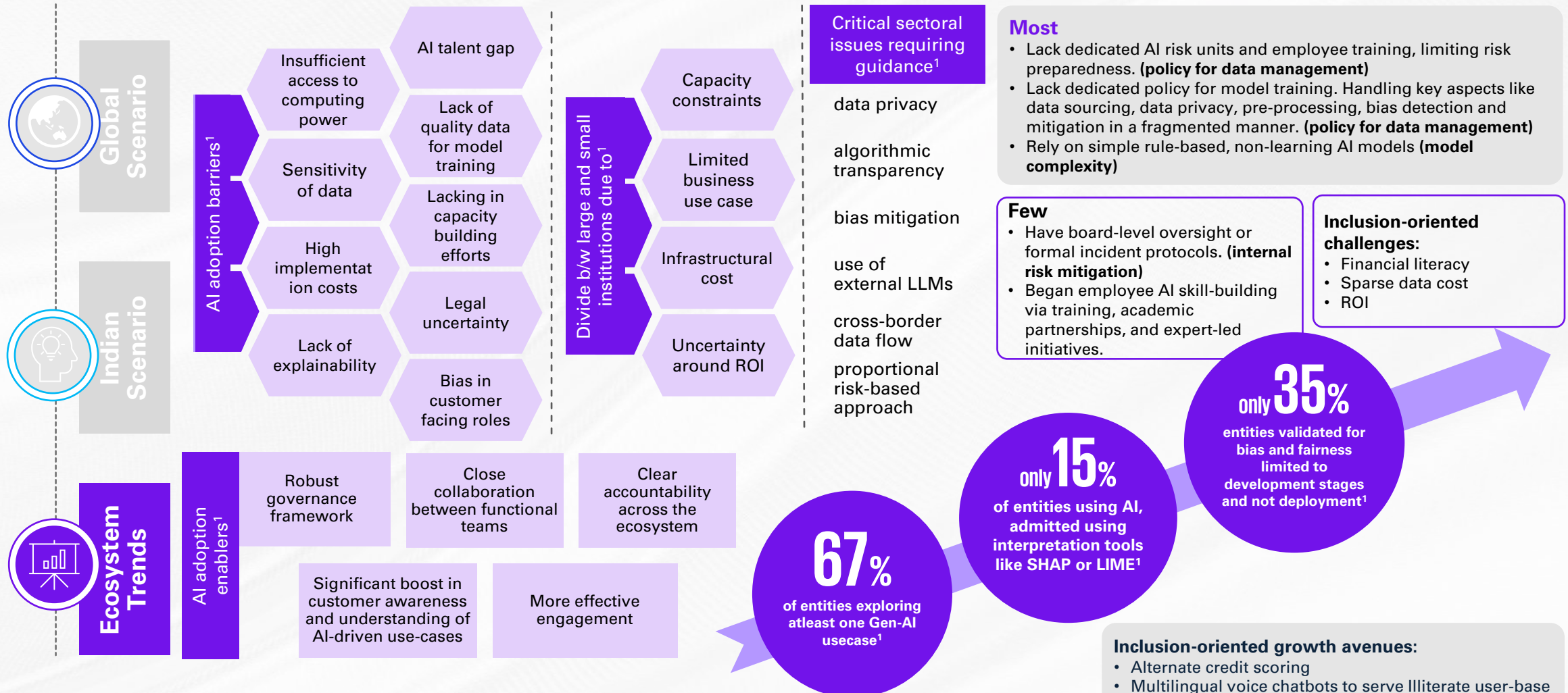
Analysing existing RBI guidelines³

Outsourcing	Cybersecurity	Lending	Consumer Protection
Outsourcing doesn't diminish liability [Cybersecurity and IT]	No specific mention of AI systems [Cybersecurity and IT]	Assess creditworthiness by economic profile [Digital Lending]	No specific mention of AI risks [Circular on consumer protection]
RBI guideline specification			
RE-AI governed internally 3 rd party AI outsourcing obligation applies.	Leverage existing Cybersecurity and IT guidelines by extending to AI systems.	Apply to AI-driven credit assessment Restrict data collection and require explicit consent.	REs should provide customers with means to challenge and seek clarifications on AI decisions.
FREE-AI Report recommendation			

3. Free-AI Committee Report, RBI, 13 August 2025 accessed on August 2025

7. Consultation Paper on guidelines for responsible usage of AI/ML in Indian Securities Markets, SEBI, 20 June 2025 accessed on August 2025

Overview: Ecosystem trends-gaps



1. FTD and DoS Survey, FinTech Department and Department of Supervision, February-May 2025, accessed on August 2025 .

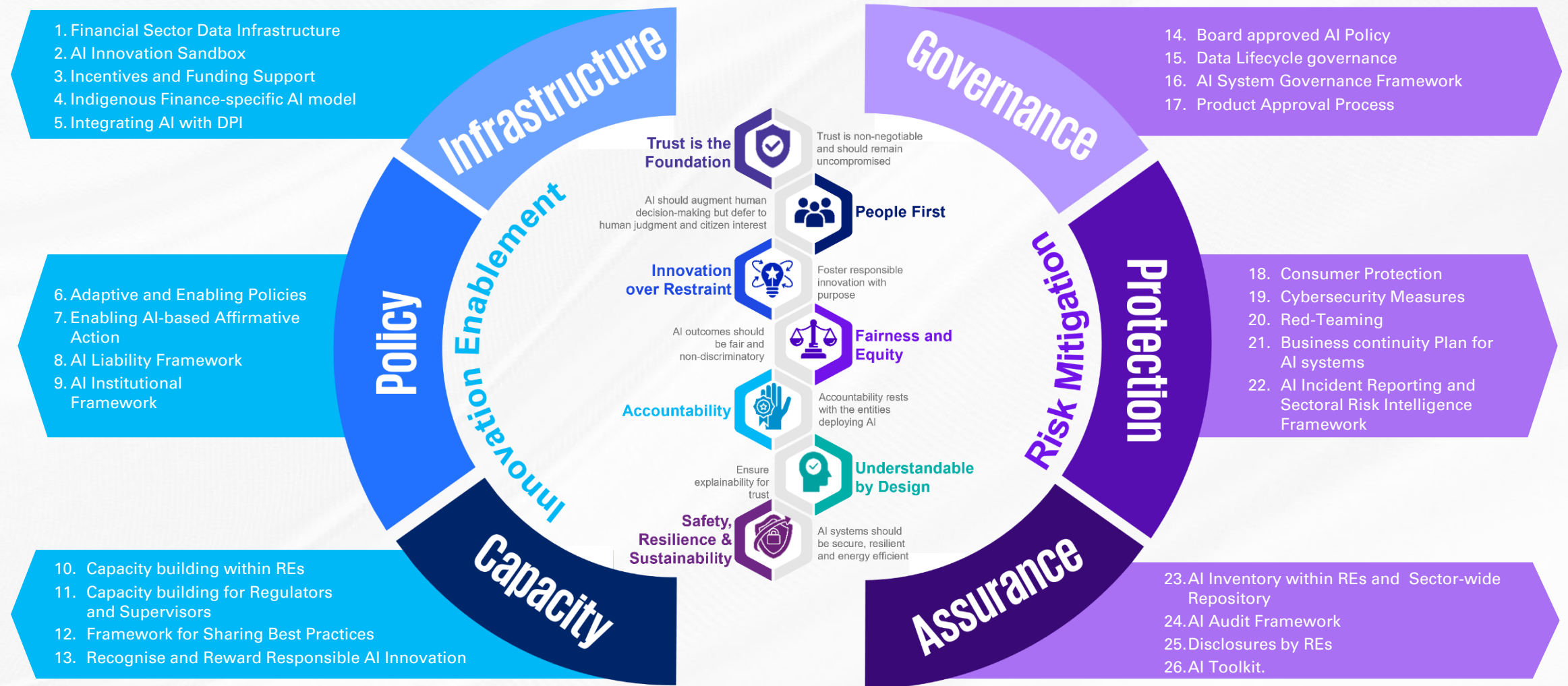
04

RBI's FREE-AI Framework:

Seven sutras and six pillars



FREE AI framework; bird's eye view¹

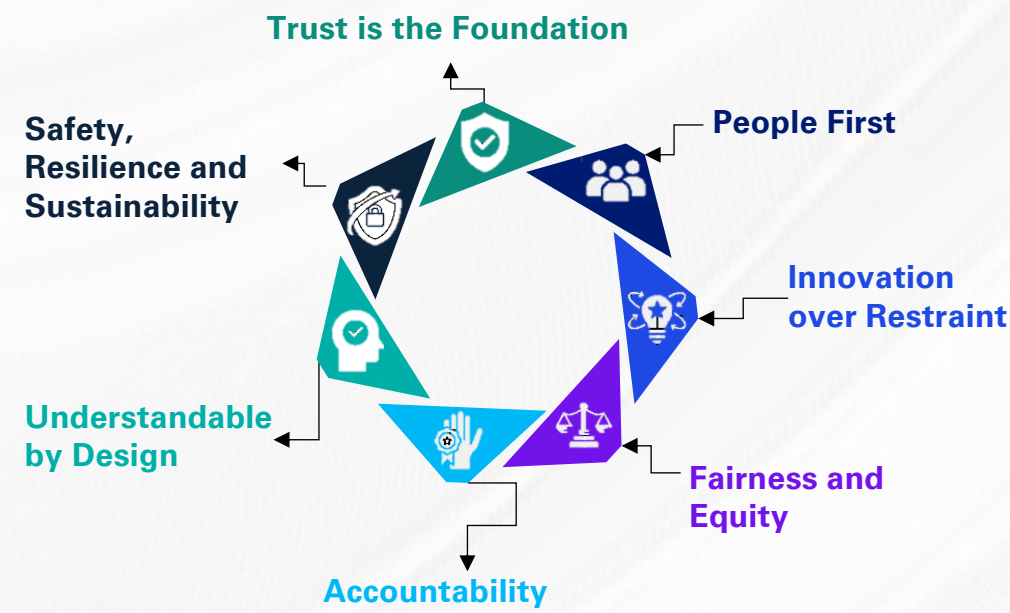


1. Free-AI Committee Report, RBI, 13 August 2025 accessed on August 2025.

FREE AI framework; the seven sutras

With the seven Sutras as the guiding light, actionable, structured, and forward-looking recommendations are made under the FREE-AI Framework.

The responsible deployment of AI within the financial sector calls for a dual focus approach; one that both fosters innovation and mitigates risks, as two complementary forces that must be pursued in tandem.



The seven sutras operate as an interconnected whole, reinforcing one another to form a robust framework for the responsible innovation and adoption of AI. True to the Sanskrit origin of the word sutra, meaning ‘thread’, these principles are to be woven through the entire lifecycle of AI systems.

1	Trust is the Foundation: Trust must be the guiding force behind all actions taken across the entire AI lifecycle. It must be viewed not as a regulatory burden but as a powerful enabler which will accelerate adoption, build confidence, and strengthen India’s competitive edge.
2	People First: While AI can help to improve efficiency and outcomes, final authority should rest with humans, who should be able to override AI, especially for societal benefit and human safety. Citizens should be made aware of AI-generated content and be informed when interacting with AI systems. Keeping human safety and interest at the core makes AI trusted.
3	Innovation over Restraint: AI should serve as a catalyst for augmentation and impactful innovation. Responsible AI innovation, that is aligned with societal values and aims to maximise overall benefit while reducing potential harm, should be actively encouraged. All other things being equal, responsible innovation should be prioritised over cautionary restraint.
4	Fairness and Equity: AI systems should be designed and tested to ensure that outcomes are unbiased and do not discriminate against individuals or groups. While AI should uphold fairness, it should not accentuate exclusion and inequity. AI should be leveraged to address financial inclusion and access to financial services for all.
5	Accountability: Entities that deploy AI should be responsible and remain fully accountable for the decisions and outcomes that arise from the use of these systems, regardless of their level of automation or autonomous functioning. Accountability should be clearly assigned. Accountability cannot be delegated to the model and underlying algorithm.
6	Understandable by Design: Understandability is fundamental to building trust and should be a core design feature, not an afterthought. AI systems must have disclosures, and the outcomes should be understood by the entities deploying them.
7	Safety, Resilience, and Sustainability: AI systems should operate safely and be resilient to physical, infrastructural, and cyber risks. These systems should have capabilities to detect anomalies and provide early warnings to limit harmful outcomes. AI systems should prioritise energy efficiency and frugality to enable sustainable adoption.

FREE AI framework; the six pillars- innovation enablement

Pillar



Infrastructure

Financial Sector Data Infrastructure:

Establish high-quality digital, public financial-sector data infrastructure. Integrate with the *AI Kosh;India Datasets* platform, established under the *INDIAai Mission*

AI Innovation Sandbox:

Establish an AI innovation sandbox to enable secure and controlled development of AI solutions for REs, FinTechs and innovators

Incentives and Funding Support:

Put appropriate incentive structures and infrastructure in place to encourage inclusive and equitable AI usage among smaller entities.

Indigenous Financial Sector Specific AI Models:

Develop indigenous AI models (including LLMs, SLMs, or non LLM models) tailored specifically for the financial sector, offered as a public good.

Integrating AI with DPI:

Establish an enabling framework to integrate AI with DPI to accelerate the delivery of inclusive, affordable financial services at scale.



Policy

Adaptive and Enabling Policies:

Regulators to periodically undertake an assessment of existing policies and legal frameworks to ensure effective enablement of AI driven innovations while addressing AI-specific risks.

Enabling AI-Based Affirmative Action:

Regulators should encourage AI-driven innovation that accelerates financial inclusion of underserved and unserved sections of society and other such affirmative actions by lowering compliance expectations as far as possible, without compromising basic safeguards

AI Liability Framework:

Regulators should adopt a graded liability framework encouraging responsible innovation. REs must continue to remain liable for any loss suffered by customers. An accommodative supervisory approach is recommended where the RE follows appropriate safety mechanisms (incident reporting, audits, red teaming etc.)

AI Institutional Framework:

Form a permanent multi-stakeholder AI standing committee under RBI to continuously advise it on emerging opportunities and risks, monitor the evolution of AI technology, and assess the ongoing relevance of current regulatory frameworks. The committee may be constituted for an initial period of five years, with a built-in review mechanism and a sunset clause.



Capacity

Capacity Building within REs:

REs to develop AI related capacity and governance competencies for the Board and C suite, along with structured and continuous training, upskilling, and reskilling programs across the broader workforce who use AI, to effectively mitigate AI risks and guide ethical as well as ensure responsible AI adoption.

Capacity Building for Regulators and Supervisors:

Regulators and supervisors to invest in training and institutional capacity building initiatives and ensure they possess adequate understanding of AI technologies to aptly match regulatory & supervisory frameworks with the evolving landscape of AI.

Framework for Sharing Best Practices:

The financial services industry to establish a framework for the exchange of AI-related use cases, lessons learned, and best practices (through bodies like IBA or SROs) and promote responsible scaling by highlighting positive outcomes, challenges and sound governance frameworks.

Recognise and Reward Responsible AI Innovation:

Regulators and industry bodies to introduce structured programs to recognise and reward responsible financial AI innovation, particularly those demonstrating positive social impact, embedding ethical considerations via recommended designs.

Recommendations

FREE AI framework; the six pillars- risk mitigation

Pillar



Governance

Board Approved AI Policy:

Craft policy explicitly articulating institution's position on AI governance, ethics and accountability, aligned with its values, obligations and risk appetite.

Create and deploy a risk classification framework for AI use-cases, to categorise as low/medium/high risk depending on factors such as impact on customers, criticality and potential for harm.

Data Lifecycle Governance:

REs to establish robust data governance frameworks, including internal controls and policies for data collection, access, usage, retention, and deletion for AI systems. These frameworks should ensure compliance with the applicable legislations, such as the DPDP Act, throughout the data life cycle.

AI System Governance Framework:

- Implement robust model governance mechanisms covering the entire AI model lifecycle (model design, development, deployment, and decommissioning, documentation, validation, and monitoring)
- Include mechanisms to detect and address model drift and degradation
- Put strong governance in place before deploying autonomous AI systems capable of independent financial decision-making.

Product Approval Process:

Ensure all AI-enabled products and solutions are brought within the scope of the institutional product approval framework and that AI-specific risk evaluations are included in the product approval frameworks.



Protection

Consumer Protection:

REs to adopt board-approved consumer protection framework focused on transparency, fairness and accessible recourse. Education campaigns essential to raise consumer awareness about their rights safe and AI usage.

Cybersecurity Measures:

REs to identify potential security risks on account of their use of AI and strengthen their cybersecurity ecosystems (hardware/software/processes) against them. REs may also make use of AI tools to strengthen cybersecurity, including dynamic threat detection and response mechanisms.

Red Teaming:

REs to establish structured red teaming processes spanning entire AI lifecycle with frequency and intensity, proportionate to assessed risk level and potential impact of application. Higher risk models to be subject to frequent and comprehensive red teaming. Trigger-based red teaming proposed for addressing evolving threats and changes.

Business Continuity Plan for AI Systems:

REs to augment existing BCP frameworks to include traditional system failures and AI model-specific performance degradation. REs to establish fallback mechanisms and periodically test fallback workflows and AI model resilience through BCP drills.

AI Incident Reporting and Sectoral Risk Intelligence Framework:

Financial sector regulators to establish a dedicated AI incident reporting framework for REs and FinTech's and encourage timely detection and reporting of AI related incidents. Said framework to adopt a tolerant, good-faith approach encouraging timely disclosure.

- The EmTech Repository of RBIH can be leveraged by expanding its scope for this purpose
- Consider developing a 'Responsible AI Adoption Score/Index' for the financial sector.



Assurance

AI Inventory within REs and Sector-Wide Repository: Regulators to periodically undertake assessments of existing policies and legal frameworks, ensuring effective enablement of AI driven innovation and AI-specific risk addressal.

AI Audit Framework: REs to implement a comprehensive AI audit framework, aligned with a board-approved AI risk categorisation, to ensure responsible adoption across the AI lifecycle, covering data inputs, model and algorithm, and the decision outputs. Audit framework to include 1.internal audit for AI application, 2.third-party audits for high-risk AI use cases and 3.periodic biennially conducted review of AI audit framework itself.

Effective AI Audit should include;

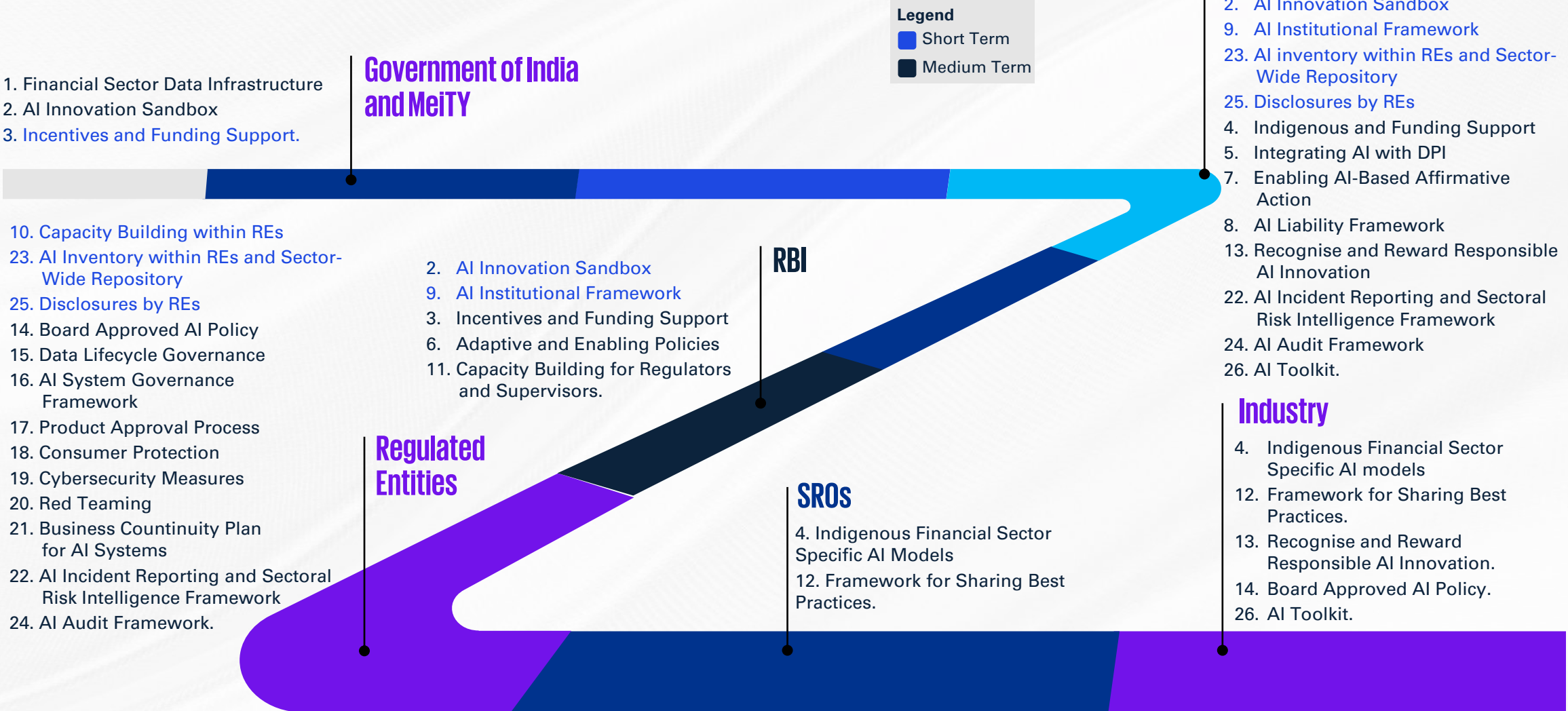
- correct and unbiased training data, decision explainability, fairness, consistency and compliance with regulations, halt mechanism for AI-driven process
- BCP for core AI systems, inspection using AI specific checklist and risk templates
- periodic assessment of policies and procedures.

Disclosures by REs: REs to include AI-related disclosures in their annual reports and websites. Regulators to specify an AI-specific disclosure framework to ensure consistency and adequacy of information across institutions. *AI disclosures should be part of annual reports* with necessary details like consumer protection measures, adoption areas etc.

AI Toolkit: 'AI Compliance Toolkit' to be used to help REs validate, benchmark, and demonstrate compliance against key responsible AI principles such as fairness, transparency, accountability, and robustness. The toolkit should be developed and maintained by a recognised SRO or industry body. Regulators to promote industry-led AI Compliance Toolkits to help REs meet regulatory standards.

Recommendations

Entity specific action items



Note: Numbers in accordance with convention used on slide 13 (titled bird's eye view)

05

Way forward



Making it happen: What REs should be doing

RBI's FREE-AI framework laid down structural intervention for smooth adoption of AI into financial sector. In the fast-evolving AI landscape, RE's should work in time-bound manner to adopt AI without compromising trust, fairness and transparency. To address the dual objectives of fostering innovation and mitigating risks, the recommendations have been organised into short-term and medium-term actions. This structure sees that immediate priorities are addressed while enabling sustainable progress.

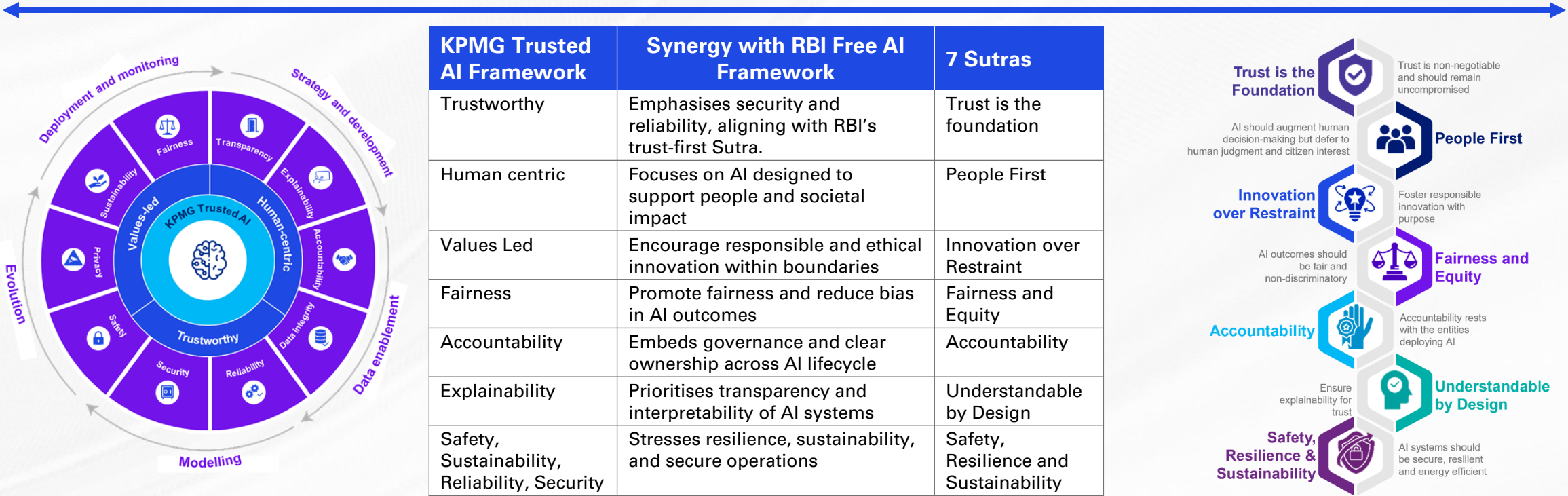
Immediate actions [Up to three months]	Capacity Building within REs	AI inventory within REs and sector-wide repository	Disclosures by REs
Medium Term Actions [Within one year]	Cybersecurity Measures	Identify potential security risks on account of their use of AI and strengthen their cybersecurity ecosystems. May use AI tools to strengthen cybersecurity, including dynamic threat detection and response mechanisms.	Board approved AI Policy
	Red teaming	Establish red teaming processes spanning AI lifecycle. It's frequency and intensity should be proportionate to the assessed risk level and potential impact of the AI application. Higher risk models to be subject to frequent and comprehensive red teaming. Trigger-based red teaming to be considered in addressing threats.	Data lifecycle governance
	Business Continuity Plan for AI Systems	Augment existing BCP frameworks to include traditional system failures and AI model-specific performance degradation. Establish fallback mechanisms and test fallback workflows and AI model resilience through BCP drills.	AI System governance framework
	AI incident Reporting and Sectoral Risk Intelligence Framework	- Financial sector regulators to create a dedicated AI incident reporting framework for REs and FinTech's (Leverage RBIH's EmTech Repository, expanding its scope for this purpose) - Encourage timely detection and reporting of AI related incidents by adopting a tolerant, good-faith approach - Develop a responsible AI adoption score/index for the financial sector.	Product approval process
	AI audit framework	- Implement a comprehensive AI audit framework, includes internal audit for AI application and third-party audits for high-risk AI use cases, periodic biennially conducted review of AI audit framework- inspection using AI specific checklists and risk templates and periodic assessment of policies and procedures - Correct and unbiased training data, decision explainability, fairness, consistency, compliance with regulations, halt mechanism, BCP, Align said audit framework with board-approved AI risk categorization, covering data inputs, models and algorithms.	Consumer protection
			- Craft policy on institution's position on AI governance, ethics and accountability, aligned with its values, obligations and risk appetite - Create and deploy a risk classification framework for AI use-cases, to categorise as low/medium/high risk depending on factors such as impact on customers, criticality and potential for harm. - Establish robust data governance frameworks, including internal controls and policies for data collection, access, usage, retention, and deletion for AI systems. These frameworks help ensure compliance with the applicable legislations, such as the DPDP Act, throughout the data life cycle. - Implement robust model governance mechanisms covering the entire AI model lifecycle (model design, development, deployment, and decommissioning, documentation, validation, and monitoring) - Includes detection and address model drift and degradation - Putting governance in place before deploying autonomous AI systems capable of independent financial decision-making. - Ensure that all AI-enabled products and solutions are in scope of the institutional product approval framework and that AI-specific risk evaluations are included in the product approval frameworks. - Adopt board-approved consumer protection framework focused on transparency, fairness and accessible recourse - Ongoing education campaigns to raise consumer awareness about safe AI usage and their rights.

KPMG in India operationalized the RBI FREE-AI framework leveraging its global trusted AI framework, meeting business objectives and values

KPMG in India has mapped RBI’s FREE-AI framework with its global trusted AI framework to synergize AI strategy with organizational goals and values. This can act as a strategic enabler for organizations seeking to deploy AI responsibly while maintaining competitive agility.

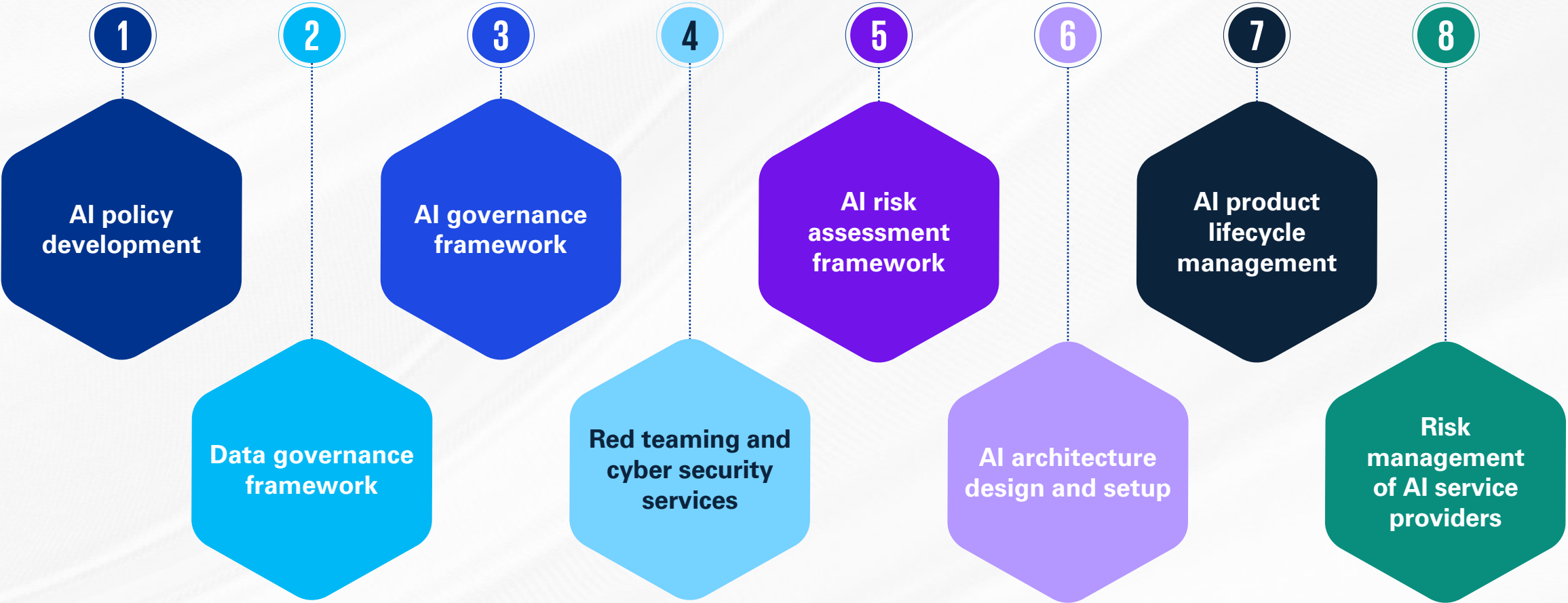
Anchored on ten guiding principles aligned with RBI FREE AI Report August ‘25 and sutras defined in the report, it integrates ethical considerations with innovation, transforms AI adoption into a sustainable, resilient, and trust-driven journey, facilitating organizations to innovate confidently in a continuously evolving landscape.

Bridging Principles



KPMG in India's responsible AI accelerators

Outlined below are the key accelerators driving progress in the Responsible AI domain



Contributors

Aakansha Gupta

Aditya Singh Rathore

Bhumika Mahajan

Jay Goyal

Kushagra Sharma

Sahil Shaikh

Shubham Sharma

KPMG in India contact:

Akhilesh Tuteja

Global Head - Cyber Security

T: +91 987102 5500

E: atuteja@kpmg.com

Atul Gupta

Partner, Head of Function - Digital Trust and Cyber

M: +91 98100 81050

E: atulgupta@kpmg.com

Kunal Pande

Partner, Co-Head - Digital Risk and Cyber

Leader - Digital Trust for FS

M: +91 98926 00676

E: kpande@kpmg.com

Rohan Padhi

Partner and Co-Lead Digital Risk and Cloud Security

M: +91 99302 24081

E: rohanpadhi@kpmg.com

kpmg.com/in



Access our latest insights on KPMG Insights Edge

Romharsh Razdan

Partner, Lead Payment Risk, Digital Risk and Cyber

M: +91 99755 96366

E: romharsh@kpmg.com

Hardik Shah

Director, Digital Risk and Cyber

M: +91 98337 00462

E: hardikshah7@kpmg.com

Follow us on:

kpmg.com/in/socialmedia



The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act on such information without appropriate professional advice after a thorough examination of the particular situation.

KPMG Assurance and Consulting Services LLP, Lodha Excelus, Apollo Mills Compound, NM Joshi Marg, Mahalaxmi, Mumbai-400 011
Phone: +91 22 3989 6000, Fax: +91 22 3983 6000.

© 2025 KPMG Assurance and Consulting Services LLP, an Indian Limited Liability Partnership and a member firm of the KPMG global organisation of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved.

The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organisation.

This document is for e-communication only.