
Automating the SOC: How AI Is Transforming Threat Detection and Response

¹ Jabulani Khumalo, ² Kim Min Joon

¹ University of South Africa, Pretoria, South Africa

² Pohang University of Science and Technology (POSTECH), Pohang, South Korea

Corresponding E-mail: jabulani.khumalo@interviauniversity.com

Abstract

Security Operations Centers (SOCs) play a critical role in protecting modern digital infrastructures against increasingly sophisticated cyber threats. However, the growing volume of security alerts, complex attack patterns, shortage of skilled cybersecurity professionals, and limitations of conventional detection mechanisms have created significant operational challenges. Artificial Intelligence (AI) and Machine Learning (ML) are transforming SOC operations by enabling intelligent threat detection, automated alert prioritization, behavioral analytics, incident correlation, and rapid security response. This research examines the integration of AI-driven automation into SOC environments, focusing on the application of supervised learning, unsupervised learning, deep learning, natural language processing, and large language models for advanced cybersecurity analytics. The study proposes an AI-enabled SOC architecture that combines Security Information and Event Management (SIEM), Security Orchestration, Automation, and Response (SOAR), endpoint telemetry, threat intelligence, and automated investigation mechanisms. An experimental evaluation framework is developed to compare conventional rule-based monitoring with an AI-assisted detection and response approach using representative security event scenarios. Evaluation parameters include detection accuracy, precision, recall, F1-score, false-positive rate, alert processing time, and incident response latency. The proposed framework demonstrates how intelligent correlation and automated triage can improve operational efficiency while maintaining human oversight over high-impact security decisions. The research further investigates challenges associated with adversarial attacks, model drift, data privacy, explainability, and automated containment. The findings establish that AI-driven SOC automation represents an important direction for advanced cybersecurity operations, particularly when machine intelligence is integrated with reliable security policies, continuous monitoring, and analyst validation.

Keywords: Artificial Intelligence, Security Operations Center, Machine Learning, Threat Detection, Incident Response, Security Automation, SIEM, SOAR, Deep Learning, Cybersecurity Analytics, Automated Threat Hunting, Advanced Data Analytics.

I. Introduction

The rapid expansion of cloud computing, distributed applications, Internet of Things (IoT) devices, remote working environments, and interconnected enterprise infrastructures has

significantly increased the complexity of cybersecurity operations. Modern organizations generate enormous volumes of network traffic, endpoint logs, authentication records, application events, and security alerts every day. Security Operations Centers are responsible for collecting, monitoring, analyzing, and responding to these events to identify malicious activities before they cause substantial damage. However, traditional SOC operations frequently depend on predefined detection rules, manual investigations, and human-driven incident prioritization. These approaches face limitations when dealing with high-volume, rapidly evolving, and previously unseen cyber threats. Artificial Intelligence has emerged as a technological approach for enhancing the analytical capabilities of SOC environments through automated pattern recognition, behavioral analysis, and intelligent decision support.

Traditional security monitoring systems generally rely on signature-based detection, static thresholds, and predefined correlation rules[1]. Although these mechanisms are effective for identifying known attack indicators, they may struggle to detect sophisticated threats that exhibit subtle behavioral changes or use previously unknown attack techniques. Advanced persistent threats, credential misuse, lateral movement, ransomware activities, and fileless malware can generate security signals that appear legitimate when examined individually. AI-based detection introduces the ability to analyze relationships between multiple events, recognize abnormal behavior, and identify deviations from established baselines. Machine learning models can process historical security data to identify patterns associated with malicious activities, while deep learning architectures can analyze sequential and high-dimensional data. This transition from isolated rule matching toward contextual and behavior-based analytics represents a significant development in modern cybersecurity operations.

The increasing workload of SOC analysts has also contributed to alert fatigue, which occurs when security professionals must investigate excessive numbers of alerts, including repetitive notifications and false positives. When analysts spend considerable time reviewing low-priority events, their capacity to investigate genuine threats may decrease. AI-driven automation can address this challenge by enriching alerts with contextual information, correlating related events, assigning risk scores, and recommending response actions. Natural language processing and large language models can further assist analysts by summarizing incident timelines, explaining detection results, extracting indicators of compromise, and generating investigation reports. Recent research on AI applications in SOC environments identifies alert analysis, detection, and response assistance as important areas of development, while also recognizing limitations involving reliability, security, and operational integration[2]. The primary objective of this research is to investigate how AI transforms threat detection and incident response through an integrated SOC automation framework. The study examines the relationship between machine learning algorithms, security event processing, automated investigation, and response orchestration. It also develops an experimental evaluation framework to assess detection performance and operational efficiency under representative cybersecurity conditions. Particular attention is given to the balance between automation and human supervision, because automated security actions can affect business continuity, access availability, and sensitive information. The research contributes a structured understanding of AI-enabled SOC operations by connecting

advanced data analytics with practical detection and response workflows. The central research proposition is that integrating AI-based analytical capabilities with established security controls can improve threat visibility, reduce repetitive investigation activities, and support faster incident handling without eliminating the essential role of human cybersecurity professionals.

II. Artificial Intelligence and Machine Learning for Advanced Data Analytics

Artificial Intelligence and Machine Learning provide the analytical foundation for transforming conventional SOC environments into intelligent security monitoring systems. A modern SOC receives heterogeneous data from firewalls, intrusion detection systems, endpoint detection platforms, identity management systems, cloud workloads, and network devices. These sources produce structured and unstructured information that must be normalized, correlated, and analyzed to identify suspicious activities[3]. Machine learning algorithms enable security platforms to discover patterns that may not be represented by predefined rules. Supervised learning techniques, including Random Forest, Support Vector Machines, Logistic Regression, and Gradient Boosting, can classify security events using labeled historical data. Unsupervised approaches, such as clustering and anomaly detection, can identify unusual behavior without requiring every possible attack pattern to be labeled in advance.

Deep learning extends these capabilities by learning complex representations from large-scale security datasets. Recurrent Neural Networks and Long Short-Term Memory networks can analyze sequences of authentication events, network connections, and process activities to identify suspicious temporal patterns[4]. Autoencoders can learn normal behavioral representations and identify deviations through reconstruction errors. Convolutional Neural Networks can be applied to transformed traffic representations and structured event features, while transformer-based models can analyze contextual relationships within security logs and incident descriptions. These techniques are particularly relevant when attacks involve multiple stages, where an initial authentication anomaly may be followed by privilege escalation, unusual process execution, and suspicious outbound communication. Nevertheless, model performance depends heavily on dataset quality, representative training examples, appropriate validation, and careful handling of class imbalance. Advanced data analytics also involves feature engineering, dimensionality reduction, event aggregation, and contextual enrichment. Security events may contain attributes such as source IP address, destination IP address, username, event timestamp, process identifier, authentication outcome, destination port, and event severity. These features can be transformed into behavioral indicators, including login frequency, unusual access times, failed authentication ratios, connection diversity, and deviations from historical activity. Threat intelligence feeds can enrich these observations with known malicious indicators, while asset inventories can establish the business importance of affected systems. AI models can combine these dimensions to generate contextual risk scores rather than treating every alert as an isolated event. This approach supports more meaningful prioritization because the same suspicious activity may have different operational implications depending on the identity, asset, and surrounding event sequence.

The application of AI to advanced cybersecurity analytics must also account for model explainability, adversarial manipulation, and data drift. A detection model that produces a high-risk score without providing understandable evidence may create difficulties for analysts who must validate and document their decisions. Explainable AI techniques, including feature importance, SHAP-based explanations, and interpretable decision rules, can help identify the factors contributing to an alert classification. Model drift occurs when normal organizational behavior changes over time, potentially reducing detection reliability if models are not monitored and updated. Furthermore, attackers may deliberately manipulate observable patterns to evade machine learning systems. Research on machine learning-based zero-day detection identifies continuing challenges in accuracy, recall, and consistency across different attack types. Consequently, AI should be deployed as a continuously evaluated analytical capability rather than a static replacement for traditional security controls.

III. Evolution of the Modern Security Operations Center

The traditional SOC architecture was primarily designed around centralized security monitoring, log collection, rule-based correlation, and analyst-led investigation. Security Information and Event Management systems collect information from multiple infrastructure components and apply detection rules to identify suspicious activities. Analysts examine generated alerts, determine their relevance, investigate supporting evidence, and coordinate incident response. Although this architecture provides centralized visibility, its operational effectiveness can decrease when event volumes exceed available analytical resources. The growing diversity of enterprise infrastructure has further complicated monitoring because security events originate from cloud platforms, remote endpoints, mobile devices, applications, and identity services. A modern SOC therefore requires a more integrated approach capable of processing information across different environments while maintaining consistent security visibility.

The integration of Extended Detection and Response platforms has expanded the monitoring capabilities of modern SOC's by combining endpoint, network, identity, email, and cloud security information. XDR systems provide broader contextual visibility than isolated security products and can support investigation across multiple security domains. AI enhances these capabilities by identifying relationships among apparently unrelated events, such as a suspicious email, an unusual authentication attempt, and an unexpected process execution. Correlation across these signals can reveal a potential attack sequence that would otherwise remain fragmented across separate monitoring systems. The MITRE ATT&CK framework can provide a structured analytical reference for mapping observed activities to adversary tactics and techniques, helping analysts organize investigations around recognizable behavioral patterns. Security Orchestration, Automation, and Response platforms introduce another important dimension by connecting detection systems with predefined response workflows. SOAR can automate repetitive actions such as gathering endpoint information, checking IP reputation, enriching alerts, creating incident tickets, and notifying responsible personnel. When integrated with AI, orchestration

workflows can use model-generated risk assessments to determine investigation priority and recommend appropriate next steps. For example, an AI model may identify unusual login behavior, while an orchestration engine collects authentication history, checks the affected account's privileges, and retrieves associated endpoint activity. A human analyst can then review the consolidated evidence and authorize containment when necessary. This combination of intelligent analytics and workflow automation reduces operational fragmentation and supports more consistent incident handling.

The modern SOC is increasingly moving toward a human-machine collaboration model in which AI handles high-volume analytical tasks while human professionals retain responsibility for complex decisions. Tier-one analysts may use AI-assisted triage to prioritize alerts, while tier-two investigators can employ automated incident summaries and attack-path reconstruction. Senior incident responders can use AI-generated hypotheses to accelerate threat hunting and identify relationships across long-running investigations. Large language models may also assist with converting technical event data into understandable reports, although their outputs require verification against original telemetry. A 2026 comprehensive survey of 216 studies on large language models in SOC operations identifies substantial research activity in detection, analysis, and response, while highlighting remaining research gaps in recovery-related functions. This evolution suggests that the future SOC will depend not only on intelligent detection but also on the effective coordination of data, automation, security expertise, and governance.

IV. AI-Driven Threat Detection and Behavioral Analytics

AI-driven threat detection transforms security monitoring from static rule evaluation toward dynamic behavioral analysis. Conventional detection systems often identify suspicious activities through known signatures, predefined indicators, and fixed thresholds. However, modern attacks may use legitimate administrative tools, compromised credentials, encrypted communication, and low-frequency activities to avoid conventional detection mechanisms. Behavioral analytics provides an alternative by establishing normal activity profiles for users, devices, applications, and network segments. Machine learning algorithms can examine deviations from these profiles and generate risk assessments based on multiple contextual variables. For instance, an account that normally accesses a limited set of internal resources during business hours may generate a high-risk event when it suddenly performs unusual authentication attempts and accesses sensitive systems from an unfamiliar location.

User and Entity Behavior Analytics represents an important application of AI in threat detection. UEBA models evaluate behavioral characteristics associated with users, endpoints, service accounts, and other entities[5]. These characteristics may include login frequency, geographic access patterns, data transfer volumes, resource utilization, and privileged activity. Statistical models and clustering algorithms can establish behavioral baselines and detect unusual deviations. A sudden increase in failed authentication attempts, unexpected access to confidential files, or unusual communication between internal systems may indicate compromised credentials or insider-related security concerns. By correlating these indicators, AI systems can identify

suspicious behavior even when individual events do not match known attack signatures. However, behavioral anomalies should be treated as investigative signals rather than automatic proof of malicious intent.

Deep learning and sequence analysis can improve detection of multi-stage attacks by examining the temporal relationships between security events[6]. A typical intrusion may involve initial access, credential discovery, privilege escalation, lateral movement, and data exfiltration. Each stage may appear relatively ordinary when considered independently, but the complete sequence can reveal malicious intent. Sequence-based models can analyze event order, timing, frequency, and transitions between system activities. Graph-based analytics can further represent relationships between identities, endpoints, processes, network connections, and resources. Such representations support attack-path analysis and enable the identification of suspicious connections between otherwise disconnected observations. The combination of sequence modeling and graph analytics provides a foundation for identifying complex threats that require contextual interpretation rather than isolated event classification.

Threat intelligence integration further enhances AI-based detection by combining behavioral evidence with external indicators and known adversary techniques. Indicators of compromise may include malicious domains, suspicious file hashes, IP addresses, and unusual process behaviors. AI systems can correlate these indicators with local telemetry and assign different levels of confidence according to evidence quality and context. MITRE ATT&CK mappings can help analysts connect observed behavior with known adversary techniques, improving investigation consistency and supporting threat hunting. However, threat intelligence should not be treated as universally reliable because indicators can become outdated, shared infrastructure may serve legitimate users, and adversaries can modify their techniques. An effective detection architecture therefore combines machine learning, threat intelligence, rule-based controls, and analyst verification to reduce dependence on any single detection mechanism.

V. Proposed AI-Enabled SOC Architecture

The proposed AI-enabled SOC architecture consists of five integrated operational layers: security data collection, preprocessing and enrichment, intelligent analytics, automated orchestration, and human-supervised incident management. The data collection layer receives telemetry from firewalls, intrusion detection systems, endpoint detection platforms, identity providers, cloud environments, servers, and applications. The preprocessing layer normalizes timestamps, standardizes event fields, removes duplicate records, and enriches observations with asset information and threat intelligence. The analytics layer applies machine learning models for classification, anomaly detection, behavioral analysis, and event correlation. The orchestration layer executes approved response workflows, while the human-supervision layer provides investigation, authorization, exception handling, and continuous feedback. This layered design separates analytical decision support from operational execution and improves the manageability of security automation[7]. The data ingestion process is designed to support both structured and unstructured security information. Structured information includes authentication logs, firewall

records, network flow data, and endpoint event identifiers. Unstructured information includes incident descriptions, analyst notes, threat intelligence reports, and security investigation narratives. Data normalization converts heterogeneous inputs into a consistent representation that can be processed by analytical models. Feature extraction identifies relevant indicators, including authentication failure ratios, connection frequency, unusual process behavior, and access deviations. Contextual enrichment connects each event with user privileges, asset criticality, historical activity, and known threat indicators. Data quality checks are necessary to identify missing values, inconsistent timestamps, duplicated events, and incomplete telemetry before model inference.

VI. Future Directions of AI-Driven SOC Automation

The future development of SOC automation is expected to involve greater integration of autonomous AI agents, advanced behavioral analytics, and security knowledge graphs. AI agents may coordinate specialized tasks such as log investigation, threat intelligence retrieval, attack-path analysis, and incident report generation[8]. Rather than depending on a single model, multi-agent architectures can distribute analytical responsibilities among components designed for different security functions. However, autonomous agents must operate within clearly defined permissions and policy constraints. Their decisions should be traceable, their data sources verifiable, and their high-impact actions subject to appropriate authorization. Research into automated cyber defense increasingly considers machine learning, deep learning, reinforcement learning, and simulation-based evaluation as connected areas of investigation. Another important direction involves the development of more adaptive detection systems capable of responding to changing organizational behavior and evolving attack techniques. Continual learning may allow models to incorporate new information while retaining previously acquired detection capabilities[9]. Federated learning may support collaborative model development across organizations without requiring the direct centralization of sensitive raw data. Privacy-preserving analytics can further reduce unnecessary exposure of operational information. However, continual and distributed learning introduce challenges involving model poisoning, inconsistent data distributions, and performance verification[10]. Future architectures should combine adaptive learning with model validation, version control, drift monitoring, and secure update mechanisms to ensure that changes improve detection rather than silently reducing reliability.

VII. Conclusion

Artificial Intelligence is transforming Security Operations Centers by enabling advanced threat detection, behavioral analytics, intelligent alert prioritization, automated investigation, and coordinated incident response. This research examined the integration of machine learning, deep learning, anomaly detection, natural language processing, SIEM, and SOAR within an AI-enabled SOC architecture designed to address the limitations of conventional security monitoring. The proposed framework demonstrates how heterogeneous security data can be normalized, enriched, analyzed, and transformed into actionable intelligence through automated workflows and human-supervised decision-making. The illustrative experimental evaluation

indicates how AI-assisted detection could improve precision, recall, and F1-score while reducing false-positive alerts, investigation time, and response initiation latency; however, these numerical outcomes require validation through actual implementation and empirical testing. The study also establishes that effective automation depends on reliable datasets, model explainability, adversarial resilience, privacy protection, continuous evaluation, and controlled execution of security actions. AI should therefore be regarded as an augmentation of human cybersecurity expertise rather than an unrestricted replacement for SOC analysts. By combining intelligent analytics with transparent governance, secure orchestration, and continuous feedback, organizations can develop more responsive and adaptive security operations capable of addressing increasingly complex cyber threats. The integration of Artificial Intelligence and Machine Learning for Advanced Data Analytics provides a foundation for the evolution of SOC environments toward intelligent, evidence-driven, and operationally resilient cybersecurity systems.

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] M. Sills, P. Ranade, and S. Mittal, "Cybersecurity threat intelligence augmentation and embedding improvement-a healthcare usecase," in *2020 IEEE international conference on intelligence and security informatics (ISI)*, 2020: IEEE, pp. 1-6.
- [3] G. Sunkara, "AI Powered Threat Detection in Cybersecurity," *International Journal of Humanities and Information Technology*, no. Special 1, pp. 1-22, 2021.
- [4] 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.
- [5] J. Uddoh, D. Ajiga, B. P. Okare, and T. D. Aduloju, "AI-based threat detection systems for cloud infrastructure: Architecture, challenges, and opportunities," *Journal of Frontiers in Multidisciplinary Research*, vol. 2, no. 2, pp. 61-67, 2021.
- [6] A. Abisoye and J. I. Akerele, "High-impact data-driven decision-making model for integrating cutting-edge cybersecurity strategies into public policy," *Governance, and Organizational Frameworks*, 2021.
- [7] T. Ahmad, "Technology convergence and cybersecurity: a critical analysis of cybercrime trends in India," *27th Convergence India Pragati Maidan*, pp. 29-31, 2019.
- [8] A. K. Soma, "Building Aether sensor network using LoRaWAN network," in *2025 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, 2025: IEEE, pp. 807-814.
- [9] A. E. Hassanien, T.-H. Kim, J. Kacprzyk, and A. I. Awad, "Bio-inspired cyber security and cloud services: Trends and innovations," 2014.
- [10] I. Wiafe, F. N. Koranteng, E. N. Obeng, N. Assyne, A. Wiafe, and S. R. Gulliver, "Artificial intelligence for cybersecurity: a systematic mapping of literature," *Ieee Access*, vol. 8, pp. 146598-146612, 2020.