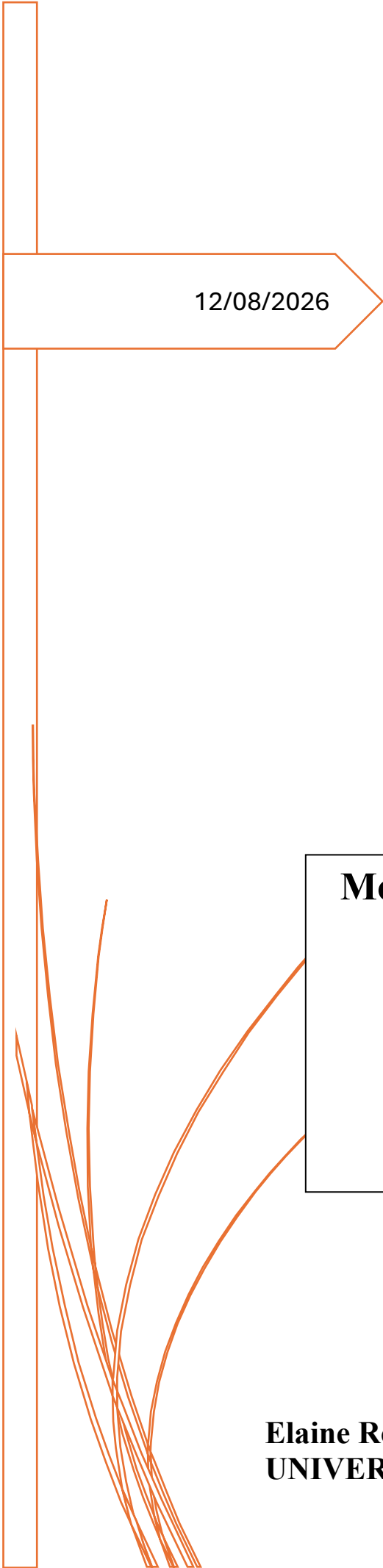




12/08/2026

**Assignment Title: Research
and case scenario in Data
Security and Compliance**

Module: Data Security and Compliance

Module code: ACCA5034

lecturer: Dr Ayman El Hajjar

Student Name: Elaine Rodovalho

Student Number: 2427102

Elaine Rodovalho(2427102)

UNIVERSITY OF WALES TRINITY SAINT DAVID

Table of Contents

Task 1 – Research in Data Security and Compliance	2
Abstract	2
1. Introduction.....	3
10. Conclusion	12
Task_1 References	13
Task 2 – Equifax Case Scenario.....	14
Executive Summary.....	14
1. Introduction and Case Foundation.....	14
8. Conclusion	22
Task_2 References	23

Task 1 – Research in Data Security and Compliance

Governance of AI and Machine Learning Systems: Managing Data Security, Ethical Risk and Regulatory Compliance

Abstract

Artificial intelligence (AI) and machine learning (ML) systems create value from large volumes of data, but they also introduce security, privacy, ethical and regulatory risks that conventional IT controls do not fully address. This paper critically evaluates how modern organisations can govern AI systems across their lifecycle so that they anticipate, protect against, detect and respond to threats. It argues that effective governance requires a combined approach: board-level accountability, risk-based regulation, secure machine-learning operations, continuous monitoring and carefully controlled automated response. Frameworks such as the NIST AI Risk Management Framework and ISO/IEC 42001 provide useful structures, while the EU AI Act and data-protection law create enforceable duties. However, frameworks alone can produce “paper compliance” if organisations lack reliable inventories, technical evidence, independent challenge and incident readiness. The paper concludes that the strongest model is continuous and evidence-based, with automation used for rapid containment but human authority retained for decisions that may affect rights, safety or access to services.

Keywords: AI governance, machine learning security, data protection, compliance, automated response, AI risk management.

1. Introduction

Artificial intelligence has moved from experimental laboratories into recruitment, lending, healthcare, education, fraud detection, customer service and cybersecurity. Unlike ordinary software, an ML model is influenced by training data, statistical relationships, model parameters, prompts, external tools and continuous feedback. This creates a wider attack surface. Sensitive information may be exposed through training data, logs, embeddings or model outputs, while inaccurate or biased decisions can harm individuals even when no traditional cyberattack has occurred.

The governance challenge is therefore not simply to secure a server. Organisations must govern the complete AI lifecycle: business purpose, data collection, model development, validation, deployment, monitoring, change and retirement. The EU AI Act applies risk-based obligations to high-risk systems, including risk management, data governance, logging, human oversight and robustness[1]In the United Kingdom, the UK GDPR and ICO guidance remain relevant whenever AI processes personal data, particularly where profiling or automated decisions affect people.

This paper addresses the research question: How can modern organisations create an environment that anticipates, protects, detects and automatically responds to data-security, ethical and compliance risks in AI and ML systems? It argues that the answer is an integrated AI governance system rather than a single security product. Such a system should combine the NIST functions of Govern, Map, Measure and Manage[2][2], an organisation-wide management system such as ISO/IEC 42001, and operational controls embedded in secure **MLOps**.

2. *Contemporary AI Threat and Compliance Landscape*****

AI risk has several connected dimensions. First, there are confidentiality risks. Training datasets may include personal, commercially sensitive or copyrighted information. Models can memorise rare records, and membership-inference attacks attempt to identify whether a particular record appeared in the training set[3].[2] Model-inversion and extraction attacks may reveal information about data or reproduce valuable model behaviour. Generative AI creates further exposure because employees may paste confidential material into public tools, a problem often described as shadow AI.

Second, integrity can be attacked. Data poisoning changes training data so that the model learns manipulated behaviour. Backdoors may remain hidden until a trigger is used. Evasion attacks modify inputs at inference time, while prompt injection attempts to override instructions or make an AI application reveal data, call unsafe tools or follow malicious content. NIST's adversarial-ML taxonomy and MITRE ATLAS show that attacks can occur at multiple lifecycle stages, not only after deployment,[4][3]OWASP also identifies risks such as prompt injection, sensitive-information disclosure, supply-chain weaknesses and excessive agency in LLM applications.

Third, AI creates ethical and compliance risk. Historical data may reproduce discrimination, proxy variables may indirectly represent protected characteristics, and performance can differ between groups. A technically accurate model may still be unlawful or unfair if people cannot understand, challenge or correct an important decision. The EDPB has emphasised accountability, fairness, transparency, lawful processing and data protection by design when personal data is used to develop or deploy AI models [5]. Poor data lineage can therefore become a privacy, bias and incident-response problem simultaneously.

Function	Main objective	Example controls	Evidence produced
Anticipate	Understand risk before harm occurs	AI inventory, impact assessment, DPIA, threat modelling, supplier review	Risk register, data sheet, model card, approval record
Protect	Reduce likelihood and impact	Data minimisation, encryption, least privilege, secure MLOps, constrained tools	Access records, signed artefacts, model registry, test results
Detect	Identify attack, drift or harmful outcomes	AI telemetry, anomaly detection, bias monitoring, red-team tests, complaints	Alerts, dashboards, audit logs, investigation tickets
Respond	Contain, recover and learn	Revoke credentials, block requests, quarantine or roll back model, human escalation	Incident record, notification decision, corrective-action plan

Table 1. Lifecycle controls for AI security and compliance

3.* Anticipating Threats and Compliance Failures*****

Anticipation begins with visibility. An organisation cannot govern systems it does not know it has. It should maintain an AI inventory covering internally developed models, third-party services, embedded AI features, datasets, prompts, model versions, owners, suppliers, intended users and affected individuals. Each use case should be classified by impact, data sensitivity, autonomy and regulatory status. A low-risk spelling assistant should not receive the same approval process as a system influencing credit, employment or medical treatment.

Before development or purchase, the organisation should complete an algorithmic impact assessment and, where personal data is likely to create high risk, a data protection impact assessment. Canada’s Algorithmic Impact Assessment demonstrates how structured questions can scale governance requirements according to impact [5]. The assessment should identify foreseeable misuse, affected groups, failure consequences, legal basis and human-oversight arrangements. It should be repeated after significant changes because a system’s risk can change when new data, tools or users are introduced.

Threat modelling must also be AI-specific. Security teams should map realistic attack paths using NIST’s adversarial-ML taxonomy, MITRE ATLAS and the OWASP LLM Top 10 [6]. Red-team exercises should test data poisoning, model extraction, prompt injection, jailbreaks,

unauthorised tool use and leakage of personal information. Supplier reviews should examine training-data provenance, retention, security testing, model updates and incident notification. Contracts are useful but weaker than technical evidence and independent testing. Good documentation supports anticipation. Datasheets for datasets record why data was collected, its composition, limitations and recommended uses. Model cards record intended use, performance, evaluation conditions and known limitations. These tools make assumptions visible, but they require ownership, version control and regular review to avoid becoming outdated or promotional.

4. *Protecting AI Data, Models and Services*****

Protection should apply defence in depth to data, models, infrastructure and people. Data minimisation is the first control: organisations should avoid collecting or retaining information merely because it may be useful later. They should establish a lawful purpose, remove unnecessary identifiers, define retention periods and separate training, validation and testing data. Encryption should protect data in transit and at rest, with keys managed separately. Highly sensitive projects may use differential privacy, secure enclaves, synthetic data or federated learning, although these can reduce accuracy or increase complexity.

Access must follow least privilege. Developers, data scientists, vendors and applications should receive only the data and model permissions needed for their roles. Strong authentication and short-lived privileged credentials reduce the chance that stolen accounts can download datasets or models. Production models should be isolated from development environments. Secrets must not be stored in notebooks, prompts or source code.

Secure MLOps extends secure software development into the model lifecycle and aligns with layered AI cybersecurity guidance [6]. Datasets, code, dependencies, prompts and model artefacts should be versioned in controlled repositories. Changes should pass peer review, security testing and validation before promotion. Model artefacts should be signed so that deployment systems can reject unauthorised versions. A model registry should record provenance, approval status, evaluation results and retirement dates. Supply-chain controls are particularly important because organisations increasingly rely on open-source models, external APIs and pre-trained components.

For generative and agentic AI, protection must include input and output controls. Applications should separate system instructions from untrusted content, restrict tool permissions, validate outputs before they reach databases or customers, and prevent the model from directly performing high-impact actions without authorisation. Content filtering and data-loss prevention reduce leakage but can be bypassed or block legitimate work. The strongest design therefore limits what the model can technically access or execute [7].

*****5. Detecting Security, Performance and Ethical Risk****

Traditional monitoring focuses on networks, endpoints and user accounts. AI governance requires these signals plus model and data telemetry. Every production system should log the model version, input source, relevant output, confidence or quality indicators, tool calls, user identity, policy decisions and errors, subject to privacy and retention limits. Logs should be tamper-resistant and linked to the security monitoring platform [8].

Detection should cover security, performance and fairness. Security analytics can identify unusual query volumes, repeated attempts to extract system prompts, abnormal token use, access from new locations or sudden downloads of model artefacts. Data monitoring can detect schema changes, missing values, unexpected categories and poisoning indicators. Model monitoring should detect accuracy degradation, concept drift, hallucination rates and differences in error rates between relevant groups. A system may remain available while silently becoming less fair or reliable, so operational uptime alone is an inadequate measure.

Continuous testing is more useful than a one-time launch assessment. Organisations can run scheduled adversarial tests, privacy-leakage tests and representative benchmark sets against each new model version. Canary prompts and synthetic records can reveal attempts to extract confidential information. Complaints should also be treated as detection signals because users may notice harmful patterns before dashboards do.

However, detection metrics have limitations. Fairness measures can conflict, ground truth may be delayed, and thresholds can generate false positives. Monitoring also creates additional personal data and may itself become intrusive. The organisation must document why each metric exists, who reviews it and what action follows; a dashboard without accountability is not governance [7].

6. *Automatically Responding to AI Incidents*****

Automated response can reduce the time between detection and containment. AI alerts should feed defined playbooks in security orchestration, automation and response systems. Confirmed credential misuse can trigger token revocation and session termination, while suspected exfiltration can block requests and isolate the affected endpoint. A compromised model can be quarantined, removed from the registry and replaced with a previously approved version. An agent behaving outside policy can have tool access disabled and be switched to read-only mode [7].

Responses should be tiered according to confidence and impact. Low-risk, reversible actions can be automatic, such as rate limiting, blocking a known malicious pattern or routing traffic to a safe fallback model. High-impact actions should normally require human approval, particularly where a response could deny a person employment, finance, healthcare or another important service. Speed is valuable, but incorrect automation can multiply harm. Human-in-the-loop controls should therefore be designed as clear decision rights, not as a vague statement that somebody can intervene.

Incident playbooks must include evidence preservation, legal assessment, notification, recovery and learning. Security, privacy, legal, technical and business owners need predefined roles. After containment, the organisation should determine whether the cause was malicious activity, poor data, model drift, supplier change or an unsuitable use case. Corrective action may involve retraining, changing thresholds, removing data, improving prompts, updating contracts or retiring the system. Lessons should update the risk assessment and control library so that response improves future anticipation [9].

7.* Governance, Ethics and Regulatory Compliance*****

Governance connects these controls and makes accountability visible, consistent with the OECD principles of robustness and accountability [6]. The board or executive committee should approve AI risk appetite and receive information about high-impact use cases, serious incidents and unresolved control gaps. An AI governance committee can coordinate security, privacy, legal, ethics, data science, procurement and business functions. Committees must not replace named ownership. Each system needs a business owner accountable for purpose and outcomes, a technical owner accountable for operation, and independent reviewers able to challenge both.

The NIST AI RMF is useful because it treats governance as a cross-cutting function and links context mapping, measurement and risk management [10]. Its voluntary and flexible nature supports different sectors, but it does not itself prove legal compliance. ISO/IEC 42001 provides a certifiable management-system approach and encourages policies, objectives, risk processes, monitoring and continual improvement. Certification can increase discipline, but may shift attention from real-world outcomes to auditable procedures.

Legal compliance should therefore be mapped to technical evidence. Under the AI Act, high-risk obligations include lifecycle risk management, suitable data governance, documentation, logging, human oversight, accuracy, robustness and cybersecurity. Data-protection law requires lawfulness, fairness, transparency, purpose limitation, data minimisation and security, while significant automated decisions may require additional safeguards [5]. International organisations should use one control framework with jurisdiction-specific mappings rather than disconnected programmes.

Culture is equally important. Employees need AI literacy on approved tools, confidential-data restrictions, escalation and model limitations. Developers and managers should be rewarded for reporting weaknesses rather than hiding them to protect a launch date. Without this culture, shadow AI and governance bypass will continue even where formal policies appear strong.

8. *Critical Evaluation and Implementation Challenges*****

A mature AI governance programme offers clear advantages. It reduces unknown systems, improves evidence for regulators, strengthens security testing and creates repeatable incident response. It can also support innovation by defining safer routes to experimentation. The best programmes use proportionality: stronger assurance is required when autonomy, data sensitivity and potential harm increase.

There are, however, important blind spots. First, governance can become a checklist exercise, with strong documentation but weak data or ignored complaints. Smaller organisations may lack specialists, testing tools and bargaining power over suppliers. Third, some risks cannot be measured precisely. Bias depends on context and affects populations, and explainability techniques may provide simplified explanations rather than the model's true reasoning. Fourth, protection controls involve trade-offs. Privacy techniques may reduce utility; strict filters may reduce accessibility; human review may introduce delay or human bias; and extensive logging may conflict with data minimisation.

The largest practical weakness is third-party opacity. A customer may not know a foundation model's training data, architecture or update process. Contractual controls, output testing and usage restrictions can reduce risk, but they cannot create complete assurance. Organisations should avoid placing an opaque external model at the centre of a high-impact decision unless the residual risk is justified and alternative controls exist.

Finally, automated response must be carefully bounded. A highly autonomous defence system may stop an attack quickly, but an attacker could intentionally trigger it to cause denial of service. False positives can also interrupt legitimate activity. Automation should therefore prioritise reversible containment, evidence preservation and escalation of uncertain cases. Governance is strongest when it accepts uncertainty rather than pretending every risk can be eliminated.

9. *Recommended Organisational Approach*****

Modern organisations should implement seven practical priorities. First, create a complete AI inventory and assign named owners. Second, classify use cases by impact and complete impact, privacy and security assessments before deployment. Third, embed data governance, access control, artefact signing, testing and approval gates into MLOps. Fourth, monitor security, quality, drift, fairness and user complaints continuously. Fifth, connect high-confidence alerts to reversible automated playbooks, with human approval for consequential actions. Sixth, maintain audit-ready evidence, including dataset documentation, model cards, logs, decisions and incident records. Seventh, test the whole governance system through red-team exercises and incident simulations, not only policy reviews.

These priorities should be measured through outcomes such as the percentage of AI systems inventoried, high-risk systems with current assessments, critical findings resolved before release, detection time, containment time, model rollback success and overdue human-review actions. This prevents governance being judged only by the number of documents produced.

10. Conclusion

AI and ML governance is now a central data-security and compliance responsibility. AI risks arise from the interaction of data, models, software, users and changing conditions. Organisations therefore need controls that operate across the full lifecycle. Anticipation requires inventories, impact assessments, threat modelling and supplier review. Protection requires data minimisation, access control, secure MLOps and constrained model capabilities. Detection requires continuous security, performance and fairness monitoring. Automated response can provide rapid containment through revocation, isolation, rollback and fallback systems, but consequential decisions must retain meaningful human authority.

No framework is sufficient on its own. The AI Act and data-protection law establish obligations, NIST provides a risk-management structure, ISO/IEC 42001 supports organisational discipline, and technical resources describe specific threats. Their value depends on evidence, implementation and independent challenge. The most effective environment is therefore risk-based, continuous and accountable: it uses automation to improve speed, while recognising that responsibility for AI outcomes remains with the organisation and its people.

Task_1 References

- [1] A. Vassilev, A. Oprea, A. Fordyce and H. Andersen, "Adversarial machine learning: A taxonomy and terminology of attacks and mitigations," 2024,
- [2] D. Casaburo, "EDPB's Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models: A Missed Chance?" vol. 11, no. 1, pp.70–75, 2025,
- [3] D. Casaburo, "EDPB's Opinion 28/2024 on Certain Data Protection Aspects Related to the Processing of Personal Data in the Context of AI Models: A Missed Chance?" vol. 11, no. 1, pp.70–75, 2025,
- [4] J. Ekeneme, C. Ucheji, C. Ezekwem and M. S. Chughtai, "Policy framework for responsible AI deployment in the national cybersecurity strategy," vol. 19, no. 10, pp.183–194, 2025,
- [5] O. A. Guirette-Barbosa, G. R. Romo-Flores, J. de Jesús Reyes-Sánchez and M. A. Ramírez-Salazar, "Artificial Intelligence and ISO 9001-2026: Transforming Quality Management Systems-Challenges, Opportunities, and Future Directions," pp.375–412, 2026,
- [6] N. AI, "Artificial intelligence risk management framework: Generative artificial intelligence profile," 2024,
- [7] E. Tabassi, "Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology," 2023,
- [8] M. Lognoul, "Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act–AI Act)," no. 3-4, pp.145–189, 2025,

Task 2 – Equifax Case Scenario

Designing a Structured and Forward-Looking Cybersecurity Programme

Executive Summary

The Equifax breach was not caused by one isolated technical error. It resulted from weak accountability, an incomplete asset inventory, ineffective patch verification, poor network segmentation and a long monitoring blind spot. This report designs a cybersecurity programme that treats security as an enterprise risk rather than an IT-only responsibility. It combines board oversight, measurable policies, NIST Cybersecurity Framework (CSF) 2.0, risk-based prioritisation, tested incident response and automated controls. The design aims to prevent known weaknesses, detect abnormal behaviour quickly, contain attacks before they spread and recover services without losing public trust.

1. Introduction and Case Foundation

In 2017, attackers exploited a known Apache Struts vulnerability in Equifax’s internet-facing dispute portal and obtained highly sensitive consumer information. The patch had been available, but it was not successfully applied to the affected system. Attackers then remained undetected for 76 days. Investigations also identified incomplete technology inventories, weak segmentation, unencrypted credentials and an expired monitoring certificate that prevented inspection of encrypted traffic [1]. The later regulatory settlement showed that cybersecurity failure creates financial, legal and reputational consequences, not only technical damage [2].

The central lesson is that a patching failure can pass through several organisational layers before becoming a breach. A forward-looking programme must therefore connect governance, people, process and technology. This report applies the lifecycle approach developed in Task 1: risks are governed before deployment, measured during operation and improved after incidents. The programme uses NIST CSF 2.0 as its main structure because its six functions—Govern, Identify, Protect, Detect, Respond and Recover—link executive decisions to operational outcomes.

2. Organisational Structure and Accountability

The proposed structure gives the board clear visibility while avoiding the conflict that occurs when the same executive is responsible for both rapid IT delivery and independent security challenge. A Board Risk and Audit Committee approves cyber risk appetite, reviews major exceptions and receives quarterly evidence on exposure, incidents and resilience. The Chief Executive Officer remains ultimately accountable. The Chief Information Security Officer (CISO) reports to the CEO and has direct access to the board, while the Chief Information Officer (CIO) remains responsible for reliable technology operations. This creates shared ownership: the CIO cannot transfer operational responsibility to security, and the CISO can escalate unacceptable risk without being overruled by delivery pressure [3].

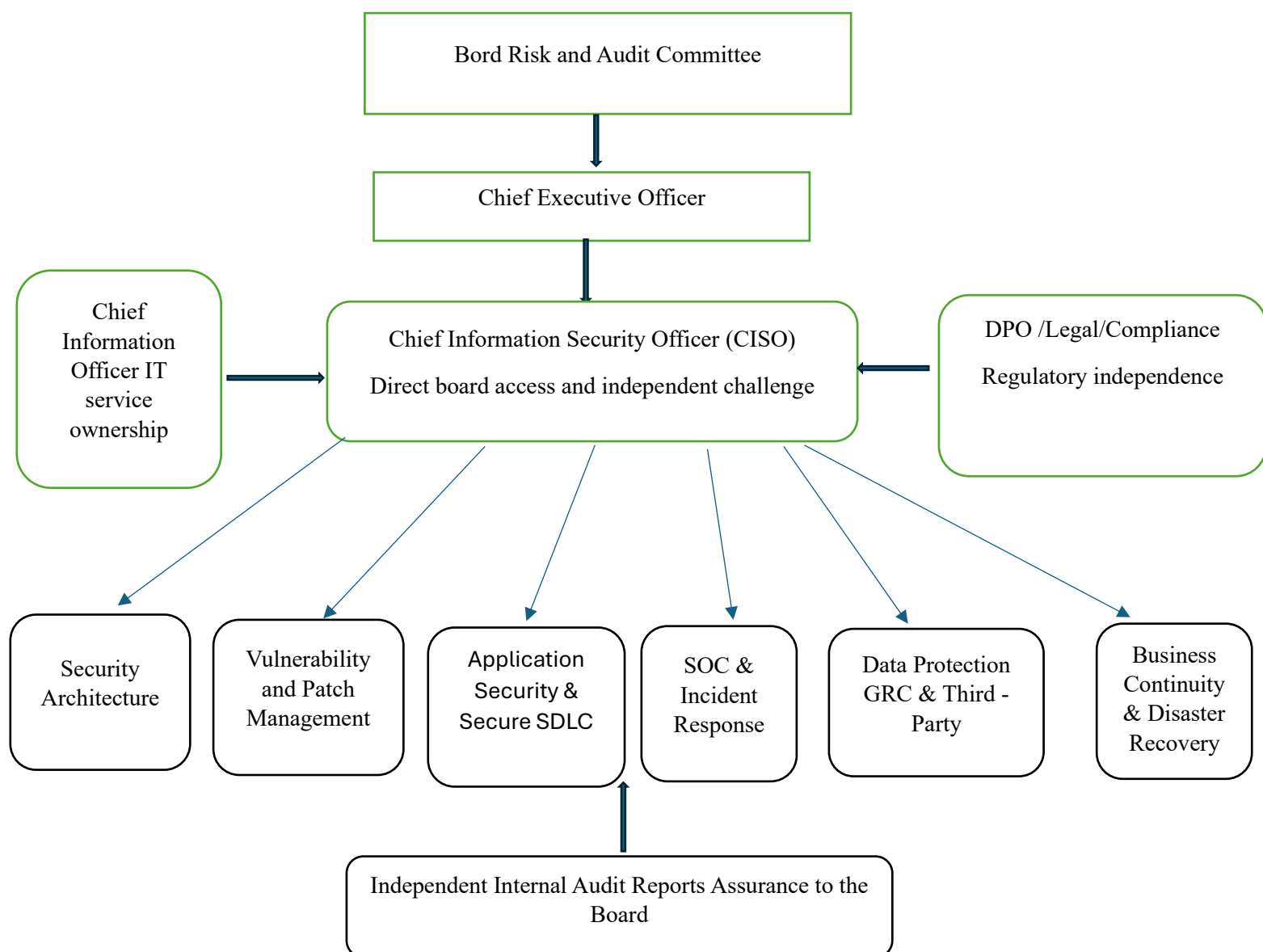


Figure 1. Proposed cybersecurity governance and reporting structure

Operational responsibility is divided into specialist teams: Security Architecture defines secure patterns; Vulnerability and Patch Management owns discovery and verification; Application Security oversees the secure software development lifecycle; the Security Operations Centre (SOC) and Incident Response team monitor and contain threats; Data Protection, Governance, Risk and Compliance (GRC) manages legal obligations and third-party risk; and Business Continuity coordinates recovery. Business system owners accept residual risk only within delegated limits. Internal Audit independently tests whether management evidence is reliable. This follows the principle of clear roles and oversight in NIST CSF 2.0 and NIST incident-response guidance [1]

Decision	Accountable role	Control against Equifax-type failure
Cyber risk appetite and funding	Board/CEO	Prevents known high risks being accepted silently.
Critical patch exception	CISO + business owner	Requires time limit, compensating control and evidence.
System and data ownership	Business owner/CIO	Every asset has an owner, criticality and support status.
Breach notification	DPO/Legal + CISO	Legal deadlines are met without delaying containment.
Independent assurance	Internal Audit	Tests the programme rather than relying on self-reporting.

Table 1. Key accountability decisions

3. Policies and Procedures

Policies must state ownership, scope, deadlines, evidence and consequences. A document that only says systems should be patched ‘promptly’ is not auditable. The vulnerability policy should combine asset criticality, exploit evidence and data sensitivity. Internet-facing vulnerabilities listed in CISA’s Known Exploited Vulnerabilities catalogue should be remediated or isolated within 48 hours; other critical issues within seven days and high issues within fourteen days. Emergency changes may be necessary but testing and rollback plans remain mandatory. Completion is proved by authenticated rescanning, not by an email stating that a patch was installed. NIST describes patch management as identifying, prioritising, acquiring, installing and verifying updates [11], while CISA recommends using known exploitation as a prioritisation input.

The secure development policy applies to internally developed software and acquired applications. It requires threat modelling for high-risk services, software-composition analysis for open-source components, static and dynamic testing, secrets scanning, peer review and security approval before production. NIST's Secure Software Development Framework supports integrating security practices throughout the development lifecycle [12]. Policies also cover asset discovery, certificate renewal, central logging, data classification, encryption, retention, privileged access and third-party security. Exceptions expire automatically and must be reviewed by the control owner and CISO.

Policy	Minimum requirement	Measure/evidence	Owner
Asset and software inventory	Daily discovery; owner, criticality, version and data class recorded.	≥98% discovery coverage; monthly reconciliation.	CIO
Vulnerability and patching	KEV/internet-facing: 48 h; critical: 7 days; high: 14 days.	Authenticated rescan; ≥95% SLA compliance.	Patch Manager
Secure SDLC	Threat model and automated testing before high-risk release.	100% critical apps pass security gate or approved exception.	AppSec Lead
Monitoring and certificates	Critical logs to SIEM; certificates centrally tracked.	100% critical log sources; renewal alert at 60/30/15 days.	SOC Lead
Data protection	Classification, minimisation, encryption and privileged-access control.	Quarterly access review; no plaintext production credentials.	DPO/Data Owner
Third parties	Risk assessment, security clauses and incident duties before onboarding.	Annual reassessment of critical suppliers.	GRC Lead

Table 2. Measurable policy register

4. Governance and Strategic Alignment

Cybersecurity strategy should support the organisation's purpose: reliable credit services, lawful processing of personal data and confidence from consumers, regulators and business customers. NIST CSF 2.0 converts these objectives into a Target Profile. Management first records the Current Profile, identifies gaps and funds improvements according to business impact. Board reporting should show risk reduction, not large counts of blocked attacks. Useful indicators include critical exposure days, patch-SLA compliance, percentage of known assets, mean time to detect and contain, privileged-access review completion and recovery-test success.

The new Govern function is especially important because it connects policy, roles, legal requirements and supply-chain risk to the remaining functions. However, a framework does not guarantee security. Organisations can create attractive maturity dashboards while weak evidence remains underneath. Metrics therefore require defined data sources, independent checking and trend analysis. The lifecycle governance principles from Task 1 are transferable: controls should be documented, traceable, continuously measured and subject to human oversight. ISO/IEC 27001 can provide the auditable management system, while the CSF provides the outcome structure [13].

CSF function	Programme application	Business outcome
Govern	Risk appetite, roles, policy, supplier oversight.	Accountability and defensible decisions.
Identify	Asset/data inventory and prioritised risk register.	Investment focuses on critical services.
Protect	Patching, MFA, segmentation, encryption, secure SDLC.	Lower probability and reduced blast radius.
Detect	SIEM, EDR/NDR, certificate and exfiltration alerts.	Shorter attacker dwell time.
Respond	Playbooks, escalation, SOAR and communications.	Faster containment and lawful notification.
Recover	Immutable backups, restoration tests and lessons learned.	Service continuity and restored trust.

Table 3. NIST CSF 2.0 alignment with business outcomes

5. Risk Management and Assessment

The programme uses a five-by-five likelihood and impact scale, but numerical scores are not treated as exact science. Priority also considers active exploitation, internet exposure, data sensitivity, control failure and the possibility of lateral movement. This avoids the Equifax weakness of treating separate technical findings as local operational issues. A known critical vulnerability on a public service containing identity data is an enterprise-level risk even if the affected application is old or considered low profile.

The first treatment option is remediation. Where patching cannot meet the deadline, the system must be isolated, protected by a tested compensating control or removed from service. Risk acceptance is temporary, named and visible to senior management. Supply-chain exposure is included because open-source dependencies can create the same vulnerability pathway as Apache Struts. Accurate asset and software inventories are therefore foundational controls: without them, scanning, patching and assurance produce false confidence [1].

Risk scenario	L	I	Score	Priority	Primary treatment
Known flaw on internet-facing legacy application	5	5	25	Critical	Patch/disable within 48 h; isolate and verify.
Incomplete asset or software-component inventory	4	5	20	Critical	Automated discovery, CMDB and ownership reconciliation.
Patch process failure or uncontrolled exception	4	5	20	Critical	SLA, escalation, rescan and expiring approval.
Flat network and excessive database connectivity	4	5	20	Critical	Segment services; least-privilege network paths.
Monitoring blind spot or expired certificate	4	4	16	High	Certificate automation and control-health alerts.
Plaintext credentials/excessive privileged access	4	5	20	Critical	Vaulting, PAM, MFA and quarterly review.
Third-party or open-source component compromise	3	5	15	High	Supplier due diligence, SBOM/SCA and contract duties.

Table 4. High-level cyber risk assessment

6. Compliance, Incident Response and Protection Controls

A multinational data organisation needs a regulatory matrix covering the jurisdictions, data types and contractual duties affected by an incident. In the United Kingdom and European Union, a notifiable personal-data breach must be reported to the relevant supervisory authority without undue delay and, where feasible, within 72 hours of awareness; high risk to individuals can also require direct communication [14]. In the United States, the response team must coordinate federal obligations, state breach laws, consumer-reporting duties and regulator commitments. The Equifax settlement illustrates that failure to use reasonable safeguards can create very large remediation and enforcement costs.

The incident response plan follows NIST SP 800-61 Rev. 3 and is integrated across the CSF functions rather than isolated within the SOC [14]. A severity-one data breach triggers immediate evidence preservation, incident command, legal and privacy participation, executive escalation and a notification clock. Playbooks cover exploited web applications, stolen credentials, data exfiltration, ransomware and third-party compromise. Quarterly tabletop exercises and annual technical simulations test decision-making, contact lists, regulator reporting, customer communication and recovery. Critical services have approved recovery time and recovery point objectives, with immutable backups and restoration tests.

Technical safeguards work as a connected system. Automated external scanning and software-composition analysis identify exposed weaknesses; EDR detects suspicious processes; network detection and SIEM correlate unusual database queries, web traffic and identity events; DLP and egress monitoring identify large data transfers; MFA and privileged-access management reduce credential abuse; encryption protects data at rest and in transit; and segmentation limits lateral movement. SIEM and SOAR can automatically enrich alerts, block malicious addresses, disable compromised accounts or isolate endpoints. Automation reduces dwell time, but disruptive actions require guardrails, confidence thresholds and human approval because false positives can interrupt critical services. Encryption also has limits: it cannot protect data when an authorised application or stolen account decrypts it. Controls must therefore be layered rather than presented as single solutions [12].

Control	Contribution	Important limitation
Continuous scanning + CMDB	Finds exposed assets and verifies remediation.	Fails when discovery credentials or scope are incomplete.
SIEM, EDR and NDR	Correlates suspicious behaviour and accelerates triage.	Poor logging and untuned rules create blind spots/noise.
SOAR containment	Blocks indicators or isolates systems in minutes.	Needs approval gates for high-impact actions.
Segmentation and PAM	Restricts lateral movement and privileged access.	Requires accurate service dependencies.
Encryption and DLP	Reduces data exposure and detects exfiltration.	Does not stop misuse by valid application access.
Backups and DR tests	Restores critical services and limits disruption.	Backups do not reverse confidentiality loss.

Table 5. How key controls reduce dwell time and impact

7. Phased Implementation and Critical Evaluation

During the first 90 days, the organisation should discover internet-facing assets, remove known exploited vulnerabilities, renew and centrally manage certificates, identify critical data stores and test the incident escalation process. During months three to six, it should deploy or improve MFA, EDR, central logging, privileged-access controls and segmentation around high-value databases. During months six to twelve, secure SDLC gates, supplier assurance, recovery exercises and independent maturity assessments should become routine.

The main danger is treating the programme as a technology purchase. Tools cannot repair unclear ownership, unsupported legacy systems or incentives that reward availability while hiding security debt. A second danger is excessive automation: automatic isolation is valuable for a confirmed malicious endpoint but unsafe when evidence is weak. Success therefore depends on accountable decisions, reliable inventories, tested procedures and transparent exceptions. Board attention should focus on remaining exposure and overdue corrective actions, not only positive compliance percentages.

Continuous assurance is required after implementation. Control owners should provide monthly evidence, while GRC samples the evidence and Internal Audit performs independent reviews. Red-team exercises should test whether an attacker can move from a public application to sensitive databases, and purple-team sessions should confirm that the SOC can observe the activity. Lessons from real incidents, near misses and failed recovery tests must update the risk register, policies and technical detection rules rather than remain in a final

report. NIST CSF 2.0 supports this improvement cycle, and NIST incident-response guidance stresses that lessons should be shared throughout response and recovery [15], [16]. The board should receive a balanced view containing both leading indicators, such as overdue critical patches, and lagging indicators, such as incident dwell time. This reduces the risk of a misleading dashboard in which high policy compliance hides one unmanaged legacy system. Annual external assurance is useful, but it should supplement rather than replace management accountability.

8. Conclusion

The Equifax breach demonstrates how a preventable vulnerability becomes a major crisis when governance, asset visibility, patch verification, monitoring and segmentation fail together. The proposed programme addresses these weaknesses through independent CISO authority, board oversight, measurable policies, NIST CSF 2.0 alignment, risk-based treatment, layered technical safeguards and tested response and recovery. Its strongest feature is traceability: each major risk has an owner, deadline, evidence source and escalation route. This does not eliminate cyber risk, but it makes risk visible, shortens attacker dwell time, limits the amount of data reachable after compromise and supports faster, lawful recovery.

Task_2 References

- [1] K. Haufe, "Information Privacy Management-A", in Proceedings of International Conference on Theoretical and Applied Computing: ICTAC 2024, pp. 129.
- [2] O. A. Guirette-Barbosa, G. R. Romo-Flores, J. de Jesús Reyes-Sánchez and M. A. Ramírez-Salazar, "Artificial Intelligence and ISO 9001-2026: Transforming Quality Management Systems-Challenges, Opportunities, and Future Directions," pp.375–412, 2026,
- [3] E. Tabassi, "Artificial intelligence risk management framework (AI RMF 1.0)," 2023,
- [4] S. Rose, O. Borchert, S. Mitchell and S. Connelly, "Zero trust architecture," vol. 800, no. 207, pp.1–52, 2020,
- [5] G. EU, "General data protection regulation," 2016,
- [6] M. Souppaya, K. Scarfone and D. Dodson, "Secure software development framework (ssdf) version 1.1," vol. 800, no. 218, pp.800–218, 2022,
- [7] A. Nelson, S. Rekhi, M. Souppaya and K. Scarfone, "Incident response recommendations and considerations for cybersecurity risk management," vol. 800, 2025,
- [8] M. Souppaya and K. Scarfone, "Guide to enterprise patch management planning," 2022,
- [9] C. Pascoe, S. Quinn and K. Scarfone, "The NIST cybersecurity framework (CSF) 2.0," 2024,
- [10] Federal Trade Commission, "Equifax to Pay \$575 Million as Part of Settlement with FTC," 2017,
- [11] Federal Trade Commission, "Equifax to Pay \$575 Million as Part of Settlement with FTC," 2017,
- [12] A. T. by Equifax, "Federal Agencies in Response to the 2017 Breach," 2018,
- [13] L. Alkarmi, A. Sarabi and M. Liu, "Estimating the Social Cost of Corporate Data Breaches," 2026,