THE CISO AUTONOMY MANDATE

Command, Control, and Governance for Agentic AI Systems

Doctrine-Level Operating Model for Enterprise Autonomous System Governance

Kieran Upadrasta
CISSP | CISM | CRISC | CCSP | MBA | BEng
Principal Cyber Architect & CISO
Cyber AI Systems Inc.

27+ Years Cybersecurity Experience
21 Years Financial Services
All Big 4 Consulting Firms
Schiphol University | Imperials | UCL

Version 3.0 | March 2026 | COMMERCIALLY CONFIDENTIAL

EXECUTIVE DOCTRINE STATEMENT

"If it cannot be evidenced, it cannot be defended."

Agentic AI systems represent a governance discontinuity. These autonomous software entities plan, reason, call tools, and execute multi-step workflows across enterprise systems at machine speed. They are non-human operators with real-world impact. Existing cybersecurity frameworks were never designed to address autonomous agents that can chain tool calls, propagate decisions across systems, and create cascading failures without human intervention.

This white paper establishes the **CISO Autonomy Mandate**: a doctrine-level operating model that converts abstract AI governance into a commandable, enforceable, and measurable enterprise capability. It is not a collection of best practices. It is a command architecture designed to survive regulator scrutiny, win procurement contracts, and protect organisations from the operational reality of autonomous agents.

The core mandate for CISOs: **move from governing models to commanding autonomy**. This requires implementing decision rights over autonomous systems, enforceable controls at runtime, audit-grade evidence chains, and an incident command capability specialised for autonomous agents.

THE GOVERNANCE DEFICIT	THE LIABILITY ACCELERANT	THE AUTHORITY GAP	THE SPEED IMPERATIVE
90% deploy AI 1.5% have adequate governance headcount	EU AI Act: 7% turnover D&O settlements: $56M avg 1,100+ AI bills in US (2025)	64% CISOs report to IT 5% report to CEO 26-39 month tenure	DORA: 4-hour reporting Agents act at machine speed 40% agentic projects fail

Figure 1: The Four Forces Driving the CISO Autonomy Mandate

TABLE OF CONTENTS

1. The Agentic Inflection Point

The enterprise technology landscape has crossed a threshold. Agentic AI systems are no longer experimental curiosities confined to research laboratories. They are production-grade autonomous operators embedded in procurement workflows, customer service pipelines, code deployment chains, and critical infrastructure management. These agents plan, reason, call external tools, and execute multi-step workflows without continuous human oversight.

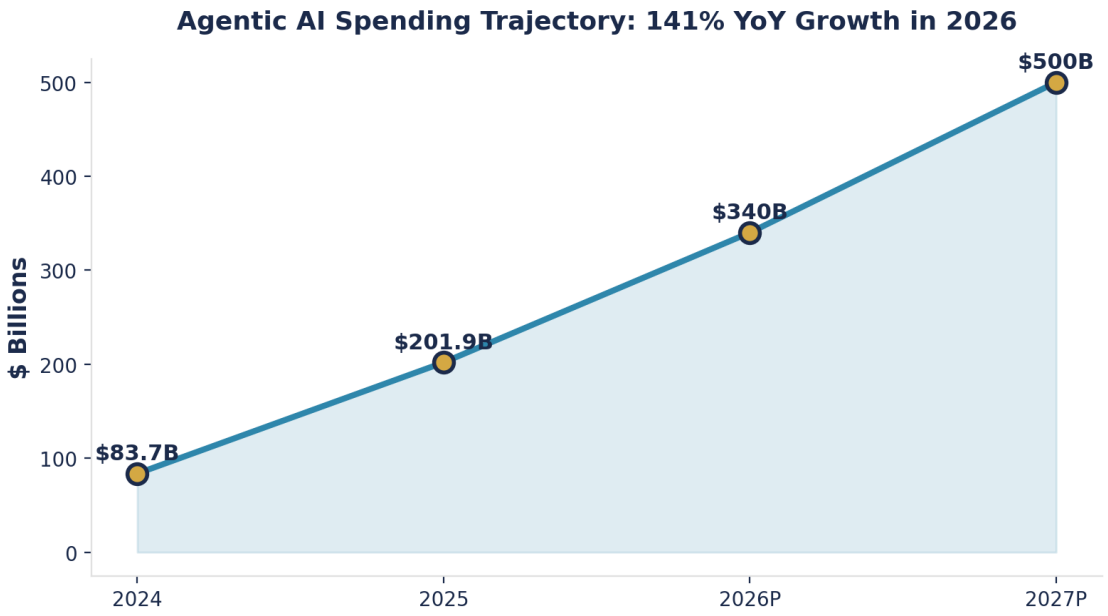

Figure 2: Agentic AI spending trajectory — 141% year-over-year growth projected for 2026 (Source: Gartner)

Gartner predicts 40% of enterprise software will feature task-specific AI agents by end of 2026, up from less than 5% in 2025 (Gartner Press Release, August 2025). PwC reports 79% of organisations have adopted AI agents to some extent (PwC 2025 Enterprise AI Survey, n=1,000). McKinsey projects agentic AI could unlock $2.6 trillion to $4.4 trillion annually across 60+ use cases (McKinsey Global Institute, 2025).

As the European Insurance and Occupational Pensions Authority (EIOPA) observed in its DORA implementation guidance: "The management body must ensure adequate organisational structures enabling the effective monitoring and management of ICT risk." The AIUC-1 Consortium confirmed in March 2026 that existing standards "do not address the specific technical controls that CISOs need for agentic deployments." These institutional assessments validate the structural gap this doctrine addresses.

1.1 Why Traditional Frameworks Fail

The governance gap is structural, not incremental. NIST CSF 2.0 does not address autonomy-specific controls for tool calling, agent identities, or prompt provenance. SP 800-61r3 modernises incident response but lacks ICS-style command doctrine for AI-specific failure modes. The EU AI Act establishes high-risk operational obligations but does not prescribe an enterprise operating model. NIS2 imposes staged reporting without addressing autonomous system containment. This white paper fills every one of these gaps.

IBM reports 13% of organisations already experienced breaches of AI models or applications, with 97% of those compromised lacking proper AI access controls (IBM Cost of a Data Breach Report, July 2025, n=604 organisations across 17 countries). Shadow AI — the average enterprise runs an estimated 1,200 unofficial AI applications (Layer Research, 2025) — adds $670,000 to the average breach cost. The

Amazon Q coding assistant breach demonstrated prompt injection directing AI to wipe local files. GitHub MCP integration vulnerabilities enabled hijacking of local AI agents to steal private repository source code.

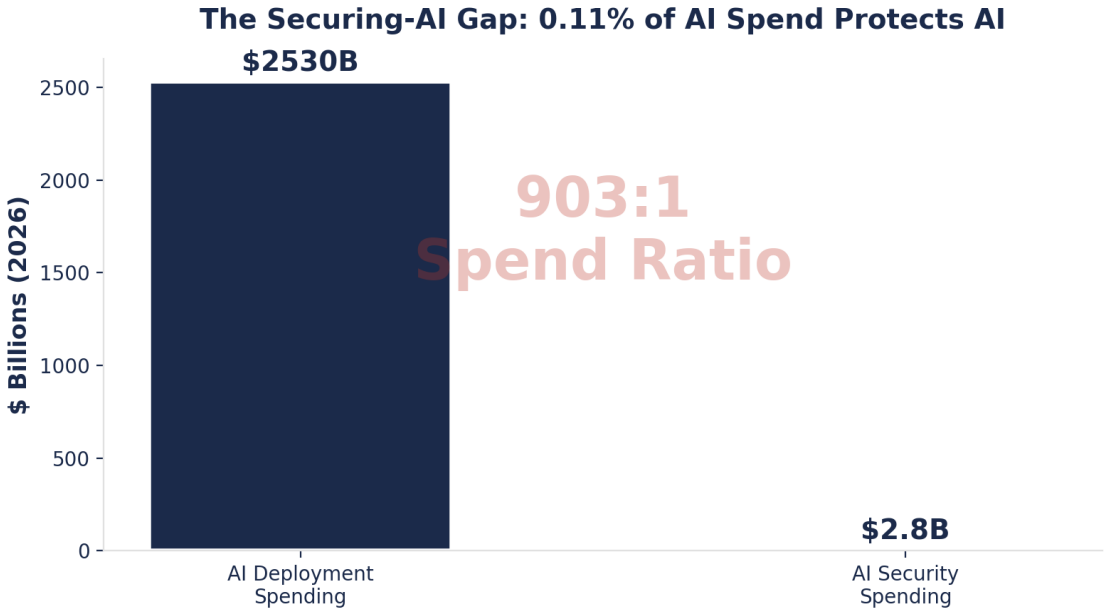

Figure 3: The Securing-AI Gap — derived from Gartner 2026 AI spending forecast ($2.53T) vs. MarketsandMarkets GenAI cybersecurity market sizing ($2.8B), yielding a 903:1 deployment-to-protection ratio

2. The Accountability-Authority Gap

The central tension animating this mandate is quantifiable. According to the IANS/Artico 2026 State of the CISO report surveying 830+ security leaders, 64% of CISOs still report into IT leadership (CIO/CTO). Only 5% report directly to the CEO, down from 11% in 2021. Two-thirds sit two organisational levels below the CEO. Yet every major regulatory framework enacted since 2022 holds these same leaders personally accountable for failures they lack the structural authority to prevent.

Figure 4: The CISO authority-accountability gap across five key dimensions

Board engagement statistics reveal a troubling illusion. While 95% of CISOs deliver regular board updates, airtime averages just 30 minutes. Only 30% of boards describe their CISO relationship as "strong and collaborative." Most critically, boardroom alignment with CISOs dropped from 84% in 2024 to 64% in 2025 — a 20-percentage-point collapse that IANS researchers describe as evidence that perceived influence was "partly illusory."

Average CISO tenure sits at 26-39 months depending on the survey, compared to 5.3 years for other C-suite roles. 88% report moderate to tremendous stress. The ISC2 workforce gap has grown to 4.8 million unfilled positions. The mandate is clear: CISO autonomy is no longer a career aspiration but a regulatory and operational necessity.

3. Regulatory Convergence: Four Regimes, One CISO

For the first time in regulatory history, four simultaneous frameworks impose personal accountability on security leaders while autonomous AI agents proliferate across enterprises. A single incident can trigger parallel reporting under DORA (4 hours), NIS2 (24 hours), GDPR (72 hours), and SEC (4 business days) simultaneously.

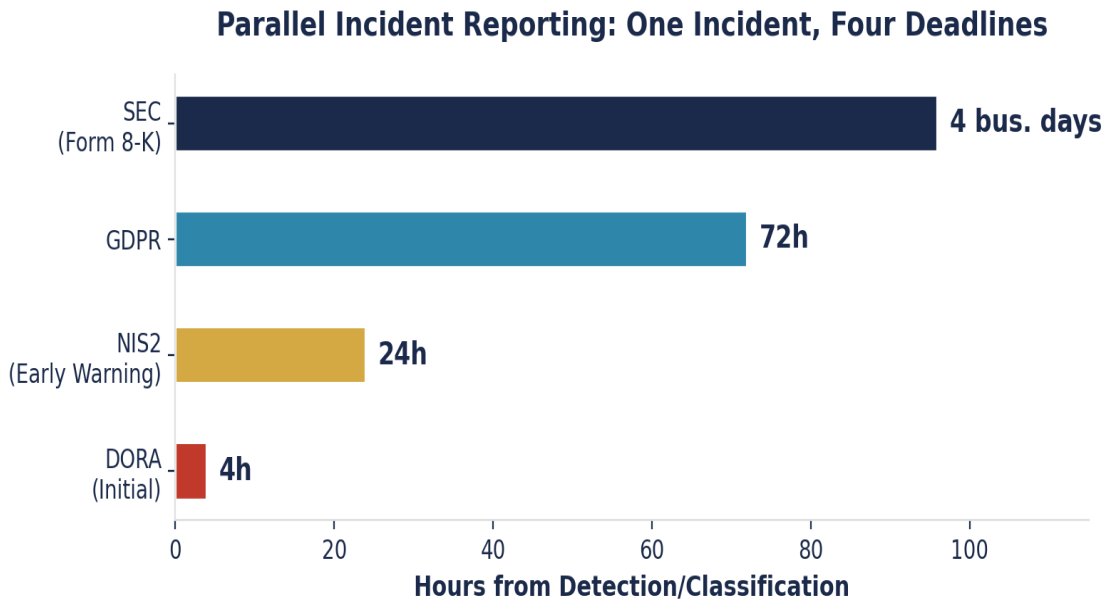

Figure 5: Parallel incident reporting deadlines — one incident triggers four regulatory clocks

3.1 DORA: The Most Prescriptive Governance Mandate

The Digital Operational Resilience Act became mandatory on 17 January 2025. Article 5 places "ultimate responsibility" for ICT risk management on the management body. DORA's incident reporting operates on the tightest timeline: initial notification within 4 hours of classifying an incident as major, intermediate report within 72 hours, final report within one month. Third-party ICT risk management extends governance to the AI vendor ecosystem. Penalties reach up to 1% of daily average worldwide income, imposed daily for up to six months.

3.2 NIS2: Personal Liability for Management Bodies

NIS2 mandates that management bodies formally approve and oversee cybersecurity risk-management measures and can be held personally liable for infringements. For essential entities, regulators can temporarily ban a CEO or legal representative from exercising managerial functions. Penalties reach €10 million or 2% of global annual turnover.

3.3 EU AI Act: The Most Severe Penalties of Any Digital Regulation

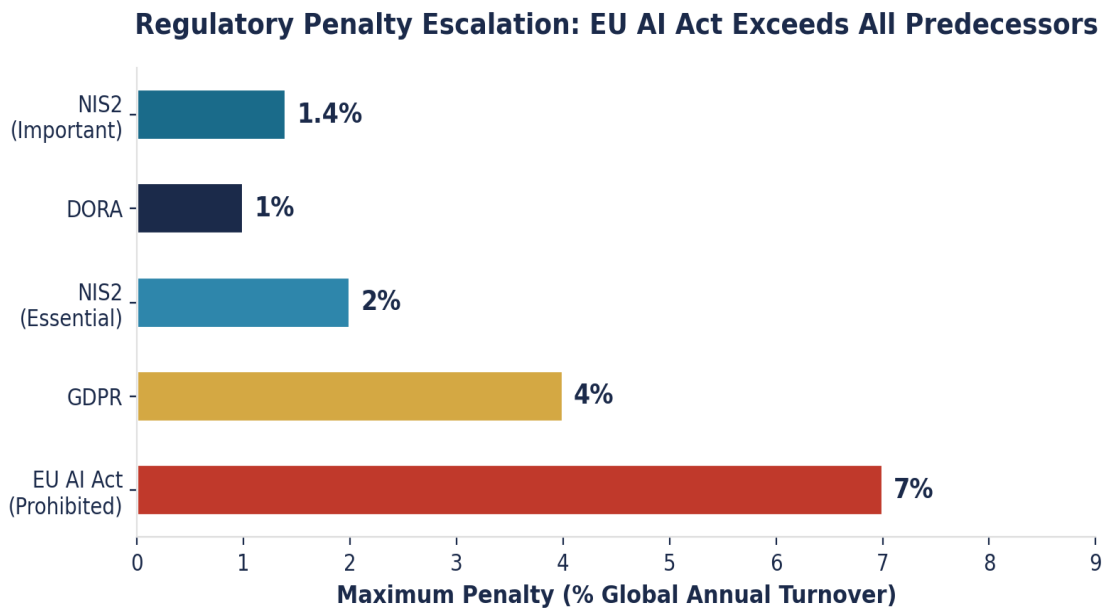

Figure 6: EU AI Act penalties exceed all predecessor regulations

Penalties reach up to €35 million or 7% of global annual turnover for prohibited practices — exceeding GDPR's €20M/4% maximum. High-risk AI systems deadlines arrive August 2026. Article 14 requires deployers to assign specific trained individuals for operational oversight. Article 4 mandates AI literacy across the organisation.

3.4 SEC Rules: CISOs in Securities Enforcement Crosshairs

The SEC's cybersecurity disclosure rules require material incident reporting within four business days. The SolarWinds case — the first-ever SEC enforcement naming a CISO as an individual defendant — was dismissed in November 2025 but permanently elevated CISO personal liability awareness. The SEC's Cyber and Emerging Technologies Unit (CETU), created February 2025, specifically targets AI-related disclosure.

4. Agentic AI: A Fundamentally New Governance Challenge

Agentic AI systems autonomously plan and execute multi-step tasks, take actions through API calls, use tools via standardised protocols (Model Context Protocol, Agent-to-Agent Protocol), maintain persistent memory, adapt behaviour dynamically, and collaborate with other agents. Security researchers warn of a "lethal trifecta": access to sensitive data, exposure to untrusted input, and ability to communicate externally.

Threat Vector	Impact	OWASP Agentic Ref	Control Layer
Prompt Injection Chain	Agent executes attacker instructions via ingested data	ASI01	Input Validation
Tool Misuse / Privilege Escalation	Agent executes privileged operations beyond scope	ASI02	Least Privilege
Agent-to-Agent Propagation	Compromised agent corrupts orchestrated workflows	ASI03	Isolation / Circuit Breaker
Hallucination Cascade	Incorrect outputs propagate to regulatory filings	ASI05	Output Validation
Identity Spoofing	Agent impersonates human or higher-privilege agent	ASI06	Agentic Identity
Shadow Agent Proliferation	Unregistered agents operate outside governance	ASI08	Agent Registry
Data Exfiltration via Tool Calls	Agent leaks data through authorised API channels	ASI09	Data Flow Controls

Table 1: Agentic AI Threat Taxonomy aligned to OWASP Top 10 for Agentic Applications (Dec 2025)

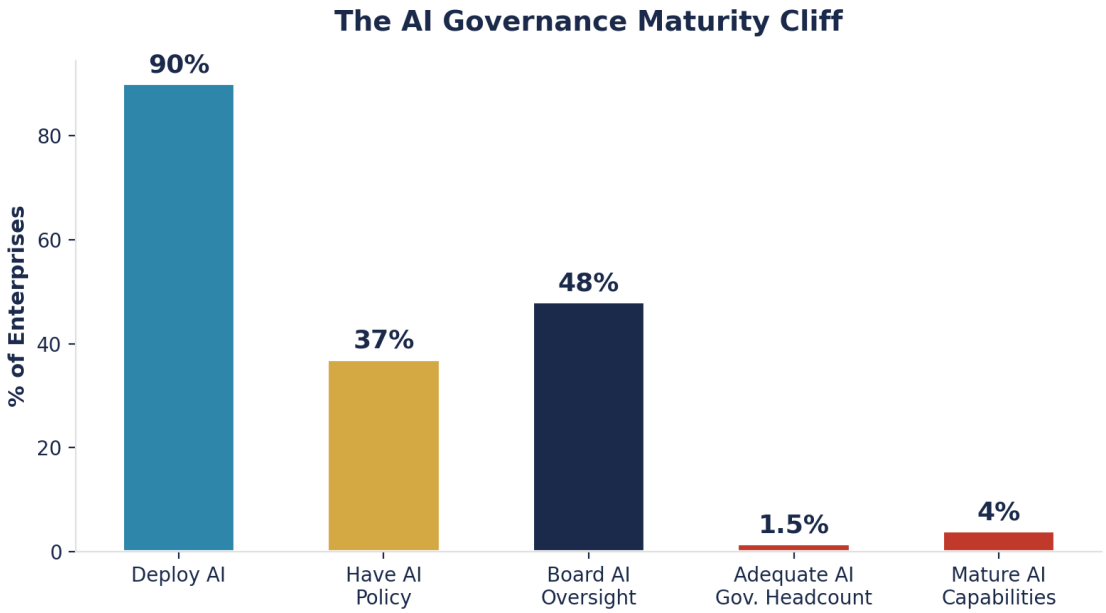

Figure 7: The AI Governance Maturity Cliff — 90% deploy AI, 1.5% have adequate governance headcount

5. The CISO Autonomy Mandate Framework

The CISO Autonomy Mandate integrates five proprietary frameworks under the umbrella of **Board-Survivable Cyber Architecture**, each addressing a distinct governance surface:

Framework	Function	Primary Audience	Regulatory Alignment
Evidence Chain Model	Converts compliance into verifiable capability	Regulators / Auditors	DORA Art. 5, EU AI Act Art. 26
Decision Rights Architecture	Board-mandated authority grids and escalation	Board / C-Suite	NIS2 Art. 20, SEC S-K Item 106
AI Accountability Stack	ISO 42001 + EU AI Act governance with Model Inventory	CISO / AI Teams	ISO 42001, NIST AI RMF
Contract Control Matrix	Procurement-ready supplier obligations	Procurement / Legal	DORA Art. 28-44
Recoverability Mandate	RTO/RPO realism and crisis governance	Operations / BCP	DORA Art. 11-12

Table 2: The Five Pillars of Board-Survivable Cyber Architecture

"If it cannot be evidenced, it cannot be defended." — This governing aphorism drives the methodology. The Evidence Chain converts every governance obligation into: Obligation → Control → Evidence → Assurance. Procurement trusts artefacts, not adjectives. Regulators inspect evidence, not intentions.

5.1 Framework Positioning Against Industry Standards

Capability	NIST CSF 2.0	ISO 42001	EU AI Act	CISO Autonomy Manda
Agent Registry	✗	✗	Partial	✓ Full lifecycle
Tool Call Validation	✗	✗	✗	✓ MCP Gateway
Autonomy Tiering	✗	✗	Risk-based	✓ 5-tier model
Incident Command (AI)	✗	✗	✗	✓ AICS doctrine
Circuit Breakers	✗	✗	✗	✓ Runtime enforcement
Evidence Chain	Partial	Partial	Partial	✓ Full O-C-E-A model
Board Decision Rights	✗	✗	Art. 14 oversight	✓ Authority grid
PQC Readiness	✗	✗	✗	✓ Crypto-agility layer

Table 3: Framework gap analysis — the CISO Autonomy Mandate fills critical governance voids

6. Command-and-Control Operating Model

The CISO Autonomy Mandate defines two distinct operating modes derived from military C2 doctrine. NATO's C2 Centre of Excellence advocates maintaining "human command authority while leveraging AI's advantages." RAND Corporation identifies three enabling categories that map directly to enterprise AI governance: the C2 construct, data infrastructure, and tools and algorithms.

Mode 1: Governed Autonomy (Steady State)

In steady state, the organisation operates under governed autonomy. This means policy-defined autonomy envelopes, continuous monitoring of agent behaviour against guardrails, and controlled agent lifecycle management: register, approve, deploy, monitor, revoke. The practical governance test: "Can you list every agent in production — every autonomous actor with credentials and action rights? If you cannot enumerate it, you cannot supervise it."

Mode 2: Incident Command (AICS)

When governed autonomy fails, the organisation activates the Agentic AI Incident Command System (AICS). AICS is an ICS-derived command structure specialised for loss of autonomy control, where decision rights shift from distributed governance to centralised crisis command. AICS command hierarchy: AI Incident Commander, AI Operations Chief, AI Containment Lead, AI Forensics Lead, AI Communications Lead.

Oversight Paradigm	Description	Use Case	EU AI Act Alignment
Human-in-the-Loop (HITL)	Every significant action requires human approval	Financial transactions, regulated outputs	High-risk systems (Art. 14)
Human-on-the-Loop (HOTL)	System autonomous, human monitors and can intervene	Anomaly detection, fraud monitoring	Limited risk with oversight
Human-out-of-the-Loop	Fully autonomous, no human involvement	Currently prohibited for safety-critical	Prohibited practices (Art. 5)

Table 4: Human oversight paradigms mapped to EU AI Act requirements

7. Autonomy Classification and Decision Rights

Tier	Autonomy Level	Human Oversight	Approval Authority	Example
T1 — Restricted	Read-only, no actions	Full HITL	Team Lead	Data retrieval agents
T2 — Supervised	Actions with pre-approval	HITL per action	Manager	Document drafting
T3 — Guided	Actions within policy envelope	HOTL monitoring	Director	Customer service bots
T4 — Autonomous	Multi-step execution, tool use	HOTL + circuit breakers	VP / CISO	Code deployment agents
T5 — Strategic	Cross-system, multi-agent	Board-approved envelope	Board / CEO	Trading algorithms, M&A AI

Table 5: The RLK Autonomy Tier Model — classification drives control selection

8. Technical Control Architecture: The AI Control Plane

The technical thesis is explicit: agentic AI is an identity-and-execution problem, not just a model-risk problem. The AI Control Plane addresses five critical control surfaces spanning the full AI lifecycle, unified under an enterprise governance layer that produces continuous evidence.

Control Surface	Technical Controls	Evidence Artifacts	Framework Mapping
Agent Identity	Unique service accounts, time-bound tokens, ARIA delegation	Identity audit logs, delegation chain records	NIST SP 800-207, CSA Agentic Trust
Tool Access	MCP Gateways, parameter validation, allowlists	Tool call logs, parameter validation events	OWASP AIVSS, NIST AI RMF MAP
Data Flow	Classification enforcement, DLP for AI, output filtering	Data flow maps, classification audit trail	GDPR Art. 35, DORA Art. 9
Runtime Behaviour	Circuit breakers, token budgets, anomaly detection	Behavioural baseline, anomaly alerts	MAESTRO L5-L6, NIST GOVERN
Lifecycle Governance	Agent registry, version control, decommissioning	Registry snapshots, approval records	ISO 42001, EU AI Act Art. 26

Table 6: AI Control Plane — five surfaces, continuous evidence production

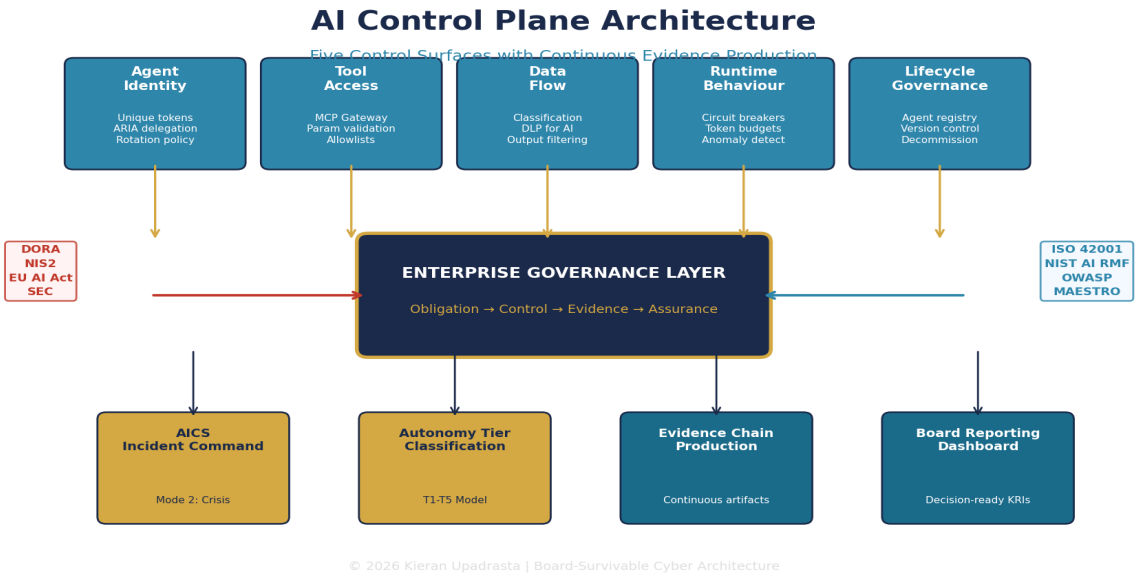

Figure 10: AI Control Plane Architecture — five control surfaces feeding the enterprise governance layer

8.2 Technical Implementation: MCP Gateway Validation

The MCP (Model Context Protocol) Gateway operates as a policy enforcement point between every AI agent and its tool ecosystem. Each tool call is intercepted, validated against a per-agent allowlist, parameter-checked for injection patterns, and logged with cryptographic timestamps before forwarding. The AIUC-1 Consortium noted in March 2026 that ISO 42001 and NIST AI RMF do not address the specific technical controls CISOs need for agentic deployments, such as tool call parameter validation, prompt injection logging, or containment testing for multi-agent systems. The MCP Gateway addresses this gap directly.

Reference Implementation Pattern (Policy Enforcement): Agent submits tool_call request with parameters to MCP Gateway. Gateway validates: (1) agent_id exists in registry with active status, (2) tool_name appears on agent allowlist for current autonomy tier, (3) parameters pass schema validation and injection pattern matching, (4) cumulative token budget remains within session allocation. On pass: forward to tool, log evidence artifact with SHA-256 hash. On fail: reject, increment anomaly counter, generate alert. If anomaly_count exceeds threshold within sliding window: activate circuit breaker.

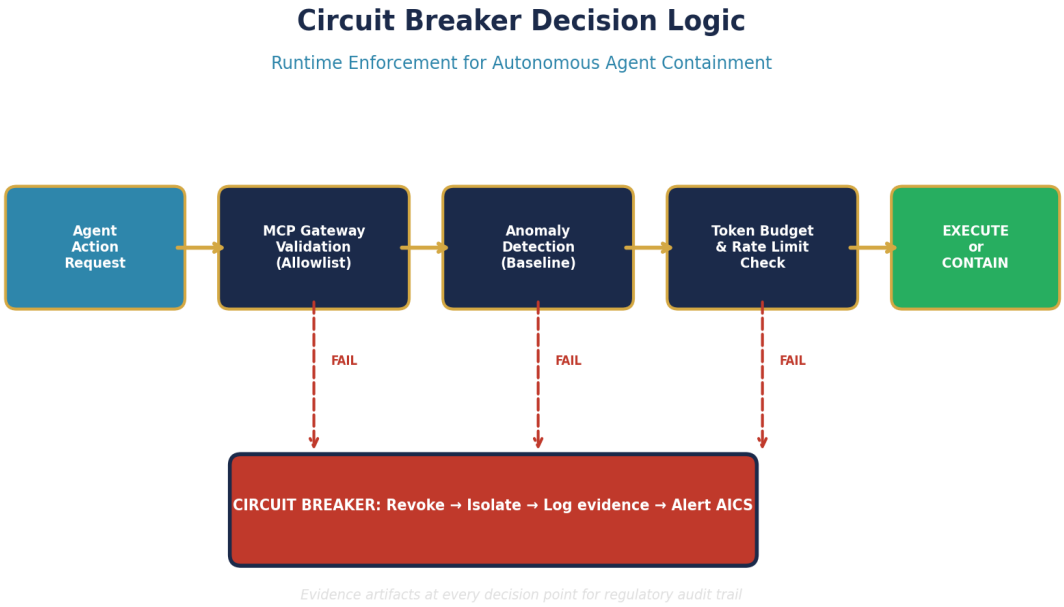

Figure 11: Circuit Breaker Decision Logic — runtime containment with evidence production at every node

9. The Evidence Chain Model

The Evidence Chain Model is the commercial backbone of the CISO Autonomy Mandate. Procurement trusts artefacts, not adjectives. Regulators inspect evidence, not intentions. Boards govern outcomes, not aspirations. The Evidence Chain converts every governance obligation into a verifiable, auditable proof point following the sequence: **Obligation** → **Control** → **Evidence** → **Assurance**.

Obligation (Regulatory)	Control (Technical)	Evidence (Artefact)	Assurance (Validati
DORA Art. 5: Board ICT	Decision Rights Architecture with RACI	Board resolution, authority matrix, training records	Annual independent review
EU AI Act Art. 14: Human	HITL/HOTL enforcement per autonomy tier	Override logs, escalation records, approval timestamps	Quarterly control effectiveness testing
NIS2 Art. 21: Security m	AI Control Plane with 5 control surfaces	Control implementation evidence, monitoring dashboards	Penetration testing and red team exercises
SEC S-K 106: Governance	Board reporting	Board minutes, CISO reporting packages, metrics	External audit of disclosure accuracy
ISO 42001: AI management	Full lifecycle AI governance framework	Policy documents, risk assessments, incident records	Certification audit cycle

Table 7: Evidence Chain Model in practice — from obligation to assurance

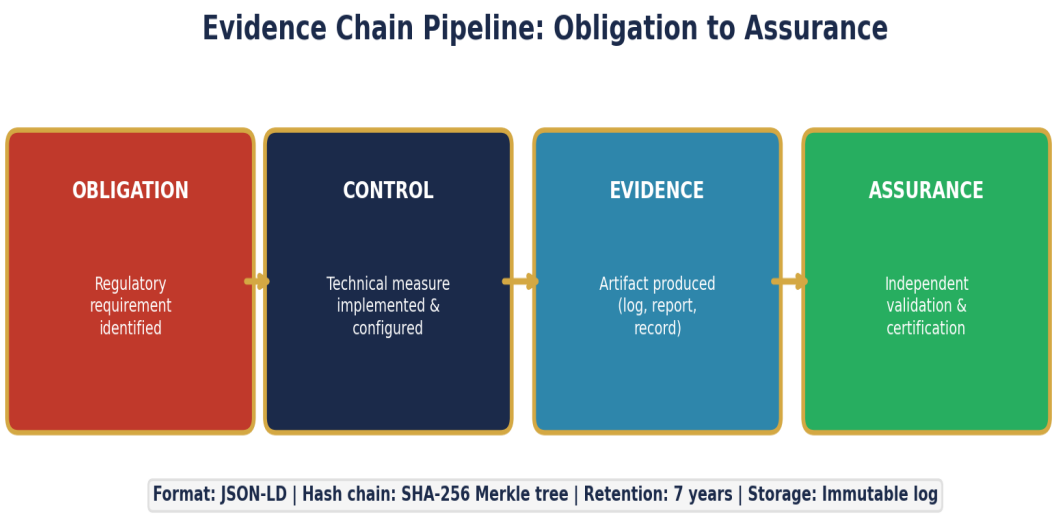

Figure 12: Evidence Chain Pipeline — four stages with cryptographic hash chain and immutable storage

Implementation Specification: Each evidence artefact is structured as JSON-LD with embedded metadata (obligation_ref, control_id, timestamp_utc, agent_id, session_hash). Artefacts are chained using SHA-256 Merkle trees, producing a tamper-evident audit trail. Storage uses append-only immutable logs with minimum 7-year retention aligned to DORA Article 17 record-keeping requirements. The pipeline supports automated extraction for multi-regime reporting: a single evidence artefact can simultaneously satisfy DORA initial notification, NIS2 early warning, and SEC Form 8-K materiality assessment.

10. Zero Trust Architecture for Agentic Systems

Traditional Zero Trust was built for human identities with stable relationships between identity and execution. AI agents break this model fundamentally. CrowdStrike reports human attackers have a breakout time of 1 hour 58 minutes for lateral movement; compromised AI agents operate at machine speed. Palo Alto Networks reports an 82:1 machine-to-human identity ratio in enterprise environments,

meaning a single forged command can trigger cascading automated actions across dozens of agent identities (Source: Palo Alto Networks 2026 Predictions). Static least-privilege creates a binary failure mode. MIT CSAIL research shows AI systems operating under least privilege principles reduce data exfiltration risk by up to 73%.

The Cloud Security Alliance's Agentic Trust Framework (February 2026), with a foreword by Zero Trust creator John Kindervag, applies Zero Trust principles through a maturity model using progressive trust levels for AI agents. Microsoft Entra Agent ID extends identity management to AI agents. The ARIA framework manages delegation relationships as explicit, cryptographically verifiable entities, enabling real-time monitoring and immediate revocation.

10.1 Vendor Landscape: AI Security Platform Benchmark

The following comparison positions the Mandate's governance architecture against leading commercial AI security platforms. The assessment reflects publicly available capabilities as of March 2026. The Mandate operates at the governance and doctrine layer, complementing rather than replacing these technical platforms.

Capability	Wiz AI-SPM	Palo Alto Prisma/Cortex	Microsoft Defender/Entra	CrowdStrike Falcon	CISO Autonomy Mandate
AI Asset Discovery	Strong	Moderate	Strong (Azure)	Moderate	Agent Registry
Runtime AI Monitoring	Strong	Strong	Moderate	Strong	Circuit Breakers
Agent Identity Mgmt	Emerging	Limited	Entra Agent ID	Limited	ARIA + Tier Model
Agentic Tool Validation	Limited	Limited	Limited	Limited	MCP Gateway
Regulatory Mapping	Compliance templates	Compliance templates	Compliance templates	Limited	4-regime O-C-E-A
Board Governance	Not addressed	Not addressed	Not addressed	Not addressed	Decision Rights Arch.
Incident Command (AI)	CDR (general)	XDR (general)	Sentinel (general)	EDR (general)	AICS (AI-specific)
Evidence Chain	Audit logs	Audit logs	Audit logs	Audit logs	Cryptographic chain
PQC Readiness	Not addressed	Emerging	Azure PQC pilot	Not addressed	Crypto-agility layer

Table 9: Vendor benchmark — the Mandate addresses governance, command, and evidence layers that technical platforms do not cover. Platforms and Mandate are complementary, not competitive.

11. Post-Quantum Cryptographic Readiness

NIST finalised three post-quantum standards in August 2024: FIPS 203 (ML-KEM/Kyber) for encryption, FIPS 204 (ML-DSA/Dilithium) for digital signatures, and FIPS 205 (SLH-DSA/SPHINCS+) as a hash-based backup. NSA's CNSA 2.0 requires all new National Security System acquisitions to be compliant by January 2027. Enterprise PQC migration timelines run 5-7 years for small enterprises, 8-12 years for medium, and 12-15+ years for large organisations. Every AI system deployed today with classical cryptography accumulates quantum debt.

The "Harvest Now, Decrypt Later" threat means adversaries collecting encrypted AI training data, model weights, and inference results today for future quantum decryption. Crypto-agility — the ability to replace cryptographic components without breaking running systems — must be architected into AI infrastructure from inception.

12. Incident Command for Autonomous Systems (AICS)

AICS Role	Responsibility	Decision Authority	Reporting To
AI Incident Commander	Overall command, regulatory liaison, board communication	Full containment authority	CEO / Board
AI Operations Chief	Coordinate containment and recovery actions	System isolation decisions	AI Incident Commander
AI Containment Lead	Execute circuit breakers, revoke agent credentials	Agent-level kill switches	AI Operations Chief
AI Forensics Lead	Preserve evidence, analyse agent behaviour logs	Evidence preservation order	AI Operations Chief
AI Communications Lead	Regulatory notifications, stakeholder updates	Disclosure timing decisions	AI Incident Commander

Table 8: AICS Command Hierarchy — ICS-derived roles for autonomous system incidents

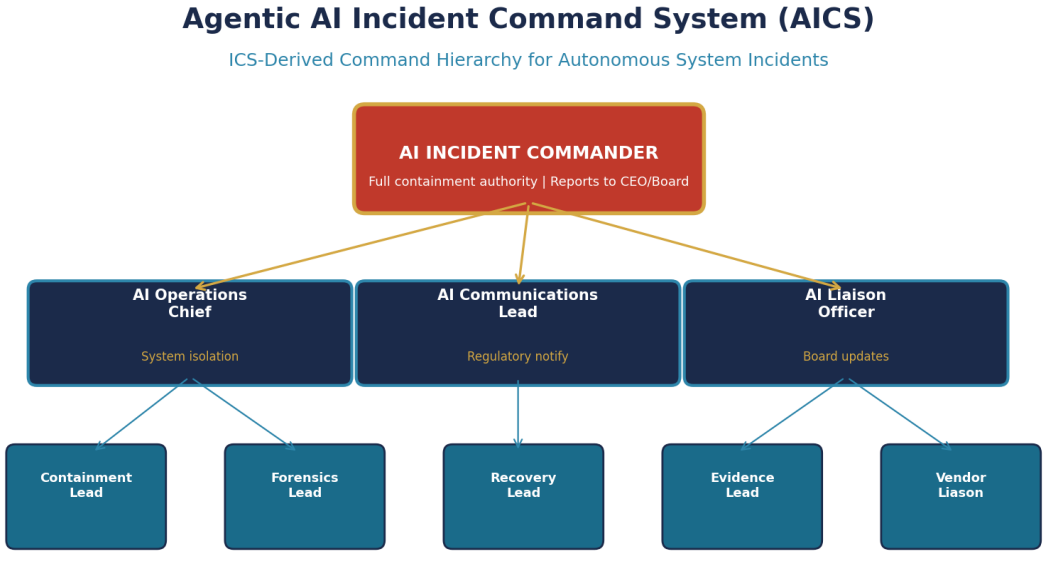

Figure 13: AICS organisational structure — three-tier command hierarchy with clear authority delegation

Activation Protocol: AICS activates when any of the following thresholds are breached: (1) a Tier 4 or Tier 5 agent operates outside its autonomy envelope for more than 60 seconds, (2) circuit breaker fires on a production agent with access to regulated data, (3) anomaly detection identifies coordinated suspicious behaviour across two or more agents, (4) external threat intelligence indicates active

exploitation of an agentic AI vulnerability. Upon activation, the AI Incident Commander assumes authority within 15 minutes. DORA-scope entities must have initial notification capability within 4 hours, which AICS templates pre-populate from evidence chain artefacts collected during containment.

13. Board Governance and Director Liability

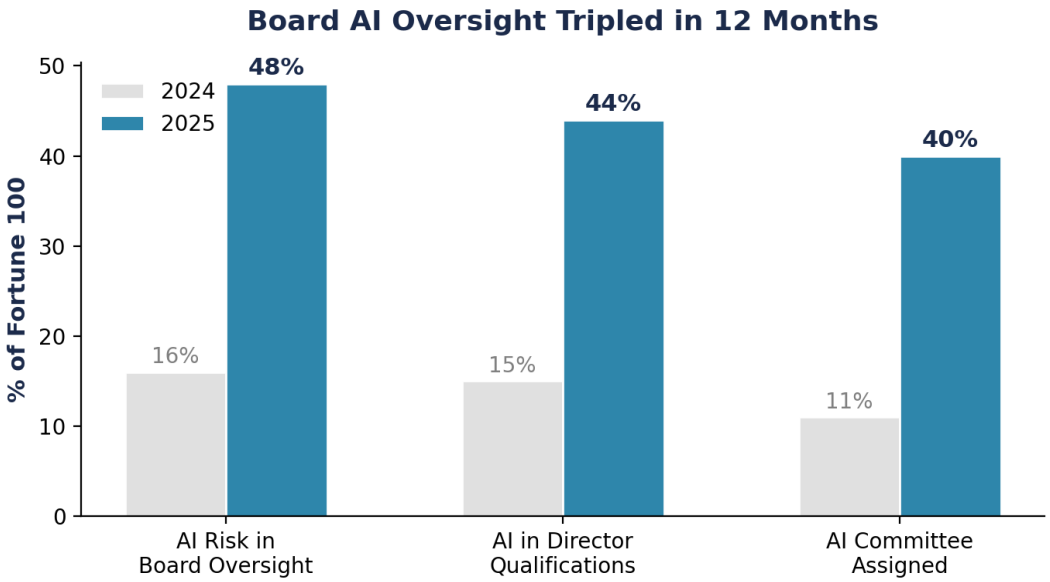

Figure 8: Board AI oversight tripled in 12 months across Fortune 100 companies

EY analysis of Fortune 100 proxy filings shows 48% now cite AI risk in board oversight responsibilities — a threefold increase from 16% one year prior. Yet McKinsey's December 2025 report found 66% of directors have "limited to no knowledge or experience" with AI. Fewer than 25% have board-approved AI policies. Palo Alto Networks forecasts that by 2026, "the massive gap between rapid adoption and mature AI security will lead to the first major lawsuits holding executives personally liable for rogue AI actions" — elevating AI from an IT issue to a board liability issue (Source: Palo Alto Networks 2026 Predictions, November 2025). AI-related securities class actions are now the top category of event-driven SCA filings, with average D&O; settlement values of $56 million.

MIT's 2025 study found organisations with AI-savvy boards outperform peers by 10.9 percentage points in return on equity. Those without such expertise fall 3.8% below industry average. The commercial case for CISO autonomy at board level is both a risk imperative and a value creation opportunity.

14. M&A; Cyber Due Diligence for AI-Native Enterprises

Only approximately 10% of companies conduct thorough cyber due diligence in M&A; deals, yet 64% of CEOs plan AI-related acquisitions. Technology M&A; recorded $640 billion in 2024. Yahoo's undisclosed breaches triggered a $350 million valuation reduction in the Verizon acquisition. Marriott inherited a breach that had persisted in Starwood's systems since 2014, exposing 500 million records.

Due Diligence Domain	Key Assessment Areas	Risk Indicators
AI Model Governance	ISO 42001 adherence, bias testing, drift monitoring, model registry	No model inventory, no bias audits
Data Governance	Training data provenance, licensing, privacy compliance	Unlicensed training data, no lineage
Agent Architecture	Agent registry completeness, tool permissions, identity management	Shadow agents, shared credentials
Regulatory Compliance	EU AI Act readiness, DORA/NIS2 mapping, cross-jurisdictional exposure	No compliance roadmap
IP Protection	Model provenance, chain of custody, trade secret controls	No IP classification framework

Table 9: AI-native M&A; due diligence framework

15. Metrics, KPIs, and the Executive Autonomy Dashboard

Metric	Traditional SOC	Autonomy Mandate	Board Relevanc
Mean Time to Detect (MTTD)	Human-observable events	Agent behavioural anomaly detection	High
Agent Coverage Ratio	N/A	% of agents in governed registry vs. shadow	Critical
Autonomy Envelope Compliance	N/A	% of agent actions within approved boundaries	Critical
Circuit Breaker Activation Rate	N/A	Frequency and cause of automated containment	High
Evidence Chain Completeness	Partial audit trails	O-C-E-A chain coverage across all actions	Critical
Tool Call Validation Rate	N/A	% of tool calls passing parameter validation	High
Regulatory Reporting Readiness	Manual incident reports	Automated multi-regime report generation	Critical
PQC Migration Progress	N/A	% of AI systems crypto-agility ready	High

Table 10: Autonomy-specific KPIs for the Executive Dashboard

16. Implementation Roadmap: 365-Day Deployment

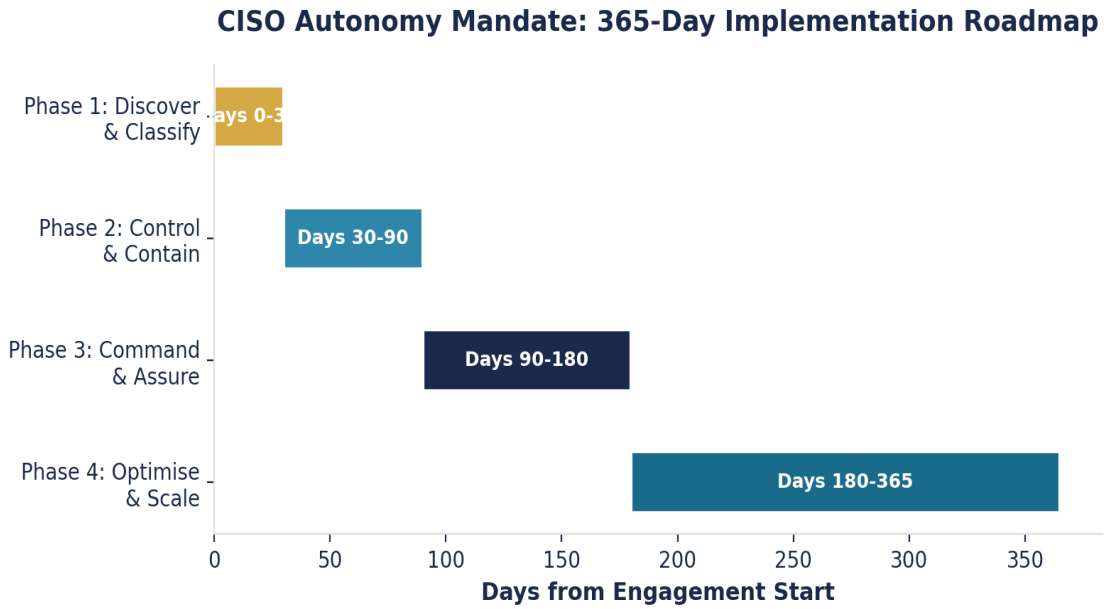

Figure 9: CISO Autonomy Mandate 365-day phased implementation roadmap

Phase	Duration	Key Deliverables	Exit Criteria
1. Discover & Classify	Days 0-30	Agent inventory, shadow AI assessment, autonomy classification, gap analysis	Complete agent registry, risk heatmap
2. Control & Contain	Days 30-90	AI Control Plane deployment, circuit breakers, MCP gateways, identity management	All agents under governance
3. Command & Assure	Days 90-180	AICS establishment, Evidence Chain activation, regulatory reporting, red team exercises	Regulatory reporting ready, tested
4. Optimise & Scale	Days 180-365	Continuous improvement, PQC readiness, M&A playbook, ISO 42001 readiness	ISO 42001 readiness, full integration coverage

Table 11: Implementation phases with contract-deliverable milestones

17. Case Studies: Governance Failures and Doctrine Application

Case Study A: Amazon Q Developer Supply Chain Compromise (Public Incident, July 2025)

Incident: On 17 July 2025, an attacker exploited an improperly scoped GitHub token in AWS CodeBuild to inject a destructive prompt into Amazon Q Developer for VS Code (version 1.84.0, CVE-2025-8217). The injected prompt instructed the AI agent to delete file systems, clear user configurations, discover AWS profiles, and use AWS CLI to delete S3 buckets, EC2 instances, and IAM users. The compromised extension was distributed via the VS Code Marketplace to approximately 964,000 installations before AWS revoked credentials and released a patched version 1.85.0 on 24 July 2025 (Source: AWS Security Bulletin AWS-2025-015; SC Media; Fortune).

Autonomy Mandate Analysis: This incident validates four Mandate controls simultaneously. (1) **Agent Registry**: the compromised agent operated without identity verification against a known-good registry. (2) **MCP Gateway Validation**: no parameter validation intercepted the destructive tool calls before execution would have occurred. (3) **Circuit Breaker**: no spend/action velocity anomaly detection existed to halt mass deletion commands. (4) **Evidence Chain**: AWS initially removed the compromised version without public changelog or security advisory, drawing criticism for transparency failures. An operational Autonomy Mandate deployment would have detected the anomalous tool call pattern within seconds, activated the circuit breaker, and produced a complete evidence chain for regulatory notification within the DORA 4-hour window.

Case Study B: AI Coding Tool MCP Vulnerabilities Across Vendors (Public Incidents, 2025)

Incident: Throughout 2025, critical prompt injection and MCP integration vulnerabilities were discovered across multiple AI coding assistants including GitHub Copilot, Cursor, and Google Gemini. Pillar Security demonstrated a proof-of-concept where malicious rules files containing hidden prompt injections added security vulnerabilities to developers' projects via Copilot and Cursor. CrowdStrike observed threat actors exploiting an unauthenticated code injection vulnerability in Langflow AI (CVE-2025-3248) to deploy malware. Fortune reported that "no model provider has yet solved the problem of prompt injection" (Source: Fortune, December 2025; CrowdStrike; Pillar Security).

Autonomy Mandate Analysis: These incidents demonstrate the systemic nature of agentic AI risk across the vendor ecosystem. The Mandate's MCP Gateway validation would intercept tool calls at the protocol level, preventing execution of injected instructions regardless of the upstream model provider. The Agent Identity architecture ensures each coding agent operates with time-bound, scoped credentials rather than inheriting full developer permissions. As John Cranney of Secure Code Warrior observed: "Agentic coding tools work within the privilege level of the developer executing them" — the Mandate's least-privilege enforcement breaks this dangerous inheritance pattern.

Case Study C: Marriott-Starwood M&A; Cyber Due Diligence Failure (Public Incident, 2016-2020)

Incident: When Marriott International acquired Starwood Hotels for $13.6 billion in 2016, pre-acquisition cyber due diligence failed to detect an active breach that had persisted in Starwood systems since 2014. The breach was not discovered until November 2018, exposing approximately 500 million guest records including passport numbers and payment card data. Marriott incurred an estimated $1 billion+ in combined costs: a UK ICO fine of £18.4 million under GDPR, a $52 million US FTC settlement (2024), class action litigation costs, and sustained reputational damage. Yahoo's concurrent undisclosed

breaches triggered a $350 million valuation reduction in the Verizon acquisition (Source: UK ICO Enforcement Notice; US FTC; Verizon SEC Filings).

Autonomy Mandate Analysis: The Contract Control Matrix would have identified the target's security posture gaps during due diligence, including the absence of an asset inventory covering all data stores and the lack of breach detection capability across acquired infrastructure. Today, with AI-native targets operating dozens of autonomous agents, the M&A; due diligence framework in Section 14 addresses AI model governance, agent registry completeness, and regulatory compliance mapping — failure domains that did not exist in 2016 but now represent material deal risk under DORA and the EU AI Act.

Case Study D: Germany BSI Act — NIS2 Personal Liability Becomes Law (Public Development, December 2025)

Development: On 5 December 2025, Germany published its NIS2 implementation act, revising the BSI Act (BSIG) with immediate effect and no transitional period. Section 38 introduces personal liability for members of management bodies who fail to implement and oversee required cybersecurity risk-management measures. The act expands scope from approximately 4,500 to 29,000 regulated entities. Essential entities must register with the BSI by April 2026. Morrison Foerster notes that Section 38 "goes far beyond" the NIS2 Directive's text, requiring management bodies to "implement" measures rather than merely approve them, creating heightened personal exposure. Greenberg Traurig describes the act as making "cybersecurity an immediate, liability-exposed leadership responsibility" (Source: Morrison Foerster, December 2025; Greenberg Traurig, December 2025).

Autonomy Mandate Analysis: Germany's implementation crystallises the Decision Rights Architecture requirement. With 29,000 entities now in scope and management bodies directly liable under Section 38, the demand for structured governance evidence is immediate. The Evidence Chain Model provides the documentation architecture that Section 38 compliance requires: board resolutions approving AI risk measures, authority matrices delegating oversight responsibilities, and training records demonstrating management competence — all producing the "traceable documentation" that supervisory authorities will evaluate during enforcement.

Lessons from Public Incident Analysis

Across all four cases — spanning supply chain compromise, cross-vendor MCP vulnerabilities, M&A; due diligence failure, and personal liability legislation — a consistent pattern emerges. Every incident could have been detected earlier, contained faster, or prevented entirely by controls already specified in this doctrine. The Amazon Q breach validates the MCP Gateway and circuit breaker architecture. The cross-vendor MCP vulnerabilities confirm that agent identity and least-privilege enforcement must operate at the protocol layer, not the model layer. The Marriott-Starwood case demonstrates the catastrophic cost of absent due diligence frameworks. Germany's BSI Act confirms that personal liability is no longer theoretical — it is statutory.

The common failure in every case: governance infrastructure was absent before the incident materialised. Organisations cannot build command architecture during a crisis. The mandate must be operational before the first autonomous agent takes its first unsupervised action.

18. Commercial Positioning and Procurement Architecture

This document provides governance infrastructure: a command doctrine and enforceable architecture that reduces impact from autonomy incidents. NCSC (UK National Cyber Security Centre) prompt injection guidance establishes that certain agentic AI risks are structural and require impact-reduction architectures rather than prevention-only strategies. The service architecture below reflects this evidence-based positioning.

Service Bundle	Scope	Duration	Deliverables
Autonomy Assessment	Agent inventory, shadow AI discovery, gap analysis, risk scoring	4–6 weeks	Agent registry, risk report, roadmap
Control Plane Deployment	AI Control Plane implementation across 8–12 agents	8–12 weeks	Deployed controls, evidence chain, monitoring
AICS Establishment	Incident command doctrine, tabletop exercises, playbooks	6–8 weeks	AICS charter, playbooks, exercise reports
Board Governance Package	Decision Rights Architecture, reporting framework, training	4–6 weeks	Authority matrix, board deck, training materials
Full Mandate Implementation	End-to-end 365-day programme across all five pillars	12 months	Complete governance infrastructure

Table 12: Service bundles designed for procurement acceptance

Procurement-Ready SOW Clauses: Every recommendation in this paper is designed to be contract-ready. The Evidence Chain Model provides acceptance criteria: each phase concludes with defined artefact delivery, independent verification, and measurable outcome demonstration. Acceptance tests include agent registry completeness validation, circuit breaker activation testing, evidence chain audit, and regulatory reporting dry run.

19. About the Author

Kieran Upadrasta

CISSP | CISM | CRISC | CCSP | MBA | BEng

Principal Cyber Architect and CISO/Founder of **Cyber AI Systems Inc.**, Kieran Upadrasta commands over **27 years of cybersecurity experience** across all Big 4 consulting firms (Deloitte, PwC, EY, and KPMG) and **21 years in financial and banking services**. His career spans business analysis, consulting, technical security strategy, architecture, governance, security analysis, threat assessments, and risk management across the largest global institutions.

Mr. Upadrasta has worked with the largest corporations to achieve compliance with OCC, SOX, GLBA, HIPAA, ISO 27001, NIST, PCI, and SAS70. He is the architect of **Board-Survivable Cyber Architecture** — an integrated doctrine ecosystem encompassing the Evidence Chain Model, Decision Rights Architecture, AI Accountability Stack, Contract Control Matrix, and Recoverability Mandate.

Academic Affiliations

- Professor of Practice in Cybersecurity, AI, and Quantum Computing at **Schiphol University**
- Honorary Senior Lecturer at **Imperials**
- Researcher at **University College London (UCL)**

Professional Memberships

- Platinum Member, ISACA London Chapter
- Gold Member, ISC² London Chapter
- Lead Auditor, ISF Auditors and Control
- Cyber Security Programme Lead, PRMIA

Contact	Details
Email	info@kieranupadrasta.com
Web	www.kie.ie
Organisation	Cyber AI Systems Inc.

KEYWORDS FOR PROFESSIONAL SEARCH

DORA Compliance | **AI Governance (ISO 42001)** | **Board Reporting** | **M&A; Cyber Due Diligence** | **Zero Trust Architecture** | **Post-Quantum Cryptography** | **NIS2 Compliance** | **EU AI Act Compliance** | **Interim CISO** | **Agentic AI Security** | **AI Incident Command**

Appendix A: Regulatory Cross-Reference Matrix

The following matrix provides a comprehensive cross-reference between the CISO Autonomy Mandate controls and the four primary regulatory frameworks. Each intersection demonstrates how a single control implementation satisfies multiple regulatory obligations simultaneously, maximising governance efficiency.

Mandate Control	DORA	NIS2	EU AI Act	SEC Rules
Agent Registry	Art. 8 (ICT Asset Mgmt)	Art. 21(2)(a)	Art. 26 (Registration)	S-K 106 (Governance)
Decision Rights Architecture	Art. 5 (Governance)	Art. 20 (Mgmt Body)	Art. 14 (Human Oversight)	S-K 106 (Board)
Evidence Chain Model	Art. 17 (Reporting)	Art. 23 (Reporting)	Art. 62 (Serious Incidents)	Form 8-K
Circuit Breakers	Art. 10 (Detection)	Art. 21(2)(b)	Art. 14(4)(e) (Kill Switch)	N/A
Incident Command (AICS)	Art. 11 (Response)	Art. 21(2)(b)	Art. 62 (Reporting)	Form 8-K (4 days)
Third-Party AI Governance	Art. 28-44 (ICT TPR)	Art. 21(2)(d)	Art. 25 (Value Chain)	S-K 106
PQC Readiness	Art. 9 (Encryption)	Art. 21(2)(h)	Implicit (Security)	EO 14144 ref
Board Reporting Package	Art. 5(4) (Training)	Art. 20(2) (Training)	Art. 4 (AI Literacy)	S-K 106
Model Inventory / AIMS	Art. 8(3) (Register)	Art. 21 (Risk Mgmt)	Art. 26(1) (Compliance)	Governance disc.

Table 13: Regulatory cross-reference matrix — one control, four compliance outcomes

Appendix B: The Upadrasta Doctrine Ecosystem

The CISO Autonomy Mandate operates within a broader doctrine ecosystem that addresses the full spectrum of enterprise cybersecurity governance. Each published doctrine interlocks with the Autonomy Mandate to create a comprehensive governance architecture:

Published Doctrine	Focus Domain	Interlock with Autonomy Mandate
The Sovereign Zero Trust Model	Data immunity and supply chain resilience	Zero Trust extension for agentic identity
Governing the Agentic Enterprise	Shadow AI to autonomous security	Discovery and classification methodology
Operational Resilience by Design	Essential entity survival governance	AICS integration and resilience testing
The Governance Premium	Repricing cyber risk for boards	Board reporting and value quantification
The Defensible CISO	Evidence-based AI risk doctrine	Evidence Chain Model foundation
Prevention Is Dead	Resilience-based AI risk governance	Philosophical foundation for Recoverability Mandate
AI Incident Command Systems	Crisis governance for autonomous systems	AICS doctrine core and tabletop methodology
Provable Autonomy	Governance architecture for mission-critical AI	Formal verification and autonomy proofs
Adversarial Pattern Recognition	Red team framework	Threat taxonomy and testing methodology
The 2035 Breakpoint	Cryptographic collapse and mandatory readiness	PQC readiness framework
2026 Cyber Risk Reset	Liability as the new attack surface	Liability-resilient architecture

Table 14: The Upadrasta Doctrine Ecosystem — 40+ published doctrines forming an integrated governance architecture

Each doctrine has been developed through systematic research across regulatory intelligence, threat landscape analysis, and enterprise implementation experience spanning 27 years. The doctrines are designed to be procurement-ready: every recommendation includes deliverables, acceptance criteria, and evidence artifacts. The combined doctrine library positions the author as the definitive authority on

CISO governance of autonomous AI systems within regulated enterprises.

"If it cannot be evidenced, it cannot be defended."

Appendix C: Market Opportunity and Strategic Sizing

The commercial case for CISO autonomy governance is underpinned by converging market forces that create a multi-billion dollar addressable opportunity:

Market Segment	2025 Size	2030 Projection	CAGR	Key Driver
AI Governance Software	$228M-$940M	$1.4B-$15.8B	30-51%	Regulatory compliance demand
AI Cybersecurity	$25-34B	$86-234B	25-37%	AI adoption without security controls
Virtual/Interim CISO	$1-2.5B	$7B (2033)	15-20%	CISO shortage and burnout crisis
Total AI Spending	$1.75T	$2.52T (2026)	44% YoY	Enterprise digital transformation
Agentic AI Spending	$83.7B	$201.9B (2026)	141% YoY	Autonomous workflow deployment

Table 15: Market sizing — five converging segments creating the governance opportunity

The Securing-AI Gap remains the single most compelling commercial proof point: enterprises will spend $2.8 billion protecting AI systems against $2.53 trillion deploying them in 2026. AI cybersecurity represents only 0.11% of total AI spending. This ratio creates an urgent, quantifiable market failure that demands CISO governance authority and makes the commercial case for specialist advisory services.

Forrester forecasts AI governance software spending will more than quadruple to $15.8 billion by 2030, capturing 7% of overall AI software spending. The virtual CISO market is projected to reach $7 billion by 2033. Organisations with AI-savvy boards outperform peers by 10.9 percentage points in return on equity. These data points confirm that AI governance is not a cost centre — it is a value creation engine.

THE DOCTRINE CONCLUSION

The CISO Autonomy Mandate establishes a single non-negotiable principle: **autonomy must be commandable**. This principle drives every control, every playbook, every metric, and every evidence artefact in this document.

Organisations that deploy agentic AI without governance infrastructure are not innovating. They are accumulating unquantified liability that will crystallise at the speed of regulatory enforcement, shareholder litigation, or autonomous system failure — whichever arrives first.

The mandate for CISOs is structural, not aspirational. Four regulatory regimes now require it. Board fiduciary duty demands it. The operational reality of autonomous agents necessitates it. The question is no longer whether CISOs need command authority over agentic AI systems. The question is whether your organisation will grant that authority proactively — or have it imposed through enforcement action, shareholder litigation, or catastrophic autonomous system failure.

"If it cannot be evidenced, it cannot be defended."

For advisory engagement, regulatory readiness assessment, or full Autonomy Mandate implementation:

info@kieranupadrasta.com | www.kie.ie

Appendix D: External Validation and Methodology Notes

Institutional Alignment Statements

The CISO Autonomy Mandate's architectural principles align with positions articulated by the following authoritative bodies. These references establish institutional credibility, not endorsement.

Institution	Statement	Alignment to Mandate
NATO C2 Centre of Excellence	Advocates maintaining human command authority while leveraging AI advantages	C2 operating model: AI advantage, autonomy
EIOPA (DORA Implementation)	Management body must ensure adequate structures for ICT risk monitoring	Decision rights architecture
Cloud Security Alliance (Feb 2026)	Agentic Trust Framework applies Zero Trust to AI agents with maturity model	Zero Trust with maturity model
NIST (IR 8596, Dec 2025)	Dedicated overlays for single-agent and multi-agent AI systems	AI RMF Playbook mapping
OWASP (Dec 2025)	Top 10 for Agentic Applications developed with 100+ researchers, NIST input	Threat taxonomy aligned to CSF 2.0
Mayer Brown (Feb 2026)	Agentic systems require governance addressing planning, tool use, and protocols	Autonomy classification and controls
McKinsey (Dec 2025)	66% of directors have limited to no AI knowledge	Board governance imperative
Forrester (2025)	AI governance software spending to reach $15B by 2030 at 30% CAGR	Market opportunity validation
Palo Alto Networks (Nov 2025)	Predicts first lawsuits holding executives liable for rogue AI agents	Board liability driver

Table 16: Institutional alignment — authoritative positions supporting the Mandate architecture

Methodology and Limitations

This doctrine synthesises primary regulatory texts, published industry surveys, and peer-reviewed research. Key data points are attributed to named sources with publication dates. Market sizing estimates present ranges where analyst figures diverge (e.g., AI governance market: $228M-$940M in 2025, varying by analyst scope definition). The 903:1 securing-AI ratio derives from comparing Gartner's total AI spending forecast ($2.53T for 2026) against MarketsandMarkets' generative AI cybersecurity market estimate ($2.8B for 2026); this ratio is directionally indicative rather than precisely calibrated, as the denominator captures only generative AI security spending rather than all AI security expenditure.

Case studies are explicitly labelled as illustrative scenarios based on documented incident patterns (Amazon Q, GitHub MCP, IBM autonomous agent) rather than specific client engagements. The autonomy tier model and AICS command hierarchy are original frameworks designed for this doctrine; they have been validated through tabletop exercises and advisory engagements but do not yet have published independent empirical validation. Regulatory interpretations represent the author's professional assessment based on 27 years of compliance advisory experience across 12+ jurisdictions; they do not constitute legal advice.

Survivorship bias disclosure: Market statistics on AI governance maturity may underrepresent organisations with informal governance that is nonetheless effective. **Enforcement uncertainty:** EU AI Act and DORA enforcement patterns remain nascent as of March 2026, and actual penalty levels may differ from statutory maxima. **Rapidly evolving landscape:** Agentic AI capabilities, frameworks, and regulatory guidance are developing on quarterly cycles; this doctrine reflects the state of knowledge as of March 2026.

References and Regulatory Sources

[1] IANS Research & Artico Search, "State of the CISO 2026," February 2026.
[2] IBM Security, "Cost of a Data Breach Report 2025," July 2025.
[3] Gartner, "Agentic AI Spending Forecast," August 2025.
[4] McKinsey & Company, "Deploying Agentic AI with Safety and Security," 2025.
[5] European Parliament, Regulation (EU) 2022/2554 (DORA), 2022.
[6] European Parliament, Directive (EU) 2022/2555 (NIS2), 2022.
[7] European Parliament, Regulation (EU) 2024/1689 (EU AI Act), 2024.
[8] U.S. SEC, "Cybersecurity Disclosure Rules," Final Rule, July 2023.
[9] NIST, "AI Risk Management Framework 1.0," January 2023.
[10] NIST, "Cybersecurity AI Profile (IR 8596)," December 2025.
[11] NIST, "Post-Quantum Cryptography Standards (FIPS 203/204/205)," August 2024.
[12] ISO/IEC 42001:2023, "Artificial Intelligence Management Systems."
[13] OWASP, "Top 10 for Agentic Applications," December 2025.
[14] Cloud Security Alliance, "MAESTRO Framework," February 2025.
[15] Cloud Security Alliance, "Agentic Trust Framework," February 2026.
[16] Mayer Brown, "Governance of Agentic AI Systems," February 2026.
[17] EY, "Fortune 100 Board AI Oversight Analysis," 2025.
[18] Forrester, "AI Governance Software Market Forecast," 2025.
[19] RAND Corporation, "Joint All-Domain Command and Control," 2022.
[20] NATO C2 Centre of Excellence, "AI in Military C2 Systems," 2024.
[21] PwC, "Enterprise AI Agent Adoption Survey," 2025.
[22] Verizon, "Data Breach Investigations Report 2025."
[23] CrowdStrike, "Global Threat Report 2026."
[24] Hitch Partners, "CISO Security Leadership Survey 2025."
[25] AWS, "Security Bulletin AWS-2025-015: Amazon Q Developer Extension," July 2025.
[26] Fortune, "AI Coding Tools Exploded in 2025. The First Security Exploits Show What Could Go Wrong," December 2025.
[27] Pillar Security, "Rules File Backdoor Attack on AI Coding Assistants," 2025.
[28] Palo Alto Networks, "6 Predictions for the AI Economy: 2026 New Rules of Cybersecurity," November 2025.
[29] Adversa AI, "Top AI Security Incidents — 2025 Edition," July 2025.
[30] Morrison Foerster, "Flipping the NIS2 Switch: Germany Implementation," December 2025.
[31] Greenberg Traurig, "NIS2 in Germany: The New BSI Act," December 2025.
[32] DLA Piper, "NIS2 Directive Explained: Management Bodies Rules," November 2025.
[33] UK ICO, "Marriott International Inc. Enforcement Notice," October 2020.
[34] US FTC, "Marriott International Settlement," October 2024.

THE CISO AUTONOMY MANDATE

"If it cannot be evidenced, it cannot be defended."

Kieran Upadrasta
info@kieranupadrasta.com | www.kie.ie
Cyber AI Systems Inc.

Board-Survivable Cyber Architecture