SISA

certin
Enhancing Cyber Security in India

CSIRT-Fin

DIGITAL THREAT REPORT 2025-26

For the Banking Financial Services and Insurance (BFSI) Sector

A collaborative effort of SISA, CERT-In & CSIRT-Fin.

Table of Contents

Point of View by Secretary MeitY

Finance sector, not just in India but across the globe, is witnessing a fundamental re-platforming of how financial services are designed, delivered and operated. While digital transformation, cloud adoption, artificial intelligence, and real-time payment systems have streamlined processes and enhanced customer experience, they have also dissolved traditional boundaries and expanded the threat surface into a distributed mesh across partners, platforms, models and infrastructure layers.

As the sector continues to adopt these emerging technologies, cyberattacks are growing faster, more sophisticated and increasingly systemic. Adversaries are no longer merely breaching perimeters; they are manipulating trust chains across biometric identity systems, AI-driven decisioning engines, real-time payment rails and third-party ecosystems. A cyberattack on a financial institution today can have disastrous consequences for customer confidence and financial stability. Given the deep interconnections among institutions, such attacks can cascade across the ecosystem, resulting in exponential and often irreversible losses.

Thus, efficient detection, response and rapid recovery from cyber incidents are essential to limit financial stability risks. Considering the interconnectedness of financial entities, regulators, technology partners and customers, and the borderless nature of cyber incidents, cyber risk is no longer limited to an entity's owned systems. Entities that were not the primary target may also be affected. Hence, sector and national-level coordination become far more important.

CERT-In and CSIRT-Fin play a critical role by coordinating with global and national financial organisations, regulators, national CERTs and government stakeholders in rendering timely cyber incident response. Through continuous monitoring, threat intelligence sharing and forensic analysis, these agencies help safeguard the resilience of the financial ecosystem.

CERT-In and CSIRT-Fin have noticed a clear acceleration in the pace and sophistication of cyberattacks targeting the financial industry. Malicious actors are using advanced tactics, techniques and procedures to get beyond traditional defences. The landscape is changing with generative artificial intelligence, large language models and machine-speed payment systems. Adversaries are exploiting gaps between systems, institutions and workflows. Notable findings also show that some firms struggle to translate compliance into real-world resilience, where controls that pass periodic assessment fail under adversarial pressure.

The report prepared by CERT-In, CSIRT-Fin and SISA offers an in-depth analysis of the evolving cyber threat landscape, focusing on the tactics, techniques and procedures employed by threat actors targeting the BFSI sector. It provides a comprehensive overview of how malicious actors exploit vulnerabilities across identity, AI, APIs, payment logic and supply chains. The report outlines practical recommendations across people, process and technology, including controls to strengthen defences, close the compliance-security translation gap, and build continuous assurance.

The report's timely insights will help organizations safeguard their assets by enhancing their security postures and preparing for breaches before they occur. It also promotes sector-wide collaboration, allowing institutions to learn from each other's experiences and strengthen the BFSI sector as a whole through a collective response to emerging cyber threats.

S. Krishnan

I.A.S. Secretary,
Ministry of Electronics & Information Technology, (MeitY),
Government of India

Leadership Reflections

When we released the previous edition of the Digital Threat Report, we did so with a clear belief that the payments ecosystem was approaching a turning point. Threats were becoming more sophisticated, more interconnected, and more capable of exploiting the speed and scale of digital finance.

At the time, many of the highlighted risks such as social engineering, credential theft, supply chain compromise and cloud exploitation, were still viewed as emerging or episodic.

Today, they are operational realities.

Over the last year the distance between innovation and exploitation has narrowed dramatically. Threats that once took years to mature are now operationalized in months, sometimes weeks. Attackers are moving faster, adapting faster, and monetizing faster than many institutions were built to respond to.

For the BFSI ecosystem, this matters profoundly. The BFSI industry is built on trust. The trust the person initiating a transaction is genuine. The trust that the system processing it is acting as intended. The trust that money will move securely, instantly, and irreversibly through a network of institutions that depend on one another. When trust is weakened, the impact is not limited to one company or one breach. It can ripple across customers, partners, markets, and economies.

This is why I believe cybersecurity can no longer be viewed as a technical control function operating at the edge of the business. It must become central to how institutions grow, innovate, and lead.

At SISA, our journey has been shaped by forensics. We have spent years studying what happens when systems fail, when assumptions break, and when trust is tested in the real world. Those experiences have taught us clarity and a simple truth: every breach leaves behind a lesson. If we are willing to learn fast enough, those lessons can turn disruption into foresight. Much of that thinking is reflected in the perspectives and recommendations contained in this report.

My hope is that this report helps boards, CEOs, CISOs, regulators, and technology leaders make that shift, from defending yesterday's perimeter to building tomorrow's trust architecture. The future of payments will belong to institutions that can move fast without losing trust, innovate boldly without losing control, and grow digitally without losing resilience.

That is the challenge before us. It is also the opportunity.

"The future of payments will belong to institutions that can move fast without losing trust, innovate boldly without losing control, and grow digitally without losing resilience."

Dharshan Shanthamurthy

Founder & CEO, SISA

Executive Summary

The BFSI ecosystem is entering a new era of cyber risk. Threats that once moved through a relatively predictable lifecycle, from research to experimentation to scaled exploitation, are now accelerating directly into operational impact. The time between emergence and monetization is shrinking. For Banking, Financial Services, and Insurance (BFSI) institutions, where transactions are real-time, irreversible, highly regulated, and deeply interconnected, this compression creates disproportionate consequences.

Cybersecurity risk in financial services is no longer limited to data theft or isolated breaches. It now extends to transaction integrity, customer trust, third-party dependencies, decision systems, operational continuity, and confidence in the digital infrastructure that underpins economic activity itself. Many of the most consequential attacks no longer resemble traditional intrusions. They appear as legitimate sessions, approved payments, manipulated workflows, compromised vendors, or normal user behavior until damage is already underway.

The Digital Threat Report 2025-26 examines how the threat landscape evolved through 2025-26 and what that signals for the years ahead. A central finding of this year's report is that six of the seven forward-looking predictions made in the previous edition have already reached full-scale realization. This is not merely a validation of forecasting accuracy. It is evidence of acceleration. Adversaries are innovating faster, scaling faster, and exploiting trust faster than many institutional control environments were designed to withstand.

Based on observed casework, market developments, and broader threat intelligence patterns, the report identifies three structural shifts now reshaping cyber threats across the payments ecosystem:

01

AI & Human Deception

where deepfakes, adversarial AI, and machine-scale impersonation are eroding trust in human judgment.

02

Software & Systems

where supply chain compromise, logic abuse, and AI-enabled attacks are weakening trust in how systems are built and behave.

03

Infrastructure & Economy

where cyber risk increasingly impacts financial continuity, critical infrastructure, and long-term cryptographic trust.

Looking ahead, the report outlines the forces most likely to shape 2026 and beyond. Collectively, these trends point to a threat environment that will become more adaptive, less visible, and increasingly embedded within everyday business operations.

The next phase of cyber threats is expected to center on attacks against trust, customer identity, transaction integrity, intelligent decision systems, partner ecosystems, and the autonomous technologies institutions are rapidly adopting. As a result, many future compromises may not resemble traditional breaches at all, but rather legitimate activity manipulated for malicious outcomes.

In essence the report aims to:

- Provide an executive-level assessment of how cyber threats are evolving across the global BFSI and payments landscape
- Explain why traditional threat maturity models are breaking down and why speed of realization now matters as much as sophistication
- Translate fragmented threat activity from observed casework into structural themes leaders can act upon
- Help boards, CISOs, leaders, technology teams, and regulators understand where trust is now most vulnerable
- Identify the threat trends most likely to influence security investment and operating models through 2026 and beyond
- Offer a practical 18-month roadmap along three horizons mapped against the six BFSI Operating Layers

The central conclusion is clear. Cybersecurity in BFSI must evolve into a continuous business discipline focused on preserving trust, ensuring operational continuity, and enabling secure growth in an increasingly hostile digital economy.

The report is based on a synthesis of various sources, including:

Direct Observations from SISA's DFIR Investigations

Drawing on unique cases and insights gained from digital forensics and incident response (DFIR) projects handled by SISA over the past year.

Observations of CSIRT-Fin and CERT-In

Based on a comprehensive analysis of cyber incidents affecting the Indian financial system, with actionable recommendations for enhancing cyber maturity, data protection, backup strategies, and recovery measures.

Cybersecurity Reports and Data Pointers

Incorporating findings from vulnerability databases and observed trends in malware and exploit usage.

Research and Analysis

Leveraging research on AI's impact on cybersecurity, including adversarial machine learning, deepfake technology, and malicious use of large language models.

Introduction

The BFSI Threat Surface Reimagined

A shift is underway in BFSI today and it is not a digitisation of existing processes. It is a fundamental re-platforming of how financial services are designed, delivered, and operated. Identity is no longer a gate but a continuous authentication plane spanning biometrics, behavioural signals, identity wallets, and cross-system session tokens.

Credit, suspicious transactions, and onboarding decisions are no longer human-reviewed but model outputs, executed at machine speed, across pipelines where the training data, feature inputs, and inference environment are all potential attack surfaces. Payments are instant, irreversible, and increasingly initiated by non-human actors operating within automated financial workflows.

The infrastructure underpinning all of this has shifted simultaneously. Monolithic core banking systems have given way to cloud-native, containerised, event-driven architectures. DevSecOps pipelines, programmable financial logic through smart contracts and tokenisation frameworks, ISO 20022-aligned streaming data architectures, and open banking API ecosystems have collectively dissolved the boundary between what is "inside" the institution and what is not.

The result is a threat surface that is no longer a perimeter. It is a distributed, continuously shifting mesh of relationships, across partners, platforms, models, APIs, and infrastructure layers - none of which an institution fully owns or controls.

What This Breaks in Security Models

Conventional security architectures were designed for a world where control was centralised and trust was bounded. That design does not hold in the current operating environment.

When user journeys span multiple partner platforms and embedded ecosystems, session context and security controls fragment across entities that operate under different risk models. When identity becomes the primary control plane for access and transactions, built on biometrics, continuous authentication signals, and cross-system token chains - a single identity compromise no longer affects one account. It provides broad, persistent, cross-system access.

When AI models are making consequential decisions in real time, the attack surface shifts from the transaction itself to the inputs that shape the model's behaviour - training pipelines, feature stores, inference-time prompt context, and adversarial input construction.

Post-facto detection, the backbone of traditional SOC operations, becomes structurally ineffective when transactions are irreversible and execute faster than any detection pipeline can respond.

Compliance monitoring that relies on control signals becomes weaponisable because attackers can suppress logs, manipulate alert thresholds, and maintain the appearance of a clean posture while operating inside the environment.

How Attackers Have Adapted

The adversarial community has not missed this transition. Attack patterns in BFSI have evolved in direct correspondence with the architectural shifts underway.

Modern financial attacks are moving from direct compromise to trust-chain manipulation across biometric onboarding, partner apps, AI decisioning, real-time payments, APIs, programmable finance, and third-party ecosystems. Attackers are exploiting the gaps between systems, institutions, and workflows, where no single control owner has full visibility.

The result is a new threat model where a breach is no longer isolated to credentials or transactions, but embedded across identity, AI, APIs, payment logic, and supply chains.

The convergence of real-time irreversibility, AI-driven decisioning, programmable financial logic, continuous identity planes, and ecosystem-dependent operations creates a threat environment where speed, complexity, and interdependency all compound simultaneously. Regulatory compliance frameworks are catching up, but they are catching up to an architecture that is itself still evolving.

The lag between structural change and security model adaptation is the window that sophisticated threat actors are actively exploiting. The table below maps these shifts, showing how changes in operating models across the BFSI operating layers are redefining security assumptions and shaping how attacks are executed.

Structural Shifts in BFSI and Emerging Security Risks

BFSI Operating Layers	Shift in Operating Model	Resulting Operating Reality	What This Breaks in Security Models
Customer & Identity Layer	Banking moves from controlled channels to partner apps, platforms, and ecosystems	Loss of control over user interface and session context	Security controls no longer consistently enforced across journeys
	Shift from static KYC to biometrics, passkeys, identity wallets, continuous authentication	Identity becomes the primary control plane for access and transactions	Identity compromise provides broad, cross-system access
	AI moves from support to execution across onboarding, credit and servicing	Trust shifts from human decisions to model outputs	Model behavior becomes a new attack surface
Transaction & Financial Logic Layer	Payments, credit, and onboarding move from batch to instant execution	No latency for detection or recovery	Post-facto detection becomes ineffective
	Financial processes move from systems into code (tokenization, smart contracts, programmable payments)	Logic becomes embedded, automated, and difficult to reverse	Exploitable logic persists at scale
Data & Intelligence Layer	Shift from siloed data to ISO 20022, streaming pipelines, and real-time analytics	Sensitive data continuously exposed across systems	Data exposure moves from storage to flow
	Continuous analytics and AI-driven incidents, credit, and risk scoring	Decisions depend on data integrity and model accuracy	Manipulating inputs alters system outcomes
Platform & Infrastructure Layer	Open banking, API marketplaces, and fintech integrations expand system boundaries	Trust extends beyond institutional control	Security dependency on external systems increases
	Shift to cloud, containers, DevSecOps, and event-driven platforms	Infrastructure becomes dynamic and continuously changing	Traditional control models struggle to keep pace
Ecosystem & Dependency Layer	BFSI increasingly relies on SaaS, fintechs, cloud, and supply chain partners	Risk becomes indirect and systemic	Breaches propagate across interconnected systems
Control & Assurance Layer	Shift from periodic audits to real-time compliance, RegTech, and automated controls	Compliance becomes signal-driven rather than assurance-driven	Systems can appear compliant while compromised

The Acceleration of Cyber Threats in Payments

The proof of the shift is already in motion. In last year's Digital Threat Report seven forward-looking predictions were made about the BFSI threat landscape of which six out of seven have already reached full-scale realization.

This is a signal of acceleration. The traditional lifecycle of cyber threats from research to experimentation to scaled exploitation is breaking down. Threats are no longer following a linear maturity curve from research to exploitation. Instead, they are collapsing directly into real-world impact often at systemic scale.

Even in the one area not fully realized/quantum risk adversaries have already adapted their strategies indicating that threat execution begins before technological maturity is reached.

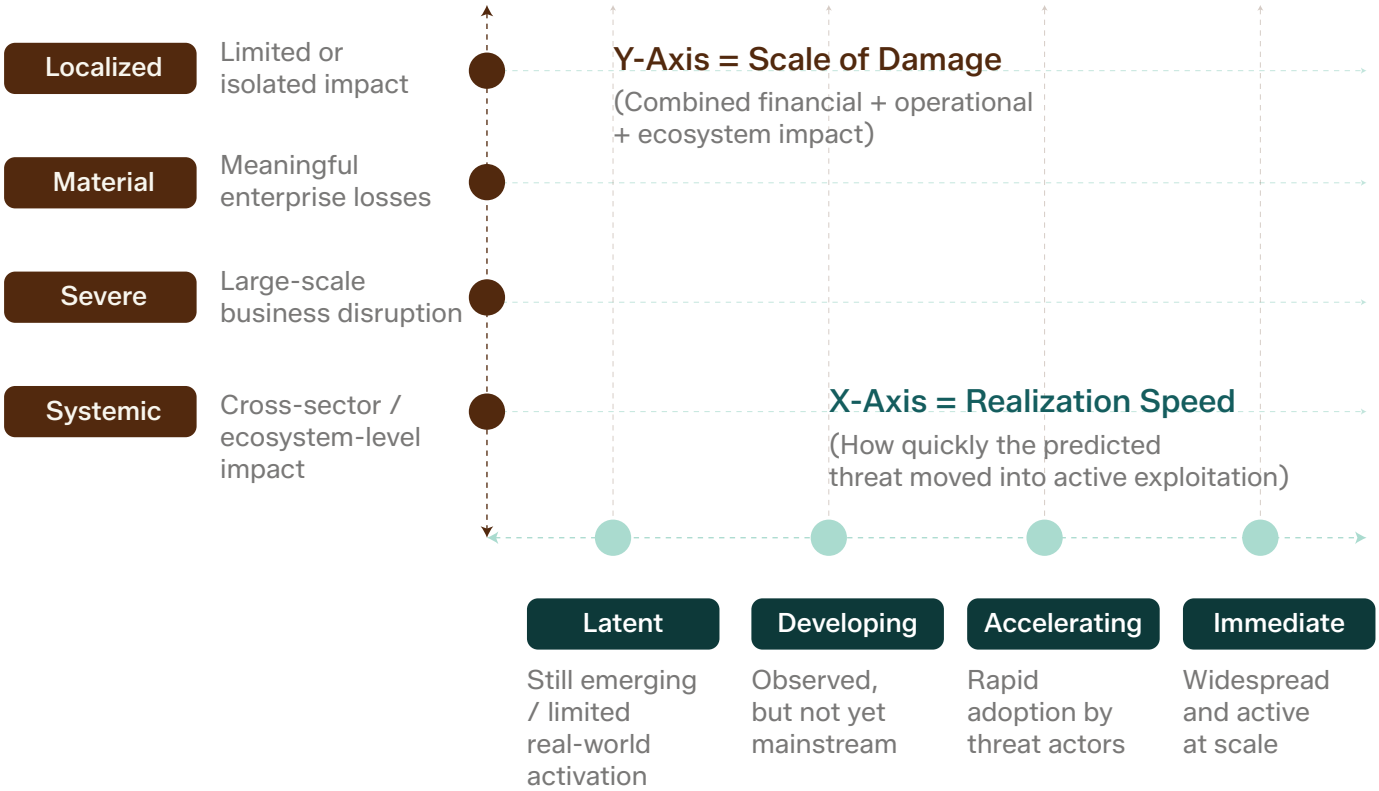
Threat Velocity in Motion: Speed, Scale, and Impact in 2025

The heatmap evaluates specific threat vectors observed this year. The analysis is based on two dimensions: how quickly they are moving from forecast to active exploitation and the scale of damage they create once realized. Viewed independently, the threats vary in pace, severity, and maturity.

2025-26 Threat Realization Matrix: Predictions vs Reality

Scale of Damage	Latent	Developing	Accelerating	Immediate
Systemic	Cryptographic / Quantum Risk			Deepfake / Synthetic Identity Zero-Day Exploit
Severe		Adversarial AI / Model Evasion	Supply Chain / Third-Party Ransomware (Multi-Extortion)	Social Engineering/ BEC Credential Theft/Session Hijack Mobile / UPI Fraud Cloud Misconfiguration Exploit AI-Driven Attacks Crypto Incidents IoT
Material		Application Race Conditions	API / Business Logic Abuse Adversarial LLMs LLM Prompt Hacking	Phishing (Traditional) IDOR / Broken Object Insider Threat Insecure Coding / SDLC Gaps
Localized		Hardware / Sensor Side Channel		

Realization Speed



From Individual Threat Signals to Structural Risk Themes

The threat vectors highlighted above are not isolated events; when assessed collectively, a deeper pattern emerges.

They cluster into three broader structural shifts that are redefining cyber risk for BFSI organizations. What appears as separate incidents on the matrix increasingly reflects interconnected failures in trust across people, software, and critical systems.

Three Emerging Clusters of Cyber Risk

01

AI & Human Deception

- Social Engineering / BEC
- Phishing (Traditional)
- Credential Theft / Session Hijack
- Deepfake / Synthetic
- Adversarial LLMs

02

Software & Systems

- API / Business Logic Abuse
- IDOR / Broken Object Auth
- Application Race Conditions
- Insider Threat
- Supply Chain / Third-Party
- Cloud Misconfiguration Exploit
- Insecure Coding / SDLC Gaps
- LLM Prompt Hacking
- Adversarial AI / Model Evasion
- AI-Driven Attacks
- Zero-Day Exploit

03

Infrastructure & Economy

- Mobile / UPI Fraud
- Ransomware (Multi-Extortion)
- Cryptographic / Quantum
- Crypto Incidents
- IoT
- Hardware / Sensor Side-Channel

From Social Engineering to Synthetic Reality

1.6x

India BFSI cyber-attacks vs global average[1]

2x+

Incidents in 4 yrs - 1.4M (2021) → 2.9M (2025)[1]

- Deepfake impersonation is now industrialised - real-time executive video fakes, adversarial LLMs, polymorphic variants that outpace detection
- Social engineering and BEC attacks have intensified, credential theft and session hijacking have become dominant initial access vectors
- Attacks are no longer handcrafted - they are generated, iterated, and deployed at machine speed
- Phishing is now context-aware and indistinguishable from legitimate comms - BFSI is the most targeted sector

Implication: *This marks a structural shift from "deception as a tactic" to "deception as infrastructure". Organizations can no longer rely on identity verification mechanisms rooted in human perception. Trust must move from what is seen or heard to what can be behaviorally validated beyond human interpretation.*

The Collapse of Trust in How Systems Are Built and Behave

263 days

Mean time to identify and contain a data breach[1]

Hours

Zero-day exploit window - compressed from weeks or months

3 layers

Code, AI model supply chain, and runtime - all now compromised

Build-time threats

- Poisoned package repos, malicious CI/CD pipeline actions
- Tampered AI model files & manipulated dependencies - risk now extends into the AI layer
- Attackers prioritising indirect access via vendors, SaaS & cloud integrations
- Insecure SDLC, cloud misconfigurations & secrets exposure - persistently high

Runtime threats

- API misuse, IDOR, OTP reuse, race conditions - gaps standard OWASP reviews miss
- Business logic, auth & transaction flow abuse - only visible under real-world parallel or abnormal execution
- Mobile & UPI fraud via device compromise, token interception, weak backend validation
- LLM prompt injection - now operational, not theoretical; systems act outside intended boundaries

Compounding factor - AI-driven attack execution

- Attackers using the same AI tooling as developers - automating recon, exploit generation & orchestration
- Exploitation now occurs on or before public disclosure of zero-days
- Mid-size organisations most exposed - digitised aggressively, contain slowly
- Insider threats amplified by privilege sprawl and trusted-access misuse

Implication: *Systems can no longer be trusted in* ***how they are built or how they behave at runtime.*** *The deterministic assumption — build it securely and it behaves predictably — has been broken. This demands a shift from point-in-time validation to continuous verification of both system integrity and system behavior.*

From Digital Risk to Systemic Exposure

Immediate

Monetization after compromise - impact lands before detection

Fleet-wide

Single IoT/firmware compromise now propagates across thousands of devices

Already begun

Quantum threat lifecycle - *harvest-now, decrypt-later* strategies active

Financial incidents layer

- Crypto now a fully operational financial layer - anonymity, rapid monetisation, frictionless cross-border movement
- Ransomware more precise - backup destruction routinely precedes encryption
- No delay between control and cash - impact is immediate and often complete before detection

Physical infrastructure

- IoT & embedded systems - persistent entry points into larger environments
- Firmware compromise & OTA update manipulation - single point triggers fleet-wide breach
- Digital to physical risk/ where systems control physical processes, cyber becomes operational failure
- Hardware & sensor attacks/ rapidly emerging in mobile and embedded environments

Cryptographic trust

- Post-quantum standards now finalised - transition window is open
- Harvest now decrypt later/ encrypted data stolen today will be readable once quantum capability matures
- Long-lived financial data at highest risk - the threat is real-time even if decryption is deferred

The final stage of attack impact has also evolved. Once control is established, monetization is immediate. There is little delay between compromise and financial impact. In many cases, by the time an attack is detected, the impact has already occurred.

***Implication:** This represents a transition from enterprise-level cybersecurity to **systemic risk management.** The challenge is no longer just to defend against breaches, but to ensure continuity of trust across systems that underpin economic activity itself.*

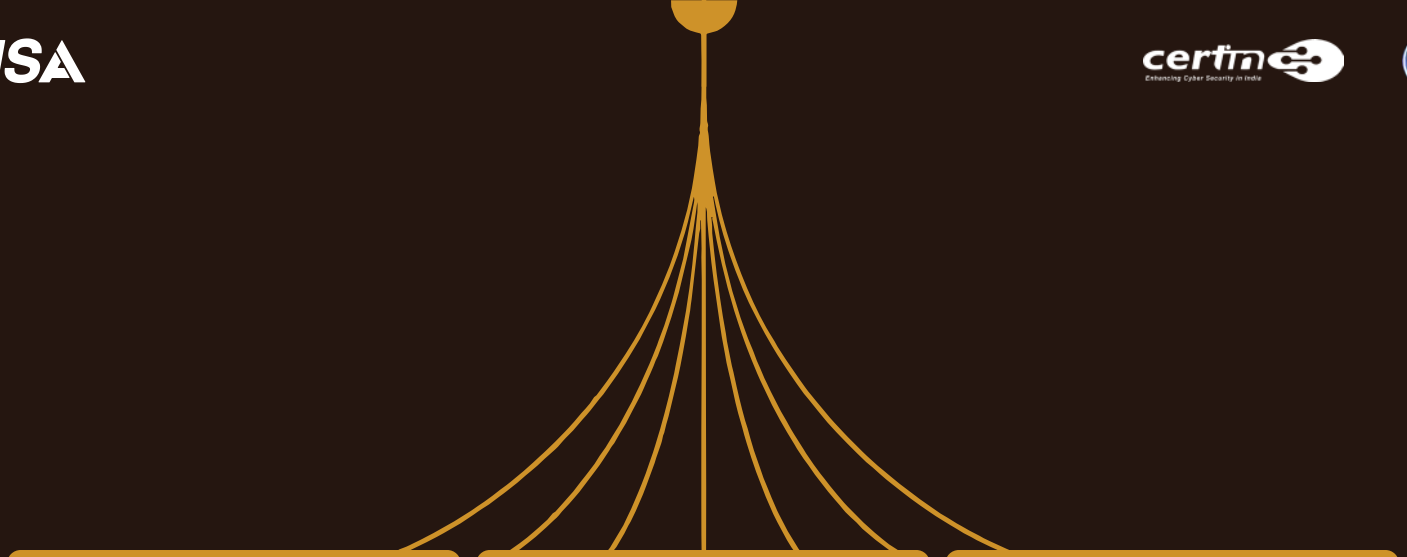
The Compliance-Security Translation Gap

Observations from compliance assessment, security testing, cyber defence, and forensic engagements across the BFSI ecosystem in 2025-26.

The same control environments that successfully complete attestation cycles are, in a meaningful number of cases, the same environments that subsequently require forensic investigation. Attestation proves controls exist at a point in time. It doesn't necessarily prove control survivability under adversarial pressure. Compliance frameworks are baseline assurance mechanisms: a defensible floor of control coverage at a point in time.

Compliances effectively ask whether the control existed during the assessment window while adversarial assurance asks whether the control holds when identity, runtime, cloud, and application layers are actively abused.

The translation gap lies in how control environments evolve between attestation cycles. Three patterns recur with sufficient consistency to be considered structural.



01	02	03
Implementation Drift: When the control exists, but operations have moved away from intent	**Scope Boundary Failures:** When the control works, but the attacker operates outside its design boundary	**Framework Evolution Lag:** When technology moves faster than the frameworks governing it
Observed in		
Machine and AI identities with excessive privilege, secrets spread across code and CI/CD pipelines, internal traffic left unencrypted, sensitive payment data persisting in runtime memory, and cryptographic keys weakened at the point of generation.	TDE protecting data at rest but not data in use, VPN securing the tunnel but not the endpoint, MFA validating access but not preventing session hijack, key rotation evidencing compliance while DEK risks persist, and privileged reviews excluding system or service accounts.	Privileged containers, hostPath mounts, tenant-side cloud misconfigurations, build-time security that does not reflect runtime risk, AI model abuse, dependency poisoning, tampered model weights, and machine identities operating with administrator-equivalent privilege.
What it means		
The control may pass assessment, but its real-world implementation no longer delivers the security outcome the framework intended.	The environment may be compliant, but the protection envelope is narrower than the real attack surface.	Organizations may be compliant with current frameworks while still exposed to architecture - level risks that have not yet been fully absorbed into mainstream control requirements.
Remediation pathway		
Continuous control validation, stronger operational discipline, secrets governance, runtime testing, and identity lifecycle reviews beyond human users.	Layered control design, adversarial assurance, red-team validation, behavioural monitoring, ZTNA/VDI models, and session-level risk detection.	Forward-map existing control intent to emerging architectures, extend assurance beyond attestation scope, and continuously validate cloud, container, AI, DevSecOps, and machine identity risks.

Control Intent vs Operational Reality

Control area	What attestation demonstrates	Where it does not hold under adversarial conditions	Evidence needed for real assurance
Encryption of data	Enabled keys under custody; rotation evidenced	Data-in-use unprotected; plain text retrieved post-authentication	Proof of memory-level protection, runtime encryption controls, session-level monitoring, and validated inability to retrieve sensitive plain text during active sessions
Key rotation	KEK rotation evidenced; audit trail clean	Encryption keys unmanaged; auto-open wallets nullify rotation	Demonstrated cryptographic separation, enforced HSM controls, tested key revocation impact, and validation that compromised systems cannot transparently access rotated keys
Remote access (VPN)	Encrypted tunnel; authenticated access	Endpoint ungoverned; ZTNA and VDI address; VPN does not	Device posture enforcement, endpoint telemetry validation, session-level anomaly detection, and evidence of access restriction from unmanaged or compromised endpoints
Privileged accounts	User account reviews completed; rotation enforced	Service/application accounts and system accounts excluded; high forensic incidence	Complete inventory of privileged non-human identities, behavioral monitoring, credential usage tracing, just-in-time privilege controls, and forensic attribution capability
MFA	Deployed at access boundaries	Session unprotected against replay, hijack, AITM, consent phishing	Phishing-resistant MFA validation, session token protection, replay attack testing, and detection telemetry for anomalous session behavior
Cloud posture	CSP attestation, image signing; change control	Tenant-side IAM drift, exposed storage, federation misconfiguration	Continuous IAM drift detection, exposed asset monitoring, federated trust validation, and adversarial testing of tenant misconfiguration exposure
Container & pipeline	Build-time signing; SCA; provenance	Runtime drift, post-deployment escalation largely unmonitored	Runtime integrity monitoring, container behavior analytics, privilege escalation detection, and evidence of runtime policy enforcement post-deployment
AI/ML workloads	Encryption-at-rest verified	Model poisoning, adversarial probing, inference-time exposure	Adversarial testing results, model integrity validation, inference monitoring, prompt abuse detection, and controls for training data provenance
Software supply chain	Vendor due diligence; SBOM where required	Dependency poisoning, maintainer compromise, transitive risk	Continuous dependency monitoring, integrity verification of updates, behavioral analysis of third-party components, and validated response workflows for supply chain compromise scenarios

Closing The Translation Gap

The following should be treated as minimum continuous assurance expectations for high-risk BFSI environments, not optional maturity initiatives:

01

Run quarterly post-attestation adversarial validation on three high-criticality controls.

02

Map consequential forensic findings - institutional or peer, against the current control inventory.

03

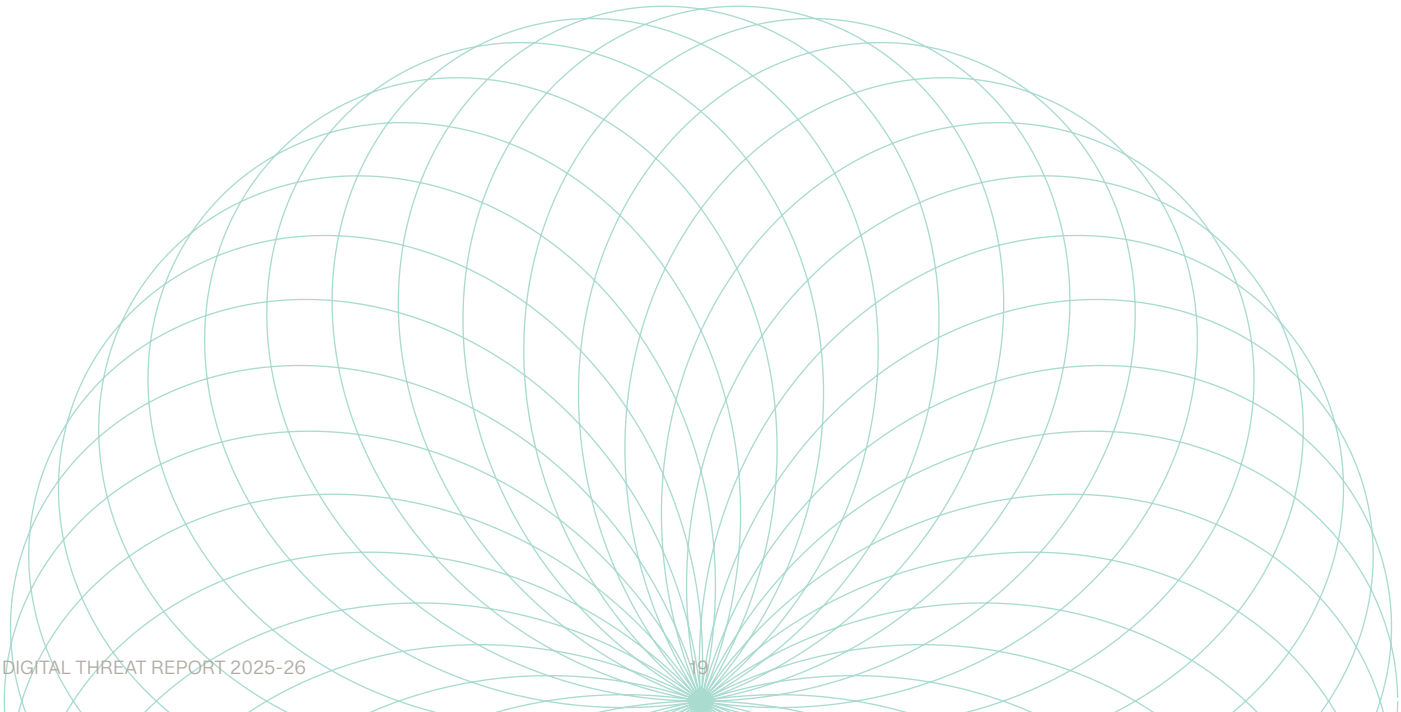
Build a cross-framework control dashboard for the top three concurrent frameworks.

04

Convert highest-frequency repeat findings from perpetual exception requests into engineering backlog.

05

Treat attestation as one signal within a continuous assurance posture, not as the assurance itself.



AI Asymmetry: How Frontier Models Are Reshaping the Cyber Threat Landscape for BFSI

BFSI has never had the luxury of obscurity. Among the most targeted sectors on the planet, its defences have evolved accordingly: layered architectures, mature threat intelligence functions, and significant regulatory oversight. That hard won defensive premium is now being eroded.

Over the past year, frontier AI models capable of autonomous vulnerability discovery, exploit chaining, and weaponisation have moved from theoretical concern to documented reality.

In November 2025, Anthropic disclosed GTG-1002, the first reported large scale AI orchestrated cyber espionage campaign, a Chinese state linked operation in which the AI executed 80 to 90 percent of the operation independently against around 30 global targets, including financial institutions.

In April 2026, the disclosure of Claude Mythos Preview, a frontier model that has identified more than 23,000 potential vulnerabilities, of which over a thousand confirmed findings are rated high or critical severity, confirmed the trajectory. This is a structural shift, not an incremental one. What once demanded specialist teams over weeks can now be executed at machine speed by actors with far more modest capability and budget.

This is the essence of AI asymmetry: offensive capability is scaling faster than the regulatory, defensive, and operational frameworks designed to contain it.[2,3]

The Industrialisation of Cyber Exploitation

The most consequential change is not that attacks are more sophisticated. It is that they are becoming systematic. AI moves adversaries from manual, specialist operations to repeatable, automated attack pipelines, with vulnerability discovery, exploit chaining, and payload delivery industrialised more like software manufacturing than tradecraft. GTG-1002 demonstrated this directly: the agent scaled, firing thousands of requests per second across roughly 30 organizations.[4]

Two shifts define what follows. The first is capability proliferation, and it is already underway rather than approaching. AI has shifted attacker economics: automation makes targeted campaigns cheaper to run, which increases the likelihood any given organization is probed. Ransomware groups and financially motivated actors that once depended on purchasing exploits from broker ecosystems can increasingly generate them independently, at marginal cost, against bespoke targets.

The second shift is reach. Frontier models have demonstrated autonomous vulnerability discovery across operating systems, browsers, and, directly relevant to digital asset operations, smart contracts, where frontier models produced working attacks against 207 of 405 historical smart contract exploits, totalling $550 million in simulated stolen funds. For BFSI this is acute: banking environments carry decades of legacy systems, SaaS dependencies with opaque CVE histories, and operational technology touching payment and settlement infrastructure/ precisely the surface these models systematically dissect.[5,6]

Critically, these capabilities are not confined to the most advanced models. The capability appeared in smaller open models first, at a fraction of the cost. Even as more capable successors arrive, and they will, tooling already in circulation is sufficient to cause significant harm, and it is already in adversarial hands.[7]

The 12 Month Horizon

The trajectory points clearly in one direction. As capability proliferates through open weight releases, API access, and adversarial fine tuning, the barrier to AI augmented attacks continues to fall. Three scenarios, mapped to this report's BFSI Operating Layers, are already in motion.

Parallel, multi vector campaigns become the norm. AI removes human bandwidth as a constraint, and adversaries run simultaneous discovery across network, messaging, clou, and vendor environments without fatigue. Incident response built on sequential attack assumptions is structurally unprepared.

The time to exploit window compresses sharply, touching the Control and Assurance Layer. Independent evaluators note current results are upper bound under scoped conditions rather than fully autonomous end to end, but the direction is unambiguous: attacker velocity rises while institutional patch velocity remains constrained by change management and compliance gates.

The software delivery pipeline becomes both target and battleground, across the Platform and Infrastructure Layer. As adversaries gain the ability to discover vulnerabilities at machine speed, the CI/CD pipeline emerges as a decisive control point in two directions.

A compromised pipeline propagates malicious or vulnerable code into production at the same speed it ships legitimate code, while a well instrumented pipeline with security testing, build provenance, and secrets governance embedded becomes the fastest available route to remediate at scale. The institutions that win this exchange are those that have moved application security from a periodic gate to a continuous capability inside the pipeline itself.

Supply chain and third party risk are systematically targeted, across the Ecosystem and Dependency Layer. Where core banking, payment, and identity infrastructure is heavily vendor dependent, the CVE history of security vendors themselves becomes a strategic liability, not merely an operational concern.

What BFSI Institutions Must Do Now

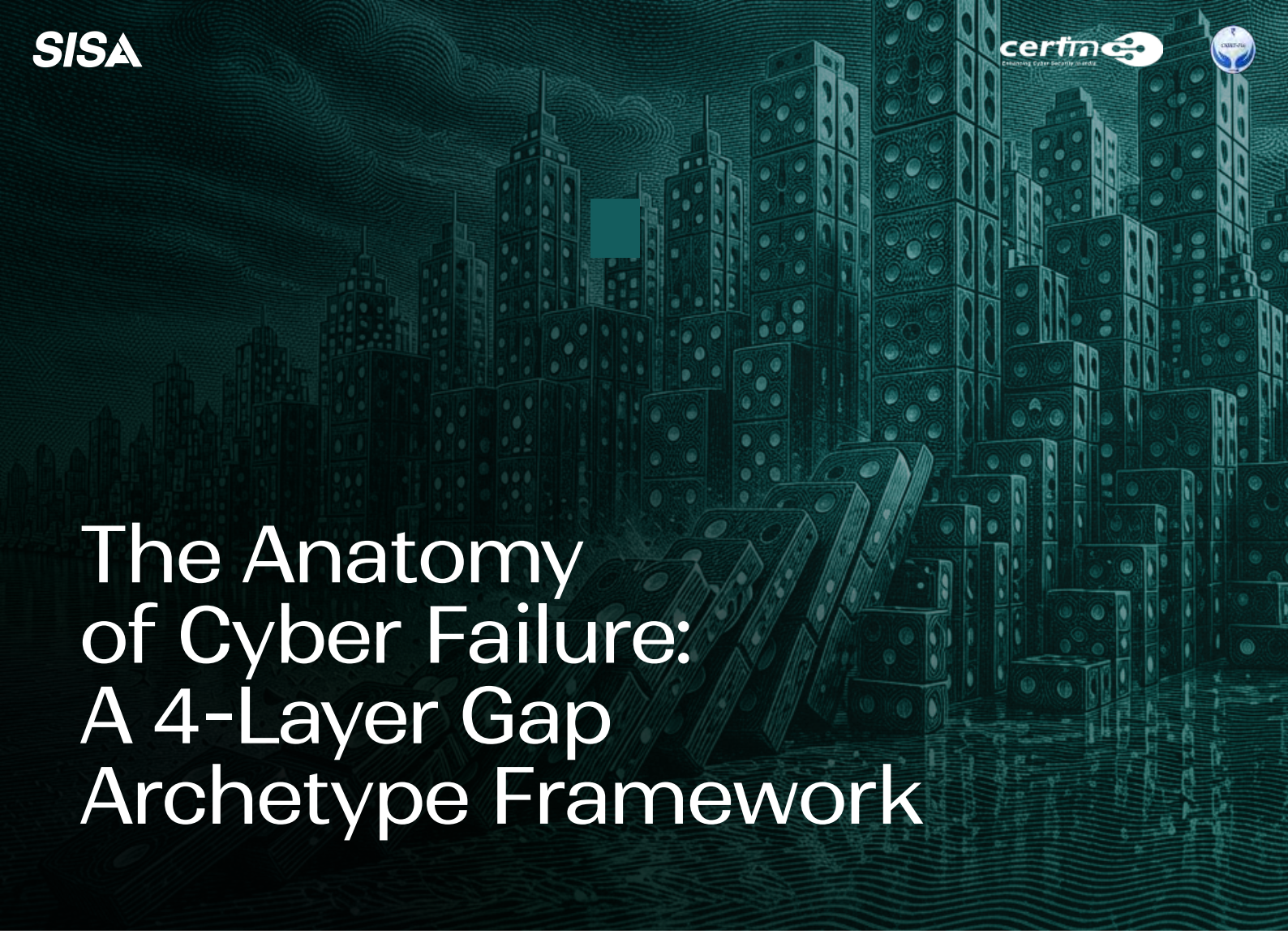
The response maps to the three horizons of this report's Implementation Roadmap.

Immediate/ asset visibility and pipeline hygien, across the Platform and Infrastructure Layer. Institutions cannot prioritise remediation for infrastructure they have not inventoried. A continuously updated inventory covering legacy systems, unmanaged endpoint, accounts without multifactor authentication, and unprotected remote access is the prerequisite for everything that follows. In parallel, secrets must be removed from code, configuration, and CI/CD variable, and static and software composition analysis must run inside the pipeline rather than as periodic reviews. Vulnerability and patch operations must be restructured around automated virtual patching and accelerated remediation.

Over six months/ detection beyond signatures. AI-generated exploits frequently lack prior patterns in threat feeds. Institutions should test whether managed detection can identify anomalous behaviour with no precedent in existing intelligence, and stress test incident response against simultaneous multi vector scenarios.

Structurally, zero trust as operating model rather than roadmap, extended to the pipeline itself. GTG-1002 is instructive: the agent operated from within the perimeter using valid entitlements, and did not need to hack the network because it used the keys it already held. Continuous verification of every user, device, and identity, both human and non human, constrains the lateral movement available to an adversary with an initial foothold. The same principle applies to the build environment: signed commits, tamper-evident pipelines, and build-to-runtime integrity attestation ensure that what runs in production is what the pipeline actually shipped. AI agents themselves are now privileged identities requiring ephemeral, task scoped credentials and continuous monitoring.[8]

AI asymmetry is not a forecast. It is a documented and accelerating reality, and the institutions that restructure detection, patching, identity, and the software delivery pipeline around it now will be the ones still standing when the capability becomes ubiquitous.



The Anatomy of Cyber Failure: A 4-Layer Gap Archetype Framework

Now that we have examined the threat vectors observed this year and the repeat control gaps across BFSI environments, one thing becomes clear: these failures are not isolated.

Cyber breaches are often explained as isolated failures whether it is misconfigurations, missed alerts, or control breakdowns. In reality, they are rarely singular events. They are the result of multiple gaps aligning across systems, from how applications are designed, to how controls are enforced, to how signals are generated, and how response is executed.

This perspective is grounded in real-world evidence observed across hundreds of SISA engagements, including security assessments across complex enterprise environments, red team exercises simulating adversarial behaviour, AI model testing under adversarial conditions, and forensic investigations following real breaches.

Across these engagements, a consistent reality emerges: cyber failures occur when gaps across systems are not connected. Adversaries move seamlessly across boundaries, exploiting gaps between systems, controls, and signals. What appears as isolated weaknesses at the domain level becomes a coordinated failure at the system level.

To bring this together, **The Anatomy of Cyber Failure Framework** structures these breakdowns into four interconnected layers - a systemic model of how design, enforcement, signal, and response gaps combine to enable modern breaches.

Design Gaps
where systems are built for correctness, not adversarial reality

Enforcement Gaps
where controls exist, but fail under real-world conditions

Signal Gaps
where the system does not generate or interpret the right signals

Response Gaps
even when signals exist, action fails

The Anatomy of Cyber Failure: A 4-Layer Gap Archetype Framework

Design Gaps	Enforcement Gaps	Signal Gaps	Response Gaps
Systems built for correctness, not adversarial reality	*Controls exist, but fail under real-world conditions*	*The system does not generate or interpret the right signals*	*Even when signals exist, action fails*
Implicit Trust Assumptions (NIST SP 800-207)	**Inconsistent Control Coverage**	**Telemetry Blind Spots**	**Detection-to-Response Latency**
Insecure Design Patterns (OWASP A04:2021)	**Privilege Sprawl** (NIST AC-6/CIS 6)	**Detection Silos**	**Manual-Only Containment**
Missing Threat Models (MITRE ATT&CK coverage gap)	**Race & Concurrency Flaws** (CWE-362)	**Misaligned Detection Logic**	**Playbook Coverage Gap**
Business Logic Flaws (OWASP API1:2023 / API6:2023)	**Control Bypass Paths**	**Alert Fatigue & Signal Suppression**	**Forensic Reconstruction Gap**
Insufficient Input Validation (CWE-20, CWE-639)	**Authentication-Without-Continuous-Verification**	**Behavioural Baseline Gaps**	**Recovery Process Immaturity** (NIST CSF Recover)

Failure Chain 1: The Trusted Entry Breach

Design Gaps	Enforcement Gaps	Signal Gaps	Response Gaps
Systems built for correctness, not adversarial reality	*Controls exist, but fail under real-world conditions*	*The system does not generate or interpret the right signals*	*Even when signals exist, action fails*
Implicit Trust Assumptions (NIST SP 800-207)	**Control Bypass Paths**	**Detection Silos**	**Detection-to-Response Latency**
Missing Threat Models (MITRE ATT&CK coverage gap)	**Authentication-Without-Continuous-Verification**	**Misaligned Detection Logic**	**Manual-Only Containment**
Insecure Design Patterns (OWASP A04:2021)	**Inconsistent Control Coverage**	**Behavioural Baseline Gaps**	**Playbook Coverage Gap**
Business Logic Flaws (OWASP API1:2023 /API6:2023)		**Alert Fatigue & Signal Suppression**	**Recovery Process Immaturity** (NIST CSF Recover)
		Telemetry Blind Spots	

What Real Cases Revealed

Across seemingly unrelated incidents involving QR-code phishing campaigns, AI model evasion, and unauthorised UPI onboarding fraud, a consistent pattern emerged:

The breach did not begin with exploitation of a hardened perimeter. It began with trust that was never revalidated.

- Emails originating from **internal relay systems** were implicitly trusted, allowing phishing payloads to bypass perimeter inspection controls.

- AI fraud detection systems were manipulated using **adversarial probing** because model outputs and confidence signals were treated as reliable internal decision mechanisms rather than externally exploitable attack surfaces.

- **Mobile onboarding** workflows trusted independently validated authentication elements without verifying the integrity of the full trust sequence across device, token, OTP, and transaction context.

- Emulator checks, device signals, and onboarding patterns were treated as sufficient verification controls despite being easily bypassed under adversarial conditions.

In each case, the system behaved as designed. But the design itself assumed that trusted inputs, trusted environments, and trusted behaviours would remain trustworthy under attack.

Insight: Trust is not broken at a single point. It is progressively exploited across identity, workflow, behavioural, and system boundaries. The breach begins where trust is assumed and succeeds because it is never continuously revalidated across systems, signals, and response workflows.

The Failure Stack: Consolidated Gap Analysis

Design Gaps 01

Trust was designed into the wrong places

Internal email relay paths, AI risk-score outputs, device-reported values, OTPs, and onboarding tokens were treated as reliable trust signals, even when they could be manipulated or relayed by an attacker.

Threat models did not reflect adversarial abuse paths

QR phishing through internal channels, AI model probing, emulator-assisted onboarding, token relay, and dual-device submission behaviour were not meaningfully modelled before deployment.

Controls validated isolated events, not the full attack sequence

OTP, token, mobile number, device, and onboarding checks were assessed independently, allowing a malicious workflow to appear legitimate when viewed one step at a time.

High-risk workflows lacked adaptive trust reassessment

New onboarding, AI-assisted approvals, legacy user environments, and post-authentication activity were not designed with step-up verification, continuous session integrity, or contextual escalation.

Business logic assumed normal user behaviour

Fraud and onboarding flows were built around expected usage patterns, not adversarial sequencing, replay, probing, or manipulation of otherwise valid control steps.

Enforcement Gaps 02

Bypass paths existed inside trusted channels

Internal mail relays, legacy authentication paths, onboarding interfaces, and AI submission flows allowed malicious activity to move around perimeter or primary control enforcement.

Authentication was treated as a one-time event

MFA, OTP delivery, token validation, and device checks established trust at a point in time, but did not continuously verify session integrity or transaction legitimacy.

AI decision systems lacked enforcement guardrails

Adversarial inputs could test and exploit approval thresholds without rate limiting, probe detection, adaptive escalation, or suppression of feedback signals.

Coverage varied across environments

External traffic, internal relay traffic, primary users, legacy users, frontend activity, and backend onboarding flows were protected unevenly, creating exploitable control gaps.

Controls existed but did not consistently block abuse

Emulator checks, developer-option flags, onboarding thresholds, probe detection, and AI escalation logic either failed open, were inconsistently enforced, or did not trigger hard stops.

Signal Gaps 03

Telemetry did not capture adversarial behaviour

AI probing, feature perturbation, dual-device token submissions, emulator abuse, and suspicious onboarding patterns were not visible as real-time behavioural signals.

Signals were fragmented across systems

Identity alerts, mailbox events, OTP logs, onboarding telemetry, SMS delivery, broker systems, backend APIs, and transaction validation data were not correlated into a single attack narrative.

Detection logic looked for incident outcomes, not incident formation

Monitoring depended on completed incident, complaints, regulatory alerts, or known indicators rather than early-stage behavioural anomalies.

Behavioural baselines were underdeveloped

Inbox rule creation after suspicious sign-ins, repeated onboarding attempts, millisecond-level token submissions, and AI probe patterns were not benchmarked as abnormal.

Weak signals were recorded but not operationalised

SPF/DKIM/DMARC failures, suspicious mailbox activity, onboarding irregularities, and adversarial interaction traces existed, but were deprioritised or surfaced only during forensic review.

Response Gaps 04

Response began after compromise, not during execution

Phishing, fraudulent onboarding, and account takeover activity were detected after complaints, regulatory alerts, or external identity triggers rather than internal behavioural detection.

Containment was manual and delayed

There were no automated actions to revoke sessions, freeze accounts, suspend onboarding, block suspicious workflows, or interrupt attack progression in real time.

Incident response playbooks did not match the attack patterns

QR-code phishing via internal relay, AI model evasion, token relay onboarding abuse, and cross-system fraud correlation lacked predefined response workflows.

Investigation required heavy forensic reconstruction

Teams had to manually correlate users, sessions, devices, OTPs, onboarding events, AI submissions, and transactions after the fact to understand scope and sequence.

Recovery models were immature for trust-abuse incidents

Organisations lacked operational recovery patterns for AI compromise, onboarding trust failure, behavioural correlation gaps, and fraud chains spanning multiple systems.

Failure Chain 2: The Privilege Escalation Breach

Design Gaps	Enforcement Gaps	Signal Gaps	Response Gaps
Systems built for correctness, not adversarial reality	*Controls exist, but fail under real-world conditions*	*The system does not generate or interpret the right signals*	*Even when signals exist, action fails*
Implicit Trust Assumptions (NIST SP 800-207)	**Inconsistent Control Coverage**	**Telemetry Blind Spots**	**Detection-to-Response Latency**
Insecure Design Patterns (OWASP A04:2021)	**Privilege Sprawl** (NIST AC-6 / CIS6)	**Detection Silos**	**Manual-Only Containment**
Missing Threat Models (MITRE ATT&CK coverage gap)	**Control Bypass Paths**	**Misaligned Detection Logic**	**Playbook Coverage Gap**
Business Logic Flaws (OWASP API1:2023 /API6:2023)	**Authentication-Without-Continuous-Verification**	**Behavioural Baseline Gaps**	**Forensic Reconstruction Gap**
			Recovery Process Immaturity (NIST CSF Recover)

What Real Cases Revealed

Across incidents involving payment skimming infrastructure compromise and large-scale cryptocurrency exchange theft, a consistent pattern emerged: The breach did not escalate because of a single vulnerability. It escalated because access, once obtained, was not constrained, segmented, or continuously verified.

- Former employee **VPN credentials** remained active long after departure, providing unauthorised access into production environments handling cardholder data.
- Compromised **developer endpoints** enabled attackers to move laterally across container orchestration environments into sensitive signing infrastructure.
- **Transaction validation logic** and production systems could be modified without independent verification or cryptographic attestation.
- Excessive permissions, **weak segmentation**, and over-trusted internal communications allowed attackers to move from peripheral access points into core transaction-processing systems without triggering meaningful resistance.
- Across the cases observed, the initial compromise was not the defining failure. The defining failure was that the environment allowed trusted access to progressively expand across infrastructure, privileges, systems, and transaction layers without containment.

Insight: *Privilege escalation is rarely a single event. It is the cumulative failure to restrict, validate, segment, and continuously verify access as it moves across systems. The breach succeeds when trusted access becomes trusted control.*

The Failure Stack: Consolidated Gap Analysis

Design Gaps 01

Infrastructure trust was overextended

Internal cluster communications, service interactions, privileged admin access, and transaction validation environments were treated as trustworthy once authenticated, without strong mutual authentication, continuous validation, or independent verification.

Identity governance was not designed for persistence risk

Former employee access, service account privileges, VPN credentials, and administrator roles relied on operational cleanup rather than automated deprovisioning, least privilege, and lifecycle enforcement.

Payment systems lacked architectural isolation

Web, API, administration, payment-processing, and signing environments were insufficiently segmented, allowing compromise in broader infrastructure to reach high-trust payment functions.

Business logic depended too heavily on application-layer controls

Multi-signature workflows, withdrawal thresholds, and transaction validation rules could be weakened or bypassed through code modification, validation-rule manipulation, or structured low-volume withdrawals.

Threat models did not account for infrastructure-led payment manipulation

Lateral movement from developer systems into production signing infrastructure, signing-key exposure, web shells, reverse shells, payment-page injection, and transaction-rule tampering were not modelled as realistic attack paths.

Enforcement Gaps 02

Privilege sprawl enabled lateral movement

Overly permissive service accounts, former employee VPN access, broad administrator rights, and weak boundaries between developer, web, API, and signing environments created paths into critical infrastructure.

Production controls could be bypassed directly

Code changes were made outside version control, secure development, and change-approval workflows, allowing insecure payment-processing behaviour to be introduced directly into live environments.

Authentication did not translate into continuous verification

Once users, services, or cluster interactions were authenticated, access remained trusted without behavioural checks for privilege misuse, unusual service interactions, or sensitive infrastructure access.

Control coverage was uneven across the stack

Hardening, patching, logging, monitoring, and access enforcement varied across web, administration, API, payment-processing, and signing environments, leaving exploitable weak points.

Critical safeguards failed open or lacked independent enforcement

Endpoint controls did not prevent trojanised applications or persistence, while transaction validation logic could be modified without cryptographic verification, dual control, or independent approval.

Signal Gaps 03

Critical infrastructure telemetry was missing or incomplete

Web access logs, API logs, network flows, audit trails, secrets access, and signing-key usage were either unavailable, inconsistently retained, or not monitored in real time.

Detection logic focused on large events, not structured abuse

Monitoring prioritised uptime, availability, known failures, and large individual withdrawals, allowing incremental withdrawal manipulation below thresholds to remain operationally invisible.

Privilege misuse was not visible as an attack signal

Developer-to-production access, unusual service account behaviour, validation-rule changes, suspicious code modifications, and infrastructure manipulation were not treated as high-risk behavioural indicators.

Persistence behaviours were not baselined

Web shell deployment, reverse-shell activity, payment-page injection, unusual production code changes, and access from developer systems into signing infrastructure were not modelled as abnormal operational behaviour.

Incident and infrastructure signals remained disconnected

Endpoint activity, network movement, admin access, transaction manipulation, blockchain activity, and withdrawal behaviour were monitored separately instead of being correlated across the attack chain.

Response Gaps 04

Detection came from outside the organisation

Compromise was identified through external card-brand notifications, third-party blockchain analytics, or post-incident signals rather than internal infrastructure monitoring and behavioural detection.

Containment depended on manual intervention

There were no automated workflows to revoke compromised credentials, isolate affected infrastructure, halt suspicious withdrawals, disable malicious transaction logic, or neutralise persistence mechanisms in real time.

Attackers had already progressed before response began

By the time investigation started, persistence had been established, transaction logic had been modified, signing infrastructure had been exposed, and wrongful activity had already occurred.

Playbooks did not cover infrastructure-level payment compromise

Response procedures were not predefined for payment-page injection, signing-key compromise, transaction validation tampering, blockchain custody compromise, or infrastructure-led incident manipulation.

Forensic and recovery confidence was limited

Missing logs, fragmented telemetry, and incomplete retention prevented full reconstruction of dwell time, initial access, attacker movement, transaction impact, and system integrity, extending recovery timelines.

Failure Chain 3: The Logic Abuse Problem

Design Gaps	Enforcement Gaps	Signal Gaps	Response Gaps
Systems built for correctness, not adversarial reality	*Controls exist, but fail under real-world conditions*	*The system does not generate or interpret the right signals*	*Even when signals exist, action fails*
Business Logic Flaws (OWASP API1:2023 /API6:2023)	**Race & Concurrency Flaws** (CWE-362)	**Detection Silos**	**Detection-to-Response Latency**
Missing Threat Models (MITRE ATT&CK coverage gap)	**Inconsistent Control Coverage**	**Misaligned Detection Logic**	**Manual-Only Containment**
Insecure Design Patterns (OWASP A04:2021)	**Control Bypass Paths**	**Behavioural Baseline Gaps**	**Playbook Coverage Gap**
Insufficient Input Validation (CWE-20, CWE-639)	**Privilege Sprawl** (NIST AC-6 / CIS 6)	**Alert Fatigue & Signal Suppression**	**Forensic Reconstruction Gap**
			Recovery Process Immaturity (NIST CSF Recover)

What Real Cases Revealed

Across incidents involving OTP race-condition exploitation, insecure payment processing logic, and pervasive object-level authorization failures, a consistent pattern emerged: It did not succeed because attackers bypassed security controls. It succeeded because the application logic itself could be manipulated while remaining technically valid.

- **OTP workflows** validated transactions correctly under normal execution but failed under concurrent adversarial conditions.

- Payment applications enforced **transaction limits**, duplicate detection, and **authorization sequencing** inconsistently across edge-case and exception paths.

- APIs accepted authenticated requests and returned sensitive records without validating whether the requesting user was actually authorized to access the referenced object.

- Security testing validated that legitimate transactions succeeded, but not that malicious or **adversarial transaction flows** failed.

In each case, the system performed the exact functions it was designed to perform. The failure was that the design itself did not account for how adversaries manipulate workflow sequencing, transaction states, authorization boundaries, and business logic interactions.

Insight: *Logic abuse does not break the application. It abuses the application's own intended behaviour. The breach succeeds when business logic, authorization flow, and transaction integrity controls are treated as functional requirements rather than adversarial trust boundaries.*

The Failure Stack: Consolidated Gap Analysis

Design Gaps 01

Transaction workflows were designed for expected behaviour, not adversarial sequencing

OTP validation, transaction authorization, duplicate-payment checks, and exception paths assumed cooperative user behaviour rather than repeated, concurrent, malformed, or deliberately sequenced requests.

Critical transaction steps lacked atomic integrity

OTP validation and consumption were implemented as separate operations, allowing transaction state to be manipulated under race conditions or parallel execution.

Sensitive identifiers were exposed too directly

Predictable object IDs, direct reference access, and weak contextual validation allowed attackers to enumerate or access records beyond their legitimate scope.

Threat models missed logic-layer abuse

Race-condition exploitation, concurrent transaction execution, object enumeration, payment exception-path manipulation, and attackers with full API visibility were not adequately modelled or tested.

Authorization was not enforced as a central design principle

APIs relied on user-supplied object identifiers and endpoint-level checks instead of consistent object-level ownership validation and centralized authorization enforcement.

Enforcement Gaps 02

Concurrency controls failed under parallel execution

Multiple authorization requests could be processed before OTP state, transaction state, or database records were locked and updated, breaking transaction integrity.

Controls worked in standard flows but failed at the edges

Duplicate-payment detection, authorization sequencing, transaction limits, and exception workflows were not consistently enforced across all paths, states, and edge conditions.

Authentication was mistaken for authorization

Authenticated users could access or manipulate objects outside their legitimate scope because object-level access control was not treated as a separate enforcement boundary.

APIs allowed workflow manipulation

Manipulated identifiers, repeated submissions, concurrent requests, weak throttling, and verbose error handling created bypass paths for enumeration and transaction abuse.

Privilege expanded implicitly across workflows

Once users entered authenticated transaction flows, access scope was not continuously revalidated across identifiers, objects, payment steps, and state transitions.

Signal Gaps 03

Detection logic focused on system success, not transaction legitimacy

Monitoring validated uptime, processing success, and completed transactions, but did not assess whether the transaction sequence should have been allowed.

Business logic abuse was not baselined

Rapid sequential submissions, concurrent authorization attempts, repeated OTP use, object enumeration, edge-case exploitation, and abnormal payment patterns were not modelled as suspicious behaviour.

Signals were logged but not correlated

OTP events, transaction authorization, API access, object requests, payment gateway reconciliation, and fraud indicators remained fragmented across systems.

Enumeration and workflow abuse looked technically valid

Repeated identifier changes, unusual transaction paths, and exception-flow exploitation were captured as normal application activity rather than elevated as behavioural risk.

Security monitoring missed logic-layer exploitation

Assessments and alerts focused on authentication, injection, availability, and operational anomalies while underweighting authorization abuse, state manipulation, and adversarial workflow testing.

Response Gaps 04

Exploitation was detected only after impact

Fraudulent transactions, OTP reuse, authorization abuse, and object-level access violations surfaced after exploitation, reconciliation, forensic review, or formal assessment.

Containment was not built into transaction workflows

There were no automated controls to block OTP reuse, pause duplicate transactions, throttle enumeration, suspend suspicious workflows, or interrupt concurrent abuse in real time.

Response playbooks did not cover business logic attacks

Race-condition exploitation, object-level authorization abuse, transaction-state manipulation, and payment exception-path abuse lacked predefined investigation and containment workflows.

Forensic reconstruction was difficult and incomplete

Limited behavioural telemetry, fragmented logs, and poor visibility into transaction sequencing made it hard to determine vulnerable workflows, exploitation scope, and transaction impact.

Recovery was retrospective rather than operationalized

Remediation required manual code review, reconciliation, and post-incident control redesign instead of predefined validation mechanisms, continuous testing, and integrated detection improvements.

Failure Chain 4: The Invisible Attack Chain

Design Gaps	Enforcement Gaps	Signal Gaps	Response Gaps
Systems built for correctness, not adversarial reality	*Controls exist, but fail under real-world conditions*	*The system does not generate or interpret the right signals*	*Even when signals exist, action fails*
Missing Threat Models (MITRE ATT&CK coverage gap)	**Control Bypass Paths**	**Telemetry Blind Spots**	**Detection-to-Response Latency**
Implicit Trust Assumptions (NIST SP 800-207)	**Inconsistent Control Coverage**	**Detection Silos**	**Playbook Coverage Gap**
Insecure Design Patterns (OWASP A04:2021)	**Authentication-Without-Continuous-Verification**	**Misaligned Detection Logic**	**Forensic Reconstruction Gap**
		Behavioural Baseline Gaps	**Recovery Process Immaturity** (NIST CSF Recover)

What Real Cases Revealed

Across incidents involving fileless endpoint evasion, covert command-and-control infrastructure, encrypted exfiltration channels, and payment-skimming compromise with missing forensic evidence, a consistent pattern emerged: The attack did not succeed because defenders lacked security controls. It succeeded because the attack operated outside the organization's visibility model.

- **Payloads executed entirely in memory** without generating disk artifacts, bypassing endpoint monitoring dependent on file-based indicators.
- Command-and-control traffic blended into trusted **CDN and HTTPS traffic**, while IPv6 communication operated completely outside monitored network visibility.
- **DNS-over-HTTPS** encrypted exfiltration activity inside legitimate HTTPS sessions, eliminating conventional DNS visibility entirely.
- Critical web access, API, audit, and network flow logs required to reconstruct payment compromise activity simply did not exist.

Across the cases analyzed, the organizations had security tooling deployed, active monitoring programs, and operational security teams. But the telemetry architecture itself was incomplete. The environments could monitor expected attack surfaces, expected protocols, and expected indicators. They could not observe attacks deliberately engineered to exist outside those assumptions.

Insight: *Invisible attack chains do not evade controls by overpowering them. They evade controls by operating where telemetry does not exist, where signals are fragmented, and where monitoring assumptions no longer align to adversarial behaviour. The breach succeeds when visibility itself becomes the attack surface.*

The Failure Stack: Consolidated Gap Analysis

Design Gaps 01

Threat models missed stealth execution paths

Fileless malware, memory-only execution, DNS-over-HTTPS, CDN-based command-and-control, web shells, reverse shells, and payment-page injection were not modelled as likely attack paths.

Trusted channels were over-trusted

HTTPS, CDN traffic, internal infrastructure, and IPv4-dominant visibility models were assumed safe or sufficient, leaving encrypted, IPv6, and covert channels under-monitored.

Monitoring was built for known indicators

Architecture relied on file artifacts, malware signatures, and standard operational logs rather than behavioural, memory-level, and forensic-grade observability.

Forensic visibility was not designed in

Web, API, audit, network flow, IPv6, encrypted DNS, and endpoint memory telemetry were incomplete or inconsistent across environments.

Enforcement Gaps 02

Controls were bypassed through invisible execution

Fileless payloads avoided disk-based endpoint controls, while memory-resident activity produced few traditional indicators.

Trusted protocols became evasion channels

DNS-over-HTTPS, HTTPS, CDN traffic, and browser channels enabled covert command-and-control without triggering security enforcement.

Coverage was uneven across infrastructure

IPv4 traffic was monitored more comprehensively than IPv6, and logging varied across web, application, API, network, and system layers.

Integrity protections failed at the application layer

Web shells, malicious code changes, and payment-page injection bypassed application integrity checks and operational oversight.

Long-lived activity was not revalidated

Covert sessions inside trusted channels persisted without behavioural reassessment, proactive hunting, or escalation.

Signal Gaps 03

Critical telemetry was absent

IPv6 traffic, encrypted DNS, memory execution, API logs, audit logs, network flows, and persistence activity were either missing or incomplete.

Traditional SIEM logic had little to detect

Fileless attacks left no disk artifacts, and covert activity did not map cleanly to known IOC-based detection rules.

Signals remained fragmented

Endpoint, application, network, and SIEM data were not correlated into a unified attack timeline.

Detection logic was outdated

Monitoring focused on malware signatures, operational failures, and known compromise indicators instead of stealth, persistence, evasion, and low-noise exfiltration.

Behavioural baselines were missing

Long-duration sessions, unusual protocol use, CDN-based C2, low-volume exfiltration, and suspicious process behaviour were not treated as anomalies.

Response Gaps 04

Detection came too late

Compromises were identified through external notification or red-team disclosure, not internal monitoring.

Attackers had already established dwell time

By investigation start, persistence, exfiltration, and covert infrastructure activity were already underway.

Playbooks did not match the attack type

Fileless malware, DNS-over-HTTPS evasion, CDN-based C2, memory-resident compromise, payment-page injection, and logging failure lacked predefined response workflows.

Forensic reconstruction was weak

Missing logs and absent telemetry made it difficult to determine entry path, dwell time, persistence timeline, attacker movement, and exfiltration scope.

Recovery lacked assurance

Remediation depended on retrospective estimates rather than telemetry-backed validation of what was compromised and whether persistence was removed.

The Next Wave: Seven Cyber Shifts That Will Reshape BFSI

Looking beyond the threats that defined 2025, the more important question for leaders now is what changes in 2026 and beyond. What is emerging is a threat landscape shaped by the convergence of forces: artificial intelligence weaponized at scale, identity becoming the primary battleground, application logic exploited beyond the reach of conventional controls, and a regulatory environment racing to keep pace with adversarial innovation.

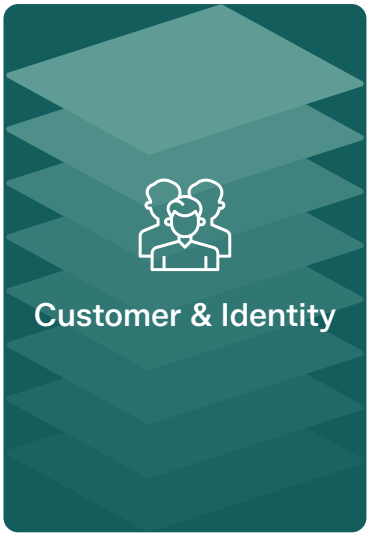
Drawing on observed threat activity across the ecosystem, the following outlook highlights the trends most likely to shape the evolving threat landscape and the strategic security priorities that should follow.

Trend	Attack vector What attackers are doing	Key shift What has changed	Defensive priority Where cybersecurity programs must focus
Identity Inheritance & Session Compromise	• AiTM frameworks, session token theft, OAuth consent abuse & federated identity manipulation - now commodity tooling • Non-human identities (service accounts, API keys, machine credentials) targeted at same intensity as human users • Consent-phishing abuses legitimate OAuth flows to obtain durable access without triggering auth alerts	• Attack surface has moved from login boundary to full authenticated lifecycle • MFA bypassed post-authentication, not broken - session token is the new credential • Long-lived tokens = primary persistence mechanism as SSO & open banking APIs expand	• Extend monitoring from auth events to continuous session integrity • Machine-identity lifecycle governance: same rigour as human accounts • Behavioural baseline deviation detection across sessions & non-human identities
AI-Scale Social Engineering & Synthetic Identity Problem	• Millions of AI-personalised phishing emails per half-year; quishing (QR codes) bypasses MFA by routing victims to AiTM infrastructure • Voice cloning used to impersonate executives & authorise wire transfers / credential resets • Deepfake-enabled fraud surging in Indian & SE Asian markets - targeting KYC pipelines, video auth & remote onboarding • AI-generated identity documents & behavioural profiles engineered to pass automated KYC	• Cost, language & credibility barriers removed - operations now run automatically at millions of targets simultaneously • Content is grammatically perfect & contextually accurate - detection logic built for imperfect human-authored attacks loses effectiveness • Campaigns combine AI content, breached data, behavioural mimicry & timing precision across multiple channels	• Out-of-band callback verification for sensitive transactions • Liveness & synthetic-media detection at KYC / video auth ingestion points • Shift detection from content-quality signals to behavioural & procedural signals
Business Logic & API Exploitation on Real-Time Rails	• Race conditions on real-time rails: OTP exploitation & parallel transaction triggering produce fraud bursts before detection correlates • IDOR patterns across endpoints - authorisation at one boundary does not propagate to adjacent systems • Mobile wallet logic flaws, broken auth in mobile banking apps; smart contract & tokenisation platform exploits - asset loss with no reversal	• Target is how systems function, not how they are coded - no malware signature, no network anomaly • Activity appears legitimate; typically discovered weeks or months after financial harm is realised • Same logic-abuse pattern now expanding into digital asset custody & programmable finance	• Adversarial business logic testing as discrete discipline, separate from infrastructure pen-testing • Protocol-level invariant enforcement at payment-network level • Continuous transaction behavioural baselines - flag deviation even when individual events appear authorised
Adversarial Attacks on AI Systems & Agentic Identity Risk	• Decision boundary probing: iterative submissions reverse-engineer fraud & credit model logic to craft inputs that pass at minimum cost • Prompt injection: malicious instructions embedded in emails, documents & web content processed by enterprise AI agents • Attackers deploy purpose-built adversarial LLMs for polymorphic malware & fraudulent document creation — structurally unique each iteration, defeating signature detection	• AI models are now active attack targets - a model not red-teamed for adversarial robustness can underperform the rule set it replaced • Agentic AI operates with admin-level privilege under service-account-grade governance - the document is the carrier, the agent is the executor • Three patterns converging: AI systems targeted, exploited as execution environments, and weaponised by attackers	• Adversarial robustness testing before AI deployment; defined reassessment cadence • Treat AI agents as privileged identities under continuous behavioural monitoring • Prompt-injection-aware content ingestion controls; suppress model decision rationale from applicant-facing surfaces

Trend	Attack vector What attackers are doing	Key shift What has changed	Defensive priority Where cybersecurity programs must focus
Cloud Control Plane & SaaS-Layer Compromise	• OAuth token abuse, federated identity misconfiguration & IAM lateral movement dominating breach analysis 2025–2026 • SaaS-to-SaaS chains: compromise of one app yields access to others via inter-application trust • Build automation, CI/CD & config management platforms exploited - actions blend into normal operational noise	• Perimeter is now defined by cloud & SaaS configuration, not own infrastructure • CSP attestation covers provider side only - tenant-side IAM, OAuth scopes & inter-service trust sit outside conventional pen-test scope • Control-plane compromise delivers estate-wide reach from a single misconfiguration	• Continuous cloud attack surface management at cloud-rate-of-change cadence • Tenant-side IAM posture management as continuously-assessed discipline • OAuth scope review, app-to-app trust mapping & least-privilege enforcement in SaaS layer
Supply Chain-Delivered Ransomware & Data Extortion	• Vectors: compromised software updates, malicious OSS packages, CI/CD pipeline compromise, MSP lateral movement • 2025: single vendor breach → confirmed impact across 70+ US banks & credit unions; 1M+ customer records exposed • April 2026: two major US banks posted simultaneously on ransomware leak site after shared vendor compromise	• Encryption stage abandoned by many groups/ pure data exfiltration + extortion is faster, harder to detect, equally coercive • Attackers bypass hardened institutions by compromising trusted providers with sector-wide access • One vendor compromise now produces concurrent impact across many institutions simultaneously	• SBOMs, provenance verification & runtime integrity attestation alongside vendor due diligence • Continuous monitoring of third-party API exposure, separate from periodic vendor review • IR planning must contemplate simultaneous multi-institution compromise via shared vendors
Pre-Disclosure Exploitation & Patch-Window Collapse	• N-day exploitation compressed to hours for high-value internet-facing infrastructure after CVE publication • Pre-disclosure exploitation: underground markets & independent discovery increasingly precede vendor patch availability • High-trust targets: management appliances, identity providers & edge devices - high privilege, historically under-invested in security	• Monthly/quarterly patch cycles are operationally insufficient - attack tempo is now hours, not weeks • Challenge is no longer faster patching alone it is detection & containment before a patch exists • Proportion of incidents involving pre-disclosure exploitation is structurally growing	• Real-time asset inventory of exposed services & components • Virtual patching & runtime protection for containment during pre-patch windows • Exploitability-weighted prioritisation - active exploitation evidence & exposure, not CVSS score alone

Suggestions to Policymakers and Supervisory Authorities

The recommendations below are framed in that spirit - as the next iteration of supervisory direction, addressing the threat realities that have emerged inside the architectures regulation itself has enabled. Each is anchored to one of the six operating layers of BFSI discussed above.

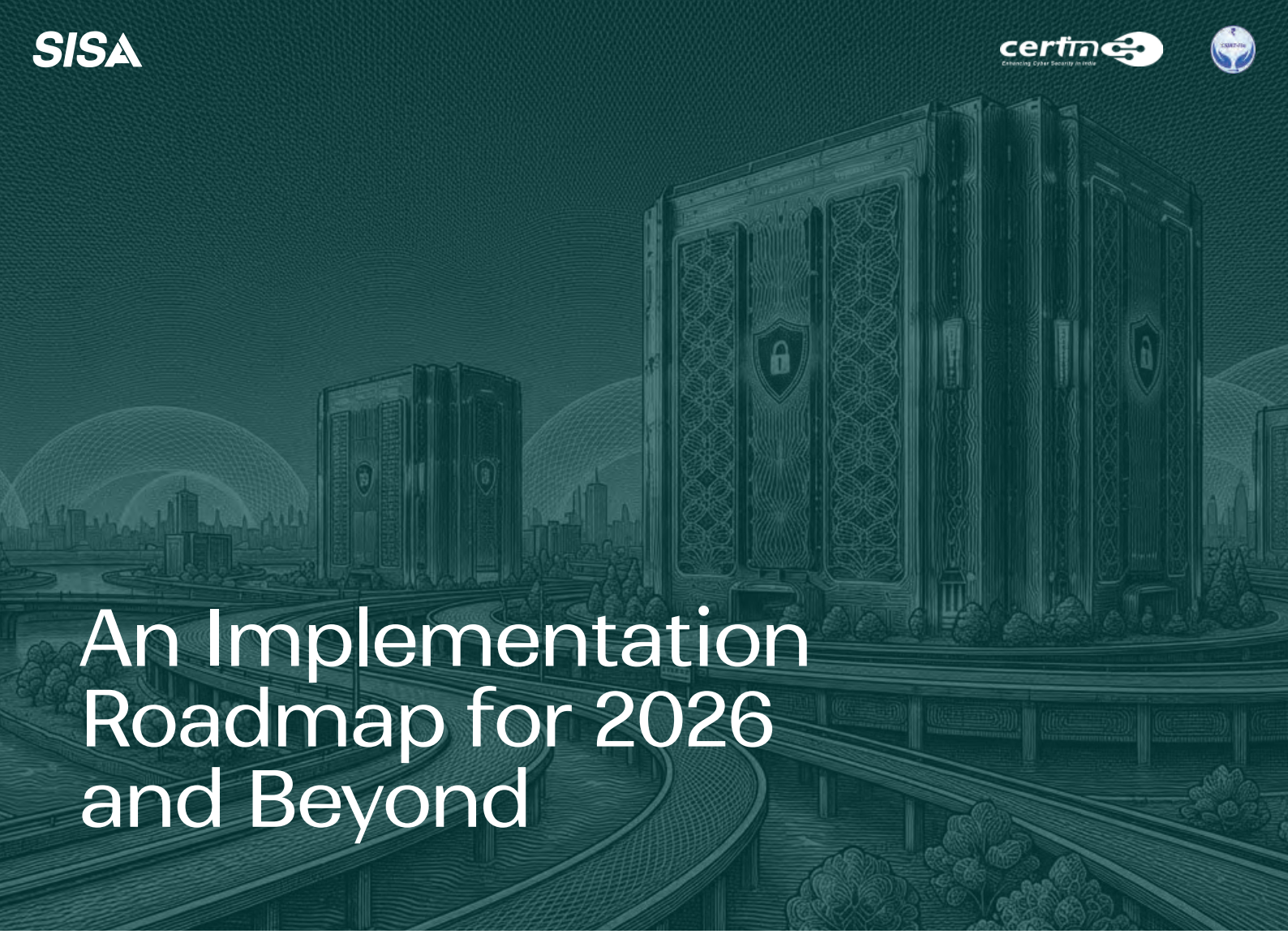
Layer	Priority Actions
	Mandate active liveness detection as the baseline for digital onboarding Require continuous session assurance (behavioural biometrics, device posture, token binding) for high-privilege sessions Extend identity governance to non-human identities - service accounts, machine identities, agentic AI - on the same attestation and rotation perimeter as human identities Enable cross-database identity verification for high-assurance account openings where legally feasible

Layer	Priority Actions
	Set minimum standards for real-time fraud signal correlation across payment rail participants Define protocol-level invariants for instant payment infrastructure to detect OTP race conditions and idempotency violations at network level Establish adversarial business logic testing as a supervisory expectation, distinct from infrastructure penetration testing Extend supervisory frameworks to programmable financial infrastructure - tokenisation, smart-contract settlement, digital asset custody - recognising irreversibility
	Require adversarial robustness testing before AI deployment in consequential financial decisions, with defined reassessment cadence Treat model weights, training data, and ML dependencies as regulated software supply chain components Mandate human-in-the-loop controls for agentic AI actions above defined financial thresholds, with full audit trails Require suppression of model confidence signals and decision rationale from applicant-facing interfaces
	Require continuous control monitoring against a defined institutional baseline, supplementing periodic attestation Mandate runtime integrity attestation in container and orchestration environments - monitoring post-deployment drift, privilege escalation, and hostPath access Clarify shared-responsibility expectations in cloud, emphasising tenant-side IAM and service-to-service trust Transition institutions above defined thresholds from periodic vulnerability assessment to continuous attack surface management

Layer	Priority Actions
Ecosystem & Dependency	Require software supply chain attestation (SBOMs, provenance verification, runtime integrity) as a complement to vendor due diligence
	Mandate continuous monitoring of third-party API exposure, recognising the API perimeter changes more frequently than the vendor inventory
	Establish cross-institutional information sharing protocols for supply chain compromise indicators, modelled on fraud-signal sharing in established payment networks
	Set supervisory expectations for transitive dependency risk across cloud, identity, and SaaS supply chains
Control & Assurance	Define a transition path from periodic attestation to continuous control monitoring for institutions above defined thresholds, with clear milestones
	Harmonise cross-framework assurance requirements to reduce overhead without weakening control intent
	Align incident reporting cadence with compressed attack timelines, particularly for AI-enabled fraud and supply chain compromise
	Set formal expectations for CISO governance - reporting line, independent escalation authority, board visibility - scaled to institutional size and systemic importance
	Require post-quantum cryptography migration planning for high-longevity data, with defined deployment milestones
Cross-Cutting	Harmonisation - coordinate through FSB, BIS, IOSCO and regional supervisory colleges to preserve rigour while reducing assurance overhead
	Information sharing - extend and standardise threat indicator sharing with appropriate confidentiality protections; one of the highest-leverage supervisory interventions available
	Public-private partnerships - deepen operational integration on the CSIRT-Fin / CERT-In model: real-time indicator exchange, joint advisories, and coordinated incident response

Supervisory Priorities for 2027

Operating Layer	Current Position	Direction for 2027	Measurable Outcome
Customer and Identity	MFA and KYC established	Continuous session assurance; active liveness; NHI parity with human identity	• Session hijack rate • % NHIs under PAM governance
Transaction and Financial Logic	Real-time rails operational; fraud reporting mandated	Protocol-level invariants; adversarial logic testing; programmable infrastructure assurance	• Fraud logic abuse detection rate • Adversarial testing coverage
Data and Intelligence	AI transparency and bias testing	Adversarial robustness testing; probe detection; AI supply chain integrity	• AI probe detection rate • Model integrity validation frequency
Platform and Infrastructure	Cloud guidance; periodic VAPT	Continuous control monitoring; runtime attestation; continuous attack surface management	• Runtime drift detection SLA • Exposed asset remediation time
Ecosystem and Dependency	Third-party risk frameworks	Software supply chain attestation; transitive dependency visibility; signal sharing	• SBOM coverage • Third-party compromise detection time
Control and Assurance	Periodic attestation; defined reporting windows	Continuous assurance pathway; cross-framework harmonisation; compressed reporting; PQC milestones	• Continuous assurance coverage % • PQC migration milestone attainment



An Implementation Roadmap for 2026 and Beyond

The final section translates that analysis in the preceding sections into a practical eighteen-month roadmap.

The roadmap below is structured along three horizons - the first six months, six to twelve months, and twelve to eighteen months - and is mapped against the six BFSI Operating Layers introduced earlier in the report. Institutions should treat the roadmap as a reference architecture for prioritisation rather than as a prescription; the specific sequence and emphasis will vary by institutional context, size, and maturity.

The Underlying Principles

Three principles inform the entire roadmap and should guide investment decisions across all three horizons.

01

Resilience over prevention

Assume compromise is inevitable; invest in the ability to detect, contain, respond, and recover faster.

02

Identity as the highest-leverage control

Prioritise human, non-human, and session identity assurance as the common control layer across emerging BFSI threats.

03

Continuous over periodic

Move from point-in-time checks to continuous validation, monitoring, and assurance as attacks become faster and always-on.

Horizon 1: First Six Months - Foundational Controls

The first six months should focus on closing the most immediate operational gaps surfaced by 2025-26 incident data. These are actions that should be in execution by mid-to-late 2026 in any BFSI institution operating at scale.

BFSI Operating Layers	Priority Actions	Strategic Outcome
Customer and Identity	Deploy phishing-resistant MFA (FIDO2 / WebAuthn) for all administrative, treasury, and high-privilege user accounts. Establish a non-human identity inventory across service accounts, machine credentials, API keys, and integration accounts.	Eliminate phishable credentials in the highest-risk user population. Establish visibility into the identity surface that has historically operated outside review cycles.
Transaction and Financial Logic	Implement adversarial business logic testing for real-time payment, mobile wallet, and open banking API surfaces. Establish protocol-level controls for OTP race condition, parallel transaction triggering, and idempotency violations on real-time rails.	Close the business logic exposure that produces direct financial harm without conventional security signatures.
Data and Intelligence	Establish adversarial robustness testing as a formal gate in the AI model deployment lifecycle for fraud, credit, AML, and KYC systems. Suppress or obfuscate model confidence signals from applicant-facing interfaces.	Close the AI decision-boundary exposure that is currently invisible to most security monitoring.
Platform and Infrastructure	Implement hardware security module-based secrets management for signing, custody, and payment systems. Establish runtime container integrity monitoring for production workloads handling cardholder or sensitive customer data. Extend telemetry coverage to IPv6 and DNS-over-HTTPS traffic.	Eliminate software-accessible signing key material. Surface post-deployment configuration drift in containerised environments. Close detection blind spots in encrypted egress.
Ecosystem and Dependency	Establish a Software Bill of Materials capability for all internally developed and third-party software entering production. Implement runtime monitoring of third-party API exposure.	Establish baseline visibility into the supply chain attack surface that is currently inferred from periodic vendor review.
Control and Assurance	Deploy XDR or equivalent extended detection capability with AI-driven triage and automated containment playbooks. Establish Mean Time to Contain as the primary security operations metric, in place of or alongside Mean Time to Detect.	Close the detection-to-response latency that converts initial access into material financial harm.

Horizon 2: Six to Twelve Months - Continuous Capability

The 6–12 month horizon focuses on the transition from periodic to continuous capability across the most critical control areas. The actions in this horizon depend on the foundational work in Horizon 1 being substantially complete.

BFSI Operating Layers	Priority Actions	Strategic Outcome
Customer and Identity	Implement continuous session assurance through behavioural biometrics, device posture verification, and token-binding for high-privilege and high-value sessions. Extend non-human identity governance - rotation, attestation, access review - to parity with human identity governance.	Close the session hijacking gap that defeats authentication-event-based detection. Bring NHI risk under the same governance discipline as human identity risk.
Transaction and Financial Logic	Establish transaction-time fraud signal correlation across participants on real-time rails where institutional consortium arrangements permit. Implement behavioural baselining for transaction patterns at customer and merchant level.	Detect coordinated fraud patterns that are invisible at the individual institution level. Identify behavioural deviation in authorised transactions.
Data and Intelligence	Implement probe-pattern detection for AI-integrated application surfaces - iterative submissions, high-velocity queries, systematically varying feature inputs. Treat AI agents as privileged non-human identities under continuous behavioural monitoring.	Detect adversarial probing before evasion profiles are established. Bring agentic AI under the privileged identity governance perimeter.
Platform and Infrastructure	Implement micro-segmentation across cloud and container environments. Establish tenant-side IAM posture management as a discrete continuous capability. Implement runtime attestation of software provenance and dependency integrity in production.	Limit lateral movement blast radius. Surface cloud configuration drift continuously rather than at periodic audit.
Ecosystem and Dependency	Implement continuous attack surface management with real-time inventory of internet-exposed services and components. Establish cross-institutional information sharing protocols for supply chain compromise indicators where sectoral arrangements exist.	Close the gap between the actual attack surface and the institution's view of its attack surface. Shorten institutional response window for supply chain compromise.
Control and Assurance	Define and operationalise continuous control monitoring against the institutional control baseline for the highest-priority control set. Establish quarterly post-attestation adversarial validation as a standing programme.	Begin the structural transition from point-in-time attestation to continuous assurance. Identify control gaps that pass periodic attestation but fail adversarial conditions.

Horizon 3: Twelve to Eighteen Months - Architectural Maturity

The 12-18 month horizon addresses the longer-cycle architectural investments that require substantial coordination, vendor engagement, or migration planning. These are the investments that position institutions for the threat reality emerging beyond 2026.

BFSI Operating Layers	Priority Actions	Strategic Outcome
Customer and Identity	Deploy passwordless authentication across the full privileged user population. Extend continuous session assurance to consumer-facing high-value transactions where customer experience trade-offs permit.	Eliminate credential-based attack vectors in the privileged user population. Extend continuous trust validation to customer-facing transaction flows.
Transaction and Financial Logic	Implement blockchain-aware transaction monitoring for institutions with custody, settlement, or digital asset exposure. Extend adversarial business logic testing to programmable financial infrastructure including tokenisation and smart contract platforms.	Close the monitoring gap for cryptographic-layer manipulation and structured threshold evasion. Bring programmable financial infrastructure under the same assurance discipline as conventional payment infrastructure.
Data and Intelligence	Establish AI supply chain integrity controls — provenance verification for model weights, training datasets, and ML library dependencies. Implement prompt-injection-aware content ingestion controls for agentic AI systems.	Bring AI development dependencies under the same supply chain governance as software dependencies. Close the prompt injection vector at the content ingestion boundary.
Platform and Infrastructure	Complete the architectural transition to identity-centric access control across cloud, SaaS, and on-premise estates. Establish runtime integrity attestation at scale for production container and orchestration environments.	Replace network-perimeter access control with identity-as-perimeter across the full estate. Achieve continuous integrity visibility across the production workload fleet.
Ecosystem and Dependency	Operationalise transitive dependency visibility - the institution's exposure through its vendors' vendors. Establish cross-institutional incident response coordination capability for supply chain compromise affecting multiple institutions simultaneously.	Surface the supply chain exposure that is structurally invisible at the direct vendor level. Reduce response time for systemic supply chain compromise.
Control and Assurance	Initiate post-quantum cryptography migration for high-longevity data with defined deployment milestones. Establish crypto-agility - the architectural capability to update cryptographic algorithms without rebuilding systems. Complete the institutional transition to continuous control monitoring as the primary assurance posture.	Address *harvest-now-decrypt-later* exposure for data with retention obligations extending beyond cryptographic horizon. Position the institution to update cryptographic primitives without major architectural rebuild. Achieve continuous assurance posture aligned with where supervisory direction is evolving.

Cross-Horizon Capability: Adversarial Assurance

One capability cuts across all three horizons and warrants articulation as a continuous discipline rather than a horizon-specific action: *adversarial assurance.*

Adversarial assurance is the discipline of validating that controls hold under realistic attacker conditions - not just that they are present and operating during the assessment window. It encompasses red-team exercises, purple-team validation, adversarial business logic testing, adversarial AI testing, and continuous attack surface management. It is the operational complement to periodic compliance attestation, and the structural answer to the compliance-security translation gap documented earlier in this report.

Institutions should establish adversarial assurance as a standing programme reporting into the same governance structure as the compliance programme, with comparable cadence, comparable resourcing, and findings treated as equal-weight inputs to the institutional risk register. Three actions span all three horizons:

Establish adversarial assurance as a defined function with cadence and governance equivalent to compliance attestation

Build the institutional muscle for converting adversarial findings into engineering backlog rather than perpetual exception requests

Develop the cross-discipline collaboration between security testing, engineering, compliance, and security operations that adversarial assurance requires to be effective

Implementation Considerations

Five considerations should guide how institutions execute against the roadmap.

01

Sequence matters more than scope

Institutions attempting to execute against all eighteen actions in parallel will execute none of them well. Horizon 1 actions establish the foundation on which Horizon 2 and 3 actions depend. Institutions should resist the pressure to commence Horizon 2 actions before Horizon 1 is substantively complete.

02

Identity investments produce the highest leverage

Across the eighteen actions, those addressing identity - phishing-resistant MFA, NHI governance, continuous session assurance, identity-centric access control, passwordless authentication - close the greatest number of threat vectors simultaneously. Where investment priority must be made among Horizon 1 actions, identity is the defensible first call.

03

Vendor concentration is a strategic variable

Several actions in the roadmap can be implemented through existing vendor relationships, through new specialist vendors, or through capability built in-house. Institutions should make these decisions deliberately rather than by default, recognising that vendor concentration creates supply chain exposure and that capability built in-house creates institutional dependency on talent retention.

04

Regulatory alignment reduces execution friction

Several actions in the roadmap align with the supervisory direction articulated in the recommendations section of this report. Institutions executing against actions that align with emerging regulatory expectations face less execution friction than those executing against actions that do not.

05

Resilience metrics, not control metrics

The maturity of the institutional security programme should be measured by resilience outcomes - Mean Time to Contain, recovery time objectives, business continuity under active compromise - rather than by control inventory or compliance attestation. Institutions whose internal security metrics remain calibrated to control coverage will struggle to recognise when their programme is mature in inventory terms but immature in resilience terms.

Conclusion

The underlying message of this report is not that cyber threats are increasing. That has been true for years. The real message is that the foundations on which financial institutions have traditionally managed trust are being rewritten.

For decades, security programs were built around clear boundaries: inside and outside, trusted and untrusted, user and attacker, fraud and cybersecurity, technology risk and business risk. Those boundaries are now dissolving and the emerging probabilities are endless. This is why the next chapter of resilience cannot be built by adding more tools to yesterday's architecture. It requires a shift in mindset, from static control to continuous assurance, from isolated defense to ecosystem trust, from prevention alone to operational resilience.

For the BFSI sector, the stakes are uniquely high. Finance runs on confidence. Every payment cleared, every account opened, every digital interaction approved depends on trust working silently in the background. When that trust weakens, the impact is measured not only in losses, but in hesitation, disruption, and systemic friction.

The pace of this change is equally important. Threats that once evolved over years are now reaching practical deployment in months. Capabilities once associated with elite actors are becoming widely accessible. The gap between innovation and exploitation continues to narrow. Institutions therefore face a dual challenge: defending against current threats while preparing for categories of risk that are still emerging.

This is why resilience must now be viewed differently. It requires cybersecurity to move closer to the center of enterprise decision-making. Boards will need clearer visibility into cyber exposure as a business risk. Technology leaders will need to build security into architecture, not layer it on afterward. Security teams will need to operate with greater convergence. Regulators and institutions will need to collaborate more closely on systemic resilience. And organizations will need to measure success not only by prevention, but by continuity, adaptability, and trust preserved.

The institutions that lead the next decade will not necessarily be those with the largest budgets or the broadest security stacks. They will be those that understand this structural shift early and respond decisively. They will treat identity as a dynamic control plane, software as a continuously assured environment, ecosystems as shared exposure, and trust as a strategic asset.

Cybersecurity, therefore, is no longer protecting the business from the sidelines. It is becoming one of the core disciplines through which the future of finance will be built. That future is already taking shape. The time to prepare for it is now.

Acknowledgements

We extend our sincere appreciation to the forensic investigators, analysts, cybersecurity professionals, payment security specialists, and industry leaders whose expertise and frontline experience helped shape the insights presented in this report.

This edition draws upon forensic investigations, threat intelligence observations, incident response engagements, technical assessments, and continuous interactions across the global payments and BFSI ecosystem. The perspectives of institutions, regulators, technology partners, and practitioners confronting evolving threats in real time have been invaluable in helping us understand not only where risk exists today, but where it is headed next.

A huge thanks to officers of CSIRT-Fin (Financial Computer Security Incident Response Team) and CERT-In (Indian Computer Emergency Response Team), whose contributions have been instrumental in the creation of this report. Their ability to synthesize findings, provide insights, and bring this analysis to life underscores the incredible talent, depth and dedication within the respective teams.

We would also like to acknowledge the teams across SISA whose deep domain expertise in forensics-led cybersecurity, compliance, digital payments security, and emerging technologies contributed significantly to the development of this report.

As cyber threats become increasingly dynamic and interconnected, collaborative intelligence remains essential. We hope this report contributes meaningfully to the broader effort of strengthening resilience, preserving trust, and securing the future of the BFSI sector.

References

1. Cybersecurity in the Age of AI: Building a Synchronous BFSI - DSCI-BCG Report, May 2026
2. 2026 Data Breaches: Cybersecurity Incidents - PKWARE®
3. Every AI Frontier Model is Now a Cyber Threat. So What Can You Do About It? | Rubrik
4. The State Of Ransomware 2026 | BlackFog
5. Significant Cyber Threats of 2026: Comprehensive Outlook, Cybersecurity Insiders
6. Cybersecurity Trends in Banking: 5 Biggest Threats | DeskAlerts, May 2026
7. AWS Marketplace: Cyber Security in BFSI Market Growth Forecast and Analysis to 2030
8. Get the 2026 Guide to Banking Cyber Threats: TekClarion, Jan 2026
9. SISA Sappers - Digital Forensics & Incident Response (DFIR) Research
10. SISA Offensive Security Intelligence

SISA

SISA is a global leader in cybersecurity for the payment ecosystem, operating at the intersection of AI, cybersecurity, and payments. Trusted by leading brands and financial institutions across 40+ countries, SISA secures over 1000 organizations by helping them anticipate threats, strengthen resilience, and protect critical payment infrastructure with its deep tech capabilities and forensic led insights by SAPPERS. Powered by real-world breach intelligence, SISA enables organizations to stay ahead of evolving cyber risks. Follow SISA on LinkedIn for the latest insights on cybersecurity, payment security innovation, and emerging technologies.

CERT-In

CERT-In is the national agency for responding to computer security incidents as and when they occur. In the Information Technology Amendment Act 2008,CERT-In has been designated to serve as the national agency to perform the following functions in the area of cyber security:

- Collection,analysis and dissemination of information on cyber incidents.
- Forecast and alerts of cyber security incidents.
- Emergency measures for handling cyber security incidents.
- Coordination of cyber incident response activities.
- Issue guidelines, advisories, vulnerability notes and whitepapers relating to information security practices,procedures, prevention, response and reporting of cyber incidents.
- Such other functions relating to cyber security as may be prescribed. Refer **www.cert-in.org.in** for more details.

CSIRT-Fin

Computer Security Incident Response Team in Finance sector (CSIRT-Fin), is a nodal sectoral CSIRT which provides Incident Prevention and Response services as well as Security Quality Management Services to the entities of the Indian financial sector. It manages cyber incidents and coordinate responses across banking, securities market infrastructure, insurance, and pension funds entities.

It carries out the following roles related to the cyber security in financial sector:

i. Collection, analysis & dissemination of information on cyber incidents.
ii. Forecast and alerts on cyber security incidents.
iii. Emergency measures on cyber security incidents.
iv. Coordination for cyber incident response activities.
v. Issue guidelines, advisories, vulnerability, and white papers relating to information security.
vi. Monitor sectoral efforts in the financial sector towards maintaining dynamic and modern cyber security architecture, developing awareness amongst regulated entities and public in general.
vii. Such other functions relating to cyber security in the financial sector, as may be prescribed.