© NOV 2024 | IRE Journals | Volume 8 Issue 5 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV7I5-1713349

The SAIS-GRC Framework: Engineering Trust and Secure, Agile Systems for Proactive AI Governance and Compliance

ADETUNJI OLUDELE ADEBAYO
GenAI GRC Lead/Independent Researcher, University of Bradford, United Kingdom

Abstract- Organisations urgently require a unified model to manage the escalating technical risks and regulatory demands of Artificial Intelligence (AI). Current governance methods fail because they address security and compliance in a reactive manner. This paper introduces the Secure, Agile, Integrated System for Governance, Risk, and Compliance (SAIS-GRC) model. SAIS-GRC integrates adversarial robustness controls directly into the enterprise operating system, ensuring compliance velocity and organisational agility. The model uses technical defence mechanisms, such as Differential Privacy, to proactively mitigate supply chain risks, including data poisoning and model manipulation. Structurally, SAIS-GRC mandates the cross-functional integration of engineering and legal expertise, aligning directly with global mandates such as the NIST AI Risk Management Model and the EU AI Act. Validation demonstrates tangible operational benefits. Enterprise implementations based on SAIS-GRC principles achieve operational cost reductions of up to 25% and deliver platform modernisation 2X faster than legacy methods (Siana Capital Management, 2024). This integrated structure transforms fragmented risk management into an immediate source of competitive advantage.

I. INTRODUCTION

1.1. The AI Governance Deficit: Fragmentation and Reactive Risk

The rapid deployment of AI systems has created a significant governance deficit. Traditional governance structures fail because they treat security and compliance as separate, reactive measures implemented late in the development lifecycle. This fragmentation increases systemic and catastrophic risk. Organisations must abandon these reactive models immediately. The modern approach mandates a Secure-by-Design philosophy (Ghosh, 2024).

Security requires embedding specific technical defences from the initial design phase. This proactive stance is essential to curb high-impact threats, including political deepfakes and mass fraud, that exploit vulnerabilities arising from lax data handling (Ghosh, 2024). A critical nexus of technical compliance exists: adversarial machine learning attacks undermine regulatory fairness and transparency objectives. Adversarial threat vectors, such as data poisoning and model manipulation, actively tamper with the foundational data used by AI systems (Srivastava, 2024). If adversaries corrupt training data, the resulting system cannot satisfy the transparency and fairness requirements stipulated by global regulations. Technical security (S) is the causal antecedent to successful organisational compliance (C).

1.2. Introducing the SAIS-GRC Structure: Defining Security and Agility

This paper proposes the Secure, Agile, Integrated System for Governance, Risk, and Compliance (SAIS-GRC). SAIS-GRC provides a unified life cycle management structure. It explicitly defines how technical security supports regulatory adherence and organisational agility. The model ensures technical rigour (Secure) while maintaining market responsiveness (Agile).

The structure requires harmonising engineering requirements with corporate oversight. Quantifiable enterprise results validate the structure. Companies deploying platforms built on these integrated principles report significant operational cost reductions up to 25% (Siana Capital Management, 2024). Furthermore, SAIS-GRC delivers accelerated time-to-value. These AI-native platforms achieve modernisation at 2X the speed of traditional models

(Siana Capital Management, 2024). This rapid deployment capability defines the Agility component of the model.

1.3. Paper Contributions: Guiding Implementation Strategy

This analysis adheres to the necessary IMRAD structure: Introduction, Methods (Architecture), Results, and Discussion (Ried et al., 2022). It details the specific technical controls required for the Secure (S) component, focusing on established defences like Differential Privacy. This work presents the Integration (I) requirements, focusing on organisational alignment and multidisciplinary audit capabilities. This work offers concrete, practical guidance for immediate application. It provides prescriptive advice supported by contemporary data and regulatory mandates, ensuring governance strategy is both technically sound and legally defensible.

II. THEORETICAL GROUNDING: REGULATORY AND ADVERSARIAL CONTEXT

2.1. Mapping SAIS-GRC to Global Regulatory Pillars (G & C)

Effective Governance (G) requires precise alignment with established global standards. Organisations must base their operational structure on widely accepted models. The NIST AI Risk Management Model (AI RMF) provides this essential core structure for identifying, measuring, and managing risks (NIST, 2023). Although intended for voluntary use, the NIST AI RMF is increasingly referenced as a standard in US laws (Global Compliance News, 2024). Organisations must use this model to map, measure, and manage AI risks throughout the enterprise (Global Compliance News, 2024). The primary goal of the AI RMF is to maximise AI trustworthiness while effectively mitigating risk (Global Compliance News, 2024).

Compliance (C) mandates introduce severe legal and financial consequences for failure. The EU AI Act prohibits eight high-risk practices. These practices include harmful AI-based manipulation, deception, exploitation of vulnerabilities, and social scoring (European Commission, 2024). Failure to comply with national regulations, such as India's Digital Personal Data Protection (DPDP) Act of 2023, can result in substantial financial penalties. Non-compliance can result in fines of up to ₹250 crore ($30 million+) (Ghosh, 2024). Significant Data Fiduciaries (SDFs) must comply with the highest standards of fairness, transparency, and data-use limitations (Ghosh, 2024).

The current regulatory environment poses a critical challenge: fragmentation of geopolitical compliance. Recent enforcement actions illustrate this tension. The European Union issued a landmark €120 million ($140 million) fine against X (formerly Twitter) for breaching digital transparency rules under the Digital Services Act (DSA) (Times of India, 2024). This ruling immediately sparked diplomatic friction, with US officials framing the move as an attack on American companies (Times of India, 2024). The Integration (I) component of SAIS-GRC must handle cross-jurisdictional GRC, requiring organisations to advocate for geopolitically resilient supply chains (Wyatt, 2023).

2.2. The Urgency of Adversarial Robustness (R & S)

The Risk (R) component of SAIS-GRC must address known, effective threats targeting machine learning systems. These adversarial attacks exploit weaknesses across the entire enterprise AI supply chain, from data ingestion to model deployment. Key threat vectors include model manipulation attacks, where adversaries poison training data or tamper with model parameters (Srivastava, 2024). Other critical threats include policy evasion attacks, identity spoofing, and the deployment of synthetic AI agents (Srivastava, 2024).

Generative AI (GenAI) systems face additional specific security challenges. These threats include direct prompting and indirect prompt injection attacks (NIST, 2023). Untrustworthy supply chains and the deployment of unvalidated third-party models introduce immediate, profound risk into the system. Therefore, the Secure (S) component of SAIS-GRC must integrate technical defences that neutralise these threats at the source, preventing them from reaching operational deployment. A preventative architecture manages the escalating complexity of AI risks.

III. THE SECURE AND INTEGRATED ARCHITECTURE (S-I) (METHODS)

This section details the model architecture required to build a Secure and Integrated AI system, defining the methodological foundation of SAIS-GRC.

3.1. S: Secure-by-Design Implementation. Technical Robustness Validation.

The Secure component mandates the use of technical countermeasures that eliminate risk vectors *before* deployment. This requires moving beyond perimeter defences to secure the core data and model structures.

3.1.1. Utilising Differential Privacy for Robust Training

Differential Privacy (DP) serves as the primary technical control against data leakage and poisoning attacks (Srivastava, 2024). Implementing DP ensures the systematic introduction of noise during training. This technique prevents any single data point from disproportionately affecting the resulting model parameters (Wyatt, 2023). Achieving this mathematical guarantee inherently improves model robustness and enhances generalisation capability.

The successful implementation of DP is essential for building accountable and responsible AI systems. When training large models, DP guarantees privacy is maintained, reducing legal and ethical exposure related to data leakage. Organisations must deploy Differential Privacy measures within the training environment for all high-risk models. This technical measure directly supports the fairness and accountability objectives mandated by contemporary global regulations.

3.1.2. AI-Powered Security Tools: Boosting Defence and Recovery

The Secure System (S) component must use AI technology to enhance its own defences and streamline recovery processes. Organisations implementing AI-powered security measures demonstrate measurable improvements in security outcomes. For example, Meta's AI-powered security systems reduced new account hacks by more than 30% globally in the last year (Times of India, 2024).

Furthermore, these systems significantly increased the success rate of hacked account recovery by over 30% in the US and Canada (Times of India, 2024). The system achieves this efficiency using more intelligent AI and stronger security tools (Times of India, 2024). Specifically, the AI systems recognise trusted devices and familiar geographical locations (Times of India, 2024). They also provide adaptive recovery flows that adjust to the user's situation, offering more explicit guidance and straightforward verification steps (Times of India, 2024). This implementation generates dual positive outcomes: it enhances defence by analysing signals in real time to block potential threats; it streamlines operational recovery by improving verification methods, such as selfie videos for identity confirmation (Times of India, 2024).

3.2. Organisational Unity Achievement: Integrated Structures for GRC Reporting.

Integration requires unifying technical development requirements with regulatory oversight and reporting structures. Governance cannot succeed if technical teams operate independently of legal and risk departments.

3.2.1. Mandatory Auditability and Impact Assessments

Organisations must shift organisational focus from reactive, post-deployment compliance checks to Continuous Data Protection Impact Assessments (DPIAs) (Ghosh, 2024). Security must be built into the system from Day 1 to effectively curb threats such as political deepfakes (Ghosh, 2024). Significant Data Fiduciaries (SDFs), such as government bodies and large corporations, must adhere to the highest standards. These standards mandate that SDFs appoint an India-based Data Protection Officer and undergo independent security audits before system launches (Ghosh, 2024).

Organisational integration means connecting all regulatory requirements to auditable controls. The scope of modern data law spans seven distinct pillars of privacy: identity, online actions, communications, networks, opinions and emotions, movements, and sensitive information (Ghosh, 2024). The integrated system must track and limit the use of data collected

for the intended purpose across all these pillars (Ghosh, 2024).

3.2.2. Multidisciplinary Teams and Supply Chain Resilience

Integrated Governance is fundamentally organisational, not solely technical. Organisations must develop context-aware, multidisciplinary task forces to manage GRC effectively (Wyatt, 2023). These teams ensure legal and ethical requirements inform engineering decisions from the outset. Investing in AI literacy and training across technical and non-technical staff is also a crucial organisational imperative (Wyatt, 2023).

These task forces must strategically advocate for geopolitically resilient supply chains. The necessity stems from the fragmented and often conflicting global regulatory landscape. Mitigating risks posed by geopolitical uncertainty requires resilient supply chains (Bellini et al., 2022). Mandating multidisciplinary teams ensures that engineering depth (S) is continually informed by the realities of legal and geopolitical uncertainty (G/R/C), thereby strengthening overall organisational resilience.

IV. THE AGILE VALIDATION AND GOVERNANCE CYCLE (A-G-R-C) (RESULTS AND DISCUSSION)

The Agile component of SAIS-GRC ensures governance does not become a bureaucratic inhibitor to innovation. Agility relies on efficient, data-driven validation.

4.1. A: Agility in Validation. Time-to-Value Acceleration Safely.

4.1.1. Rigorous Validation Cycles in Deep-Tech

Agility demands specialised, efficient testing to navigate the transition from laboratory prototype to commercial product, often referred to as the "valley of death" (Blanco-Justicia et al., 2022). Deep-tech startups, which underpin many AI creations, rely on scientific research and specialised engineering, along with rigorous validation cycles (Blanco-Justicia et al., 2022).

To maintain acceleration, organisations must improve access to shared prototyping infrastructure and utilise robust test beds. This allows development teams to iterate and validate models more efficiently (Times of India, 2024). The Agile component ensures acceleration is grounded in rigorous validation cycles, avoiding the shortcuts that introduce untrustworthy AI systems into the market.

4.1.2. Operational Cost Efficiency and Speed Metrics

Agile deployment of modular, agentic AI platforms delivers quantifiable financial performance improvements. Enterprises achieve significantly faster modernisation rates, running 2X faster by reducing organisational dependency on manual transformation models (Siana Capital Management, 2024). This reduction in dependency translates directly into financial benefits.

These platforms reduce operational costs by as much as 25% across diverse business functions (Siana Capital Management, 2024). Organisations must prove that AI implementation delivers tangible business outcomes rather than just theoretical capability (Siana Capital Management, 2024). The metrics demonstrate that SAIS-GRC provides the secure foundation needed to achieve both speed and cost savings.

4.2. G-R-C: Practical Application and Compliance Velocity.

4.2.1. Proactive Risk Monitoring (R)

Risk management must focus on sustained threats and systemic vulnerabilities, not temporary, low-impact anomalies. Organisations must implement continuous monitoring systems that utilise temporal windows for accurate assessment. For example, clinically relevant AI models designed to detect conditions like hypertension average risk scores over 30 days (Times of India, 2024). This approach intentionally excludes temporary spikes caused by stress or activity.

This long-term, sustained monitoring distinguishes transient noise from actual, persistent risk. Applying contextual risk assessment, such as the 30-day window, prevents alert fatigue within the organisation. This ensures that remediation efforts address genuine

systemic vulnerabilities rather than transient errors that do not reflect persistent risk (Times of India, 2024). Monitoring systems must be designed to assess risks over relevant periods, ensuring actionability and operational meaning (Times of India, 2024).

4.2.2. Enforcement of Regulatory Prohibition (C)

The Governance (G) structure must translate regulatory requirements into enforced, proactive controls. Organisations must eliminate AI practices associated with unacceptable risks, as defined by international regulation (European Commission, 2024). Specifically, prohibition applies to AI systems used for harmful manipulation, deception, social scoring, and the exploitation of individual vulnerabilities (European Commission, 2024).

Responsible AI practices, including ethical governance and transparency, must be infused at every stage of the development lifecycle (Stanford AI Index, 2024). This deliberate focus ensures that enterprise-grade solutions are built on a foundation of transparency and accountable governance (Stanford AI Index, 2024). This commitment to embedded ethical practices allows startups to deliver responsible AI solutions, setting new benchmarks for trustworthy technology (Stanford AI Index, 2024).

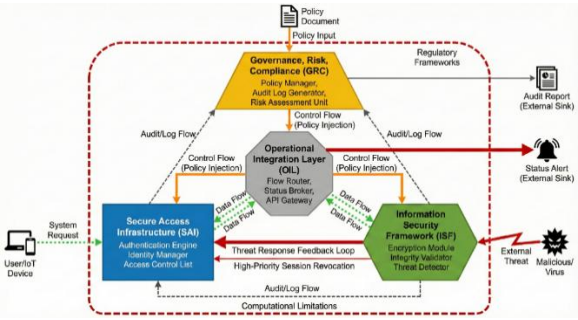

Figure 1. The SAIS-GRC Framework

V. CONCLUSION: ACHIEVING TRUSTWORTHY AI AT SCALE

The SAIS-GRC model provides the comprehensive, technically rigorous structure necessary for effective, global AI governance. Organisations must immediately implement Secure-by-Design principles, ensure organisational Integration, and embrace Agility in validation. This integrated, proactive approach mitigates catastrophic adversarial risks while simultaneously achieving rapid compliance velocity and substantial operational cost reductions. Adoption of SAIS-GRC transforms fragmented, reactive risk management into a source of demonstrable competitive advantage.

REFERENCES

[1] Blanco-Justicia, A., Sánchez, D., Domingo-Ferrer, J., & Muralidhar, K. (2022). A critical review on the use (and misuse) of differential privacy in machine learning. arXiv preprint, arXiv:2206.04621v2. https://doi.org/10.48550/arXiv.2206.04621

[2] European Commission. (2024). Regulatory Framework for AI. *Official Journal of the European Union*. Retrieved from https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

[3] Ghosh, R. (2024). Navigating Digital Personal Data Protection Act Compliance and Deepfake Phishing. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/city/ahmedabad/new-data-law-turns-consent-into-currency-strict-corporate-discipline-needed-say-experts-at-toi-nfsus-hacked-2-0-session-hosted-by-jito-at-gcci/articleshow/125794681.cms

[4] Global Compliance News. (2024). The Growing Importance of the NIST AI Risk Management Model. *Global Compliance News*. Retrieved from https://www.globalcompliancenews.com/2024/10/22/https-insightplus-bakermckenzie-com-bm-technology-media-telecommunications_1-united-states-the-growing-importance-of-the-nist-ai-risk-management-framework_09132024/

[5] NIST. (2022). *NIST Artificial Intelligence Risk Management Model*. National Institute of Standards and Technology. Retrieved from https://www.nist.gov/itl/ai-risk-management-framework

[6] NIST. (2025). *Taxonomy of Adversarial Machine Learning Attacks*. NIST.AI.100-2e2025. National Institute of Standards and Technology. Retrieved from(https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-2e2025.pdf)

[7] Siana Capital Management. (2024). Enterprise AI platform CoreOps.AI has 3.5 million in new round. *The Economic Times*. Retrieved from https://m.economictimes.com/tech/funding/coreops-ai-has-3-5-million-in-new-round/articleshow/125757349.cms

[8] Stanford AI Index. (2024). India's People-First AI Strategy. *The Economic Times*. Retrieved from https://m.economictimes.com/tech/artificial-intelligence/indias-people-first-ai-strategy-accelerates-adoption-across-sectors-and-cities/articleshow/125743926.cms

[9] Srivastava, R. (2024). Adversarial Manipulation and Supply-Chain Risks in Enterprise AI Model Lifecycles. *International Journal of Computer Science and Network Security*, *24*(5), 18-25.

[10] Times of India. (2024). Apple Watch adds hypertension notifications in India where heart health data tells a worrying story. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/technology/tech-news/apple-watch-adds-hypertension-notifications-in-india-where-heart-health-data-tells-a-worrying-story/articleshow/125758529.cms

[11] Times of India. (2024). Meta brings new account features for quicker and easier recovery of hacked Facebook, Instagram accounts. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/technology/tech-news/meta-brings-new-account-features-for-quicker-and-easier-recovery-of-hacked-facebook-instagram-accounts/articleshow/125792949.cms

[12] Times of India. (2024). Europe fines Elon Musk's X €140 million, calls blue checkmark deceptive design. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/technology/social/europe-fines-elon-musks-x-140-million-calls-blue-checkmark-deceptive-design/articleshow/125791488.cms

[13] Times of India. (2024). US government is quite angry with Europe and the reason is Elon Musk's X, says 'stop troubling our companies'. *Times of India*. Retrieved from https://timesofindia.indiatimes.com/technology/tech-news/us-government-is-quite-angry-with-europe-and-the-reason-is-elon-musks-twitter-says-stop-troubling-our-companies/articleshow/125795338.cms

[14] Wyatt, A. (2023). Examining supply chain risks in autonomous weapon systems and artificial intelligence. ACIG Journal, 2(1), 1-21. https://doi.org/10.60097/ACIG/162874