KPMG

RBI advisory on AI-ACT&RS

AI-Accelerated Cyber Threats and Related Safeguards

Published by Reserve Bank of India (RBI), Department of Supervision, Central Office Cyber Security and IT Risk (CSITE) Group

June 2026

kpmg.com/in

KPMG. Make the Difference.

Contents

Introduction

Introduction to the advisory

The Reserve Bank of India (RBI) has issued two advisories on AI-Accelerated Cyber Threats and Related Safeguards (AI-ACT&RS) for:
(i) all Commercial Banks, including Small Finance Banks and Payments Banks *on 27th April 2026*
(ii) for all authorised Non-Bank Payment System Operators (PSOs) *on 1st June 2026.*

Applicability

- Public / Private Sector Banks
- Private Sector Banks
- Foreign Banks in India
- Small Finance Banks
- Payments Banks
- Non-Bank Payment System Operators

Advisory's rationale

While Regulated Entities (REs) under RBI continue to increase adoption of AI tools and applications within their ecosystem to support innovation and efficiency, threat actors too can parallelly use AI to accelerate reconnaissance, exploitation, and social engineering

As AI-driven attacks increase speed, scale, and sophistication beyond traditional controls, *enhanced cyber-resilience* is required to address *AI-accelerated* threats to critical systems.

Objectives

- Strengthen technology stacks against AI-accelerated cyber threats
- Ensure safe and governed use of AI and generative AI by regulated entities
- Improve preparedness, detection, response, and resilience for AI-enabled attacks
- Address AI-driven social engineering and impersonation risks
- Supplement existing CIS protections to address emerging AI-related risks.

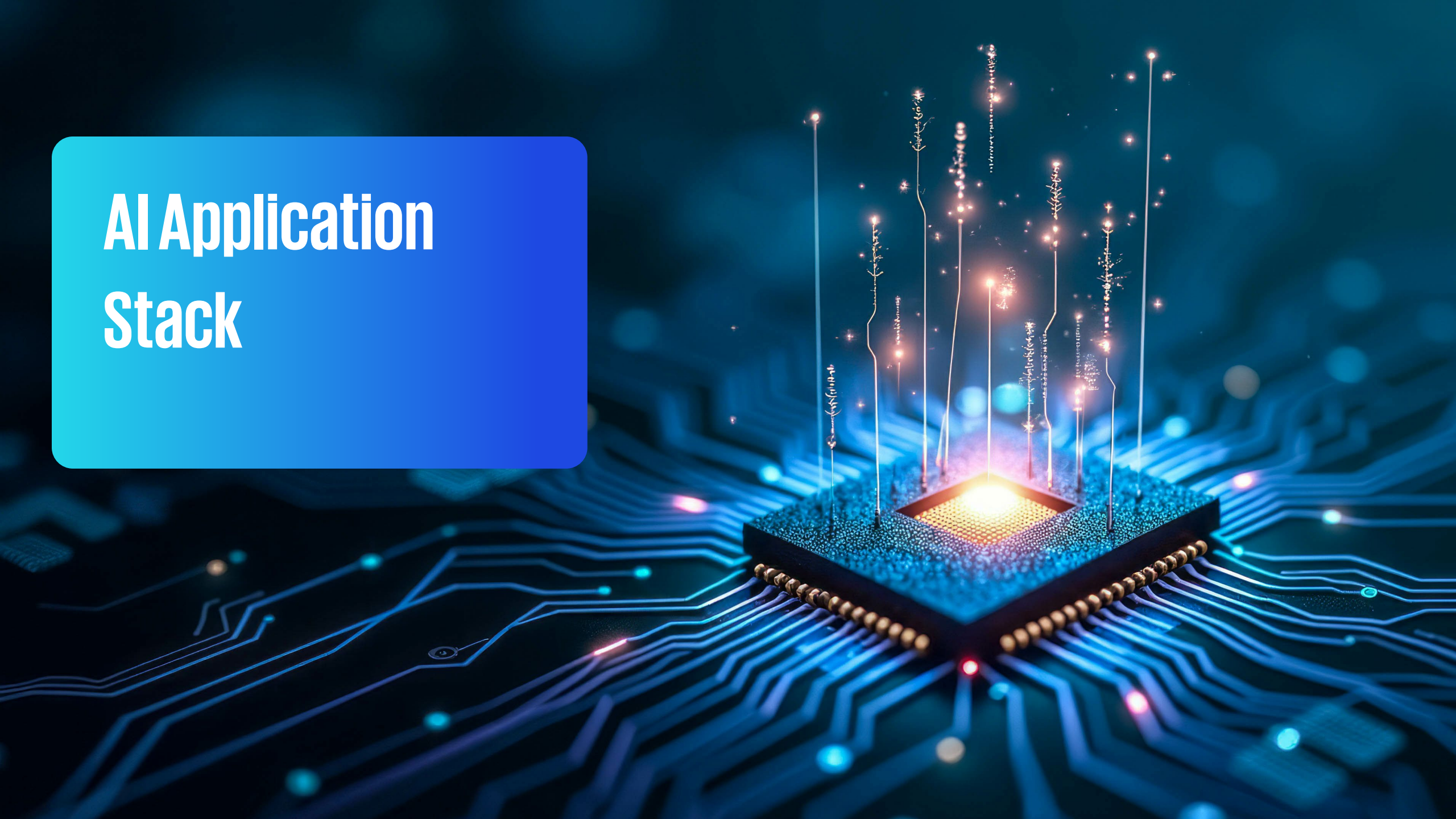
Advisory's structure

The advisory addresses AI-related risks from a cyber-security perspective. The key focus areas are:

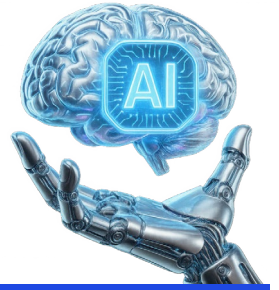
I — **Defending Technology Stacks against AI-Accelerated Cyber Threats**
Clause 4.1 – 4.6
Strengthening REs technology stacks against AI-accelerated cyber threats

II — **AI Governance and Security Framework**
Clause 5.1 – 5.13
Ensuring appropriate safeguards for REs own use, deployment, integration, and sourcing of AI and generative AI tools and systems

AI Application Stack



Mapping the Advisory's Controls to AI Application Stack



Leveraging its deep understanding and experience, KPMG in India has mapped AI Application Tech-Stack with the RBI Advisory's Controls Providing a structured view to identify control gaps across the AI lifecycle,

AI Tech Stack	Applicable Clause 4 Controls	Applicable Clause 5 Controls	
User and Channel Layer Web Apps, Mobile Apps, Chatbots	**4.5** AI-Enabled Social Engineering and Impersonation	**5.9.b** Human oversight and override in customer-facing AI	**5.12** AI Usage Policy and Controls
AI Application and Orchestration Layer RAG | Workflows | Business Logic	**4.4** Monitoring, Detection, Response and Testing	**5.2** Classification, Ownership and Lifecycle Management **5.3** Secure Development, Integration and Configuration	**5.6** Performance Monitoring and Output Validation **5.7** AI Change Management
AI Model Layer Models used behind the application	**4.3** Vulnerability and Patch Management	**5.2** Classification, Ownership and Lifecycle Management **5.5** AI-Specific Threat Controls	**5.11** Third-Party AI Risk Management **5.7** AI Change Management
Data Layer Training Data | Prompts | Outputs | Logs	**4.4** Monitoring, Detection, Response and Testing	**5.2** Classification, Ownership and Lifecycle Management	**5.4** Access Control and Data Protection
Middleware Layer APIs | Integrations	**4.3** Vulnerability and Patch Management	**5.3** Secure Development, Integration and Configuration	
Platform and Infrastructure Layer Cloud | VMs | Containers | OS | Networks	**4.3** Vulnerability and Patch Management **4.3** Monitoring, Detection, Response and Testing	**5.9** Resilience and Continuity	
Access, Identity and Control Plane IAM | Privileged Access		**5.4** Access Control and Data Protection	**5.10** AI Agents and Privileged Access
Monitoring, Logging and Resilience Layer SIEM/SOAR | IR | Backup | Recovery	**4.4** Monitoring, Detection, Response and Testing	**5.8** Logging, Traceability and Forensic Readiness	**5.9** Resilience and Continuity

Clause 4

Defending Technology Stacks against AI-Accelerated Cyber Threats

Defending Technology Stacks against AI-Accelerated Cyber Threats (1/2)

A structured approach to strengthening technology stacks against emerging risks from frontier and AI-enabled cyber threats

	What is Required ?	Potential outcome	Potential risks of non-implementation
	4.1 Risk Assessment and Preparedness		
1	• Ongoing awareness of the evolving AI-driven cyber risk landscape • Periodic, structured AI-focused cyber risk assessments across CIS • Review and strengthening of preparedness measures based on risk assessments.	Provides an up-to-date, risk-informed view of AI-enabled cyber threats, enabling proactive preparedness, faster detection, and improved response to AI-accelerated attacks.	• Blind spots to emerging AI-enabled threats • Delayed detection and response to sophisticated attacks • Increased likelihood of significant cyber incidents impacting CIS.
	4.2 Architecture and Cyber Resilience		
2	• Adoption of ‘Assume-breach’ and ‘Zero-trust’ principles • Strong baseline cyber hygiene and hardened configurations • Secure-by-design, threat modelling, and security testing for critical system changes.	Strengthens cyber resilience by limiting attack propagation, reducing exposure to exploitable weaknesses, and embedding security into system architecture and change lifecycles.	• Increased attack blast radius and lateral movement • Prolonged detection and recovery timelines • Greater impact from advanced or AI-accelerated attacks.
	4.3 Vulnerability and Patch Management		
3	• Accelerated vulnerability assessment and patching for critical and exposed systems • Up-to-date asset inventory, dependency mapping, attack-surface visibility, and SBOMs • Controls over open-source and third-party software components • Formal recording, approval, and tracking of remediation exceptions.	Ensures timely identification, prioritisation, and remediation of vulnerabilities across critical and exposed systems, reducing the window of exposure to rapidly exploitable weaknesses and improving visibility of affected assets and components.	• Extended exposure to known and exploitable vulnerabilities • Increased risk of compromise through open-source and third-party components • Uncontrolled exceptions leading to high-impact cyber incidents.

Defending Technology Stacks against AI-Accelerated Cyber Threats (2/2)

	What is Required ?	Potential outcome	Potential risks of non-implementation
	4.4 Monitoring, Detection, Response and Testing		
4	• Continuous monitoring for AI-enabled and AI-accelerated attack patterns • Periodic tuning of detection tools using threat intelligence • Updated incident response playbooks for AI-accelerated scenarios • Regular breach simulations, red-team exercises, and cyber drills.	Improves detection of anomalous activity, enables faster and coordinated incident response, and validates preparedness against AI-accelerated cyber attacks.	• Missed or delayed detection of advanced attacks • Ineffective or slow incident response • Untested readiness for AI-driven attack scenarios.
	4.5 AI-Enabled Social Engineering and Impersonation		
5	• Review and strengthen controls against AI-enabled phishing, vishing, deepfake-based impersonation, and targeted social engineering • Employee awareness and simulation exercises covering emerging AI-enabled social engineering scenarios.	Improves organisational resilience against AI-enabled social engineering by reducing susceptibility of employees, privileged users, vendors, and customers to impersonation and deception attempts.	• Increased success of AI-enabled phishing and impersonation attacks • Higher risk of fraud, unauthorised access, and financial loss • Elevated reputational and customer trust risks.
	4.6 Leveraging AI Tools for Cybersecurity/IT Activities		
6	• Policies and controls to prevent data leakage through external or public AI tools • Use of AI tools for cyber/IT activities only through approved and controlled arrangements, with safeguards aligned to data sensitivity.	Enables safe and controlled use of AI tools for cybersecurity and IT operations while protecting sensitive customer, financial, and security information from unauthorised exposure.	• Leakage of sensitive or regulated information through AI tools • Unauthorised or uncontrolled use of external AI services • Increased regulatory, data protection, and security risks.


Clause 5
AI Governance and Security Framework

AI Governance and Security Framework (1/4)

As AI and generative AI adoption increases, Regulated Entities must ensure that such technologies are deployed responsibly, with appropriate safeguards to prevent cyber, data protection, and operational resilience risks.

	What is Required ?	Potential outcome	Potential risks of non-implementation
	5.1 AI Governance		
1	• Formulation of AI Governance and Security Framework • Periodic risk assessment covering AI-specific threat vectors • Update and Review framework measures at regular intervals.	Provides controlled and accountable use of AI across the organization, reducing security, data, and resilience risks.	• Uncontrolled use of AI across the organization • AI-specific security and data risks • Increased Cyber, data protection, and regulatory exposure.
	5.2 Classification, Ownership and Lifecycle Management		
2	• Risk-based classification of AI systems • Defined ownership and accountability • Lifecycle governance from approval to decommissioning.	Risk-aligned controls over AI systems based on criticality, data sensitivity, and impact with clear accountability and end-to-end governance across the AI system lifecycle.	• High-risk AI systems may operate without appropriate controls • Lack of clear ownership and accountability • Unmanaged AI lifecycle.
	5.3 Secure Development, Integration and Configuration		
3	• Adoption of secure development, integration, and configuration practices for AI systems and components. • Maintenance of an inventory of material AI systems and their integrated environment.	Supports a secure-by-design approach to the development and integration of AI systems and provides visibility of AI systems and their dependencies.	• Lack of visibility into AI systems and dependencies • Unmanaged vulnerabilities, insecure components, or outdated configurations remaining in production.

AI Governance and Security Framework (2/4)

	What is Required?	Potential outcome	Potential risks of non-implementation
4	**5.4 Access Control and Data Protection** • Implementation of access management, encryption, data loss prevention and monitoring and data protection controls for AI systems handling sensitive data.	Enables controlled access to AI systems and protection of sensitive data across the AI lifecycle, including prompts, outputs, logs, and integrations.	• Unauthorized access to AI systems and sensitive data • Data leakage or unintended exposure of sensitive information.
5	**5.5 AI-Specific Threat Controls** • Implementation of safeguards to mitigate AI-specific threats such as prompt injection, data poisoning and model manipulation, etc. • Validation of AI-generated outputs before use in production.	Reduces exposure to AI-specific attack techniques. It also prevents unsafe or incorrect AI-generated outputs from being deployed in production or security-sensitive environments.	• AI systems may be compromised through AI-specific attacks • AI-generated outputs may introduce vulnerabilities • Sensitive data may be exposed or misused.
6	**5.6 Performance Monitoring and Output Validation** • Continuous monitoring of material AI systems • Validation of AI outputs against expected outcomes.	Continuous monitoring of material AI systems Validation of AI outputs against expected outcomes.	• Incorrect or misleading AI outputs • Undetected model drift or anomalies.
7	**5.7 AI Change Management** • Implementation of change management controls for material changes to AI systems and it's associated components, including third party AI services.	Provides controlled, traceable changes to AI systems by enforcing review, approval, testing, and rollback mechanisms for modifications to models, data, configurations, and integrations.	• Provides controlled, traceable changes to AI systems by enforcing review, approval, testing, and rollback mechanisms for modifications to models, data, configurations, and integrations.

AI Governance and Security Framework (3/4)

	What is Required ?	Potential outcome	Potential risks of non-implementation
	5.8 Logging , Traceability and Forensic Readiness		
8	• Logging of AI inputs, outputs, model decisions, access events, and administrative actions • Traceability and forensics- readiness for material AI systems with integration intro enterprise monitoring system.	Provides end-to-end visibility and traceability of AI system activity, enabling monitoring, investigation, and accountability.	• Delayed incident response • Reduced forensic capability • Limited monitoring of AI activities.
	5.9 Resilience and Continuity		
9	• Implementation of resilience arrangements, including fallback, recovery, incident response and continuity measures for AI Systems.	Ensures continuity and recovery of critical AI systems during disruptions by providing defined fallback, recovery, and continuity arrangements, and human-override mechanisms.	• Disruption of critical or customer-facing services • Delayed response to abnormal AI behavior.
	5.10 AI Agents and Privileged Access		
10	• Define AI-enabled tools or agents with system access as privileged non-human identities. • Apply least privilege, scoped access, approval gates, complete logging, and periodic access reviews.	Establishes controlled and auditable use of AI agents with elevated access, reducing unintended actions across production, security, and enterprise systems.	• Excessive or uncontrolled privileges for AI systems • Uncontrolled automated workflows triggered by AI.

AI Governance and Security Framework (4/4)

	What is Required ?	Potential outcome	Potential risks of non-implementation
5.11 Third-Party AI Risk Management			
11	• Subject third-party AI services and AI-enabled products to third-party risk assessment • Implementation of compensating controls wherever required.	Provides clear risk visibility and control over third-party AI dependencies, particularly in relation to data handling and model behavior. It supports safe adoption of external AI services.	• Sensitive data shared with third-party AI services can be misused • Limited audit rights or exit arrangements.
5.12 AI Usage Policy			
12	• Establish policies governing employee and third-party use of AI/generative-AI tools • Implementation of controls to monitor AI usage.	Enables controlled use of AI tools across the organization, aligned to approved use cases and data-handling expectations, supporting timely identification of policy deviations or misuse.	• Unauthorized use of AI tools • Unmonitored AI usage • Policy violations may go undetected.
5.13 Audit, Testing, Training and Awareness			
13	• Periodic audit and security testing of material AI systems • Provision of role-based training/awareness to employees, senior management and board on use of AI systems/tools.	Provides independent assurance over AI security and robustness, enabling early identification of weaknesses and stronger user accountability.	• Exploitable AI behaviors. • Employees and decision makers misuse or over-relay on Ai systems • Lack of awareness among employees on the use of AI Systems.

Making it happen:

What REs should be doing

KPMG in India operationalized RBI's Advisory controls, meeting business objectives and values

KPMG in India has mapped controls of RBI's ACT&RS Advisory with its global trusted AI framework to synergize AI strategy with organizational goals and values. This can act as a strategic enabler for strengthening technology stack against AI-related threats and help implement governance and security frameworks addressing ecosystem-wide related developments.

Trusted AI Framework Tenet	RBI Advisory Controls	Insight
Fairness	5.6 – Performance Monitoring and Output Validation; 5.12 – AI Usage Policy and Controls	The focus on output review and controlled AI usage supports consistent and context-appropriate outcomes, helping organisations promote balanced AI-assisted decision-making.
Transparency	5.6 – Performance Monitoring and Output Validation; 5.8 – Logging, Traceability and Forensic Readiness	Monitoring and traceability expectations reinforce transparency by improving visibility into how AI systems operate and how outputs flow through processes.
Explainability	5.6 – Performance Monitoring and Output Validation	Review of AI outputs prior to operational use aligns explainability with practical decision workflows and human checkpoints.
Accountability	5.1 – AI Governance; 5.2 – Classification, Ownership and Lifecycle Management; 5.7 – AI Change Management	Governance, ownership, and structured change practices anchor AI systems to defined roles and responsibilities across the lifecycle.
Data Integrity	5.4 – Access Control and Data Protection; 5.2 – Classification, Ownership and Lifecycle Management	The linkage between AI controls and data classification and protection highlights disciplined data handling as a foundation for dependable AI outcomes.
Reliability	5.6 – Performance Monitoring and Output Validation; 5.9 – Resilience and Continuity	Continuous monitoring and resilience planning position reliability as an ongoing operational capability.
Security	4.2 – Architecture and Cyber Resilience; 4.3 – Vulnerability and Patch Management; 4.4 – Monitoring, Detection, Response and Testing; 5.5 – AI-Specific Threat Controls	AI is positioned as part of the broader cyber ecosystem, integrating AI-specific risks into established resilience and defence practices.
Safety	5.9 – Resilience and Continuity; 5.9 (b) – Human Oversight and Override	Human oversight and fallback arrangements support measured adoption of AI in critical and customer-facing environments.
Privacy	5.4 – Access Control and Data Protection; 4.6 – Leveraging AI Tools for Cyber/IT Activities	Safeguards around sensitive data and AI tool usage reinforce privacy-by-design considerations alongside AI adoption initiatives.

How KPMG in India can help

Building on its work in financial services, KPMG in India supports organisations in preparing for an AI-ready future.

Our AI Labs have been certified for ISO 42001:2023, the international standard for Artificial Intelligence Management Systems, we are the first among leading professional firms in India to attain this certification.

By aligning capabilities to business needs, we help organisations navigate the evolving AI threat landscape.

KPMG in India's Capabilities

Regulatory alignment

Helping clients achieve alignment with regulatory expectations through KPMG's deep understanding of AI capabilities and risk across the sector, and leveraging its' study across circulars like –

1. RBI's AI-Accelerated Cyber Threats and Related Safeguards
2. RBI's FREE AI Framework Report
3. India AI Governance Guidelines

Strengthening AI Governance

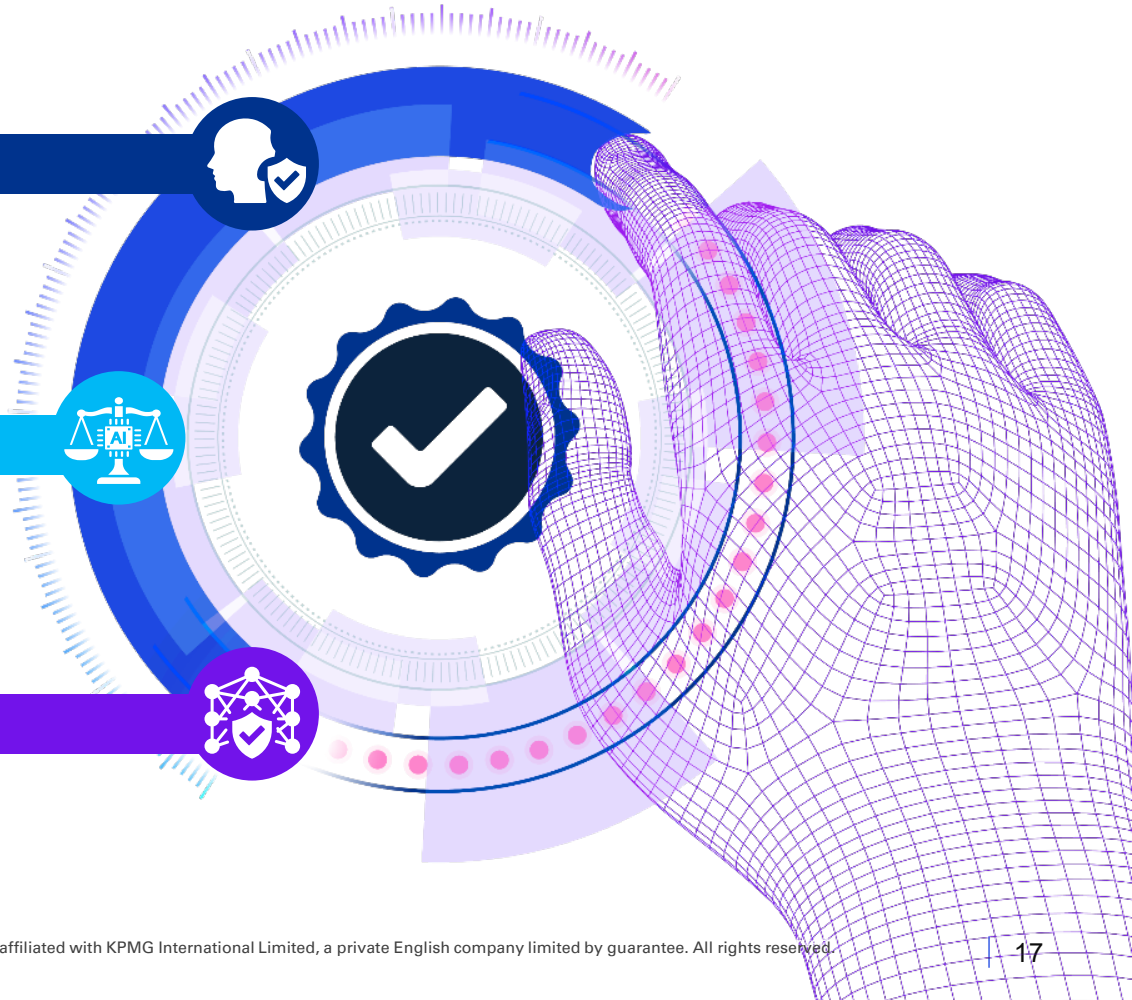
Create a robust governance structure through –

1. AI policy and procedure development
2. AI governance framework
3. AI risk assessment framework
4. AI efficacy assessments and reviews

Incorporate AI within client's ecosystem

Find the right AI-use-case fit and assist in AI solution's implementation –

1. Ecosystem assessment for AI use-case best-fit
2. AI architecture design and setup
3. Risk management of AI service providers



Contributors

- Aakansha Gupta
- Bhumika Mahajan
- Kushagra Sharma
- Madhuri Gangaramani
- Sahil Shaikh
- Shubham Sharma

KPMG in India contacts:

Akhilesh Tuteja
Partner & National Leader,
Clients and Markets
KPMG in India
E: atuteja@kpmg.com

Kunal Pande
National Leader – Cyber, Risk
and Compliance Services
KPMG in India
E: kpande@kpmg.com

Rohan Padhi
Partner, Cyber, Risk and
Compliance Services
Lead AI
E: rohanpadhi@kpmg.com

Romharsh Razdan
Partner, Cyber, Risk and
Compliance Services
Lead Payment Risk
E: romharsh@kpmg.com

kpmg.com/in

Access our latest insights on KPMG Insights Edge

Follow us on:

kpmg.com/in/socialmedia

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.