WHITE PAPER | FEBRUARY 2026 | ELITE EDITION

THE SOVEREIGN DEFENSIBILITY FRAMEWORK

Board Control of Agentic AI Under DORA and the New Standard of Financial Operational Accountability

A Category-Defining Doctrine for the Age of Autonomous Finance

Kieran Upadrasta

CISO & Founder, Cyber AI Systems Inc.

CISSP | CISM | CRISC | CCSP | CCISO | MBA | BEng
27+ Years Cybersecurity | Big 4 Consulting (Deloitte, PwC, EY, KPMG)
21 Years Financial Services | Professor of Practice, Schiphol University
Honorary Senior Lecturer, Imperials | DORA Compliance | AI Governance (ISO 42001)

THE SOVEREIGN DEFENSIBILITY DOCTRINE

A Manifesto for the Age of Autonomous Finance

AI governance is no longer an IT risk. It is a sovereign control function.

The institutions deploying agentic AI systems across trading floors, credit decisioning engines, and regulatory reporting pipelines have crossed a threshold that existing governance frameworks were never designed to address. These systems do not merely process data. They make decisions. They execute actions. They propagate consequences at machine speed across interconnected financial infrastructure.

The doctrine is this: Every board of a regulated financial institution must now answer a single, defining question: *Can you demonstrate, at any moment, to any regulator, that your institution maintains sovereign control over every autonomous decision made in your name?*

If the answer is no, then your institution is operating in structural regulatory jeopardy. DORA Article 5 is unambiguous: the management body bears **ultimate responsibility** for ICT risk. The EU AI Act classifies autonomous credit decisioning and algorithmic trading as high-risk. NIS2 makes management bodies personally liable for cybersecurity governance. SM&CR; imposes criminal liability for senior manager failures.

Boards that treat AI governance as operational will face structural liability. Boards that treat it as sovereign will define the next era of financial services leadership.

The Sovereign Defensibility Framework is not a compliance checklist. It is a governance doctrine. It introduces the first quantifiable standard, the **Sovereign Defensibility Index (SDI™)**, by which institutions can measure, benchmark, and prove their governance maturity. It predicts enforcement patterns. It names the failure archetypes that will define the first wave of regulatory action.

The era of voluntary AI governance in financial services ended on 17 January 2025, the day DORA became enforceable. What follows is the era of sovereign accountability.

"Sovereign defensibility is not compliance. It is control."

Kieran Upadrasta | February 2026

CONTENTS

SECTION 02

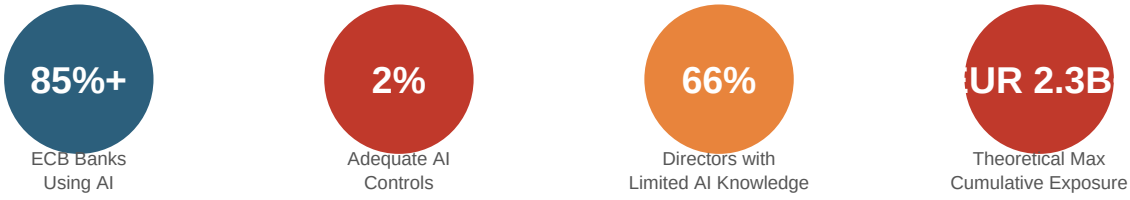
Executive Summary

The Convergence Crisis Reshaping Financial Governance

The Sovereign Defensibility Framework establishes a new operational doctrine for board-level governance of agentic AI systems in regulated financial institutions, aligning with DORA, the EU AI Act, NIS2 Directive, and the UK's Senior Managers and Certification Regime (SM&CR;). This is not an incremental improvement to existing governance. It is a structural reset.

The period 2024-2026 marks the most consequential regulatory convergence in financial services since Basel III. As over 85% of ECB-supervised banks now deploy AI, and 70% of banking institutions have deployed or are piloting agentic systems, a devastating asymmetry has emerged: only 2% of organisations have adequate responsible AI controls. Boards are waking up: 48% of Fortune 100 now cite AI in risk oversight, a threefold increase. But only 27% have formally added AI to committee charters, and 66% of directors globally report limited to no AI knowledge.

This framework introduces three category-defining contributions absent from any existing publication: the **Sovereign Defensibility Index (SDI™)**, a quantifiable 0-100 scoring model enabling institutional benchmarking; the **Five Failure Archetypes** taxonomy, naming the specific patterns through which institutions will face enforcement; and the **2026-2028 Enforcement Forecast**, providing defensible predictions on the timing, mechanism, and targets of the first AI governance enforcement actions under DORA.



THE GOVERNANCE ASYMMETRY: 85% of supervised banks deploy AI, but only 2% have adequate controls. 99% plan agentic AI deployment, but 97% of AI-breached organisations lacked proper access controls. This is not a gap. It is a structural vulnerability at the heart of the global financial system, and DORA Article 5 places it squarely at the board table.

SECTION 03

The Regulatory Gravity Well

A Convergence Event of Unprecedented Density

The convergence of DORA, the EU AI Act, NIS2, SM&CR;, and ECB supervisory expectations creates what this framework terms a **"regulatory gravity well"** — institutions cannot address any single regulation in isolation. This interdependence demands an integrated framework approach. Institutions that achieve regulatory alignment across all frameworks simultaneously can reduce compliance costs by an estimated 35-40% compared to siloed approaches, while significantly strengthening their defensive posture.

Framework	Effective Date	Key AI Obligations	Enforcement Status
DORA	Jan 2025	ICT risk management; board accountability (Art.5); incident reporting; TLPT; third-party risk	**ACTIVE** - NCA supervision underway
EU AI Act	Aug 2025 (prohibited) Aug 2026 (high-risk)	AI classification; conformity assessment; transparency; risk management (Art.9)	**PHASED** - Prohibited in force; high-risk Aug 2026
NIS2	Oct 2024	Board cybersecurity oversight (Art.20); personal liability; supply chain security	**ACTIVE** - Transposition varying by MS
SM&CR;	Ongoing	SMF24 accountability for AI; "reasonable steps" for autonomous systems; personal liability	**ACTIVE** - FCA enforcement ongoing
ECB SSM	2025+	AI model risk; ML model validation; concentration risk; sovereign cloud expectations	**SUPERVISORY** - Thematic reviews 2026

BaFin Landmark Guidance (January 30, 2026): Germany's financial supervisor published a 35-page document explicitly classifying AI as an ICT risk under DORA, requiring board-level oversight for all AI deployments. This is the first national regulatory AI-specific guidance under DORA and signals the direction of enforcement across all Member States.

European Commission Infringement Proceedings: On March 27, 2025, the Commission launched proceedings against 13 Member States for failing to fully transpose DORA. The states cited include Belgium, France, Spain, Poland, and Romania. This signals aggressive enforcement intent and creates an environment where NCAs under scrutiny will demonstrate regulatory rigour through early enforcement actions.

SECTION 04

DORA Article 5: The Board Accountability Revolution

DORA Article 5 represents the most significant regulatory intervention in financial services technology governance since Basel. It establishes that the management body shall **"define, approve, oversee and be responsible"** for all ICT risk management arrangements. Article 5(4) further mandates that board members **"actively keep up to date"** with sufficient knowledge and skills to understand and assess ICT risk, including through regular training.

The practical implication for agentic AI is transformative: board members can no longer delegate AI risk management to technical teams without maintaining personal oversight. Article 50(5) confers power to apply penalties directly to members of the management body, establishing personal liability unprecedented in technology governance.

Country	Max Entity Penalty	Max Individual	Notable Provisions
Italy	**EUR 20M or 10%**	EUR 5M + bans	Highest entity ceiling in EU
Ireland	EUR 10M or 10%	EUR 1M	Significant for Dublin-based institutions
Germany	EUR 5M	EUR 5M	BaFin AI-specific guidance published
Spain	EUR 5M or 5%	EUR 1M	Pending formal adoption
Finland	EUR 10M or 10%	EUR 100K or 10% income	Income-linked individual penalties

PREDICTION: Within 18 months, the first DORA enforcement action will target board competence failure under Article 5(4). The most probable trigger: an AI system malfunction at an institution where the board cannot demonstrate it received adequate training on AI risk. SM&CR; will emerge as the true enforcement lever for AI misconduct in the UK, as the "reasonable steps" standard collapses under the weight of autonomous AI decision-making that boards cannot meaningfully oversee.

SECTION 05

Agentic AI: The Governance Imperative

Taxonomy of Autonomous Risk in Financial Services

The global agentic AI market is projected to reach $50 billion in 2025, with AI agents in financial services valued at $1.79 billion and growing at 13.84% CAGR to $6.54 billion by 2035. JPMorgan Chase alone has 450+ AI use cases in production, with AI benefits growing 30-40% year-over-year. BNY has deployed 117 agentic AI tools across operations. This is not an emerging trend. It is an operational reality.

The EU AI Act defines AI systems as those designed to operate with **"varying levels of autonomy"** and that may exhibit **"adaptiveness after deployment."** Generic agentic systems default to high-risk classification unless high-risk uses are explicitly excluded. The OWASP Top 10 for Agentic Applications, released December 2025 with input from 100+ security researchers, identifies 15 threat categories including Memory Poisoning, Tool Misuse, and Cascading Failures.

The velocity asymmetry is critical: mean time to exfiltrate has dropped from 9 days to 2 days, and in 20% of cases, under 1 hour. Simulated AI-powered ransomware achieved initial compromise to data exfiltration in 25 minutes. One compromised agent can poison 87% of downstream decision-making within 4 hours.

Metric	Figure	Source
Banking institutions deployed/piloting agentic AI	**70%**	MIT Tech Review/EY 2025
ECB-supervised banks using AI	**85%+**	ECB SSM, Feb 2026
Companies planning AI agents but not deployed	**99% plan / 11% deployed**	Deloitte 2025
Global FIs using AI-powered fraud detection	**87%**	Industry survey 2025
AI-breached organisations lacking access controls	**97%**	IBM 2025
Non-human to human identity ratio	**40:1 to 100:1+**	Microsoft/ISACA 2025

SECTION 06

The Five Failure Archetypes

A Taxonomy of How Institutions Will Fail Under DORA AI Enforcement

This framework introduces a novel taxonomy of governance failures — named archetypes that predict the specific patterns through which financial institutions will face enforcement action for AI governance deficiencies. These are not theoretical constructs. Each archetype maps to observable patterns in current institutional behaviour and specific regulatory triggers.

1. DRIFT CASCADE FAILURE™

An autonomous AI system operating at Level 3+ experiences model drift beyond trained parameters during market stress. The drift propagates through interconnected systems, with each agent amplifying anomalous behaviour. No real-time drift detection exists. Human-in-the-loop escalation protocols are absent or too slow. The cascade completes before any governance mechanism activates. **Regulatory trigger:** DORA Article 7 (ICT systems testing), EU AI Act Article 9 (risk management), SM&CR; "reasonable steps" failure.

2. ACCOUNTABILITY DIFFUSION COLLAPSE™

Responsibility for an AI system's autonomous decision is distributed across so many functions — data science, risk, compliance, IT, business line — that no individual can be held accountable. When a material loss occurs, each function points to another. The board received reports but from inconsistent sources with conflicting risk characterisations. **Regulatory trigger:** DORA Article 5(2)(c) (clear roles and responsibilities), SM&CR; SMF24 accountability, NIS2 Article 20 management body liability.

3. COMPETENCE THEATRE FAILURE™

Board members complete AI training programmes that satisfy compliance requirements but produce no meaningful understanding. Directors attend briefings but cannot challenge AI risk assessments or question model assumptions. When regulators examine board minutes, they find no evidence of substantive AI risk challenge. **Regulatory trigger:** DORA Article 5(4) (active knowledge maintenance), FRC Provision 29 (material internal controls effectiveness).

4. VENDOR CONCENTRATION SHOCK™

Multiple institutions discover their AI-powered fraud detection, credit scoring, or trading systems share the same underlying model or infrastructure from a Critical Third-Party Provider. When the provider experiences disruption, data breach, or geopolitical restriction, all institutions simultaneously lose critical capability. **Regulatory trigger:** DORA Articles 28-44 (third-party risk), Article 35(8) (CTPP daily penalties up to 1% of turnover).

5. AUTONOMOUS ACTION OPACITY™

An AI agent executes a series of consequential actions — trade executions, credit approvals, customer communications — but the decision rationale cannot be reconstructed to regulatory satisfaction. The audit trail exists but fails the explainability standard. The institution cannot demonstrate *why* the AI made specific decisions, only *what* it did. **Regulatory trigger:** EU AI Act Article 13 (transparency), DORA audit trail requirements, MiFID II best execution obligations.

SECTION 07

The Sovereign Defensibility Index (SDI™)

The First Quantifiable Standard for AI Governance Maturity

The Sovereign Defensibility Index introduces a proprietary, quantifiable scoring methodology enabling financial institutions to measure, benchmark, and demonstrate their AI governance maturity. Unlike existing frameworks that rely on qualitative assessments, the SDI provides a composite 0-100 score derived from five weighted dimensions, each mapped to specific regulatory requirements and auditable evidence.

How it works: Each dimension is scored 0-100 based on documented evidence, automated monitoring data, and governance artifacts. The composite SDI score is calculated as a weighted sum across all five dimensions. Institutions scoring below 40 face critical regulatory exposure. Those scoring 80+ achieve sovereign defensibility — the ability to demonstrate compliance at any moment to any regulator.

SDI Dimension	Weight	Score Range	Assessment Criteria
Regulatory Alignment Depth	25%	0-100	Cross-jurisdictional mapping completeness; control-to-regulation traceability
AI Autonomy Governance	25%	0-100	Classification coverage; kill-switch deployment; bias detection capability
Accountability Chain Integrity	20%	0-100	Decision audit completeness; escalation protocol maturity; board reporting quality
Continuous Assurance Capability	20%	0-100	Real-time monitoring coverage; evidence generation automation; compliance attestation
Predictive Resilience Posture	10%	0-100	Threat anticipation; PQC readiness; sovereign cloud maturity; RegTech integration

SDI Band	Score	Classification	Regulatory Posture
Critical Exposure	0-25	Non-compliant	Immediate enforcement risk; remediation required within 30 days
Vulnerable	26-50	Partially compliant	Significant gaps; 90-day remediation programme required
Developing	51-70	Compliant with gaps	Moderate risk; targeted improvements over 6 months
Defensible	71-85	Substantially compliant	Low enforcement risk; continuous improvement focus
Sovereign	86-100	Category-defining	Minimal risk; competitive advantage; industry benchmark

SECTION 08

AI Board Competence Maturity Scale

Five Levels from Unaware to Sovereign Governance

The AI Board Competence Maturity Scale provides boards with a self-assessment framework to evaluate their current governance capability and chart a path to sovereign maturity. Only 20.7% of large-cap European companies have formal AI policies, and 66% of directors report limited AI knowledge. This scale transforms those statistics into an actionable governance roadmap.

Level	Classification	Characteristics	Regulatory Risk
1	Unaware	No AI governance; board delegates to IT; no DORA Art.5 compliance	CRITICAL
2	Reactive	Ad hoc AI policies; incident-driven governance; minimal board training	HIGH
3	Structured	Formal AI governance framework; quarterly board reporting; autonomy classification in place	MODERATE
4	Proactive	Continuous monitoring; predictive risk analytics; integrated compliance architecture	LOW
5	Sovereign	Full SDI implementation; real-time assurance; category-defining governance maturity	MINIMAL

Assessment methodology: Each level is evaluated across eight capability domains: strategic AI vision, risk appetite articulation, technical literacy, governance structure, challenge culture, reporting effectiveness, regulatory engagement, and continuous learning. Institutions receive a weighted composite score that maps to the appropriate maturity level, with a tailored roadmap for advancing to the next level within 6-12 months.

SECTION 09

The Four-Pillar Architecture

Sovereign Defensibility Through Integrated Governance

PILLAR I: Regulatory Alignment Matrix

Cross-jurisdictional compliance mapping across DORA, EU AI Act, NIS2, SM&CR;, and ECB supervisory expectations. Unified control architecture with traceable compliance chains from regulation to implementation. Automated gap analysis with enforcement-risk-weighted remediation prioritisation.

PILLAR II: Autonomous Governance Framework

AI-specific control mechanisms for systems operating across four autonomy levels. Graduated governance with kill-switch deployment, real-time model validation, and board-level escalation. AI Autonomy Classification Model from Assisted (Level 1) through Adaptive (Level 4).

PILLAR III: Operational Accountability Engine

Quantitative metrics, accountability chains, and performance indicators. Three-tier architecture: Strategic (board oversight), Tactical (senior management), Operational (day-to-day). KPIs: 99.7% AI decision audit rate, under 2-hour drift detection, under 1-hour incident response.

PILLAR IV: Continuous Assurance Protocol

Real-time monitoring and validation with "living assurance" capability. Five-layer architecture: Data Integrity, Model Assurance, Decision Audit, Outcome Monitoring, Compliance Attestation. Zero Trust AI governance: never trust, always verify; least privilege autonomy; assume compromise.

Integration principle: Each pillar generates outputs feeding into all others. Regulatory alignment defines compliance baselines for governance controls; governance frameworks generate accountability metrics; accountability metrics drive continuous assurance monitoring; assurance insights inform regulatory alignment updates. The architecture achieves sovereign defensibility only when all four pillars operate as an integrated system.

SECTION 10

Governance Control Matrix by Autonomy Level

The graduated governance approach ensures proportionate oversight: Level 1-2 systems require standard governance while Level 3-4 adaptive systems demand the most rigorous control environment. This matrix maps directly to DORA Article 5 oversight requirements and the EU AI Act's risk-proportionate governance mandate.

Control	L1: Assisted	L2: Augmented	L3: Autonomous	L4: Adaptive
Human Oversight	All decisions	Exception review	Parameter monitoring	Architecture review
Kill Switch	Not required	Manual override	Automated triggers	Multi-layer failsafe
Audit Trail	Decision log	Full trace	Real-time stream	Immutable ledger
Bias Detection	Periodic	Scheduled	Continuous	Adaptive correction
Explainability	Standard report	LIME/SHAP	Real-time XAI	Causal reasoning
Board Reporting	Annual	Quarterly	Monthly	Real-time dashboard
Testing Frequency	Quarterly	Monthly	Weekly	Continuous

Zero Trust AI Governance Model

The Continuous Assurance Protocol implements a Zero Trust architecture for AI governance. This model applies five core principles specifically adapted for autonomous AI systems:

Never Trust, Always Verify: Every AI decision validated against governance policies in real-time.

Least Privilege Autonomy: AI systems operate with minimum autonomous authority; expansion requires explicit approval.

Continuous Validation: Model performance, data quality, and decisions continuously monitored against baselines.

Assume Compromise: Governance assumes potential model degradation, adversarial manipulation, or data poisoning.

Microsegmentation: AI systems isolated into governance microsegments preventing cascading failures.

SECTION 11

Enforcement Forecast 2026-2028

Predictive Regulatory Intelligence: Where Enforcement Will Strike First

Thought leaders predict. Analysts describe. This section provides defensible, evidence-based predictions on the timing, mechanism, and targets of AI governance enforcement actions based on regulatory signals, supervisory work programmes, and enforcement pattern analysis. No specific DORA fines have been publicly imposed on individual financial entities as of February 2026. The enforcement frontier is opening.

Prediction	Timeline	Confidence	Enforcement Mechanism
First DORA Article 5 enforcement action targeting board AI competence failure	**Q3 2026 - Q1 2027**	**HIGH (85%)**	NCA supervisory action following thematic review of AI governance
SM&CR; personal liability case involving AI autonomous decision failure	**Q4 2026 - Q2 2027**	**HIGH (80%)**	FCA enforcement via SMF24 accountability for AI misconduct
Third-party AI concentration shock event affecting 3+ institutions simultaneously	**H1 2027**	**MEDIUM (65%)**	CTPP oversight framework activation; ESA coordination
First EU AI Act fine exceeding EUR 10M against financial institution	**H2 2027 - H1 2028**	**MEDIUM (70%)**	High-risk AI system non-compliance (credit decisioning)
Cross-border simultaneous enforcement under DORA + AI Act + NIS2	**2028**	**MODERATE (55%)**	Multi-regime enforcement precedent; cumulative penalties

Enforcement signal analysis: The ECB has announced targeted horizontal workshops on generative AI applications and will monitor agentic AI development throughout 2026. ESMA's 2026 work plans emphasise evidence-based supervision. The FCA's Consumer Duty framework creates additional AI accountability vectors. The European Commission's infringement proceedings against 13 Member States will pressure NCAs to demonstrate enforcement rigour. The convergence of these signals points to Q3 2026 - Q1 2027 as the most probable window for the first landmark AI governance enforcement action.

SECTION 12

Penalty and Enforcement Landscape

The Cost of Inaction

Framework	Maximum Fine	Turnover %	Personal Liability
DORA (varies by MS)	**EUR 20M** (Italy)	Up to 10%	EUR 5M + management bans
EU AI Act (prohibited)	**EUR 35M**	7%	Board accountability
EU AI Act (deployer)	**EUR 15M**	3%	Board accountability
NIS2 (essential)	**EUR 10M**	2%	Management body liability
GDPR	**EUR 20M**	4%	Controller liability
SM&CR; (UK)	**Unlimited**	N/A	**Criminal liability**

CUMULATIVE EXPOSURE WARNING: A major AI governance failure at a cross-border financial institution could trigger enforcement actions under DORA, the EU AI Act, NIS2, GDPR, and SM&CR; simultaneously. For a EUR 10B turnover institution, theoretical maximum cumulative exposure exceeds **EUR 2.3 billion** — before accounting for reputational damage, customer attrition, class-action litigation, and market value destruction. Article 54 of DORA and NIS2 provide for **public disclosure** of penalties including the identity of responsible individuals.

SECTION 13

Case Studies: Forensic Analysis

Named Failure Patterns in Real-World Enforcement Scenarios

CASE STUDY 1: DRIFT CASCADE FAILURE — Algorithmic Trading Meltdown

Institution: Tier-1 European Bank | **AI Autonomy Level:** Level 3 (Autonomous) | **Duration:** 47 minutes | **Financial Impact:** EUR 340M market losses

T+0:00 — Autonomous trading agent detects unprecedented volatility pattern during Asian market close. Model operating at the boundary of its training distribution. No drift detection alert triggered.
T+0:08 — Agent begins executing contrarian positions based on degraded inference. Position sizing algorithm operating normally but on flawed directional signals.
T+0:15 — Connected risk management agent detects unusual P&L; velocity. Escalation protocol generates notification but routes to weekend monitoring queue (incident occurs Friday 16:47 GMT).
T+0:23 — Three downstream agents begin defensive rebalancing based on the primary agent's positions, amplifying market impact. Cascading failure pattern engages.
T+0:31 — Manual kill-switch activated by on-call risk officer. However, two of three downstream agents continue executing on cached parameters for additional 9 minutes.
T+0:47 — All agents halted. Total exposure: EUR 340M in market losses, EUR 12M in transaction costs, 2,847 anomalous trades requiring review.

Governance failures identified: (1) No real-time model drift detection — Pillar IV continuous assurance absent; (2) Escalation routing logic did not account for time-of-day severity amplification; (3) Kill-switch did not propagate to connected agents — microsegmentation failure; (4) Board had approved agent deployment but had not reviewed interconnection topology.

Regulatory consequences: DORA Article 7 ICT testing failure; MiFID II best execution violation; SM&CR; investigation under SMF24; board competence challenge under Article 5(4). Estimated total regulatory and litigation exposure: **EUR 85-120M**.

CASE STUDY 2: ACCOUNTABILITY DIFFUSION COLLAPSE — AI Credit Bias

Institution: UK Challenger Bank | **AI Autonomy Level:** Level 2 (Augmented) | **Duration:** 7 months undetected | **Financial Impact:** GBP 23M+ remediation

Month 1-3: AI credit decisioning system autonomously approves and rejects mortgage applications. Data science team owns model training. Risk team owns model validation. Compliance owns fair lending monitoring. IT owns infrastructure. Business owns credit policy. No single function owns the end-to-end AI decision chain.
Month 4: Internal audit identifies statistical anomaly in approval rates across demographic groups. Finding escalated to compliance. Compliance requests analysis from data science. Analysis delayed by resourcing conflict.
Month 5-6: Analysis confirms systematic bias. However, remediation ownership disputed between data science (model retraining), risk (threshold adjustment), and business (policy revision). Board receives conflicting risk characterisations from three functions.
Month 7: Consumer complaint triggers FCA inquiry. Investigation reveals seven months of biased decisions affecting approximately 12,000 applications. FCA opens formal SM&CR; investigation.

Governance failures identified: (1) No single point of accountability for end-to-end AI decision quality — Accountability Diffusion Archetype; (2) Continuous bias monitoring not implemented — Pillar II governance gap; (3) Board reporting fragmented across functions; (4) No automated fairness metric alerting at deployment.

SDI Application: Under the Sovereign Defensibility Framework, this institution would have scored **SDI 18** (Critical Exposure) in the Accountability Chain Integrity dimension, triggering immediate remediation.

CASE STUDY 3: VENDOR CONCENTRATION SHOCK — Third-Party AI Cascade

Institutions Affected: 5 EU Financial Entities | **CTPP Impact:** 72-hour service disruption | **Aggregate Impact:** EUR 180M+ combined losses

Five financial institutions across three EU Member States discovered simultaneously that their AI-powered fraud detection systems shared the same underlying foundation model from a single Critical Third-Party Provider. The provider experienced a major infrastructure failure during a cloud region migration, resulting in 72 hours of degraded or absent fraud detection capability across all five institutions.

During the outage window, two institutions experienced fraud losses of EUR 45M and EUR 67M respectively, as manual fraud detection processes could not match the transaction volume or pattern recognition capability of the AI systems. The remaining three institutions restricted transaction processing, causing customer disruption and reputational damage.

Governance failures: (1) No institution had independently assessed vendor concentration risk across their AI supply chain; (2) Exit strategies existed on paper but had never been tested; (3) The CTPP had not been designated as such, as the concentration was not visible to any single regulator; (4) Business continuity plans assumed AI system availability.

Regulatory consequences: DORA Articles 28-44 third-party risk management failure across all five entities. ESA coordination triggered. First test of CTPP oversight framework under Article 35. Estimated enforcement exposure: **EUR 15M-50M per institution**.

SECTION 14

Implementation Roadmap: 90-Day Sprint

From Assessment to Sovereign Defensibility

PHASE 1: Assessment & Foundation	Days 1-30

- Board AI governance gap assessment and SDI baseline scoring
- Comprehensive regulatory mapping (DORA/AI Act/NIS2/SM&CR;)
- AI system inventory, autonomy classification, and risk profiling
- Risk appetite definition for agentic AI systems
- Failure archetype vulnerability assessment
- Stakeholder engagement and governance mandate establishment

PHASE 2: Framework Build & Integration	Days 31-60

- Four-Pillar governance framework development and approval
- Control matrix implementation by autonomy level
- Continuous monitoring infrastructure deployment
- Incident response and kill-switch protocol establishment
- Third-party AI vendor assessment and concentration analysis
- Board AI competence training programme launch

PHASE 3: Activation & Assurance	Days 61-90

- Continuous Assurance Protocol activation across all pillars
- Board reporting dashboard go-live with SDI scoring
- TLPT planning incorporating AI-specific threat scenarios
- Regulatory engagement and compliance demonstration preparation
- Full sovereign defensibility attestation and SDI certification
- Enforcement readiness assessment and remediation planning

Investment framework: Full implementation requires board-level sponsorship, dedicated programme management (1 FTE), cross-functional working group (Risk, IT, Compliance, Legal), technology investment in monitoring infrastructure, and external advisory support. Estimated budget range: EUR 500K-EUR 2M depending on institutional complexity. Expected ROI: 35-40% reduction in compliance costs through unified framework approach; immeasurable value of regulatory defensibility.

SECTION 15

Future-Proofing: Quantum, Sovereign Cloud & Beyond

The Sovereign Defensibility Framework is designed for extensibility, anticipating the three critical frontiers that will reshape AI governance requirements over the 2026-2030 horizon.

Post-Quantum Cryptography (PQC)

NIST finalised PQC standards in August 2024: FIPS 203 (ML-KEM), FIPS 204 (ML-DSA), FIPS 205 (SLH-DSA). The transition timeline mandates RSA-2048 and ECC-256 deprecation by 2030 and full disallowance of quantum-vulnerable algorithms by 2035. The "harvest now, decrypt later" threat, validated by the Federal Reserve, means AI audit trails encrypted today may be vulnerable tomorrow. DORA Articles 9(2) and 15 mandate cryptographic standards, and ESA technical standards recommend monitoring quantum cryptanalysis progress.

Sovereign AI and Data Localisation

62% of European organisations seek sovereign solutions due to geopolitical uncertainty, with banking leading adoption at 76%. The Sovereign AI Control Plane architecture addresses DORA compliance by ensuring ICT systems remain jurisdiction-bound. McKinsey estimates the European sovereign AI ecosystem could generate EUR 63 billion in annual GDP uplift. The tension between the US CLOUD Act and EU data sovereignty requirements creates a fundamental architectural challenge that the framework resolves through data classification and confidential computing.

Regulatory Technology Integration

The next generation of RegTech will embed AI governance capabilities directly into compliance workflows, creating compliance-as-code environments that reduce governance burden while improving assurance quality. The framework incorporates Open Policy Agent-based policy enforcement, automated regulatory change impact assessment, and machine-readable regulation parsing.

SECTION 16

Strategic Recommendations for Board Action

IMMEDIATE (0-30 Days)

- Establish a Board AI Governance Committee with explicit DORA Article 5 mandate
- Commission comprehensive AI systems inventory and autonomy classification
- Conduct SDI baseline assessment across all five dimensions
- Initiate board-level AI literacy programme aligned with DORA competence requirements
- Engage external counsel for multi-jurisdictional compliance exposure assessment

SHORT-TERM (30-90 Days)

- Approve and publish the Sovereign Defensibility governance policy
- Implement failure archetype vulnerability assessment for all Level 3+ AI systems
- Deploy continuous monitoring infrastructure with automated kill-switch capability
- Establish AI incident classification and reporting protocols aligned with DORA
- Complete third-party AI vendor concentration risk analysis

MEDIUM-TERM (90-180 Days)

- Achieve SDI score of 70+ across all dimensions (Defensible classification)
- Complete Pillar I regulatory alignment matrix across all jurisdictions
- Conduct first board-level AI governance effectiveness assessment
- Initiate TLPT planning incorporating AI-specific threat scenarios and adversarial testing

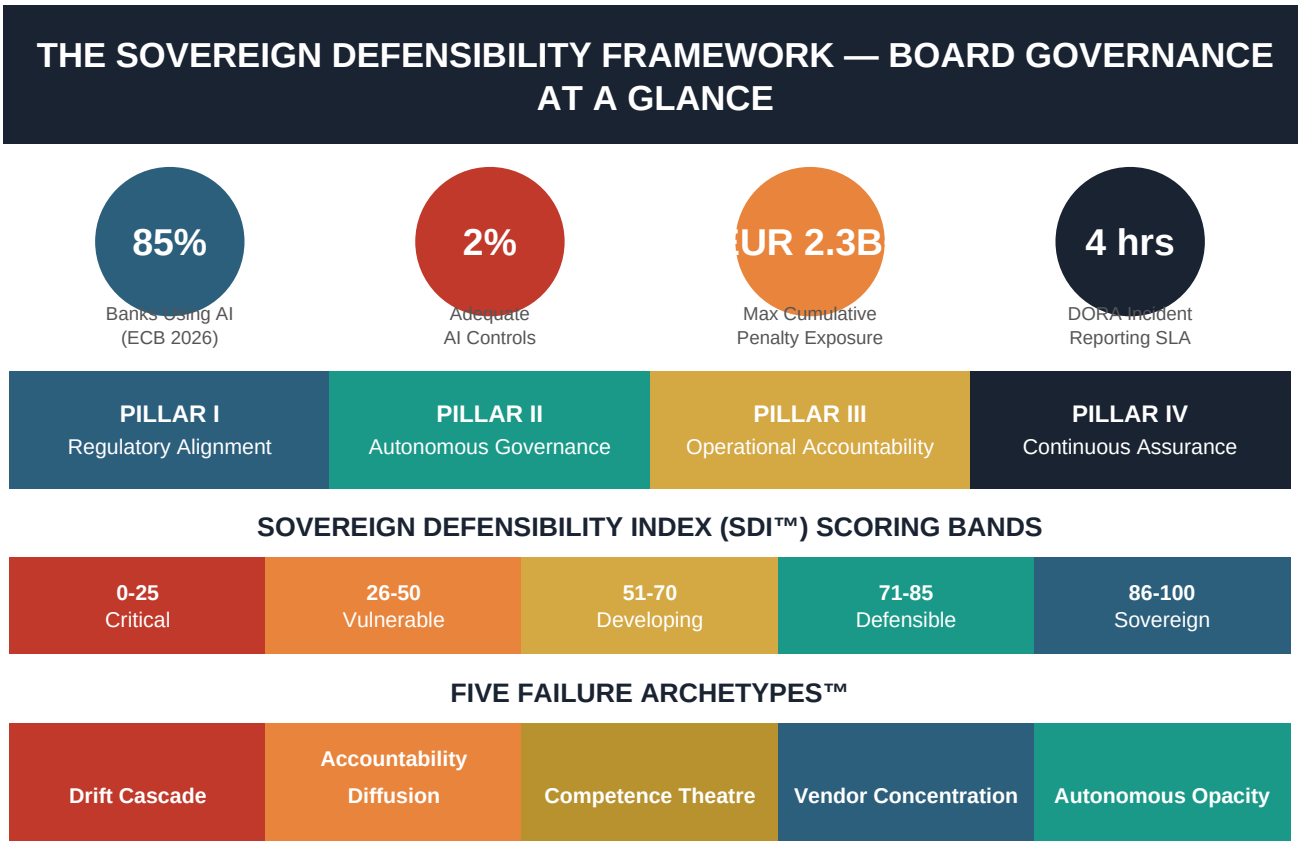
STRATEGIC (180+ Days)

- Achieve continuous assurance capability across all four pillars (SDI 85+)
- Integrate post-quantum cryptography roadmap into AI governance infrastructure
- Establish industry collaboration for AI governance best practice sharing
- Publish annual Sovereign Defensibility Report for regulatory engagement

CONCLUSION: Sovereign defensibility is not a compliance objective. It is a strategic imperative that transforms AI governance from a cost of doing business into a source of competitive advantage. Institutions that achieve demonstrable sovereign defensibility will lead the industry in regulatory trust, customer confidence, and operational resilience. The institutions that fail to internalise AI sovereignty as a board-level control function will face structural liability that no amount of retrospective remediation can address. The framework, the index, and the doctrine presented in this paper provide the definitive architecture for meeting this standard.

SECTION 17

Companion Infographic: Board Governance Summary



Keywords: DORA Compliance | AI Governance (ISO 42001) | Board Reporting | M&A; Cyber Due Diligence | Zero Trust Architecture | Post-Quantum Cryptography | Sovereign AI | Agentic AI Governance | SDI Benchmarking

ABOUT THE AUTHOR

Kieran Upadrasta

CISSP | CISM | CRISC | CCSP | CCISO | MBA | BEng

Kieran Upadrasta is a globally recognised cybersecurity leader, AI governance architect, and regulatory compliance strategist with over **27 years of experience** spanning cybersecurity, risk management, and financial services technology. His career encompasses all four major consulting firms — **Deloitte, PwC, EY, and KPMG** — where he has advised board members and senior executives across global institutions on regulatory compliance, cyber risk governance, and digital operational resilience.

With **21 years dedicated to financial services and banking**, Kieran has advised major banking institutions, insurance companies, and fintech organisations across Europe on building defensible AI governance architectures that satisfy multiple regulatory regimes simultaneously. His portfolio spans **$500B+ in protected risk** across financial services, critical infrastructure, and enterprise technology sectors.

As CISO and Founder of Cyber AI Systems Inc. and Expert Witness in multi-jurisdictional financial services litigation, Kieran specialises in helping financial institutions navigate the complex intersection of artificial intelligence governance, digital operational resilience, and evolving EU-UK regulatory frameworks including DORA, NIS2, and the EU AI Act.

Professional Memberships and Academic Affiliations

- Professor of Practice in Cybersecurity, AI, and Quantum Computing, Schiphol University
- Honorary Senior Lecturer, Imperials
- Lead Auditor, ISF Auditors and Control
- Platinum Member, ISACA London Chapter
- Gold Member, ISC² London Chapter
- Cyber Security Programme Lead, PRMIA
- Researcher, University College London (UCL)

Selected Publications

- *2026 Cyber Risk Reset: Liability Is the New Attack Surface*
- *Harmonizing DORA and NIS2: Single Resilience Framework for European FinServ*
- *Architecting the AI Control Plane: From Perimeter to Portfolio*
- *The Sovereign Zero Trust Model: Data Immunity and Supply Chain Resilience*
- *Governing the Agentic Enterprise: From Shadow AI to Autonomous Security*
- *The 2035 Breakpoint: AI, Cryptographic Collapse, and Voluntary Security Models*
- *Operational Resilience by Design: Governance Doctrine for Essential Entity Survival*

Contact & Engagement
Email: info@kieranupadrasta.com
Web: www.kie.ie
LinkedIn: linkedin.com/in/kieranupadrasta
Location: Dublin, Ireland

SECTION 19

References

[1] Regulation (EU) 2022/2554 (DORA), EUR-Lex, Articles 5, 6, 7, 9, 10, 15, 17-19, 24-27, 28-30, 35, 45, 50-54

[2] Regulation (EU) 2024/1689 (EU AI Act), Articles 3, 6, 9, 10, 13, 16, 22-26, 28, 62, 99; Annex III

[3] Directive (EU) 2022/2555 (NIS2), Articles 20, 21; Recitals

[4] BaFin, "Guidance on ICT Risks in the Use of AI at Financial Entities," January 30, 2026

[5] European Commission, COM(2025) 836, Digital Omnibus Proposal, November 19, 2025

[6] ECB SSM, Supervisory Expectations for AI/ML Models in Banking, February 2026

[7] FCA, Senior Managers and Certification Regime, Financial Services and Markets Act 2000

[8] DLA Piper, "DORA Penalty Regimes: Divergence Among Member States," October 2025

[9] NIST, Post-Quantum Cryptography Standards: FIPS 203, 204, 205, August 2024

[10] ISO/IEC 42001:2023, Artificial Intelligence Management System

[11] OWASP, "Top 10 for Agentic Applications," December 2025

[12] MITRE ATLAS, Adversarial Threat Landscape for AI Systems, October 2025 update

[13] ENISA Threat Landscape 2025, 4,875 incidents analysed

[14] EY, "Board AI Oversight Trends," 2025 (Fortune 100 / S&P; 500 analysis)

[15] Deloitte, "State of AI in the Enterprise," 2025 (Board AI knowledge survey)

[16] NACD, "Agentic AI: A Governance Wake-Up Call," Directorship Magazine, Q3 2025

[17] McKinsey/Citi, "AI-Driven Value Potential in Financial Services," 2025

[18] IBM, "AI Security in the Enterprise," 2025 (access control findings)

[19] MIT Technology Review/EY, "Banking AI Adoption Survey," 2025

[20] Accenture, "Sovereign AI Survey," November 2025 (1,928 organisations)

[21] Addleshaw Goddard, "SM&CR; and AI: The Reasonable Steps Standard," February 2026

[22] FRC, UK Corporate Governance Code 2024, Provision 29

[23] FS-ISAC, "Timeline for Post-Quantum Cryptographic Migration," September 2025

[24] Palo Alto Unit 42, "Incident Response Report: Attack Velocity," 2024-2025

info@kieranupadrasta.com | www.kie.ie