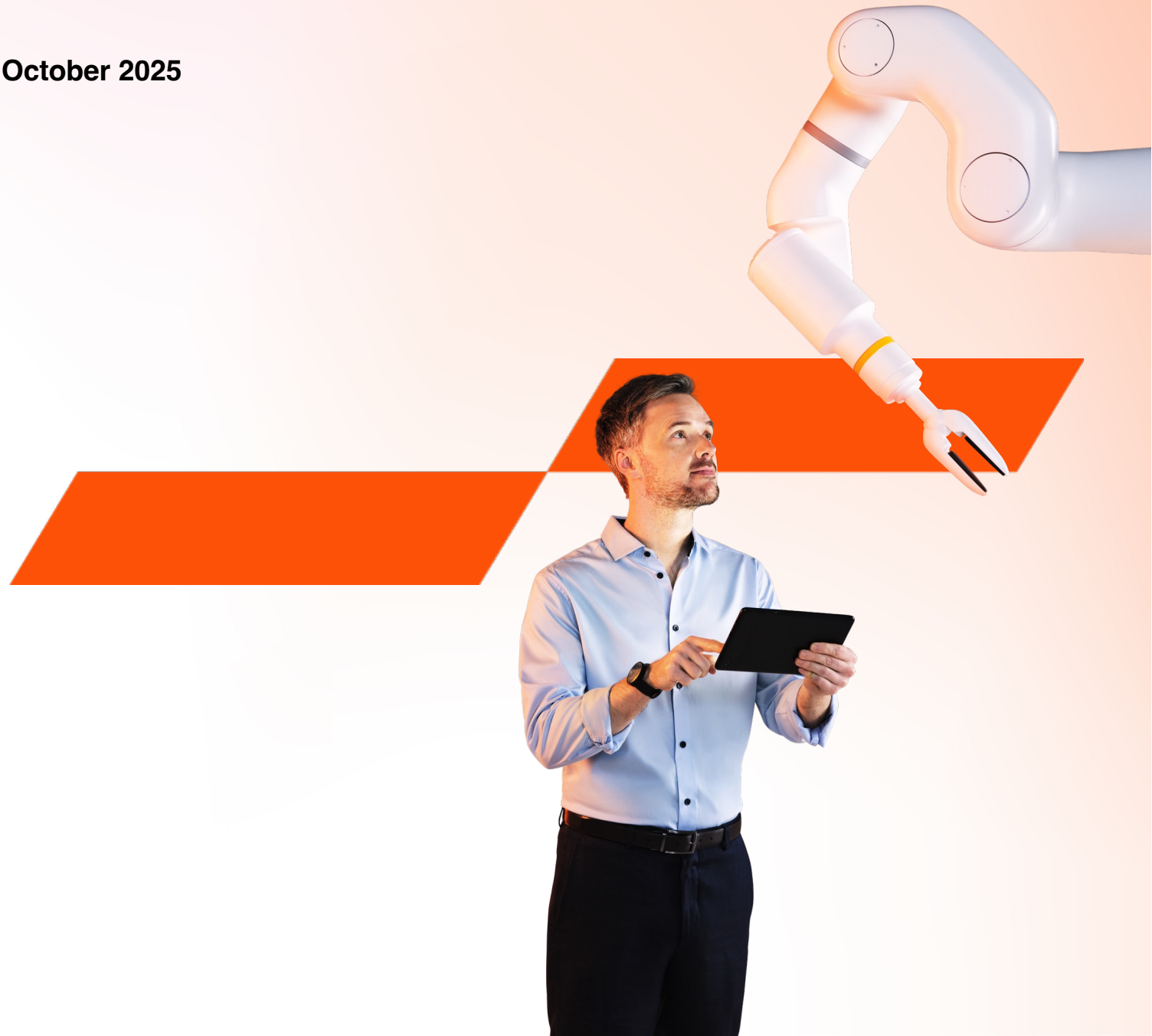
pwc

AI trust and governance: Enabling innovation with responsibility

October 2025



From concept to critical asset: The AI adoption timeline

From being an aspirational concept half a decade ago to becoming a strategic imperative for enterprises across the market spectrum – the evolution of artificial intelligence (AI) has been nothing short of remarkable. Initially limited to niche academic and translation tasks, it now plays a vital role within complex business operations, enhancing efficiencies and unlocking new revenue streams. For today's executives – from the board to chief information security officers (CISOs) – AI is no longer a theory but a transformative force reshaping traditional decision-making processes, pushing forward operational agility and optimising risk management.

There have been distinct phases in the evolution of AI:

AI: The journey from hype to enterprise value

Pre-2020

Foundation phase

AI could understand and respond to text but was limited mostly to research and simple detection or translation tasks.

2020–2022

Breakthrough phase

AI capabilities expanded rapidly with tools like ChatGPT, GitHub Copilot and DALL-E becoming popular.

AI gained practical use in writing, image creation, coding and answering questions.

2022–2023

Democratisation and adoption

AI became easy to use – just type a question and get smart answers.

Businesses started using it for customer service, marketing and reports.

It became part of everyday software and websites.

2023–2024

Enterprise integration

Companies began embedding AI into critical industry workflows – including legal, banking and healthcare – to improve accuracy, customer experience and operational efficiency. AI moved from novelty status to a fundamental work tool.

2024–2025

Governance and regulation era

Governments and organisations recognised the need for AI regulation and ethical oversight. New laws and frameworks emerged to prevent misuse and protect people, with companies creating dedicated AI governance teams to manage responsible deployment.

Anatomy of AI risks and implementation challenges: Navigating risks and framework adoption in India

As AI becomes deeply woven into strategic decision-making and operational frameworks, the stakes of its deployment have never been higher. The widespread adoption of AI calls for vigilant oversight and a proactive understanding of where this technology can falter.

The table below offers an insightful lens into real-world misfires – moments when AI systems have stumbled, causing unexpected consequences. By exploring these incidents, leaders can anticipate risks, reinforce safeguards, and steer AI innovations with confidence and care.

Anatomy of AI failures: AI incidents, risks and regulatory impacts

Incident scenario	Incident details	Risk	Type of risk	Dimension
How did employees inadvertently leak confidential data using AI tools?	Employees used AI tools to debug and summarise internal documents, inadvertently leaking confidential data. **Impact:** Generative AI (GenAI) tools were subsequently banned internally to prevent further leaks.	Sensitive data may be exposed during model training or inference.	**Regulatory**	Data privacy and confidentiality
How did a chatbot spread illegal information about housing laws in one of the biggest cities in the world?	A chatbot spread illegal information about housing and businesses violating the law. **Impact:** The chatbot was shut down, and policies were updated to require source verification in AI outputs.	Outputs may violate industry rules and propagate misinformation.		Regulatory and legal non-compliance
How did a globally recognised AI tool cause a market dip with false statements?	During its launch demo, a GenAI tool gave incorrect information about the largest space telescope, undermining research potential and stakeholder confidence. **Impact:** A major tech company's stock dropped significantly due to these false claims.	AI can generate plausible but incorrect or misleading content.	**Accuracy**	Hallucinations and inaccuracy

Incident scenario	Incident details	Risk	Type of risk	Dimension
How did a chatbot crash during Black Friday sales cause transaction failures?	On Black Friday, an AI retail chatbot failed from token limits and server overload, unable to handle customer queries or transactions, forcing a temporary switch to human agents. **Impact:** The retailer lost sales and suffered reputational damage.	AI models may fail under real-world conditions or scale poorly.	**Accuracy**	Performance risk
How was a major airline's chatbot failure linked to lost sales and missed opportunities?	An airline booking chatbot provided incorrect information on bereavement fare discounts, conflicting with company policy. **Impact:** The court held the airline responsible for the chatbot's statements.	Processing and computation are opaque, making adverse outcomes hard to justify.		Lack of explainability
How was personal data exposed due to weak passwords in an AI platform?	The data of millions of job applicants at a multinational fast food chain was breached through vulnerabilities in third-party AI tools used to screen and manage job applications. **Impact:** Personal data exposure raised serious privacy concerns.	AI systems can be vulnerable to phishing, deepfakes or malicious injections.	**Security**	Cybersecurity threats
How did bias in AI recruiting tools affect female candidates?	An AI recruiting tool biased against female candidates downgraded resumes containing 'women', leading to its discontinuation. **Impact:** The tool was scrapped after the bias was discovered.	AI may reflect or amplify societal biases present in training data.	**Ethical/ societal**	Model bias and fairness
How did failure to monitor AI-driven KPIs impact business outcomes?	Lack of monitoring AI-driven business KPIs caused an investment to miss promised outcomes. **Impact:** Significant investment losses occurred due to untracked AI performance.	Lack of monitoring can hide underperformance or inefficiencies.	**Business**	RoI and operational efficiency

Executive evolution: How AI is reshaping the C-suite

AI is fundamentally transforming the roles and responsibilities of C-suite executives across organisations. The integration of AI capabilities – from data analytics to autonomous decision making – is redefining traditional functions, enabling more strategic, proactive and data-driven leadership.

Below is a snapshot of AI's impact across key C-suite roles, highlighting the transition from legacy processes to AI-enabled operations:

C-suite role transition from legacy processes to AI-enabled operations

Role	Executive evolution	AI-driven transformations
Chief financial officer (CFO)	From retrospective analysis to real-time financial intelligence	• **Intelligent document scanning:** AI reviews internal and competitor annual reports and filings for benchmarking. • **Forecasting and scenario planning:** Generates real-time financial forecasts and simulates what-if scenarios. • **Revenue and expense Insights:** Analyses transactional data to uncover cost-saving opportunities, and more
Chief risk officer (CRO)	From reactive risk management to predictive and proactive oversight	• **Risk modelling:** AI generates simulations for credit, market and operational risks. • **Early warning systems:** Detects anomalies and emerging threats in real time. • **Regulatory stress testing:** Automates scenario generation for compliance.
Chief compliance officer (CCO)	From manual compliance checks to automated, real-time governance.	• **Horizon scanning:** Scans regulatory websites, summarises changes and highlights implications. • **Control testing:** Reviews control evidence, detect gaps and recommends remediation. • **Fraud and AML monitoring:** Analyses unstructured data and identifies patterns and potential risks.

Role	Executive evolution	AI-driven transformations
Chief information security officer (CISO)	From IT security to AI-integrated cyber defence	• **Threat detection:** Identifies AI-specific risks like model evasion and data poisoning. • **Security audits:** Automates vulnerability assessments across AI systems. • **Governance:** Collaborates with legal and compliance to ensure AI safety.
General counsel	From contract review to strategic AI governance	• **Contract analysis:** AI reviews and flags risks in legal documents. • **IP protection:** Monitors AI-generated content for copyright and patent issues. • **Regulatory alignment:** Tracks global AI laws (e.g. the EU AI Act) and ensures compliance.
Chief marketing officer (CMO)	From campaign management to hyper-personalised engagement	• **Content creation:** Generates ads, social media posts and email campaigns. • **Customer insights:** Analyses behaviour to tailor messaging. • **Brand voice modelling:** Ensures consistency across channels.
Chief people officer (CPO)	From HR operations to talent strategy and culture shaping	• **Recruitment:** AI screens resumes and drafts job descriptions. • **Employee engagement:** Analyses feedback and sentiment. • **Learning and development:** Creates personalised training paths.
Chief business officer (CBO)	From business operations to AI-driven growth strategy	• **Market analysis:** AI identifies new opportunities and competitive threats. • **Partner evaluation:** Assesses strategic fit and risk in M&A or alliances. • **Operational efficiency:** Optimises workflows and resource allocation.

The infusion of AI into the C-suite is not merely a technological upgrade – it represents a fundamental shift in leadership paradigms. By harnessing AI's capabilities, executives are evolving into proactive strategists who leverage data-driven insights to anticipate risks, optimise operations and innovate with confidence. Embracing this transformation is essential for organisations seeking to navigate the complexities of the modern business landscape and secure sustainable growth in an AI-driven future.

Bringing order to AI evolution and global governance overview

As AI technologies rapidly evolve and integrate into global economies, establishing clear governance frameworks is essential to ensure their responsible and ethical use. This section highlights key regulatory and governance approaches from the European Union, the United States, Singapore and the National Institute of Standards and Technology (NIST) AI Risk Management Framework.

How the world is governing AI

United States of America — Active ✓

Intent

- Promote innovation while ensuring safety and fairness.
- Regulate federal use of AI.
- Address national security and civil rights concerns.

Key provisions

- **Executive Order (Oct 2023):** Federal agencies must assess AI risks and ensure transparency.
- **AI Bill of Rights:** Non-binding framework emphasising data privacy, algorithmic fairness, and user consent.
- **State-level laws:** Texas and California lead with bills on AI disclosures and consumer protection.

Implications

- Federal contractors must comply with AI risk management protocols.
- Companies face increasing state-by-state compliance complexity.
- Push toward self-regulation and ethical AI boards.

Singapore — Active ✓

Intent

- Promote responsible innovation.
- Support small states in AI governance.

Key provisions

- **Model AI Governance Framework for Generative AI (MGF-Gen AI):** Defines nine principles (e.g. transparency, accountability) for responsible GenAI use, backed by global tech leaders.
- **Global AI Assurance Pilot:** Tests GenAI systems for safety and reliability, aiming to set global standards for AI certification.
- **Regional collaboration:** Joint efforts with Japan and Asia-Pacific partners to assess GenAI performance across languages, cultures and harm categories.

Implications

- Singapore is positioning itself as a global AI governance hub.
- Companies must align with technical testing standards and transparency norms.

NIST AI Risk Management Framework

Active ✓

Goal

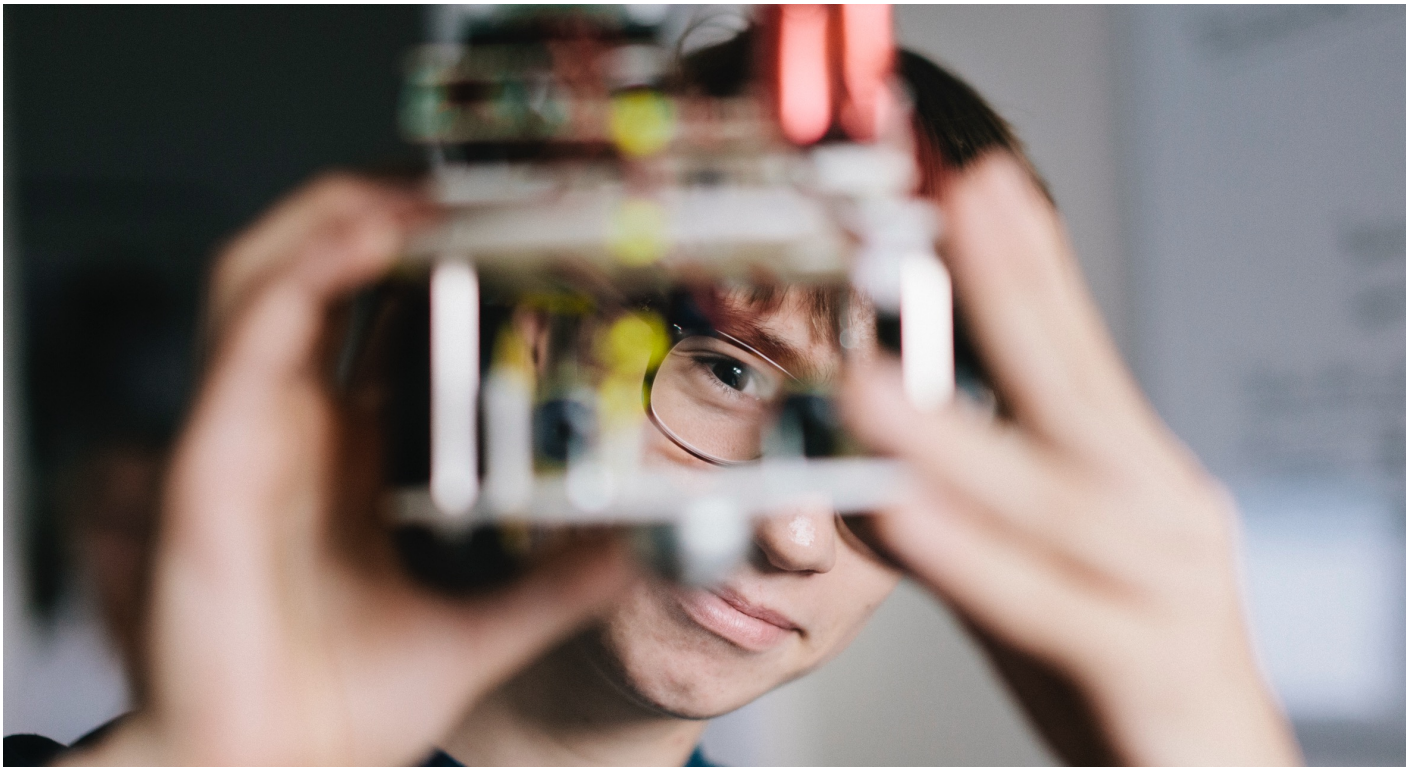
- Understand and manage risk related to AI systems.
- Ensure AI systems are fair, transparent, safe and secure, and accountable.
- Promote the development of AI in a manner that fosters trust and societal benefit.

Key aspects covered

- **Govern:** Establish clear AI risk governance structures, policies and roles.
- **Map:** Identify and map the risks associated with AI systems, including potential impacts on privacy, fairness, accountability, safety and security.
- **Measure:** Develop and implement metrics and methods to evaluate the risks and performance of AI systems.
- **Manage:** Ensure continuous monitoring, auditing and improvement of AI systems post-deployment.

Benefits of adoption

- Adopted by leading industry players globally.
- The NIST framework provides more comprehensive guidance on the risks associated with AI and risk management practices.



Sources: https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf - **NIST AI Framework**
https://aiverifyfoundation.sg/wp-content/uploads/2024/05/Model-AI-Governance-Framework-for-Generati… - **MGF-Gen AI**
https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2025/singap… - **Joint testing by SG government**
https://www.federalregister.gov/documents/2023/11/01/2023-24283/safe-secure-and-trustworthy-develop… - **Executive Order**
https://bidenwhitehouse.archives.gov/ostp/ai-bill-of-rights/- **AI Bill of Rights**

This multi-jurisdictional patchwork signals a global commitment to aligning AI development with human values and societal needs, crafting guardrails that preserve innovation without sacrificing trust

RBI's FREE–AI Committee Report

In the dynamic landscape of financial services, the integration of AI brings immense opportunities as well as significant risks. Recognising this dual-edged nature, the Reserve Bank of India (RBI) established the **Framework for Responsible and Ethical Enablement of Artificial Intelligence (FREE-AI) Committee**.[1] This pioneering initiative aims to equip regulated financial entities with a robust principled foundation for adopting AI technologies that are trustworthy, secure and aligned with societal values.

FREE-AI articulates **seven core principles**, termed as the 'Seven Sutras', which are designed to ensure that AI innovations are **fair, reliable, compliant and sustainable**. Complementing these are pragmatic, medium- and short-term recommendations that guide financial institutions (FIs) in implementing governance, capacity building, risk mitigation and consumer protection measures.

The following infographic captures the core philosophy and actionable pillars of the RBI's FREE-AI Committee Report – a vital blueprint for fostering innovation that inspires confidence in India's financial ecosystem.

RBI's recommendations for AI governance in FIs

Foster innovation

Infrastructure

I. Financial sector data infrastructure
II. AI innovation sandbox
III. Incentives and funding support
IV. Indigenous financial sector specific AI models
V. Integrating AI with digital public infrastructure (DPI)

Policy

I. Adaptive and enabling policies
II. Enabling AI-based affirmative action
III. AI liability framework
IV. AI institutional framework

Capacity

I. Capacity building within regulated entities (REs)
II. Capacity building for regulators and supervisors
III. Framework for sharing best practices
IV. Recognise and reward responsible AI innovation

Core principles

Seven sutras

Risk mitigation

Governance

I. Board-approved AI policy
II. Data lifecycle governance
III. AI system governance framework
IV. Product approval process

Protection

I. Consumer protection
II. Cybersecurity measures
III. Red teaming
IV. Business continuity plan (BCP) for AI systems
V. AI incident reporting + sectoral risk intelligent framework

Assurance

I. AI inventory within REs + sector-wide repository
II. AI audit framework
III. Disclosures by REs
IV. AI toolkit

Source: RBI's FREE-AI Committee Report

1. https://rbidocs.rbi.org.in/rdocs/PublicationReport/Pdfs/FREEAIR130820250A24FF2D4578453F824C72ED9F5D5851.PDF

Key recommendations for REs: Charting a responsible AI adoption path

To navigate the complexities of AI adoption in the financial sector responsibly, the RBI's FREE AI Committee outlines actionable recommendations for REs. The recommendations are segmented across four critical pillars – Capacity building, Governance, Protection, and Assurance – each with specific tasks and timelines to help entities develop resilient, trustworthy and compliant AI ecosystems.

Key recommendations for REs across the four pillars of the RBI's FREE-AI Committee

Pillar	Key recommendation	Implementation timeline
Capacity building	Develop AI capacity and governance competencies organisation-wide, including **board and C-suite**, for responsible AI adoption.	**Medium term**
Governance	• Develop a **board-approved AI policy** covering governance, accountability, risk appetite, safeguards, auditability, consumer protection, AI disclosures, model lifecycle and liability, including a risk classification framework based on customer impact, criticality and harm potential. Integrate AI risks into overall risk management with oversight by the risk management or equivalent committee; REs may form an AI adoption committee. • **Establish robust data governance** with controls and policies for AI data collection, access, usage, retention and deletion, compliant with laws such as the DPDP Act. • **Implement structured oversight across the entire AI lifecycle** – design, development, deployment and decommissioning – to ensure safe, compliant AI. Prioritise human oversight for autonomous AI in financial decisions, especially for medium and high-risk use cases. • Include AI-specific risk evaluations within **product approval frameworks.**	**Medium term**

Pillar	Key recommendation	Implementation timeline
Protection	• Establish a **board-approved consumer protection framework** emphasising transparency, fairness and accessible recourse mechanisms for customers, complemented by continuous consumer education on safe AI use and rights. • Identify AI-related security risks and strengthen cybersecurity, leveraging AI tools for dynamic threat detection and response. • Establish lifecycle-wide red teaming with frequency and intensity aligned to AI risk levels, including trigger-based exercises for evolving threats. • Augment existing **BCP frameworks to address AI model performance degradation** by establishing fallback mechanisms and regularly testing resilience through drills. • Detect and **report AI-related incidents** in a timely manner.	**Medium term**
Assurance	Maintain an **inventory of AI systems** covering AI models and algorithms, use cases and applications, dependencies, risk categorisation, and grievances. Update the inventory semi-annually.	**Short term**
	• Implement a risk-based AI audit framework aligned with **board-approved risk categories**, covering inputs, models and outputs. Conduct independent third-party audits for high-risk AI and review the framework biennially to address emerging risks and regulations. • Include **AI-related disclosures** in their annual reports and websites.	**Medium term**

SEBI's vision on responsible AI use

Recognising the increasing significance and reliance on AI/ML technologies in capital markets, SEBI has released a consultation paper on 'Guidelines for responsible usage of AI/ML in Indian securities markets',[2] inviting feedback from various stakeholders. The proposed guidelines draw on NITI Aayog's framework on AI/ML usage in their paper on 'AI for Viksit Bharat: The opportunity for accelerated economic growth'[3] as well as the International Organization of Securities Commissions' (IOSCO's) consultation paper on 'Artificial intelligence in capital markets: Use cases, risks, and challenges'[4] that looks into the application of AI/ML and its associated risks. Emphasising adaptability based on the size and scale of organisations, SEBI's white paper is structured around five key guiding principles.

SEBI's five guiding principles for responsible AI use

01

Model governance

- Establish senior-level expertise to oversee AI/machine learning (ML) deployment and governance.
- Include continuous risk assessment, especially when using third-party solutions.
- Implement strong governance practices such as audits, detailed logging, and compliance with legal and regulatory requirements.

02

Data privacy and cybersecurity

- Define clear policies for data security, cybersecurity and privacy in AI/ML use.
- Ensure personal data collection and processing comply strictly with relevant laws.
- Maintain stringent controls over data handling and protection.

03

Fairness and bias

- Ensure AI/ML models are fair and unbiased, avoiding discrimination against any group.
- Deploy robust processes and controls to detect and mitigate bias in data and algorithms.
- Provide targeted training for teams on identifying and addressing bias.

04

Testing framework

- Rigorously test AI/ML models using real-life, realistic scenarios, including shadow testing.
- Conduct comprehensive document testing processes to demonstrate functionality and fairness.
- Maintain continuous monitoring to detect unexpected or inexplicable behaviours.

05

Investor protection and disclosure

- Disclose AI/ML usage clearly to customers in understandable language.
- Provide mechanisms for customer grievance and informed decision making.
- Ensure compliance with relevant regulatory frameworks governing AI/ML deployment.

This approach signals SEBI's commitment to balancing innovation with investor confidence and market integrity.

Achieving this balance requires concerted efforts from regulators, industry consortia and organisations to enable clear guidelines, governance maturity, technical innovation, and capacity building tailored to India's unique AI ecosystem.

---

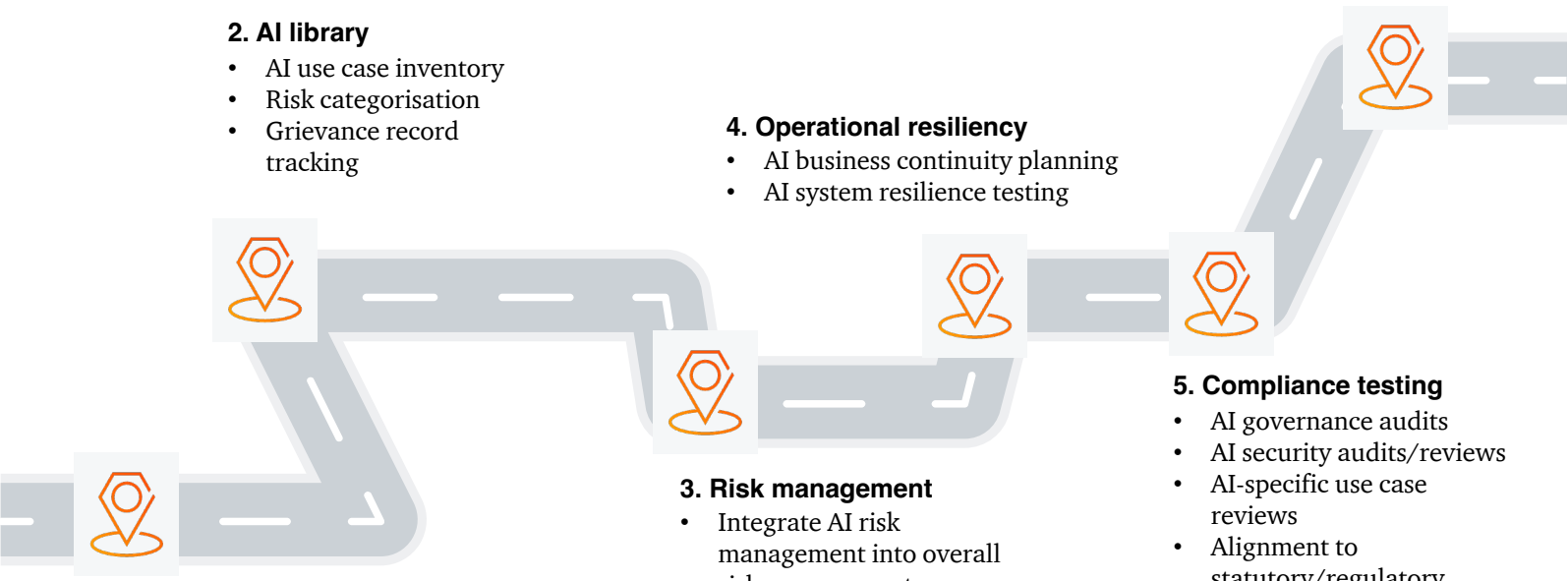
2. https://www.sebi.gov.in/reports-and-statistics/reports/jun-2025/consultation-paper-on-guidelines-for-responsible-usage-of-ai-ml-in-indian-securities-markets_94687.html
3. https://niti.gov.in/sites/default/files/2025-09/AI-for-Viksit-Bharat-the-opportunity-for-accelerated-economic-growth.pdf
4. https://www.iosco.org/library/pubdocs/pdf/IOSCOPD788.pdf

AI governance journey: A roadmap for REs in India

As AI continues to reshape the financial and regulated sectors in India, organisations face mounting pressure to adopt and govern AI responsibly amid an evolving regulatory landscape. Successfully navigating this terrain requires a structured, phased approach that integrates compliance, risk management, operational resilience and capacity building.

PwC has developed a clear, actionable AI governance journey to help REs adopt, govern and scale AI in a way that balances innovation with accountability and trust .

Roadmap to adopt and govern AI responsibly with organisations



1. Governance foundation
- AI policy and governance framework
- Data lifecycle governance framework
- Integrate AI-related risk mitigants in other frameworks
- AI model development lifecycle: Use case development design and deployment
- Design principles – fairness, explainability, accountability and security

2. AI library
- AI use case inventory
- Risk categorisation
- Grievance record tracking

3. Risk management
- Integrate AI risk management into overall risk management process
- Assess and implement cybersecurity measures
- Model validation
- AI use case approval checklist and templates

4. Operational resiliency
- AI business continuity planning
- AI system resilience testing

5. Compliance testing
- AI governance audits
- AI security audits/reviews
- AI-specific use case reviews
- Alignment to statutory/regulatory requirements (e.g. DPDPA)
- Related disclosures in annual reports
-

6. Ongoing monitoring and reporting
- AI use case – benefits tracking
- Model efficacy reviews
- AI incident management and root cause analysis (RCA)
- Performance and risk indicators monitoring **risk and control self-assessments (RCSA)**

Capacity building
- Specific awareness sessions for board and C suite
- Upskilling training on AI across 3 lines of defence (LOD)
- Awareness sessions on governance and usage

Source: PwC analysis

This roadmap equips REs with a holistic and scalable framework to govern AI responsibly. By embedding governance at every stage – from policy foundation to continuous monitoring – organisations can foster innovation securely and transparently, building trust with regulators, customers and internal stakeholders alike.

How can PwC help?

Navigating the complex and rapidly evolving landscape of GenAI requires not just technology expertise but also a deep understanding of financial services, regulatory risks and emerging frameworks. At PwC, guided by our commitment to 'trust in what matters', we bring together this domain knowledge with a governance-led approach that helps organisations adopt AI responsibly. Our focus is on aligning AI strategies with compliance, trust and long-term value creation, ensuring businesses harness AI's power while managing associated risks effectively.

Our comprehensive value proposition is built around four key pillars that address the entire AI lifecycle and governance landscape:

01

Responsible AI governance and operating model

We support organisations in crafting robust AI policies and governance frameworks. This includes developing AI model design and validation methodologies, establishing data lifecycle governance and integrating AI oversight into existing organisational frameworks such as customer service, consent management and product approval. We also provide practical tools like AI approval checklists to streamline governance processes.

02

AI model review services

Our experts assist in detailed model design reviews, and assessing data quality, inputs, outputs and algorithms. We conduct thorough code reviews to detect risks like backdoors or trojans and ensure models comply with regulatory standards. We also carry out AI use case-specific risk assessments and control testing to identify and mitigate potential vulnerabilities early in deployment.

03

AI security services

PwC helps fortify your AI environment by tailoring your cloud security posture and access controls for AI workloads and related services. Our services extend to patch management, vulnerability handling and preparation of incident response readiness for AI-related threats. We also offer red teaming exercises that simulate evolving cyberthreats, ensuring your AI systems remain resilient.

04

AI lifecycle review – design to decommission

We support the second and third lines of defence in governance, assisting with AI risk management integration, security assessments and model governance reviews. Our engagements include comprehensive AI use case reviews that evaluate risks, controls and operational impacts – all aimed at ensuring safe and compliant AI deployments.

AI governance managed services

PwC offers end-to-end managed services to support your AI governance journey through:

Development and implementation:

Building and maintenance of responsible AI governance frameworks

Implementation support:

Embedding AI governance practices into existing enterprise risk, compliance and operational frameworks

Risk assessment:

Continuous evaluation of AI model risks throughout the development lifecycle

Security monitoring and incident response:

Ongoing surveillance and swift management of AI-related security events

Capacity building:

Tailored training programmes for boards, C-suite and all LOD to foster governance maturity

By partnering with PwC, FIs will be empowered to navigate the complexities and risks of AI adoption with confidence. Our comprehensive approach ensures that AI initiatives are not only innovative but also aligned with regulatory standards and ethical expectations.

Our deep expertise across technology, risk and compliance will you to build resilient AI ecosystems that foster trust among stakeholders, enhance operational effectiveness and drive sustainable growth.

Together, we help you unlock the transformative potential of AI responsibly – so you can gain lasting business value from emerging technologies while safeguarding your organisation's reputation and future readiness.

About PwC

We help you build trust so you can boldly reinvent

At PwC, we help clients build trust and reinvent so they can turn complexity into competitive advantage. We're a tech-forward, people-empowered network with more than 370,000 people in 149 countries. Across assurance, tax and legal, deals and consulting we help build, accelerate and sustain momentum. Find out more at www.pwc.com.

PwC refers to the PwC network and/or one or more of its member firms, each of which is a separate legal entity. Please see www.pwc.com/structure for further details.

Contact us

Rounak Shah

Partner, Governance, Risk and Compliance
PwC India
rounak.shah@pwc.com

Shreyans Dudheriya

Partner, Risk Analytics
PwC India
shreyans.dudheriya@pwc.com

Amol Bhatt

Partner, Cybersecurity
PwC India
amol.bhat@pwc.com

Narmada Viswanathan

Director, Governance, Risk and Compliance
PwC India
narmada.viswanathan@pwc.com

Authors

Priyanjali Moulik

Vaishnavi Thiruvenkadam

KA/October 2025 – M&C 48923