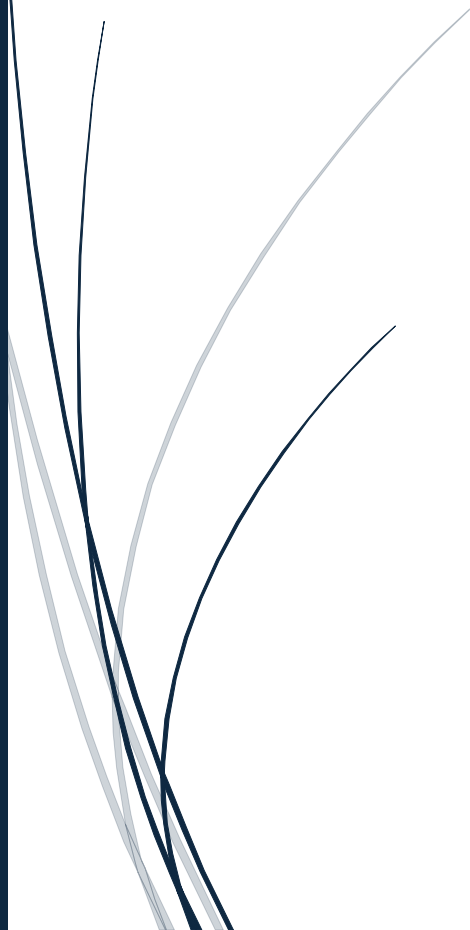


4/29/2026

Data Security and Compliance (ACCA5034) – Assessment

Lecturer: Dr. Ayman El Hajjar



Student: Andrei Stefanita Stoica (2413927)
UNIVERSITY OF WALES TRINITY SAINT DAVID –
LONDON – CANARY WHARF BRANCH

Table of Contents

1. Governance of AI and Machine Learning Systems	2
1.1 Abstract	2
1.2 Introduction	2
1.3 The Unmistakable Nature of AI Risk.....	3
1.3.1 Technical and Operational Risks	3
1.3.2 Ethical and Societal Risks.....	4
1.3.3 Regulatory and Compliance Risks	4
1.4 Foundations of AI Governance Frameworks.....	5
1.5 Lifecycle-Based Governance of AI Systems	5
1.5.1 Design and Development: Anticipating Risks	5
1.5.2 Deployment: Protecting Systems	6
1.5.3 Operation: Detecting Issues.....	7
1.5.4 Continuous Improvement: Responding to Threats.....	7
1.6 Anticipating, Protecting, Detecting and Responding: An Integrated Model	8
1.7 Challenges in Implementing AI Governance	8
1.8 Critical Evaluation of Current Approaches	8
1.9 Recommendations for Effective AI Governance	9
1.10 Conclusion	9
2. Zero Trust Security in AI Healthcare: MediTrust Case Study.....	10
2.1 Abstract	10
2.2 Introduction	10
2.3 Organizational Structure for Security Governance.....	12
2.4 Policies and Procedures	13
2.4.1 AI Usage Policy	13
2.4.2 Secure SaaS Configuration Policy	14
2.4.3 BYOD Policy for Hybrid Workforce.....	14
2.5 Governance and Strategy: NIST CSF 2.0 Alignment	14
2.6 Compliance with DORA and NIS2.....	15
2.7 Risk Management.....	16
2.8 Contingency Planning: Ransomware Scenario.....	17
2.9 Protection Mechanisms.....	18
2.10 Critical Evaluation.....	19
2.11 Conclusion.....	19
3. References.....	19

1. Governance of AI and Machine Learning Systems

1.1 Abstract

Artificial Intelligence (AI) and Machine Learning (ML) systems are increasingly being implemented in critical organizational processes creating both unprecedented opportunities and complex risks. Unlike traditional software, AI systems rely heavily on data which often is garnered through unethical ways by corporate giants under the guise of protection [1]. Furthermore, AI systems exhibit probabilistic behavior and can evolve over time, making governance significantly more challenging. This paper critically examines how modern organizations can design governance frameworks which anticipate, protect against, detect and automatically respond to risks associated with AI systems. Aligning with already established standard paradigms such as the NIST AI Risk Management Framework and emerging Regulatory approaches like the EU AI Act, the study focuses on embedding ethical and secure controls across the technological lifecycle. The paper argues that effective AI governance must evolve beyond static policies and adopt a more adaptive, automated framework and implement it within organizational processes. It concludes that organizations adopting a more proactive and lifecycle-oriented governance model are better positioned to ensure resilience, trustworthiness and regulatory compliance [2].

1.2 Introduction

Artificial intelligence started as experimental research but progressively evolved into a foundational component of modern society. Organizations now deploy AI systems in sensitive domains such as medical diagnostics, financial decision-making, fraud detection and cybersecurity. While these systems offer efficiency and innovation, at the same time, they also introduce new forms of risk which are fundamentally different from those found in traditional programs or information systems.

AI systems are inherently data-driven and often operate as opaque “black boxes”, which in turn, makes their internal decision-making processes difficult to interpret. Moreover, they are dynamic, meaning their behavior can change over time due to new data inputs or model updates. Considering how rapidly the AI race is progressing and

how much data is being inputted into training AI systems, sudden shifts in behavior are to be expected. These characteristics complicate accountability, increase vulnerability to manipulation and bias, in addition to creating further challenges for the compliance with legal and ethical standards.

Over the past years of AI governance surging as a question to the rapidly growing industry, it has become clear that AI governance has evolved into being a critical discipline. It encompasses the structures, policies, processes and technical mechanisms which ensure AI systems operate in a secure, ethical and compliant manner. Therefore, effective governance must address not only technical risks but also societal and organizational impacts such as fairness, transparency and accountability.

This paper explores the venue of how organizations can establish governance environments which thoroughly anticipate risks, protect systems, detect oddities and most importantly, respond to incidents in a timely manner. The perspective presented suggests that governance must be embedded throughout the development, deployment and operation of AI systems while still maintaining a high level of transparency to avoid introducing doubt [3] [4].

1.3 The Unmistakable Nature of AI Risk

1.3.1 Technical and Operational Risks

AI systems differ from standard software due to them relying heavily on training data and inferential statistics. This dependency introduces risks such as data quality issues, bias and the chance for the AI model to shift its way of processing which could render it dangerous to use in sensitive industries. For instance, an adequate example of an AI system drifting is the SaaSr major data blunder, where an emerging AI focused company building tools for B2B founders deleted their entire production database during a code freeze. Another example of an AI model exhibiting unintended behaviors may occur when the training data contains historical biases, subsequently perpetuating or even amplifying discriminatory outcomes. In a similar manner, model drift can occur when the underlying data distribution changes, causing performance issues over time.

Another critical risk may arise from outsider attacks. Attackers can manipulate input data to deceive AI models into incorrect predictions, data leaks and other major data catastrophes. This requires high-stake applications such as autonomous vehicles or medical systems to provide robust defense mechanisms against possible data related disasters which, in turn, could have severe consequences [3] [5].

1.3.2 Ethical and Societal Risks

The main issue with biased algorithms and AI systems is, if they are to be implemented in areas such as law enforcement or hiring, they could lead to unequal treatment of individuals raising serious ethical questions related to fairness, accountability and transparency. Furthermore, the lack of explainability in many AI models further complicates the aforementioned issues as, for instance, stakeholders may not understand how certain decisions are made.

Ethical risks are closely linked to reputational and legal consequences. Organizations who decide to deploy biased or opaque AI systems may face public scrutiny, regulatory penalties from governments and loss of trust from future potential stakeholders [3][6].

1.3.3 Regulatory and Compliance Risks

The regulatory landscape for AI is evolving as rapidly as the technology used to advance it, with governments introducing frameworks to ensure responsible AI use. The European Union's AI Act, for example, classifies AI systems based on risk levels and imposes strict requirements on high-risk applications. Compliance with such regulations requires organizations to implement robust governance mechanisms, including documentation, transparency and risk management processes. A 2025/2026 report commissioned by Tenable highlights a critical "AI exposure gap," finding that only 22% of organizations fully classify and encrypt artificial intelligence (AI) data. This leaves the remaining 78% of organizations with AI data inadequately secured and vulnerable to exploitation. Despite 89% of organizations running or piloting AI workloads as of late 2025, security measures are not keeping pace with adoption rates, suggesting this vulnerability gap will likely widen. Consequently, 34% of organizations implementing AI in their workflows have already reported experiencing an AI-related security breach [6] [7].

1.4 Foundations of AI Governance Frameworks

The foundations of AI governance frameworks rest on a set of core principles, structural elements and lifecycle-oriented controls which collectively ensure artificial intelligence systems operate in a trustworthy, ethical and accountable manner. At the heart of these frameworks are key principles such as fairness, transparency, accountability, privacy and safety, which guide how AI systems are created and used. These principles are supported by clearly defined organizational structures that assign roles, responsibilities and oversight authority across technical, legal and organizational stakeholders, ensuring that accountability is embedded throughout the system lifecycle.

A robust foundation begins with clearly defining the purpose of an AI system, ensuring it is used in ways which align with organizational goals and social expectations. In combination with a solid foundation, a strong emphasis on data governance including data quality, integrity and provenance is required, thus recognizing that reliable data inputs essentially result in trustworthy outputs.

Risk management is another crucial central pillar, it involves systematic identification, assessment and mitigation of potential harms such as bias, discrimination and system failures. Importantly, these frameworks implement lifecycle governance mechanisms including continuous monitoring, auditing and feedback loops to address the dynamic and ever-evolving nature of AI systems. Furthermore, documentation and transparency practices such as model reporting and explainability measures reinforce trust and enable external scrutiny if there need be. Additionally, alignment with emerging regulatory standards and industry best practices provides a structured baseline for compliance while promoting consistency across different contexts. Together, these foundational elements create a practical and balanced approach to guiding the responsible use of AI [4] [8].

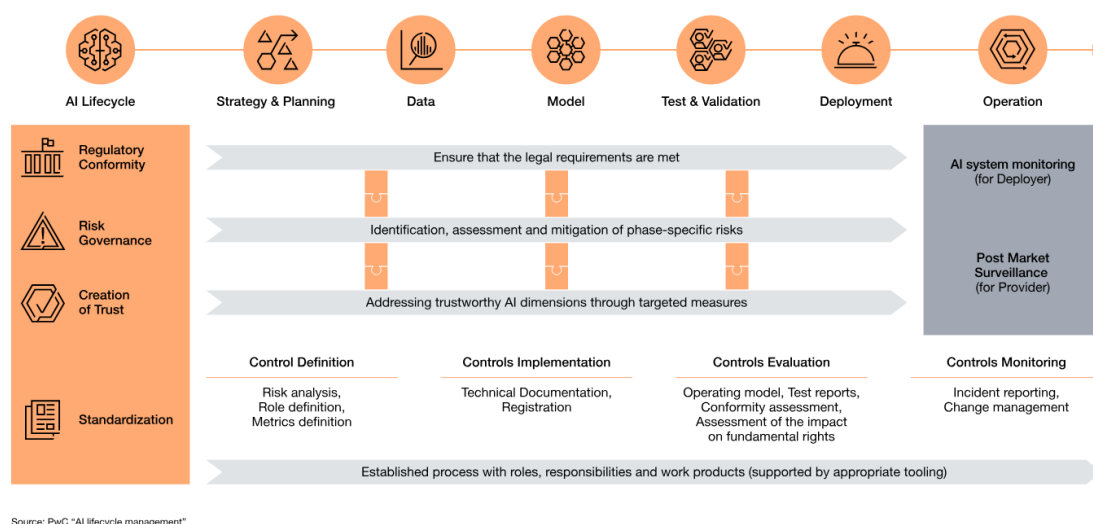
1.5 Lifecycle-Based Governance of AI Systems

1.5.1 Design and Development: Anticipating Risks

Lifecycle-based governance frameworks emphasize that risk management must begin at the earliest stages of AI system creation. Foundational guidance such as the National Institute of Standards and Technology (NIST) AI Risk Management Framework highlights that governance should anticipate, identify and mitigate risks

during design rather than reacting post-deployment. In this phase, organizations are expected to conduct structure use case-scoping, feasibility analysis and rigorous testing to evaluate bias, robustness and security vulnerabilities. These processes enable the identification of potential harms and societal impacts before systems are put in place. Industry perspectives further reinforce that implementing governance controls early enhances trustworthiness and regulatory compliance across the lifecycle. As such, design-stage governance represents a proactive approach, ensuring safeguards are built into AI systems prior to their deployment [9] [10] [11] [12].

The implementation of governance is ensured throughout the AI lifecycle

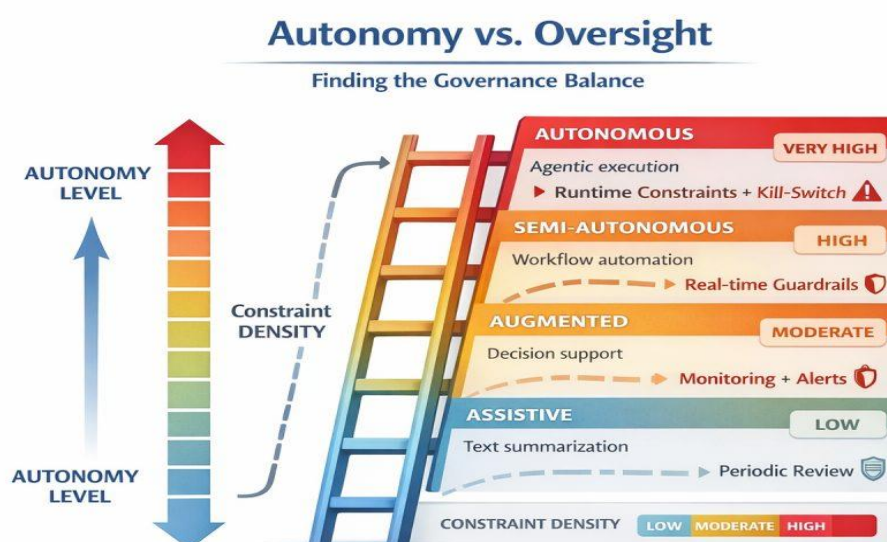


1.5.2 Deployment: Protecting Systems

During deployment, governance shifts from anticipation to protection, focusing on ensuring only validated and compliant systems are introduced into real-world environments. This stage involves enforcing approval thresholds, performance validation and regulatory compliance checks before release. Additionally, governance frameworks stress the importance of clearly defined accountability structures, specifying roles and responsibilities for system supervision and risk management. Deployment governance must also address operational risks, including vulnerabilities associated with third party dependencies, infrastructure and cybersecurity threats. The addition of these controls will benefit organizations in preventing unsafe or inadequately tested systems from scaling, therefore ensuring a secure and compliant transition from development to practical application [13] [14] [15].

1.5.3 Operation: Detecting Issues

Once AI systems are deployed, governance priorities shift toward continuous monitoring and real-time risk detection. Effective operational governance requires mechanisms to identify model drift, performance degradation, bias and other anomalies which may arise in dynamic environments. Advanced approaches incorporate anomaly detection and runtime policy enforcement to maintain system integrity and alignment with intended objectives. Public sector guidance similarly emphasizes ongoing performance monitoring and risk assessment as essential for maintaining safe and reliable AI systems over time. This phase reflects a shift from predictive risk management to continuous assurance, where governance systems must actively detect and respond to issues as they emerge in real-world situations [11] [13] [16].



1.5.4 Continuous Improvement: Responding to Threats

The final stage of lifecycle-based AI governance focuses on continuous improvement, recognizing that static management mechanisms are insufficient in rapidly evolving technological and threat landscapes. Governance frameworks highlight the importance of feedback loops, incident response processes and periodic reassessment to ensure systems remain effective, safe and compliant over time. The concept of “governance drift” emphasizes how controls can become outdated if not regularly updated, therefore necessitating recurring refinement of both models and policies. Research in MLOps further supports this approach, emphasizing continuous

development cycles driven by monitoring insights and new data. This adaptive governance model ensures organizations can respond effectively to emerging threats, thus maintaining resilience and trustworthiness across the AI lifecycle [13] [17].

1.6 Anticipating, Protecting, Detecting and Responding: An Integrated Model

An integrated model of AI governance conceptualizes management as a continuous cycle of anticipating, protecting, detecting and responding to risks across the AI lifecycle. Research emphasizes that effective governance must align technical, organizational and regulatory mechanisms into a unified system rather than relying on split interventions. Comprehensive surveys further argue that governance should integrate technical robustness, ethical accountability and real-world harm mitigation, offering a clear warning that reactive approaches undermine effectiveness. This model recognizes that AI systems operate in dynamic environments which require governance frameworks that combine ex ante controls with real time monitoring and adaptive response strategies [18] [19].

1.7 Challenges in Implementing AI Governance

Putting AI governance into practice is much harder than it appears in theory. One major issue is that, although many frameworks outline ethical principles like fairness and accountability, there is often little guidance on how to apply these ideas in real-world situations, leading to inconsistent approaches across organizations and sectors. In addition, many organizations struggle with practical barriers such as limited expertise, disconnected data systems, and unclear responsibility for managing AI risks. Laws and regulations also tend to lag behind technological developments, creating uncertainty about what standards should be followed and making compliance more difficult. Furthermore, effective governance requires collaboration between technical experts, policymakers, and ethicists, but this kind of coordination is often lacking in practice. Overall, these challenges show that AI governance is not just a technical task but a broader organizational and societal issue [20] [21].

1.8 Critical Evaluation of Current Approaches

When existing AI governance approaches are examined more closely, several weaknesses become clear. Many frameworks focus on broad ethical principles but do not provide clear ways to enforce or measure them, creating a gap between what

organizations say they will do and what they actually implement. This can result in organizations appearing responsible without making meaningful changes, sometimes referred to as “ethics washing.” In addition, governance efforts often struggle to keep up with the rapid pace of AI development, making it difficult to address new and unexpected risks as they arise. There is also limited coordination between regulators and organizations, which can lead to fragmented and inconsistent oversight. As a result, current approaches are often not strong or flexible enough to fully manage the risks associated with modern AI systems [20] [22].

1.9 Recommendations for Effective AI Governance

To improve AI governance, researchers and practitioners suggest focusing on more practical and adaptable approaches. One key recommendation is to embed governance throughout the entire lifecycle of an AI system, from design to deployment and ongoing use, rather than treating it as a one-time check. This includes setting clear responsibilities, improving data management, and ensuring that systems are regularly reviewed and tested. Organizations should also invest in building teams with a mix of technical, legal, and ethical expertise to better manage complex AI systems. In addition, continuous monitoring and feedback are essential so that problems can be identified early and addressed quickly. Overall, these recommendations highlight the need for governance systems that are not only well-designed but also flexible enough to adapt to change, helping ensure that AI remains safe, reliable, and trustworthy over time [21] [23].

1.10 Conclusion

In conclusion, while AI governance has developed significantly in terms of frameworks and guiding principles, there are still clear gaps in how effectively these ideas are applied in practice. The literature consistently shows that governance cannot rely on static rules or one-time assessments but instead must function as an ongoing process that anticipates risks, protects systems, detects issues, and responds to emerging challenges over time. At the same time, practical difficulties such as limited expertise, regulatory uncertainty and the fast pace of technological change continue to limit how well governance systems perform in real-world settings. Addressing these issues will require closer collaboration between organizations, policymakers, and researchers, as well as a stronger focus on making governance approaches more practical,

enforceable, and adaptable. Ultimately, effective AI governance depends not only on having well-designed frameworks, but on ensuring they can be consistently applied and updated, so that AI systems remain safe, ethical and trustworthy in increasingly complex environments.

2. Zero Trust Security in AI Healthcare: MediTrust Case Study

2.1 Abstract

The integration of Artificial Intelligence (AI) into healthcare systems has significantly enhanced diagnostic accuracy, operational efficiency and patient care delivery. However, these benefits are accompanied by complex challenges related to data security, governance and regulatory compliance. MediTrust Analytics Ltd, a HealthTech organization operating across multi-cloud environments, processes highly sensitive patient data, including medical records and real-time IoT inputs. As the organization expands into the European Union, it must comply with strict regulatory frameworks such as the General Data Protection Regulation (GDPR), the Digital Operation Resilience Act (DORA) and the NIS2 Directive.

This part of the paper proposes a comprehensive zero trust driven security program tailored to MediTrust's operational and regulatory requirements. The design integrates governance frameworks, risk management strategies and modern security controls aligned with the National Institute of Standards and Technology (NIST) Cybersecurity Framework (CSF) 2.0. It further incorporates AI governance principles, emphasizing lifecycle-based risk management and continuous monitoring. The study demonstrates how organizations can anticipate, protect, detect and respond to threats within AI-driven environments. It concludes that a unified, adaptive and governance driven approach is essential for achieving resilience, compliance and trust in modern healthcare systems.

2.2 Introduction

The rapid adoption of AI technologies in healthcare has transformed clinical decision-making processes, enabling predictive analytics, real time monitoring and

personalized treatment strategies. MediTrust Analytics Ltd represents a typical modern HealthTech organization which relies heavily on AI-driven systems to process large volumes of sensitive data. However, this reliance introduces new categories of risk, including vulnerabilities in training data, exposure to attacks and challenges related to explainability and accountability.

AI systems have been on a continuous upward trend in the past few years due to cost effectivity, meaning that newer technologies such as cloud software are significantly more affordable than in the past. This upward trend is predicted to progressively increase throughout the next several years.

Unlike traditional IT systems, AI systems are inherently probabilistic and data dependent. This creates uncertainties in system behavior and increases the difficulty of ensuring consistent security and compliance. Research highlights that AI systems are particularly susceptible to risks such as data poisoning, model manipulation, model drift and unintended bias, which can lead to incorrect clinical decisions and significant legal consequences. Furthermore, builders of such AI systems do not take the necessary precautions when creating AI-driven programs, which subsequently create problems of a massive scale depending on which industry they are implemented in.

AI-driven systems have proven to pose challenges in various areas such as data handling, which present significant complexities particularly regarding the systematic collection, integration and preparation of datasets. The main focus to these challenges is the necessity for robust quality assurance mechanisms including reliable alerting systems and thorough version control to ensure comprehensive data traceability [3] [17].

At the same time, the regulatory landscape is evolving rapidly out of necessity to keep up with the briskly, ever-evolving AI systems. Spanning over 100 articles, The EU AI Act introduces risk-based classification of AI systems, requiring strict controls for high-risk applications such as healthcare, biometrics and identification, critical infrastructure, education and vocational training, law enforcement, employment and justice [6]. Additionally, DORA and NIS2 impose strict requirements for operational resilience, incident reporting and supply chain security.

To address the challenges posed by AI-driven systems, organizations must adopt a proactive and integrated approach to security governance. Zero trust architecture

provides a suitable foundation as it eliminates implicit trust and enforces continuous verification across all systems. When combined with governance frameworks such as the NIST CSF and AI lifecycle management principles, zero trust enables organizations to manage risks effectively while maintaining compliance and operational efficiency. The risk management is a fundamental element in the ethical development and deployment of artificial intelligence. By implementing responsible AI practices, organizations can ensure that system design and utilization remain consistent with established societal values and intended objectives. [2] [10].

This part of the paper presents a strategic design for MediTrust's security program, which focuses on organizational structure, policies, governance and risk management.

2.3 Organizational Structure for Security Governance

The effectiveness of any security program depends on well-defined organizational structure which integrates governance, technical expertise and compliance functions. MediTrust must establish a dedicated information security unit capable of addressing both traditional cybersecurity risks and emerging AI specific challenges.

At the strategic level, the Chief Information Security Officer (CISO) is responsible for aligning security initiatives with business objectives and regulatory requirements. This role ensures that cybersecurity is not treated as a purely technical issue but as a core organizational priority [24].

The Cloud Security Architect plays a critical role in managing MediTrust's multi-cloud infrastructure. Given the organization's reliance on cloud platforms, this role ensures consistent implementation of security controls and best practices including identity and access management, encryption, cloud security monitoring and vulnerability management and providing a form of multi-factor authentication (MFA) to lower the risk of unauthorized access [25].

To address risks within the software supply chain and third-party dependencies, the DevSecOps Lead embeds security directly into the Continuous Integration and/or Continuous Delivery (CI/CD) pipeline. This role ensures that AI powered clinical tools are developed with automated security testing, improved auditing and change management on top of ensuring that software updates are "secure by design" before reaching hospital clients [26].

An AI Security Specialist is essential for addressing risks unique to AI systems. This includes ensuring the integrity of training data, monitoring model performance, continuously scanning for vulnerabilities and implementing safeguards against adversarial attacks. The growing importance of AI governance highlights the need for specialized roles focused on ethical and secure AI deployment [27].

Finally, a Compliance Officer ensures adherence to regulatory frameworks such as GDPR, DORA and NIS2. This role facilitates communication between technical teams and regulatory bodies, ensuring that compliance requirements are integrated into operational processes [28].

This structure supports lifecycle-based governance, enabling MediTrust to manage risks across all stages of AI system development and deployment.

2.4 Policies and Procedures

To govern the digital landscape of MediTrust Analytics Ltd, a robust set of policies must be established to translate the demanding goals of Zero Trust into an operational tight-knit umbrella of standards. These policies serve as the “rules of engagement for a hybrid workforce, ensuring that whether a clinical analyst is working from a hospital site or a remote home office, the security levels remain consistent and resilient against the specific threats facing this organization and the HealthTech sector as a whole.

2.4.1 AI Usage Policy

Given MediTrust’s focus on clinical decision support, a dedicated AI Usage Policy is essential to manage the ethical and security risks associated with machine learning. This policy must dictate the standards for data anonymization and integrity during the training phase to prevent the exposure of sensitive medical records. It should also establish strict requirements for AI generated clinical insights, ensuring that the technology assists rather than replaces clinical judgment while maintaining compliance with EU AI governance standards. To comply with the EU AI Act, this policy should mandate rigorous data provenance checks to ensure that training sets for diagnostic purposes are free from bias and external or internal malicious manipulation [2] [6].

2.4.2 Secure SaaS Configuration Policy

As MediTrust relies on multi-cloud and SaaS platforms to deliver services internationally, Secure SaaS Configuration policies are vital to prevent misconfigurations which lead to data leaks. Under a Zero Trust model, these procedures mandate the enforcement of least privilege access and multi-factor authentication across all software interfaces. According to the Cloud Security Alliance, securing the SaaS layer involves continuous monitoring of third-party permissions to ensure that diagnostic imaging and IoT data are only accessible to authorized personnel under verified conditions [30].

2.4.3 BYOD Policy for Hybrid Workforce

The bring your own device (BYOD) policy must address the inherent risks of a hybrid workforce accessing the MediTrust network from personal hardware. To align with NIS2 and GDPR, this policy should require the use of managed departure or mobile device management solutions which create a secure, encrypted container for corporate data on personal devices. This ensures that sensitive patient monitoring data remains isolated from personal applications, allowing the organization to verify device health before granting access to the internal cloud environment. Additionally, it is imperative that MediTrust personnel maintain rigorous device security protocols while operating in remote environments. Specifically, employees should be mandated to utilize Virtual Private Networks (VPNs) and abstain from accessing organizational data via unsecured public wireless networks. To minimize the risk of unauthorized physical access or data exfiltration, digital assets must remain supervised and encrypted at all times, thereby mitigating the threat of hardware theft and subsequent security breaches [31] [32].

2.5 Governance and Strategy: NIST CSF 2.0 Alignment

The proposed security program is aligned with the NIST Cybersecurity Framework (CSF) 2.0, which provides a comprehensive and structured approach to managing cybersecurity risks. The framework introduces six core functions: Govern, Identify, Protect, Detect, Respond, and Recover [29]:

CSF Function	MediTrust Business Goal	Security Strategy
--------------	-------------------------	-------------------

Govern	Regulatory Alignment	Establish DORA/NIS2 oversight committees
Identify	Asset Visibility	Maintain an inventory of AI models, data sets and third-party APIs
Protect	Data Confidentiality	Implement Phishing-Resistant MFA and Encryption at rest/transit
Detect	Operational Resilience	Deploy SIEM/EDR for real-time anomaly detection in AI workloads
Respond	Incident Containment	Automated playbooks for isolating compromised IoT or cloud nodes
Recover	Service Continuity	Air-gapped and immutable backups for AI training data

2.6 Compliance with DORA and NIS2

As MediTrust Analytics Ltd expands into the EU healthcare market, navigating the regulatory intersection of the Digital Operational Resilience Act (DORA) and the NIS2 Directive is a strategic necessity. While NIS2 classifies healthcare and HealthTech providers as “essential entities”, requiring high-level risk management and incident reporting, DORA introduces even more prescriptive requirements for the software supply chain. Specifically, because MediTrust provides critical AI powered services to hospitals which may fall under DORA’s scope for financial or operational stability, the company must move beyond basic security toward total operational resilience. This involves satisfying DORA’s third-party risk management, which mandates that vendors grant clients unrestricted audit rights and provide detailed mapping of sub-processors. Due to increasingly stringent propositions imposed upon HealthTech providers and possibly other areas of trade, such as the EU Cybersecurity Act (CSA 2.0), organizations are expected to master DORA compliance to survive the harsher sovereignty requirements that are coming to the broader European market [33]. Simultaneously, NIS2 Article 21 requires MediTrust to demonstrate supply chain sovereignty, ensuring that its multi-cloud infrastructure and AI training data are protected against foreign government access, often necessitating the use of sovereign cloud solutions within the European Health Data Space [34]. To meet these regulatory

demands, MediTrust must synchronize its internal protocols with the specific timelines and liability structures dictated by European law. A primary challenge lies in the discrepancy between incident reporting windows; while NIS2 mandates a 24-hour early warning for significant threats, DORA requires an initial notification for major ICT-related incidents within a much narrower 4-hour window, demanding a highly responsive internal monitoring system [35].

Feature	NIS2	DORA
Early warning	Within 24 hours	Within 4 hours (for major incidents)
Incident notification	Within 72 hours	Within the same business day (or the next day)
Final report	After 1 month	As soon as root-cause analysis is completed

Furthermore, the company must address supply chain sovereignty by ensuring its AI training datasets and diagnostic imaging archives comply with the European Health Data Space (EHDS), which permits Member States to enforce strict localization, requiring health data to be processed exclusively within EU borders [34]. Ultimately, these frameworks signify a major shift in executive liability, moving accountability from technical departments directly to the boardroom. Senior leadership is now required to personally approve cybersecurity strategies and can be held legally and personally responsible for any systemic failures in the organization’s risk management posture [36].

2.7 Risk Management

The software supply chain, AI-specific vulnerabilities and cloud misconfigurations represent the highest risk to MediTrust’s expansion:

Risk Area	Threat Scenario	Impact (Low/Med/High)	Mitigation
-----------	-----------------	-----------------------	------------

AI Supply Chain	Poisoned training data via third-party partners	High (Clinical Errors)	Data provenance tracking
Cloud Config	Exposure of diagnostic images due to misconfigurations	High (GDPR Fines)	Continuous CSPM monitoring and automated remediation
Dependency	Vulnerable open-source libraries in AI frameworks	Medium	Automated Software Bill of Materials analysis

2.8 Contingency Planning: Ransomware Scenario

Ransomware attacks represent one of the most severe threats to healthcare organizations, particularly those reliant on AI systems and cloud-based infrastructure. In the context of MediTrust Analytics Ltd, an attack which encrypts AI training datasets and source code repositories would disrupt both operational services and ongoing model development, potentially affecting clinical decision-making systems.

An effective disaster recovery (DR) strategy must begin with rapid detection and containment. Security monitoring tools such as SIEM platforms should identify anomalous behaviors, including unusual file encryption activity or unauthorized access patterns. Once detected, infected systems must be immediately isolated from the network to prevent lateral movement. Guidance from National Institute of Standards and Technology emphasizes that containment is critical in limiting ransomware spread and reducing overall impact [37].

The recovery phase relies heavily on the availability of secure and immutable backups. MediTrust should implement a “3-2-1 backup strategy,” maintaining multiple copies of critical data across different storage environments, including offline backups. This ensures that encrypted data can be restored without paying ransom demands [38].

3-2-1 Backup Rule



Regular disaster recovery testing is also essential. Simulated ransomware scenarios enable organizations to evaluate response times, identify weaknesses, and ensure that recovery procedures are effective under real-world conditions.

2.9 Protection Mechanisms

Preventing ransomware attacks and other cyber threats requires the implementation of modern, layered security controls aligned with Zero Trust principles. These controls must provide continuous visibility, automated response, and strong identity protection across MediTrust's distributed environment.

Cloud Security Posture Management (CSPM) tools are critical for identifying and remediating misconfigurations in multi-cloud environments. These tools continuously assess cloud resources against security benchmarks and automatically correct vulnerabilities. Research from Gartner highlights that misconfiguration remains a leading cause of cloud security incidents, reinforcing the importance of CSPM solutions [39].

Security Information and Event Management (SIEM) systems provide centralized logging and real-time threat detection by aggregating data from across the organization's infrastructure. When integrated with Security Orchestration, Automation, and Response (SOAR) platforms, SIEM systems enable automated responses to detected threats, significantly reducing response times [40].

Endpoint Detection and Response (EDR) solutions enhance endpoint security by continuously monitoring device activity and detecting suspicious behaviors such as ransomware encryption patterns [41].

Phishing-resistant multi-factor authentication (MFA), such as hardware security keys or biometric authentication, provides strong protection against credential-based attacks. This is particularly important in hybrid work environments, where remote access increases exposure to phishing attacks.

Together, these controls create a layered defence strategy that supports continuous verification, rapid detection, and automated response. This aligns with Zero Trust principles and ensures that MediTrust can effectively protect its AI systems, data, and infrastructure against evolving cyber threats.

2.10 Critical Evaluation

While the proposed program offers strong security and compliance benefits, it also presents challenges, including high implementation costs and complexity. However, the benefits of improved resilience and trust outweigh these challenges.

2.11 Conclusion

A Zero Trust-aligned security program is essential for MediTrust to operate securely within the EU regulatory environment. By integrating governance, risk management, and advanced security controls, the organization can effectively manage risks and ensure compliance.

3. References

- [1] “Roblox, Reddit and Discord users compelled to use biometric ID system backed by Palantir co-founder Peter Thiel,” *Open Rights Group*, 2026. <https://www.openrightsgroup.org/press-releases/roblox-reddit-and-discord-users-compelled-to-use-biometric-id-system-backed-by-palantir-co-founder-peter-thiel/>
- [2] National Institute of Standards and Technology, “AI Risk Management Framework,” *NIST*, Jul. 12, 2023. <https://www.nist.gov/itl/ai-risk-management-framework>
- [3] N. Gill, A. Mathur, and M. V. Conde, “A Brief Overview of AI Governance for Responsible Machine Learning Systems,” *arXiv (Cornell University)*, Jan. 2022, doi: <https://doi.org/10.48550/arxiv.2211.13130>
- [4] “AI Governance: Ensuring Responsible AI Development and Deployment,” *Witness.ai*, 2025. <https://witness.ai/blog/ai-governance/>
- [5] P. Okunytė, “AI coding tool wipes production database, fabricates 4,000 users, and lies to cover its tracks,” *Cybernews*, Jul. 21, 2025. <https://cybernews.com/ai-news/replit-ai-vive-code-rogue/>
- [6] EU Artificial Intelligence Act, “The AI Act Explorer | EU Artificial Intelligence Act,” *EU Artificial Intelligence Act*, Jun. 13, 2024. <https://artificialintelligenceact.eu/ai-act-explorer/>
- [7] C. Hale, “An ‘AI Exposure Gap’ could be the most worrying security issue your business isn’t aware of,” *TechRadar*, Nov. 17, 2025. <https://www.techradar.com/pro/security/an-ai-exposure-gap-could-be-the-most-worrying-security-issue-your-business-isnt-aware-of> (accessed Apr. 29, 2026).
- [8] aign.global, “🔗 What is an AI Governance Framework? The Ultimate Guide (2025 Edition),” *AIGN*, Jul. 21, 2025. <https://aign.global/ai-governance-insights/aign-global/%F0%9F%9F%A2%F0%9F%92%A1-what-is-an-ai-governance-framework-the-ultimate-guide-2025-edition/>
- [9] “AI RMF Core,” 2025. <https://airc.nist.gov/airmf-resources/airmf/5-sec-core/>
- [10] R. Al, “AI Lifecycle Management: From Design to Production in 8-12 Weeks,” *Regulativ.ai*, Dec. 17, 2025. <https://www.regulativ.ai/blog-articles/ai-lifecycle-management> (accessed Apr. 29, 2026).
- [11] R. Khan, D. Joyce, and M. Habiba, “AGENTS SAFE: A Unified Framework for Ethical Assurance and Governance in Agentic AI,” *arXiv.org*, 2025. <https://arxiv.org/abs/2512.03180>

- [12] H. Reese, Dr. S. Becker, "AI Lifecycle Management - PwC," *PwC*, 2024.
<https://www.pwc.de/en/risk-regulatory/responsible-ai/ai-lifecycle-management-as-part-of-ai-governance.html>
- [13] R. X, "Runtime AI Governance: Control, Escalation, And Failure Patterns," *AI Competence*, Feb. 24, 2026. <https://aicompetence.org/runtime-ai-governance/> (accessed Apr. 29, 2026).
- [14] Michelle M., "AI Governance Continuous Improvement: A Comprehensive Guide," *Project Templates*, Jan. 23, 2026. <https://www.projectmanagertemplate.com/post/ai-governance-continuous-improvement-a-comprehensive-guide> (accessed Apr. 29, 2026).
- [15] R. Murphy, "Operational Risks in AI Lifecycle Management," *Grcreport.com*, 2025.
<https://www.grcreport.com/post/operational-risks-in-ai-lifecycle-management> (accessed Apr. 29, 2026).
- [16] L. Lahti, E. Aittokallio, "Saidot x Scottish Government AI Risk Awareness Session," 2025.
Accessed: Apr. 29, 2026. [Online]. Available:
<https://www.gov.scot/binaries/content/documents/govscot/publications/foi-eir-release/2025/08-c/foi-202500473823/documents/foi-202500473823---information-released---annex/foi-202500473823---information-released---annex/govscot%3Adocument/FOI%2B202500473823%2B-%2BInformation%2BReleased%2B-%2BAnnex.pdf>
- [17] M. Steidl, M. Felderer, and R. Ramler, "The pipeline for the continuous development of artificial intelligence models—Current state of research and practice," *Journal of Systems and Software*, p. 111615, Jan. 2023, doi: <https://doi.org/10.1016/j.jss.2023.111615>
- [18] B. W. Wirtz, J. C. Weyerer, and I. Kehl, "Governance of artificial intelligence: A risk and guideline-based integrative framework," *Government Information Quarterly*, vol. 39, no. 4, p. 101685, Mar. 2022, doi: <https://doi.org/10.1016/j.giq.2022.101685>
- [19] Y. Jiang *et al.*, "Never Compromise to Vulnerabilities: A Comprehensive Survey on AI Governance," *arXiv.org*, 2025. <https://arxiv.org/abs/2508.08789>
- [20] A. Batool, D. Zowghi, and M. Bano, "AI governance: a systematic literature review," *AI and Ethics*, vol. 5, Jan. 2025, doi: <https://doi.org/10.1007/s43681-024-00653-w>
- [21] atomicdust, "Challenges & Frameworks for AI Governance | WashU Law," *WashU Law*, Feb. 20, 2026. <https://law.washu.edu/news/artificial-intelligence-governance-frameworks-challenges-and-best-practices/>
- [22] A. Taihagh, "Governance of Artificial Intelligence" *Oup.com*, 2021.
<https://academic.oup.com/policyandsociety/article/40/2/137/6509315>
- [23] P. Upmann, "Challenges and Risks in Implementing AI Governance," *AIGN*, Dec. 18, 2024.
<https://aign.global/ai-governance-insights/patrick-upmann/challenges-and-risks-in-implementing-ai-governance/> (accessed Apr. 29, 2026).
- [24] P. Bowen, J. Hash, and M. Wilson, "Information Security Handbook: A Guide for Managers," *csrc.nist.gov*, Mar. 07, 2007. <https://csrc.nist.gov/pubs/sp/800/100/upd1/final>
- [25] EC-Council, "How to Become a Cloud Security Architect: A Complete Career Guide," *Cybersecurity Exchange*, Feb. 07, 2025. <https://www.eccouncil.org/cybersecurity-exchange/cloud-security/how-to-become-a-cloud-security-architect-career-and-role-guide/>
- [26] R. Mogull, J. Arlen, F. Gilbert, A. Lane, D. Mortman, G. Peterson, M. Rothman "Security Guidance for Critical Areas of Focus in Cloud Computing," *CSA*, Jul. 26, 2017.
<https://cloudsecurityalliance.org/artifacts/security-guidance-v4>

- [27] "What Is an AI Security Specialist? | ACI Learning," *Acilearning.com*, 2025. <https://www.acilearning.com/blog/ai-security-specialist-the-role-protecting-today-s-most-powerful-technologies/> (accessed Apr. 29, 2026).
- [28] "Compliance officer | Prospects.ac.uk," *Prospects.ac.uk*, 2024. <https://www.prospects.ac.uk/job-profiles/compliance-officer/>
- [29] NIST, "The NIST cybersecurity framework (CSF) 2.0," *The NIST Cybersecurity Framework (CSF) 2.0*, vol. 2.0, no. 29, Feb. 2024, doi: <https://doi.org/10.6028/nist.cswp.29>
- [30] A. Akinbosoye, Y. S. Tao, J. R. Santos, M. Law, V. Murthy, Z. Somani, P. Lanois, M. Roza "SaaS Governance Best Practices for Cloud Customers | CSA," CSA, 2026. <https://cloudsecurityalliance.org/artifacts/saas-governance-best-practices-for-cloud-customers> (accessed Apr. 29, 2026).
- [31] "Security when working remotely" ENISA, 2026. Available: https://www.enisa.europa.eu/sites/default/files/all_files/Security%20when%20working%20remotely_Train%20the%20trainer%20guide.pdf
- [32] J. Jutila "How MDM helps European businesses meet NIS requirements," *Miradore*, May 17, 2024. <https://www.miradore.com/blog/mobile-device-management-helps-you-stay-safe-against-cybercrime-and-comply-with-nis2/> (accessed Apr. 29, 2026).
- [33] J. Viault, "DORA Compliance for Tech Firms: 2026 Revenue Guide," *Kobalt.io*, Feb. 23, 2026. <https://kobalt.io/the-dora-license-to-operate-how-tech-firms-can-unblock-european-revenue-in-2026/> (accessed Apr. 29, 2026).
- [34] J. A. Wondrasek, "DORA, NIS2, and the EU AI Act Are Making Sovereign Cloud Mandatory for Some Workloads - SoftwareSeni," *SoftwareSeni*, Feb. 27, 2026. <https://www.softwareseni.com/dora-nis2-and-the-eu-ai-act-are-making-sovereign-cloud-mandatory-for-some-workloads/> (accessed Apr. 29, 2026).
- [35] Martin "heyData GmbH," *Heydata.eu*, 2026. <https://heydata.eu/en/magazine/difference-nis2-dora-compliance-guide/> (accessed Apr. 29, 2026).
- [36] M. Sharron "NIS 2 Board Liability: What Every Director Must Now Prove and Sign," *ISMS.online*, 2025. <https://www.isms.online/nis-2/enforcement/board-liability/directors/> (accessed Apr. 29, 2026).
- [37] P. Cichonski, T. Millar, T. Grance, and K. Scarfone, "Computer security incident handling guide," *Computer Security Incident Handling Guide*, vol. 2, no. 2, Aug. 2012, doi: <https://doi.org/10.6028/nist.sp.800-61r2>
- [38] J. Thakkar "3-2-1 Backup Rule: The Rule of Thumb to Solve Your Data Loss Problems," *Hashed Out by The SSL Store™*, Feb. 05, 2021. <https://www.thesslstore.com/blog/3-2-1-backup-rule-the-rule-of-thumb-solution-to-your-data-loss-problems/>
- [39] L. Pingree "The Expanding Enterprise Investment in Cloud Security," *Gartner*, 2024. <https://www.gartner.com/en/newsroom/press-releases/2024-06-05-the-expanding-enterprise-investment-in-cloud-security>
- [40] "SOAR vs. SIEM: What Is the Difference?," *Palo Alto Networks*. <https://www.paloaltonetworks.com/cyberpedia/what-is-soar-vs-siem>
- [41] Microsoft, "What Is EDR? Endpoint Detection and Response | Microsoft Security," *www.microsoft.com*, 2025. <https://www.microsoft.com/en-us/security/business/security-101/what-is-edr-endpoint-detection-response>