Artificial Intelligence.
Real Security.

Own AI Securely with SANS.

By Rob T. Lee,
SANS Institute Chief of Research and Chief AI Officer

SANS

Executive Summary

Artificial intelligence (AI) has crossed the threshold from experimentation to widespread adoption on a personal, organizational, and geopolitical level. AI marks the biggest transformative change since the invention of the internet. It is changing the way people work, the way organizations function, and the way nations are attempting to position themselves for the next several decades.

From an enterprise adoption standpoint, the last several years were marked by pilot projects and temporary integrations. However, we've now entered a new phase of permanent utilization at scale. OpenAI reports that more than 92% of Fortune 500 organizations are currently using its Generative AI (GenAI) offerings. And based on recent financial disclosures, the Big Four technology companies (Google, Amazon, Meta, and Microsoft) are projected to spend a combined $400 billion on AI capabilities in 2026.

This comes as regulatory efforts around AI have intensified. In the United States, the current administration's America's AI Action Plan calls for accelerated AI adoption across government and critical infrastructure, while introducing new requirements around vendor transparency, bias mitigation, and national security alignment. NIST's AI Risk Management Framework is designed to improve the ability to incorporate trustworthiness considerations into the design, development, use, and evaluation of AI products, services, and systems. The European Union's AI Act raised the stakes for any organization operating in the EU or serving EU markets to maintain transparent and traceable AI systems. And within the APAC region, government bodies across Singapore, South Korea, Australia, and Japan are advancing AI governance frameworks that emphasize technical robustness, human oversight, and sector-specific risk controls.

While regulation around AI security is catching up, commitments to secure implementation lag. Most organizations understand that AI is business-critical to staying competitive, but lack a clear vision for its secure implementation and safety protocols. A Gartner survey released in March 2025 revealed that only 27% of CISOs have mapped their organization's AI controls to established security frameworks. In addition, a recent AWS survey reported that 45% of IT leaders now prioritize Generative AI (GenAI) over cybersecurity in their 2025 budgets, while security tools received only 30% of the focus. This has resulted in AI advancing faster than most CISOs can secure it.

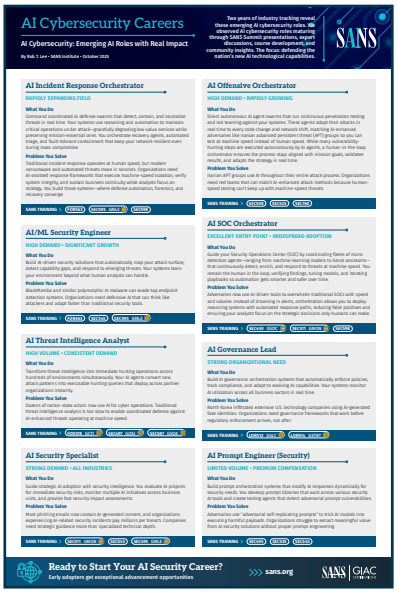
At SANS, we believe that to harness AI's full potential, we all must take decisive action to Own AI Securely. The question isn't whether AI will change the cyber threat landscape – it's whether your team will be prepared when it does. AI security must be treated as a foundational priority and not an afterthought. It isn't a future concern. It's table stakes today.

The SANS Secure AI Blueprint:

A Model for Secure AI Adoption

For more than three decades, SANS has served as a global security leader for the public and private sectors. Now, as AI adoption accelerates, we aren't just responding to the AI security challenge – we're actively shaping the standards, tools, and talent that define it. Our clear AI security leadership encompassing a comprehensive training portfolio, including courses, summits, research, and policy engagement, ensures organizations are equipped to both realize AI's promise and navigate its inherent security complexities by reducing risk.

SANS delivers strategic guidance modeled on real-world adversarial techniques, and the only unified learning path that equips business leaders, defenders, developers, and operators to securely design, deploy, and defend AI systems. This makes us uniquely positioned to serve as a partner-of-choice to prepare organizations for what comes next in a new era of security – including an array of new AI-focused job positions that will emerge in the coming years.



Download the SANS AI Security Careers poster here >

This paper introduces the ***SANS Secure AI Blueprint***, a holistic model for driving secure AI adoption structured around a three-track framework: Protect AI, Utilize AI, Govern AI. SANS distils hundreds of defensive controls into these three operational tracks that map one-to-one with primary enterprise AI stakeholders:

Protect AI:

Security teams, engineers, and AI/ML developers must work together to deploy advanced defense measures that anticipate and mitigate adversarial threats, safeguard model integrity and data privacy, and ensure continuous operational resilience in an AI-augmented threat environment.

Utilize AI:

Security teams must effectively integrate AI-enabled workflows across cyber defense operations. This calls for CISOs to equip their SOC managers, analysts, incident responders, and threat hunters with the necessary skills and specializations to accelerate defensive AI capabilities without losing operational control.

Govern AI:

Executive leadership teams must elevate governance to a board-level priority, establishing the policies, oversight, and accountability required in a rapidly evolving regulatory landscape. This requires Boards, C-suites, and business-line leaders to develop AI fluency and ensure that every AI system is documented, auditable, and aligned with global standards.

Our three-track model provides a clear strategy for Owning AI Securely to make AI innovation a cornerstone of organizational advantage. It also provides a roadmap for organizations to align with America's AI Action Plan (AAIAP) in the years to come.

America's AI Action Plan:

Why It Matters and How to Act Now

The AAIAP functions as the national guide to drive AI leadership while ensuring safety measures. It focuses on three main objectives including the fastening of innovation, the creation of AI infrastructure and power, and leadership in international AI diplomacy and security, with a core emphasis on workforce training, dependable systems, and misuse prevention. It outlines immediate actions including DARPA programs on robustness and interpretability, NIST work on deepfake forensics, grid and datacenter build-out, and export-control enforcement, which shows what Washington will fund, require, and align with allies.

The AAIAP establishes a comprehensive national vision that requires fast-paced innovation together with secure and expandable AI infrastructure, and international governance alignment. The plan requires immediate action from both public and private sectors.

Putting Policy into Practice

While the AAIAP outlines our nation's future direction, SANS acts as the implementation force that makes it a reality – bringing together public and private leaders to create field-tested controls, publish research and surveys, and expand workforce skills. SANS helps organizations fulfil AAIAP requirements through summits, hackathons, open standards with OWASP, and existing training programs in the market to reduce AI risk across development, deployment, and oversight. Through these efforts, we support national strategy implementation while providing organizations with the tools to lead confidently in the fast-changing AI environment, which minimizes risk while maximizing AI innovation and security potential.

SANS transforms the AAIAP's pillars into practical results by implementing field-tested controls for AI protection and providing safe usage training and developing enforceable governance frameworks.

Operationalizing AAIAP via SANS's Secure AI Blueprint

Protect AI:
The SANS Critical AI Security Guidelines outline six domains for Access, Data, Deployment, Inference, Monitoring, and Governance, with practical defensive measures that organizations can execute at present while matching the robustness standards, infrastructure requirements, and exportable standards recommended in AAIAP. It demonstrates the same control structure which runs throughout this paper.

Utilize AI:
SANS workforce development programs, along with summits and webcasts, work toward building the skilled workforce AAIAP emphasizes. AI content webcasts and briefs summarize the 100+ hours of freely accessible AI content.

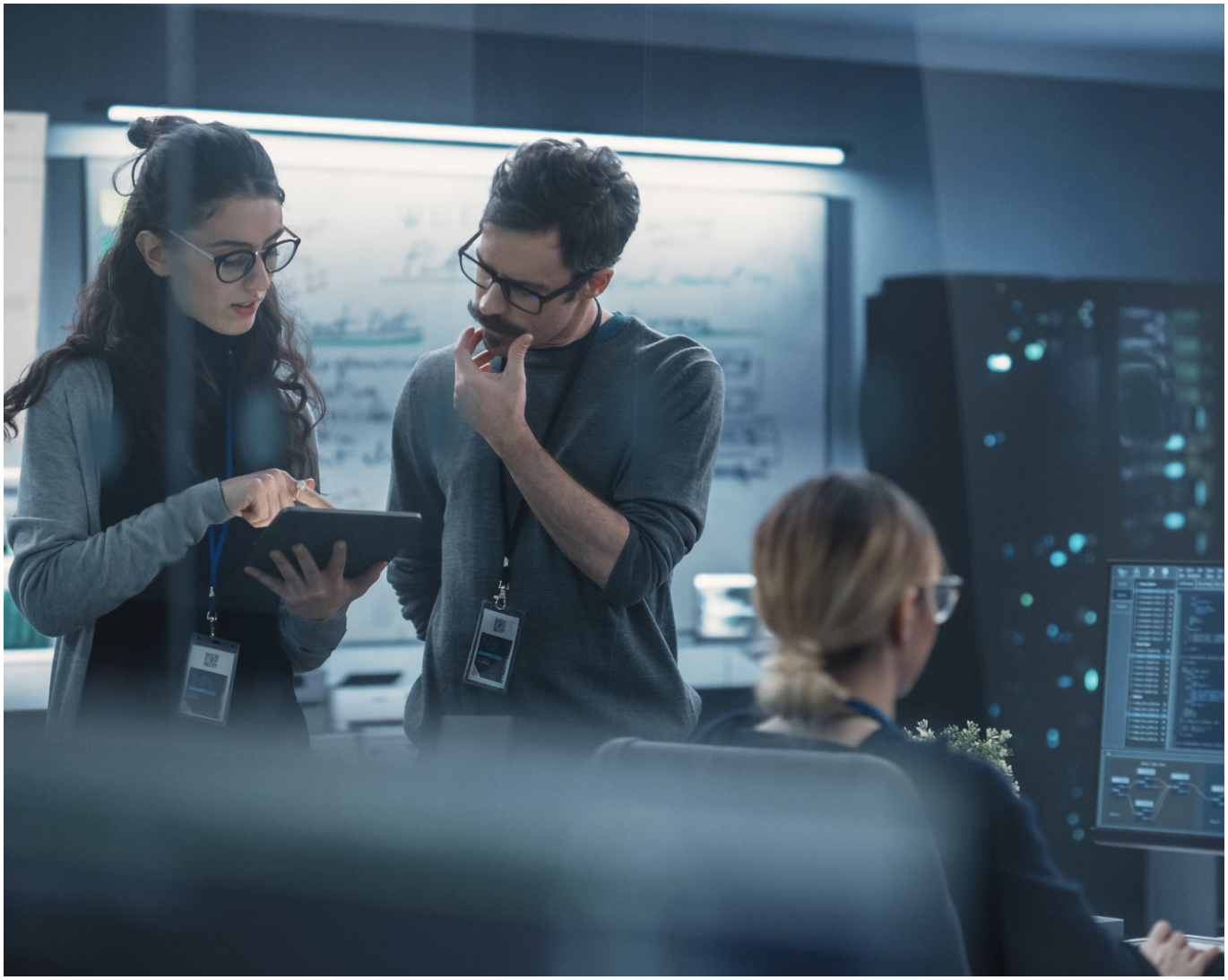
Govern AI:
SANS policy templates, such as the AI Acceptable Use Standard, and our formal partnership with OWASP AI Exchange to co-standardize AI controls, operationalize accountability and map cleanly to NIST AI RMF and the EU AI Act for global readiness .

SANS has operated based on these priorities for multiple years now. We held an AI-centered cybersecurity summit on May 31, 2023 after ChatGPT's launch, then organized AI Summits for 2024 and 2025, and added another solutions track leading to a total of four Summits that brought together public and private decision-makers on AI security. The program included a Global AI Security Hackathon which demonstrated tools at the 2025 Summit to establish direct links between research, workforce, open tooling, and real-world risk mitigation.

In addition, the SANS YouTube channel provides leaders with more than 40 AI Summit session videos which are available to the public along with additional briefings such as Demystifying the NIST AI Risk Management Framework and AIS247 executive primers. The SANS Technology Institute (SANS.edu) has released several AI research papers that focus on LLM guardrails, AI-driven security threats in IT guidelines, GenAI for memory analysis, and the security of coding assistants, developing essential evidence for practitioners.

The Way Forward

Building on our AAIAP analysis, each of the following sections in this paper offer a detailed assessment into the key components of AI security, drawing from SANS's unique vantage point melding frontline operational expertise, regulatory insight, and the most rigorous training available in the industry.



Section 1

Protect AI: A Shared Responsibility for Security, Engineering, and AI/ML Teams

Protecting AI involves securing the models, applications, and data pipelines that make AI systems work. In the same way networks, endpoints, and cloud infrastructure are treated as mission-critical assets, AI systems now require the same level of attention. The objective is simple in statement, but complex in execution: reduce risk to organizations who are utilizing AI systems by protecting those systems from attacks.

The execution complexity stems from rapid upticks in large language model (LLM) adoption across SaaS environments. According to reporting from the Cisco Talos threat intelligence team, LLM-enabled applications on corporate applications grew 4x year-over-year in 2025. Without proper safeguards, LLMs can be vulnerable to model poisoning, prompt injection attacks and data leakage. And when adoption is delayed due to a lack of safeguards, it introduces the risk of shadow AI.

Model Poisoning:

Attackers may compromise the integrity of AI systems by introducing malicious data into training pipelines, distorting outputs or embedding hidden backdoors. Without continuous model validation and strong data provenance, outcomes become unpredictable and risk grows system-wide.

Prompt Injection Attacks:

Large language models, while game-changing for efficiency, are acutely vulnerable to crafted prompts designed to manipulate, subvert, or exfiltrate sensitive data. Building durable defenses requires layered input validation, rigorous prompt management, and intelligent monitoring, all core elements championed in SANS training and operational frameworks.

Shadow AI:

When organizations delay internal AI adoption due to uncertainty or perceived risk, they often create an even greater one: employees using unsanctioned AI tools on their own. These self-directed efforts — unvetted, unmonitored, and often invisible — can expose sensitive data to third-party platforms with unknown behaviors.

Data Leakage:

The very data that powers AI, often sensitive, regulated, and proprietary, makes it a high-impact target. Errors in access control, faulty anonymization, or misconfiguration can result in severe exposure events. Comprehensive encryption, strict permission models, and auditable, end-to-end logging transforms data protection and compliance from a weakness to a point of strength.

For example, a single poisoned training file or crafted prompt can exfiltrate regulated data in minutes, demonstrated by the February 2025 "Promptware" exploit against Google Gemini. In this incident, a malicious calendar invite exploited Google's Gemini AI assistant via a hidden prompt embedded in the event title, triggering unauthorized device commands like opening smart home appliances, without any user action. The attack underscored why strict access controls, input validation, and continuous monitoring are essential to secure AI models.

The SANS Critical AI Security Guidelines:
A Baseline for Protecting AI Systems

Protecting AI systems is not about a single control or a one-time project. It requires ongoing discipline grounded in the SANS Critical AI Security Guidelines. These guidelines establish six essential categories — Access Controls, Data Controls, Deployment Strategies, Inference Security, Monitoring, and Model Security — that together form a layered defense against the full spectrum of AI-specific threats.

SANS AI Security Guidelines – Control Categories

Control Category	Description
Access Controls	Secure models, APIs, and datasets via least privilege, strong authentication, and encrypted access. Monitor for unauthorized use.
Data Controls	Protect training and augmentation data through sanitization, masking, lineage validation, and restricted access.
Deployment Strategies	Harden AI pipelines and infrastructure with sandboxing, model validation, CI/CD gating, and red teaming.
Inference Security	Prevent prompt injection and misuse with guardrails, I/O validation, limited functionality, and escalation paths.
Monitoring	Continuously log and monitor inputs, outputs, and behavior to detect drift, anomalies, and degradation.
Model Security	Apply version control, integrity checks, and anti-tamper protections to defend models from theft or manipulation

When these controls work together, they create a resilient operational environment where protections reinforce each other, and AI systems remain accurate, reliable, and defensible even under sustained adversarial pressure. Protecting AI under this framework demands close coordination between engineering, security, and risk management teams.

Securing Agentic AI Systems

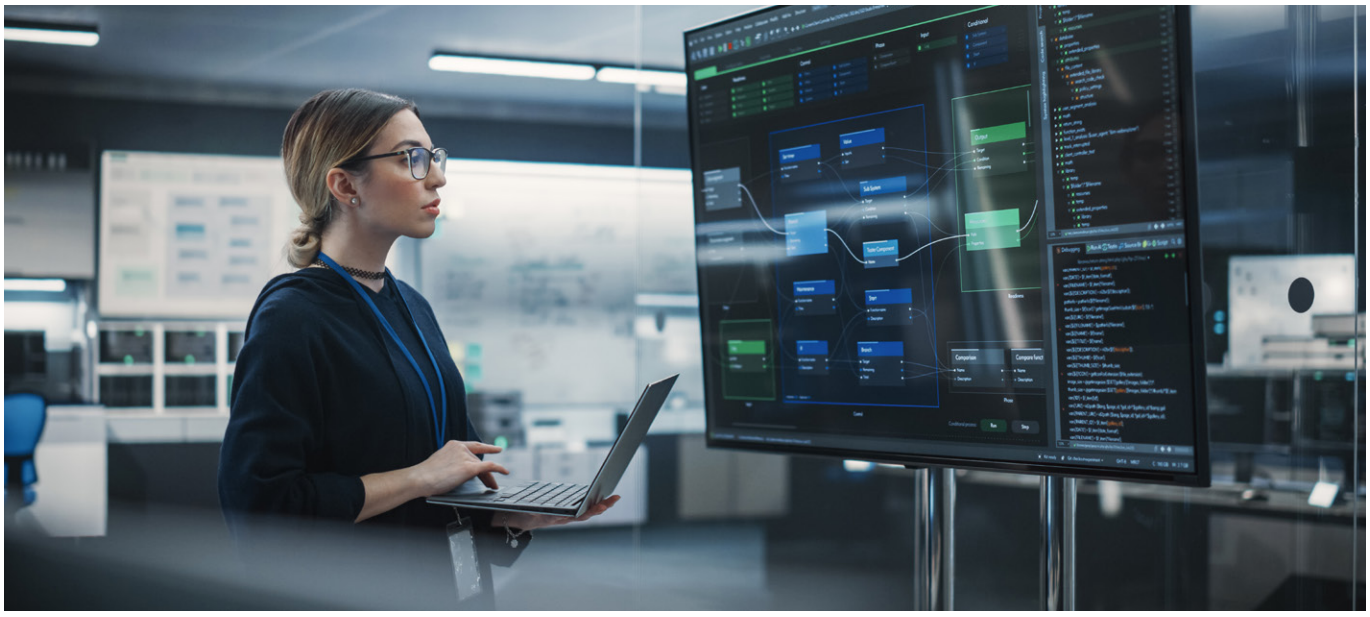
The emergence of agent-based architectures, such as multi-agent control planes (MCPs) and agent-to-agent (A2A) coordination frameworks, has introduced powerful new automation capabilities. These systems, often driven by LLMs acting autonomously, can now chain tasks, invoke tools, pull external data, and execute actions across systems without direct human intervention. With that functionality comes a new class of security and control challenges.

Notable among these is the concept of scope creep: agents that are insufficiently constrained may access unintended tools, retrieve sensitive information, or trigger behaviors well outside their intended domain. In complex environments, even minor misconfigurations can lead to major downstream effects. To securely deploy agentic AI, organizations should apply the SANS Critical AI Security Guidelines across the six control categories, treating agents as semi-autonomous actors subject to the same protections as users, systems, and APIs.

Securing Agentic AI – Control Mapping

Control Category	Agentic AI Security Focus
Access Controls	Scope agent permissions and enforce least privilege. Limit access to essential tools, datasets, and APIs with authentication and logging.
Data Controls	Restrict access to sensitive data unless required. Encrypt, sanitize, and log all data flows to prevent leakage or misuse.
Deployment Strategies	Isolate agent execution in sandboxes or containers. Red-team and validate imported code before deployment.
Inference Security	Use allowlists, runtime validation, and guardrails. Require human approval for high-risk or irreversible actions.
Monitoring	Continuously track agent behavior, API calls, and deviations. Escalate ambiguous outcomes for human review.
Model Security	Use signed, versioned models and validate them against adversarial inputs. Prevent unauthorized swaps or retraining.

While agent-based systems are not inherently insecure, their autonomy and adaptability expand the attack surface. Applying layered controls and internal trust boundaries ensures these systems remain predictable and defensible as they scale and interact with complex environments.



Why SANS:
The Global Authority on Protecting AI Systems

At SANS, we aim to add clarity and control to securing AI systems, one of the most urgent challenges across global security today. While organizations are assessing how to govern and integrate AI, we've focused on how to protect it and safeguard data by developing practical controls, engineering patterns, and training programs that enable defenders to secure models, pipelines, and agents in production.

We offer an extensive curriculum that spans offensive testing, secure deployment, and AI-specific defense engineering. Our programs teach learners how to validate LLMs, harden retrieval-augmented generation (RAG) architectures, lock down vector databases, and deploy agentic AI systems that are resilient under attack. Every course is backed by threat intelligence, live lab environments, and seasoned faculty members who operate in the field.

Leveraging LLMS: Building & Securing RAG, Contextual RAG, and Agentic RAG: Secure-by-design engineering for GenAI applications, RAG pipelines, and orchestration frameworks.

GenAI and LLM Application Security: Defense of prompt inputs, agent behavior, model endpoints, and runtime environments.

Offensive AI: Attack Tools and Techniques: Red teaming against LLMs, data pipelines, and AI-integrated apps.

Applied Data Science and AI/ML for Cybersecurity Professionals: Building and validating models with integrity, observability, and drift detection.

Our commitment to this facet of security extends even farther. We are also a new contributor to the OWASP AI Exchange, a collaborative effort between industry, academia, and the public sector to create practical, implementation-ready security guidance for AI systems. SANS's work with OWASP has already informed draft annexes of the EU AI Act and is being used by organizations aligning to ISO/IEC 27090 and NIST's AI RMF.

Protecting AI Systems with Confidence

AI systems are not secure by default, and they will be targeted by adversaries. Protecting them requires more than patches or policies. It calls for security, engineering, and AI/ML teams who understand which tools trigger accidental data leaks, how adversaries can exploit them, and how to engineer protection into every stage of the AI system development lifecycle. AI systems cannot be rushed into market without core security models, secure-by-design implementation, and proper auditing to prevent leakage. That is what SANS trains defenders to do. Partnering with us ensures your teams have the right skills to protect AI where it matters most.

Learn more about SANS's AI security offerings here >

Section 2

Utilize AI: A Competitive Requirement for Cyber Defense

Malicious AI is accelerating the cyber threat landscape at a pace that forces a complete rewrite of defensive assumptions. Adversaries are no longer limited by human speed or manual tradecraft. They are using advanced language models, generative AI (GenAI) tools, and autonomous agents to execute intrusion activities in seconds or minutes that would take a skilled threat actor hours to complete.

AI can be leveraged to create tailored phishing campaigns, manipulate audio and video for social engineering, chain exploits without human input, and adjust mid-operation to evade defenses. The result is an environment where privilege escalation, lateral movement, and data exfiltration can occur before human defenders can react, making AI-driven threats not just faster but exponentially more dangerous.

Research and operational testing from MIT's autonomous exploit chaining to Horizon3's 60-second privilege escalation show attack workflows running, conservatively, 47x faster than historical averages, with the potential for far greater speeds as these capabilities mature.

How We Arrived at "47× Faster"

- MIT autonomous agent research demonstrated privilege escalation and exploit chaining in seconds to minutes compared to hours for human operators.
- Horizon3 NodeZero testing achieved full privilege escalation in about 60 seconds.
- CrowdStrike 2023 threat hunting data reported average time from compromise to lateral movement at 79 minutes, with fastest observed breakout times around 7 minutes.
- Using 60–79 minutes as the human benchmark, AI-driven workflows complete the same steps about 120–158 times faster.
- To keep the figure conservative and credible, we halved these values and set the public number at 47× — a speedup that is already achievable with publicly available tools like Metasploit and likely much greater with APT-level capabilities.

Legacy defensive techniques struggle to keep pace with these adaptive attacks made possible by AI. The 2025 Verizon Data Breach Investigations Report found that over 75% of social engineering breaches involved pretexting or phishing content that showed signs of AI-assisted generation. In addition, DBIR research showed that AI-driven reconnaissance reduced the average time from initial compromise to privilege abuse from weeks to less than 48 hours in observed cases.

For CISOs and security teams, the call to action is clear. The only effective way to counter a new generation of AI-enabled adversaries is by fighting fire with fire. Defenders must instrument AI for detection and response as aggressively as attackers weaponize it; otherwise, mean-time-to-detect will widen. Organizations that fail to modernize Security Operations Centers (SOCs) with advanced AI capabilities will put their security posture at risk.

Despite the acceleration of AI-powered threats and collective alignment across security circles regarding AI's criticality to cyber defense, most SOCs are not equipped to match adversary velocity.

The SANS 2025 SOC Survey found that 42% of SOCs are using ML tools "out of the box" with no customization or integration into formal workflows. Moreover, GenAI tools scored just 2 out of 4 in satisfaction, making them among the lowest-performing technologies in the SOC. Several crucial gaps persist:

Lack of Visibility into AI Behaviors: Most SOCs do not monitor prompt inputs, output deviations, or agent actions—leaving blind spots vulnerable to exploitation.

No AI-Specific Playbooks or Procedures: Existing incident response plans fail to address threats unique to AI, such as adversarial prompt injection, model poisoning, or inference drift.

Inadequate Detection of AI-Driven Attacks: Traditional tools cannot reliably identify AI-generated phishing, synthetic data exfiltration, or automated lateral movement, leaving defenses open to bypass.

Telemetry and Monitoring Shortfalls: The dynamic nature of AI systems demands telemetry and monitoring frameworks that can detect drift, unauthorized modifications, and emerging anomalies in real time. Most current architectures are not designed with these requirements in mind.

These operational deficiencies impede effective defense, heightening the likelihood of undetected breaches and undermining the organization's ability to respond rapidly and decisively when attacks arise.

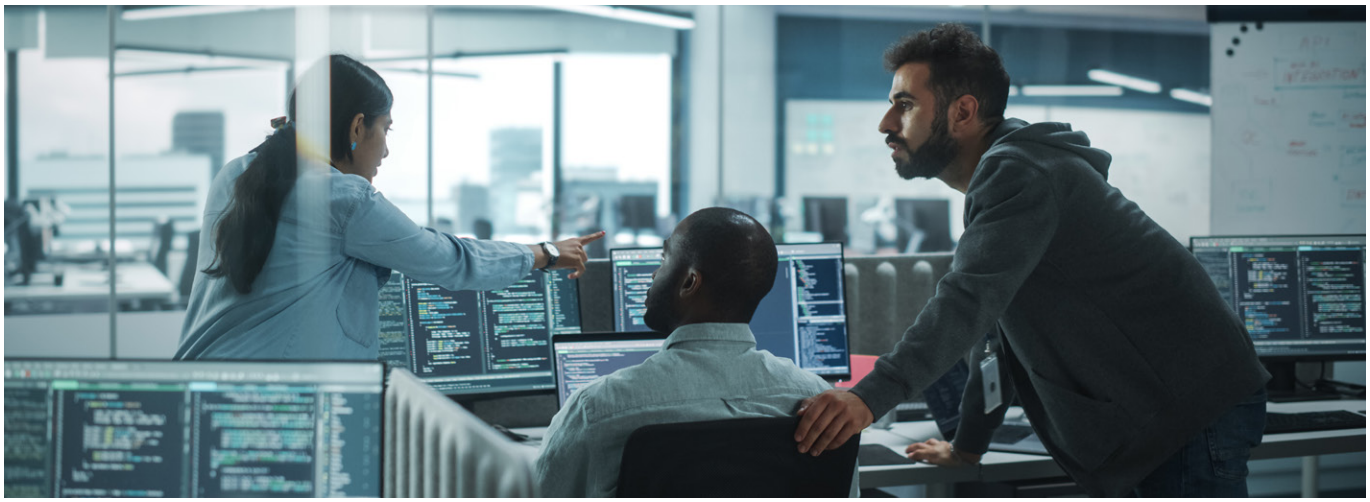
Why SANS:
Cultivating AI Readiness Inside the SOC

Bridging AI gaps inside the SOC requires a sharp shift in how security leaders design workflows and cultivate readiness across their teams. Every day, SOCs are expected to solve high-consequential problems with half the data and twice the pressure. Analysts are overwhelmed by the systems and processes in place that are meant to help them respond. Tooling is fragmented. Workflows are heavy. Context lives in five places, and alerts never slow down. What started as a fast-paced, high-impact role has, for many analysts, become a repetitive loop of alert triage and data wrangling that offers little room for an effective strategy.

Most SOC teams also run lean. In 2024, our annual SANS SOC Survey found that a majority of SOCs only consist of just 2–10 full-time analysts, a number unchanged since the survey began tracking in 2017. Meanwhile, the scope of coverage has exploded, ranging from on-prem infrastructure to cloud environments, remote endpoints, SaaS platforms, and beyond. Compounded at scale, this has led to SOC environments that hinder an organization's ability to defend itself.

To address this challenge, we need to change how SOC work is designed and executed. Enter AI implementation here. It offers a practical path forward by optimizing parts of the job that slow analysts down: the disorganization, repetitive steps, and cognitive overhead. It can streamline inefficient workflows and support skill development to facilitating more impactful team-wide oversight, opening wider avenues for making SOC work more successful against AI-enabled adversaries.

For example, AI-powered automation can act as a powerful contextual aggregator and investigative assistant. When paired with capabilities like those enabled by Model Context Protocol (MCP), language models integrate telemetry, threat intelligence, asset metadata, and user history into a single view, tailoring it to each unique situation the analyst faces. This gives analysts enriched, case-specific summaries instead of raw events. Clarity replaces guesswork. Response decisions happen faster and with greater confidence.



Skills That Position SOC Teams for Success

SANS training enables SOC teams to achieve operational AI readiness by combining role-specific instruction, live lab environments, and the latest adversary tradecraft mapped to the SANS Critical AI Security Guidelines. Training spans building and validating AI-driven detections, hardening AI applications, and integrating automation to close detection gaps.

- Building AI-driven detections. In SEC595, teams learn to design and validate machine learning models for malware and log anomaly detection, uncover covert channels, and deploy drift-aware detection pipelines that reveal threats signature-based tools miss.
- Strengthening SOC monitoring and threat-informed defense. SEC450 develops analyst efficiency through telemetry collection, SOAR scripting, and AI-agent triage. SEC511 teaches engineering detection pipelines that defend GenAI and LLM applications, hunt for prompt-injection and model-poisoning, and apply ML-enabled continuous monitoring.
- Detecting behavior on the network when signatures fail. SEC503 trains teams to use ML-based anomaly detection on NetFlow and packet data, build behavioral baselines, and detect covert channels or insider threats, even in encrypted traffic.
- Automating detection, response, and validation. SEC598 focuses on AI-powered purple-team automation, adversary emulation, and detection-effectiveness measurement to prove SOC readiness at production speed.

Every course is taught by practitioners with current field experience, ensuring SOC teams can immediately apply skills to secure AI-affected environments, detect and respond to AI-enabled attacks, and maintain continuous assurance over their defenses.

Supporting a Stronger SOC

Modernizing security operations to meet the demands of the AI era is not optional. With SANS, organizations build operational frameworks that are not just reactive but resilient, predictive, and equipped to defend against both present and future threats. The right training, expertise, and strategy turns AI from a source of risk into an advantage, enabling security teams to safeguard the enterprise with confidence and agility.

Learn more about SANS's SOC training offerings here >

Section 3

Govern AI: A Strategic Imperative for Enterprise Leaders

It's safe to say that the corporate enterprise AI hype cycle is over. For boardrooms and C-suite leaders, AI adoption is no longer viewed as a choice — it's a strategic imperative for outpacing industry competitors in digitally-driven market environments. A Dataiku and Harris Poll survey showed that 74% of CEOs globally said "they could lose their job within two years if they don't deliver measurable AI-driven business gains." Meanwhile, a Mayfield survey of Fortune 2000 IT executives found that 68% of organizations already have AI in production environments in 2025, with many of these deployments explicitly "mandated from above."

Yet many of those directives are coming from business-line leaders with minimal understanding of how to implement, scale, and govern AI effectively. In Deloitte's latest Governance of AI: A Critical Imperative for Today's Boards survey, two-thirds of respondents (66%) said their boards still have "limited to no knowledge or experience" with AI. And fewer than one in three organizations have established a comprehensive governance framework for enterprise AI. Those findings show that the pace at which AI has entered the enterprise is outstripping most organizations' ability to govern it with rigor.

When AI threats proliferate, risk related to technical vulnerabilities, data exposure, regulatory mandates, and reputational exposure multiplies.

Risk Category	**Typical Failure**	**Illustrative Incident** (2024-25)
Technical	Prompt-injection bypasses auth	UK retailer ChatBok leak (Dec 2024)
Data	Embedding of PII in vectors	APAC telco breach (Jan 2025)
Regulatory	Incomplete AIBOM	EU DPA fine, €4M (Mar 2025)
Reputation	Hallucinated policy advice	Bank chatbot PR crisis (May 2025)

In this context, a robust AI governance, risk, and compliance (GRC) program is fundamental to facilitating secure AI adoption that can deliver tangible business value. From the top down, AI governance must become a real, living discipline embedded across the organization. This requires a structured approach that connects oversight, accountability, and control, and reaches both technical and business functions to make risk visible, decisions traceable, and responsibilities clear. It's not enough to only govern internal development; governance also must cover third-party AI tools, vendor-supplied models, and unauthorized deployments operating outside standard workflows. That level of governance is becoming a prerequisite for operating in a regulatory environment that's tightening fast.

Table 1. Sample AI Security and Regulatory Frameworks

Framework Name	Country/Region	Enactment Date	Key Concern Addressed
Artificial Intelligence Act	European Union	August 2024	Establishes a risk-based classification system for AI applications[14]
ELVIS Act	United States	March 2024	Addresses unauthorized use of AI in replicating voices and likenesses[15]
Executive Order 14110: Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence	United States	October 2023	Defines national policy goals for AI governance and mandates agency actions[16]
Framework Convention on Artificial Intelligence	Council of Europe	September 2024	Emphasizes human rights and democratic values in AI development[17]
Interim Measures for the Management of Generative AI Services (生成式人工智能服务管理暂行办法)	China	August 2023	Ensures generative AI aligns with socialist values and prevents misuse[18]
Israel's Policy on Artificial Intelligence Regulation and Ethics	Israel	December 2023	Advocates for a sector-based, risk-oriented approach to AI regulation[19]
Safe and Secure Innovation for Frontier Artificial Intelligence Models Act	United States	September 2024	Mandates safety tests for powerful AI models to mitigate catastrophic risks[20]
Utah's Artificial Intelligence Policy Act	Utah, US	March 2024	Establishes liability and oversight for generative AI usage[21]

Organizations with mature governance in place have already moved toward multi-faceted GRC frameworks. Among the most important components of a well-designed GRC framework include:

1 AI Bill of Materials

Envision an AI Bill of Materials (AIBOM) as a working map. It's a live, detailed inventory of every model, dataset, API, and config in your AI environment. When something fails or when someone asks what went wrong, it's the first place you look. Done right, it gives everyone from your security team to your auditors a clear view into where models live, what data they touch, and how they behave. Without it, you're flying blind. And in a world where AI audits are becoming a when (and not an if), that's a risk most boards won't tolerate.

2 Model Registries

Model registries monitor how AI is updated and managed in production. They track lineage, versioning, model performance, and who touched what, when. More than just a source of truth, a registry creates the conditions for rollback, validation, and accountability. It's the line between experimentation and production — and without it, changes can slip through without review or documentation. For governance to function, registries are non-negotiable.

3 Executive and Cross-Functional AI Committees

Governance doesn't belong to any one department. It must be stewarded by the business. That's why high-performing organizations are forming AI governance bodies with real enforcement power, often chaired by a Chief AI Officer, Chief Risk Officer, or designated AI lead. These teams draw from legal, compliance, security, IT, and HR to align AI deployment with risk tolerance, regulatory requirements, and corporate values. This helps ensure systems are in place for safe deployment, clear escalation, and institutional accountability when things go wrong.

These components form the backbone of getting AI governance right. They take governance out of policy documents and embed it into the systems, workflows, and decisions that shape how AI is used every day. Enterprise leaders must champion this work; not just because regulators will expect it, but because the risks tied to AI are systemic and already unfolding in production environments. Without top-down ownership, governance efforts stall, oversight becomes fragmented, and accountability evaporates.

Why SANS:
Cultivating AI Fluency for Enterprise Leaders

As boards mandate faster AI adoption and regulatory bodies accelerate compliance pressures, enterprise leaders being asked a hard question about oversight: Can you prove that your organization is taking decisive action to reduce the risk of AI adoption?

This highlights the importance of AI fluency at the leadership level. Before leaders can effectively govern AI, they must inherently understand it. That doesn't mean every executive needs to know how a transformer model works, but they do need to understand what kinds of decisions AI systems are making, what data they rely on, and what risks those decisions carry. Lacking a baseline of AI fluency hinders their ability to provide meaningful oversight that adequately mitigates risk. SANS standards and training cultivate AI fluency in the areas that matter most.

Hands-On Training Mapped to Real-World Implications

SANS training empowers executive teams to rise from policy-level understanding to execution-level oversight. Whether you're approving budget, setting risk appetite, or owning incident accountability, our courses ensure that enterprise leaders are positioned to spearhead GRC programs at scale.

AIS247: AI Security Essentials for Business Leaders: Designed for senior decision-makers, this course helps boards, CIOs, CISOs, and risk management officers understand how to evaluate AI risk, set oversight strategy, and align programs to evolving regulations.

LDR512: Security Leadership Essentials for Managers: This course supports CISOs and business-line leaders in operationalizing governance across teams, tools, and policies. It covers the integration of AI risk into existing security and compliance functions.

SANS Curricula Aligned to Key Regulatory Frameworks

Regulatory Framework	Governance Focus	Mapped SANS Courses
EU AI Act	Risk classification, auditability, executive oversight	AIS247, LDR512, LDR514, SEC495, SEC545
ISO/IEC 27090	AI-specific control enforcement and lifecycle governance	SEC545, SEC535, SEC595
NIST AI RMF	Enterprise-level AI risk management	AIS247, LDR512, LDR514, SEC595
DORA	Resilience, third-party risk, operational IR	LDR514, SEC595, SEC495, SEC598

LDR514: Security Strategic Planning, Policy, and Leadership: This course focuses on enterprise-wide planning, board reporting, and third-party oversight. It helps governance owners build the structures that support defensible and ongoing AI usage.

The controls outlined in the SANS Critical AI Security Guidelines serve as the lynchpin to a strong governance, risk, and control (GRC) program that moves beyond checkbox compliance to deliver operational clarity, audit readiness, and security that holds up under pressure. Each control is directly supported by SANS's AI-focused courses for executives, policy leads, and technical owners. Adopting them drives measurable AI security without stalling innovation.

Driving a Commitment to AI Governance

When it comes to AI governance, the pressure to demonstrate oversight, defend decisions, and maintain trust ultimately falls on enterprise leaders. Their impact on the organization will be measured not just by whether they used AI, but whether they governed it effectively. SANS can help build the readiness to meet that moment. Our training fosters the knowledge, fluency, and operational strength that effective AI governance demands — at every level, in every environment, and under real-world conditions.

Learn more about SANS's cybersecurity leadership training here >

Conclusion

The global race to adopt AI is well underway, but the readiness to secure it is still trailing behind. Most organizations are deploying AI faster than they can govern, integrate, or defend it, leaving critical gaps in oversight, infrastructure, and trust. What's needed now is a stronger commitment to secure AI adoption backed by control, transparency, and accountability from the ground up.

At SANS, our position is clear: securing AI cannot be a technical afterthought. It's a prerequisite for responsible innovation. Protect the systems first. Utilize them with operational discipline. Govern them with clarity. AI can be a force multiplier, but only if it's owned securely with intent. That's the future SANS is helping organizations build toward.

Either write the rules, or adversaries write our future.

AI moves fast. Security needs to move faster by protecting, utilizing, and governing AI effectively.

Learn more about SANS's AI Leadership at sans.org/ai and contact our team today.