AI in Financial Services: An RBI-Aligned Roadmap for Security Leaders

Applies to RBI-Regulated Entities - Commercial Banks · Small Finance & Payments Banks · NBFCs · ARCs / AIFIs · Non-Bank Payment System Operators

> AI systems do things they are not supposed to. In the last 12 months, **RBI and CERT-In have made it your job to prove they don't** - with evidence, not assertions. Here is what they now expect, and where the real gaps are.

THE MANDATE HAS LANDED - 3 INSTRUMENTS IN ~12 MONTHS

RBI FREE-AI Framework

AUG 2025 · PRINCIPLES + ROADMAP

7 Sutras · 6 Pillars · **26 recommendations** spanning AI adoption and risk mitigation.

RBI AI-ACT&RS Advisory

APR 2026 BANKS · JUN 2026 PSOS

Clause 4 - defend the tech stack. **Clause 5.1–5.13** - govern your own AI.

CERT-In - Frontier AI + OEM Guidelines

APR–JUN 2026 · DIRECTIONS

AI-accelerated threat defence, VAPT/BAS, accelerated patching, 6-hour incident reporting.

THE THROUGH-LINE IN ALL THREE

- Trust is the foundation
- Human oversight, always
- Accountability stays with you (not model)
- Evidence, not assertions

THE AI IMPLEMENTATION ROADMAP - 5 PHASES A SUPERVISOR WILL INSPECT

PHASE 0 Govern

Own the risk before AI touches production

Board-approved AI policy (Rec 14) · risk classification · AI inventory (Rec 23) · data lifecycle + DPDP (Rec 15)

PHASE 1 Build & Test

Prove the AI is safe before it ships

Red-teaming across lifecycle (Rec 20) · product approval (Rec 17) · threat controls (5.5) · VAPT/BAS/SDL

PHASE 2 Deploy & Enforce

Control what the AI can access & do

AI agents & privileged access (5.10) · access control (5.4) · human oversight (Rec 16) · shadow-AI leakage (4.6)

PHASE 3 Monitor & Respond

See it, contain it, report it

Logging & forensics (5.8) · output validation (5.6) · incident reporting (Rec 22) · BCP for AI (Rec 21)

CONTINUOUS Assure & Disclose

Sustain trust; answer regulator, board & buyers

Risk-based AI audit (Rec 24) · disclosures (Rec 25) · sector repository (Rec 23) · compliance toolkit (Rec 26)

THE READINESS GAP - WHY THIS IS URGENT, NOT THEORETICAL

~33%
of REs have any board-level AI oversight

~25%
have a formal AI-incident process

18%
keep AI audit logs

14%
do real-time AI performance monitoring

Source: RBI FREE-AI survey of 600+ regulated entities. The gap between policy on paper and control in production is the exposure.

THE GAP YOUR EXISTING STACK CANNOT CLOSE

Your IAM/PAM, WAF and API gateways, DLP/CASB, SIEM and EDR are built for **deterministic** software. None can catch a **legitimately-authorised AI agent, driven by untrusted input, doing something inside its permissions but outside your policy** - the dominant 2026 failure mode (prompt injection, the "lethal trifecta", agentic data exfiltration).

IAM / PAM	WAF / API gateway	DLP / CASB	SIEM / EDR
identity authenticated, in scope	request well-formed	sanctioned tool & channel	no known-bad signature
WAVED THROUGH	WAVED THROUGH	WAVED THROUGH	WAVED THROUGH

Every layer says "yes." None can see the AI's **intent vs your policy**.

Trampolyne makes sure your AI remains within intended bounds - and proves it - from pre-deployment testing, through runtime, to audit-ready evidence.

COVERAGE MAP - TRAMPOLYNE AI PRODUCTS × ROADMAP PHASE

PRODUCT	0 · GOVERN	1 · BUILD & TEST	2 · DEPLOY & ENFORCE	3 · MONITOR & RESPOND	CONTINUOUS · ASSURE
AI Red-Teaming		●			●
Shadow AI Detection & Runtime Control			●	●	○
Enterprise AI Security & Runtime Control			●	●	○
AI Compliance Assistant				○	●

● core control & evidence Trampolyne AI provides · ○ supporting input. **Each RBI phase also demands policies, documentation and processes the RE owns** — Trampolyne AI supplies the technical controls and audit evidence that make them credible, not the entire pillar. Phase 0 (Govern) stays fully with the RE; we feed its inventory and risk classification.

WHAT EACH PRODUCT SOLVES - MAPPED TO RBI / CERT-IN

AI Red-Teaming ↗ AWS MARKETPLACE

Continuous adversarial testing - scheduled cadence + on every material change. 27+ attack classes (prompt injection, jailbreak, tenant-isolation bypass, MCP/tool-call hijack, exfiltration). Working exploits + fixes, mapped to OWASP LLM Top 10 & MITRE ATLAS. No code changes.

Supports: Rec 20 Red-Teaming · Rec 17 Product Approval · Clause 5.5 · CERT-In VAPT/BAS/SDL · Rec 8 liability

Enterprise AI Security & Runtime Control ↗

Sits inline, evaluates every request **before the model or agent acts** - RBAC/ABAC/PBAC/NGAC enforcement in milliseconds. Allow / block / modify / redact. A deterministic guardrail over a non-deterministic model. No model rewrites.

Supports: Clause 5.10 Agents & Privileged Access · Clause 5.4 Access Control · Rec 16 Human Oversight · Clause 5.8 Logging

Shadow AI Detection & Runtime Control ↗

Sits between employees and every public AI tool (ChatGPT, Claude, Gemini, Copilot, APIs, MCP). Classifies data by type / provenance / role, blocks sensitive data before it leaves, logs every interaction. Exception workflow preserves productivity.

Supports: Clause 4.6 Public-tool leakage · Clause 5.4 Data Protection · Rec 15 Data Lifecycle · DPDP Act

AI Compliance Assistant ↗

Turns enforcement + red-team + runtime evidence into review-ready answers. Continuous per-decision audit trail; checks alignment; packages evidence for regulators, board and enterprise-buyer questionnaires.

Supports: Rec 24 Audit · Rec 25 Disclosures · Rec 22 Incident Reporting · Clause 5.8 Forensic Readiness

Outcomes, not threat counts.
We stay to help you turn outputs into evidence.

We're your team.
Partners take us into their enterprise-client CISO meetings.

Enforcement + proof.
Human oversight, accountability with the RE, evidence on demand.

See if we're a fit - a 30-minute call is enough to map your AI span to the roadmap. Book a fit call →

AI Red-Teaming · Enterprise Runtime Control · Shadow AI Control · AI Compliance Assistant · How it works

Sources: RBI FREE-AI Committee Report (Aug 2025) · RBI AI-Accelerated Cyber Threats & Related Safeguards Advisory - Banks 27 Apr 2026, PSOs 1 Jun 2026 (Dept. of Supervision, CSITE) · CERT-In "Defending Against Frontier AI Driven Cyber Risks" (26 Apr 2026) & OEM/Technology-Provider Guidelines (10 Jun 2026). Anchors: DPDP Act · ISO/IEC 42001 · NIST AI RMF · OWASP LLM Top 10 · MITRE ATLAS.

Positioning collateral, not legal advice. Trampolyne AI helps you align with regulatory expectations; it does not issue certifications.