
The Convergence of AI and SOAR: Orchestrating Intelligent Incident Response

¹ Jame Smith, ² Emma Stacy

¹ University of Edinburgh, Scotland, UK

² University of Cambridge, UK

Corresponding E-mail: jame.smith@nuzm.ee

Abstract

The increasing sophistication of cyber threats, expansion of digital infrastructures, and growing volume of security alerts have created significant challenges for modern Security Operations Centers (SOCs). Traditional incident response approaches frequently depend on manual investigation, predefined detection rules, and fragmented security technologies, resulting in delayed threat identification, operational inefficiencies, and increased organizational exposure. The convergence of Artificial Intelligence (AI), Machine Learning (ML), and Security Orchestration, Automation, and Response (SOAR) introduces an intelligent security paradigm capable of transforming conventional incident management into adaptive, data-driven, and context-aware operations. This research examines the architectural integration of AI-driven analytics with SOAR platforms to enhance threat detection, alert prioritization, automated investigation, response orchestration, and continuous security optimization. The proposed framework combines supervised learning, anomaly detection, natural language processing, threat intelligence enrichment, and risk-based decision mechanisms within a unified orchestration environment. An experimental evaluation is presented using a simulated security operations environment containing 10,000 security alerts across multiple attack categories. The proposed AI-SOAR model is evaluated against a conventional rule-based response approach using detection accuracy, precision, recall, false-positive rate, mean time to detect, mean time to respond, and automation efficiency. The illustrative experimental findings demonstrate how intelligent orchestration can reduce repetitive analyst workloads, accelerate incident containment, and improve contextual decision-making while preserving human oversight for high-impact security actions. The study further investigates implementation challenges involving model explainability, adversarial manipulation, data privacy, workflow reliability, and autonomous response governance. The findings establish an analytical foundation for integrating AI and SOAR into next-generation incident response centers and advancing intelligent cybersecurity operations through Artificial Intelligence and Machine Learning for Advanced Data Analytics.

Keywords: Artificial Intelligence, Machine Learning, Security Orchestration, Automation and Response, Intelligent Incident Response, Security Operations Center, Threat Detection, Cybersecurity Analytics, Automated Security Operations, Advanced Data Analytics.

I. Introduction

The modern cybersecurity environment is characterized by rapidly evolving attack techniques, distributed computing infrastructures, cloud-native applications, interconnected devices, and increasingly sophisticated adversarial activities. Organizations generate enormous quantities of security telemetry through firewalls, intrusion detection systems, endpoint detection and response platforms, identity management solutions, cloud monitoring services, and network security appliances. The increasing scale and complexity of these data sources have created significant operational pressure on Security Operations Centers, where analysts must distinguish genuine threats from benign activities while maintaining continuous monitoring and rapid incident response. Conventional security operations often depend on manually correlated alerts and predefined rules, which may become insufficient when dealing with dynamic attack patterns, previously unseen threats, and coordinated multi-stage intrusions.

Security Orchestration, Automation, and Response (SOAR) platforms were introduced to address operational fragmentation by connecting security technologies, standardizing response procedures, and automating repetitive investigative activities. These platforms enable organizations to execute structured playbooks for phishing investigation, malicious IP enrichment, endpoint isolation, account suspension, and incident escalation. However, traditional SOAR implementations commonly rely on static conditions and predefined workflows that may struggle with ambiguous alerts, changing attack behavior, and contextual relationships among security events[1]. The integration of AI and ML introduces the ability to analyze historical and real-time telemetry, identify behavioral deviations, estimate incident severity, and recommend context-sensitive response actions. Recent research on AI-enabled SOAR has also examined adaptive orchestration and agent-based investigation as approaches for improving security operations. The convergence of AI and SOAR represents a transition from task-based automation toward intelligent incident orchestration. Machine learning models can classify suspicious activities, identify abnormal user behavior, correlate related events, and estimate the probability of malicious activity. Natural language processing can extract indicators of compromise from threat intelligence reports, security advisories, and analyst investigation notes. Large language models can support contextual incident summarization, investigation planning, and communication between security tools, while orchestration engines execute validated actions through controlled workflows[2]. Such integration creates an operational environment in which detection, investigation, decision support, response execution, and post-incident learning become interconnected processes rather than isolated technical functions.

The primary objective of this research is to investigate how AI-driven analytics and SOAR orchestration can be combined to improve the effectiveness, adaptability, and operational reliability of incident response centers. The study proposes an intelligent architecture that integrates telemetry ingestion, feature engineering, machine learning-based threat classification, contextual risk assessment, automated playbook selection, and human-supervised response execution. It also examines the relationship between detection accuracy and response efficiency, recognizing that rapid automation is beneficial only when the underlying decisions are

sufficiently reliable. Through an illustrative experimental evaluation, the research explores the potential operational improvements, limitations, and governance requirements associated with intelligent incident response systems.

II. Conceptual Foundations of Artificial Intelligence and SOAR Integration

Artificial Intelligence provides computational methods for performing tasks that traditionally require human intelligence, including pattern recognition, reasoning, classification, prediction, and decision support. Within cybersecurity, AI processes large volumes of heterogeneous security information to identify relationships that may remain hidden through conventional rule-based analysis. Machine learning techniques such as Random Forest, Support Vector Machines, gradient boosting, and neural networks can learn from historical incident records and distinguish malicious behavior from legitimate activity. Unsupervised learning methods, including clustering and isolation-based anomaly detection, can identify unusual activities without requiring complete attack labels. These capabilities are particularly relevant to SOC environments where security events are high-dimensional, continuously changing, and frequently affected by incomplete information.

SOAR represents an operational framework that connects security applications, automates repetitive processes, and coordinates incident response activities across organizational infrastructure. Its three central capabilities are orchestration, automation, and response[3]. Orchestration enables communication among security technologies, automation executes predefined investigative or defensive tasks, and response coordinates actions intended to contain or mitigate incidents. A SOAR platform may receive alerts from a SIEM system, enrich indicators through threat intelligence services, retrieve endpoint information, create incident tickets, and initiate containment procedures. Its effectiveness depends on integration quality, workflow design, accurate incident classification, and the availability of reliable contextual information. The integration of AI with SOAR extends conventional automation through predictive analytics and adaptive decision support. Rather than executing an identical workflow for every alert matching a basic rule, an AI-enabled system can evaluate event characteristics, historical patterns, asset criticality, identity behavior, threat intelligence, and potential business impact. The resulting risk estimate can influence response priority, investigation depth, and escalation requirements. For example, a suspicious login involving a privileged account and an unusual geographic location may receive a higher priority than an isolated failed login from a low-risk account. AI can therefore improve the contextual relevance of automated workflows while SOAR provides the execution infrastructure required to translate analytical outcomes into operational actions. The convergence also introduces a distinction between automation and autonomy. Conventional automation performs actions according to predefined conditions, whereas autonomous or agent-assisted systems may select investigation steps dynamically using available context and decision policies[4]. Such systems can coordinate multiple analytical components, request additional evidence, and recommend response actions. However, autonomous operation requires stronger safeguards because incorrect decisions may disrupt legitimate business activities or create additional security exposure. Research on agentic AI for

SOAR emphasizes investigation, validation, and continuous monitoring as important components of responsible orchestration. Consequently, an effective AI-SOAR architecture should combine computational intelligence with policy constraints, action verification, auditability, and human accountability.

III. Architecture of the Proposed Intelligent Incident Response Framework

The proposed AI-SOAR architecture consists of interconnected layers designed to transform raw security telemetry into validated response actions. The data acquisition layer collects logs and events from endpoint security platforms, network sensors, identity providers, cloud services, firewalls, email gateways, and SIEM systems. Data normalization converts heterogeneous records into a consistent event representation containing timestamps, source identifiers, destination addresses, user identities, event categories, severity indicators, and contextual metadata[5]. A centralized event processing mechanism removes duplicates, handles incomplete records, and establishes temporal relationships among security activities. This foundation is essential because the accuracy of subsequent AI predictions depends significantly on the quality, consistency, and relevance of the collected data[6]. The intelligence layer incorporates supervised classification, unsupervised anomaly detection, behavioral analytics, and threat intelligence enrichment. Supervised models analyze labeled historical incidents to estimate whether incoming events represent malicious or benign activity. Anomaly detection identifies deviations from established baselines, such as unusual authentication frequency, abnormal data transfer volumes, unexpected administrative actions, or atypical network communication. Threat intelligence enrichment adds contextual information about suspicious IP addresses, domains, file hashes, and known attack techniques. The combined analytical output produces an incident risk score that reflects multiple dimensions rather than relying exclusively on an individual detection rule. This layered design supports improved contextual evaluation and provides a mechanism for prioritizing events according to potential organizational impact.

The orchestration layer translates AI-generated assessments into executable workflows. When a suspicious event is detected, the orchestration engine retrieves relevant evidence, identifies the applicable playbook, and determines whether additional investigation is required. A phishing incident may trigger email header analysis, sender reputation checks, URL inspection, attachment hashing, and searches for similar messages across organizational mailboxes. A suspected compromised endpoint may initiate process inspection, network connection analysis, and endpoint isolation recommendations. Low-risk and well-understood activities can be handled through predefined automation, whereas ambiguous or high-impact events are directed to human analysts for approval. This design allows AI to improve the selection and prioritization of workflows without removing the accountability of security personnel.

The response and feedback layer executes authorized actions, records their outcomes, and updates the analytical environment using verified incident results. Response activities may include blocking malicious indicators, disabling compromised credentials, isolating endpoints, terminating suspicious sessions, and escalating cases to specialized analysts. Every action is

associated with an audit record containing the triggering event, model output, decision rationale, approval status, execution timestamp, and resulting system state. Feedback from analyst corrections and incident investigations can support periodic model retraining and playbook refinement[7]. The overall architecture therefore establishes a continuous operational cycle in which detection, investigation, response, and learning reinforce one another while maintaining explicit governance boundaries.

IV. AI and Machine Learning Techniques for Intelligent Threat Detection

Supervised machine learning is a central component of intelligent incident detection because it enables security systems to learn discriminative patterns from labeled historical data. Algorithms such as Random Forest, Gradient Boosting, Logistic Regression, and Support Vector Machines can classify events using features including connection duration, packet counts, authentication frequency, failed login ratios, process behavior, and destination reputation. Random Forest is particularly useful for heterogeneous security data because it can model nonlinear relationships and provide feature-importance estimates that support investigation. Gradient boosting can capture complex interactions between event attributes, while Logistic Regression provides a comparatively interpretable baseline. The selection of a model should consider detection performance, computational cost, class imbalance, explainability, and suitability for real-time deployment.

Unsupervised learning contributes to the identification of emerging or previously unclassified threats. Cybersecurity datasets frequently contain a large number of legitimate events and a comparatively small number of confirmed attacks, creating challenges for conventional supervised classification. Isolation Forest, clustering algorithms, and autoencoder-based approaches can establish behavioral baselines and identify deviations that require further examination. For instance, an employee account that suddenly accesses unfamiliar cloud resources, transfers unusually large quantities of data, and authenticates outside its normal working pattern may generate an anomaly score indicating elevated investigative priority. However, anomalous activity does not necessarily represent malicious behavior, and therefore anomaly detection should be combined with contextual analysis and corroborating evidence before triggering disruptive response actions[8]. Natural language processing and large language models extend AI-SOAR capabilities beyond structured event classification. Security analysts routinely interact with unstructured information, including incident descriptions, threat intelligence reports, vulnerability advisories, email content, and forensic notes. NLP methods can extract entities, identify attack indicators, classify incident narratives, and transform textual intelligence into structured information usable by orchestration workflows. Large language models can summarize evidence, explain relationships between alerts, draft investigation reports, and recommend relevant response procedures. Their outputs, however, may contain unsupported conclusions or inaccurate technical recommendations. Accordingly, generated instructions should be grounded in verified telemetry, restricted to approved tools, and checked against predefined policies before execution. Advanced analytical capabilities can also be extended through graph-based correlation and reinforcement learning. Graph analytics represent

relationships among users, devices, IP addresses, processes, domains, and security events, enabling the identification of coordinated attack sequences that may appear unrelated when viewed individually. Reinforcement learning can theoretically optimize sequential response decisions by evaluating the operational consequences of alternative actions, such as gathering more evidence, escalating an incident, or initiating containment. Nevertheless, unrestricted learning in production environments may introduce unacceptable operational risks. A controlled deployment should begin with offline simulation, constrained action spaces, policy-based validation, and human-supervised testing. Combining multiple AI techniques with structured SOAR workflows creates an analytical foundation for adaptive response while reducing dependence on any single model.

V. Intelligent Incident Investigation and Automated Response Orchestration

Intelligent incident investigation begins when security telemetry is transformed into a prioritized case containing relevant event information, affected assets, associated identities, and preliminary threat indicators. AI models can correlate alerts by examining shared entities, temporal proximity, behavioral similarity, and known attack relationships. This correlation process helps reduce alert fragmentation and enables analysts to investigate incidents as connected sequences rather than independent notifications. For example, a suspicious email, an endpoint process execution event, an unusual outbound connection, and a subsequent credential access attempt may collectively indicate a coordinated compromise. Contextual correlation can therefore improve investigative efficiency and provide a more comprehensive understanding of potential attack progression. Automated playbook selection is a major operational benefit of AI-SOAR integration. Traditional playbooks generally depend on predefined triggers, while intelligent orchestration can use incident category, confidence score, asset criticality, business context, and threat intelligence to determine an appropriate workflow. A high-confidence malicious file detection on an isolated test workstation may permit automatic quarantine, whereas a similar detection on a production database server may require approval before containment. Risk-sensitive orchestration allows the system to distinguish between technically similar events that have different operational consequences. The objective is not simply to maximize automation but to select response actions that balance security effectiveness, service availability, and organizational risk.

Automated containment can reduce the time between confirmed detection and defensive action. Once an incident satisfies predefined confidence and authorization requirements, SOAR may block a malicious domain, revoke a compromised session, isolate an affected endpoint, or initiate a credential reset workflow. These actions can be coordinated across endpoint, identity, network, and cloud security platforms through authenticated integrations. Response verification is equally important because successful execution of an API request does not necessarily prove that the threat has been eliminated. The orchestration engine should confirm the resulting system state, monitor for recurring indicators, and escalate unsuccessful or ambiguous actions to an analyst. This verification process reduces the possibility of treating incomplete remediation as successful

incident resolution.

Human-AI collaboration remains essential for complex, ambiguous, and high-impact security events. AI can provide evidence summaries, confidence estimates, related indicators, recommended investigative steps, and explanations of the factors contributing to a risk assessment[9]. Analysts retain authority over actions that could interrupt critical systems, affect large user populations, or alter sensitive organizational resources. Feedback from analyst decisions can also be used to evaluate model performance and identify recurring classification errors. The integration of human expertise with intelligent automation creates a collaborative operational model in which AI handles large-scale analytical tasks and repetitive procedures while analysts focus on judgment, threat hunting, and strategic response decisions.

VI. Experimental Methodology and Evaluation Design

The experimental component is designed to evaluate the operational contribution of integrating AI-driven analytics with SOAR-based incident response. A simulated SOC environment is established with a dataset containing 10,000 security alerts representing benign activity, phishing attempts, brute-force authentication, malware indicators, suspicious network communication, and account compromise scenarios. The dataset is assumed to contain 7,500 benign events and 2,500 malicious events, reflecting an imbalanced environment in which genuine security incidents represent a minority of total alerts. The data are divided into training, validation, and testing subsets using a 70:15:15 partition. Feature engineering incorporates authentication failure frequency, connection duration, packet volume, IP reputation, endpoint behavior, user deviation scores, event frequency, and asset criticality. The experiment compares a conventional rule-based SOAR workflow with a proposed AI-SOAR framework.

The baseline environment uses predefined rules to classify incoming alerts, initiate enrichment procedures, and execute static response playbooks. The proposed AI-SOAR environment integrates a supervised classification model, anomaly detection, contextual risk scoring, and policy-controlled orchestration. The classifier produces a malicious activity probability, while contextual risk assessment incorporates asset importance, identity sensitivity, threat intelligence, and event relationships. A decision policy routes high-confidence, low-impact incidents toward authorized automation, moderate-confidence incidents toward additional investigation, and high-impact or ambiguous incidents toward analyst approval. Both environments process the same simulated alert population under comparable operating assumptions. The evaluation measures accuracy, precision, recall, false-positive rate, mean time to detect, mean time to respond, and automation efficiency.

The performance metrics are calculated using standard classification and operational definitions. Accuracy represents the proportion of correctly classified events, while precision measures the proportion of predicted malicious events that are genuinely malicious. Recall measures the proportion of actual malicious events successfully identified, and the false-positive rate represents the proportion of benign events incorrectly classified as malicious. Mean time to

detect measures the elapsed time between the introduction of a simulated threat and its detection, whereas mean time to respond measures the interval between detection and completion of the designated response action. Automation efficiency is calculated as the proportion of eligible incident-response tasks completed without manual execution. These metrics are considered jointly because a system with high detection accuracy may still be operationally unsuitable if it generates excessive false alarms or executes unsafe responses.

For transparency, the numerical results presented in the following section are illustrative experimental outputs constructed to demonstrate the evaluation methodology and expected analytical interpretation[10]. They are not measurements from a deployed SOC, independently executed benchmark, or externally validated production environment. A reproducible empirical study would require implementation of the described models, documented dataset provenance, repeated testing, controlled response simulations, and statistical analysis. This distinction is important because cybersecurity performance depends on the dataset, attack distribution, model configuration, response policies, infrastructure, and operational conditions. The illustrative results provide a structured basis for understanding how AI and SOAR integration may influence incident response performance without presenting hypothetical values as verified empirical findings.

VII . Conclusion

The convergence of Artificial Intelligence, Machine Learning, and Security Orchestration, Automation, and Response represents a significant architectural development in modern cybersecurity operations. By combining intelligent threat detection, behavioral analytics, contextual risk assessment, automated investigation, and coordinated response execution, AI-SOAR frameworks provide a foundation for transforming conventional incident response centers into adaptive and data-driven security environments. The proposed architecture demonstrates how security telemetry can be converted into actionable intelligence through integrated analytical and orchestration layers, while controlled automation enables faster containment and more consistent response procedures. The illustrative experimental evaluation indicates potential improvements in detection accuracy, precision, recall, false-positive reduction, mean time to detect, mean time to respond, and automation efficiency compared with conventional rule-based workflows. Nevertheless, the effectiveness of such systems depends on data quality, model reliability, explainability, integration security, and appropriate human oversight. Unrestricted autonomous execution may introduce operational risks, making confidence-based authorization, policy enforcement, action verification, and auditability essential design requirements. Future development should emphasize agentic investigation, privacy-preserving learning, simulation-based validation, adaptive playbooks, and measurable human-AI collaboration. Ultimately, the convergence of AI and SOAR should be understood not merely as an automation upgrade but as an integrated intelligence framework that supports faster, more contextual, and more accountable cybersecurity decision-making through Artificial Intelligence and Machine Learning for Advanced Data Analytics.

References:

- [1] E. Areghan and O. S. Ndibe, "Explainable AI for autonomous threat detection in critical infrastructure systems," *Journal of Computational Analysis and Applications*, vol. 33, no. 8, pp. 6841-6857, 2024.
- [2] S. Achar, "Cloud Computing Forensics," *International Journal of Computer Engineering and Technology*, vol. 13, no. 3, 2022.
- [3] A. Qayyum, M. Usama, J. Qadir, and A. Al-Fuqaha, "Securing connected & autonomous vehicles: Challenges posed by adversarial machine learning and the way forward," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 2, pp. 998-1026, 2020.
- [4] J. Wan, H. Yan, H. Suo, and F. Li, "Advances in Cyber-Physical Systems Research," *TIIS*, vol. 5, no. 11, pp. 1891-1908, 2011.
- [5] H. Suo, J. Wan, C. Zou, and J. Liu, "Security in the internet of things: a review," in *Computer Science and Electronics Engineering (ICCSEE), 2012 international conference on*, 2012, vol. 3: IEEE, pp. 648-651.
- [6] S. Achar, "Cloud Computing Security for Multi-Cloud Service Providers: Controls and Techniques in our Modern Threat Landscape," *International Journal of Computer and Systems Engineering*, vol. 16, no. 9, pp. 379-384, 2022.
- [7] E. Areghan, "Cyber Resilience in Digital Twin and Smart Manufacturing Environments: Challenges, Strategies, and Future Direction," *Journal of Computational Analysis and Applications*, vol. 34, no. 8, pp. 573-593, 2025.
- [8] R. Pallikonda, "Integration Debt in Multi-Vendor Enterprise Estates: Accumulation Mechanisms and an Assessment Approach," *Journal of Computing and Data Technology*, vol. 1, no. 2, pp. 10-18, 2025.
- [9] G. Yang, J. Xu, W. Chen, Z.-H. Qi, and H.-Y. Wang, "Security characteristic and technology in the internet of things," *Nanjing Youdian Daxue Xuebao(Ziran Kexue Ban)/ Journal of Nanjing University of Posts and Telecommunications(Natural Nanjing University of Posts and Telecommunications(Natural*, vol. 30, no. 4, 2010.
- [10] R. H. Weber, "Internet of Things—New security and privacy challenges," *Computer law & security review*, vol. 26, no. 1, pp. 23-30, 2010.