Integrating LLM-Based Security Copilots with SIEM Platforms for Explainable Threat Detection and Automated Incident Investigation

Jamal-Deen Abdulai
Department of Computer Science, University of Ghana, Accra, Ghana

Abstract

Security information and event management (SIEM) platforms give security teams a central view of events across the enterprise, but turning that view into decisions still depends on analysts writing queries, reading logs and piecing incidents together by hand. Large language model (LLM)-based security copilots promise to change this by letting analysts work with SIEM data in natural language and by automating much of the investigative work. This article proposes an architecture for integrating LLM security copilots with SIEM platforms to deliver explainable threat detection and automated incident investigation. The copilot sits alongside the SIEM. It explains detections using feature attribution, retrieves context through retrieval-augmented generation, and investigates incidents through a reason-and-act cycle of queries and tool calls. The article pays particular attention to enterprise SAP landscapes, where trust boundaries, identity context and data privacy shape what a copilot may see and do. Drawing on nine studies, including six published between 2020 and 2022 on SIEM, security operations centres, explainable intrusion detection, large language models and retrieval-augmented generation, together with work on SAP security, privacy governance and real-time enterprise analytics, it sets out an integration architecture, an investigation workflow and a governance model. It concludes that copilots add most value when every answer they give is traceable to SIEM evidence.

Keywords: security copilot; large language models; SIEM; explainable AI; incident investigation; retrieval-augmented generation; SAP security; data privacy

1. Introduction

The SIEM is the central platform of most security operations centres. It collects, aggregates, stores and correlates events from firewalls, endpoints, identity systems, cloud services and business applications, giving analysts a single place to detect and respond to attacks (González-Granadillo et al., 2021). Yet the SIEM's value depends on the people using it. Analysts must write queries in specialised languages, interpret correlation alerts, pivot across data sources and reconstruct what happened. In a busy SOC, these tasks consume most of the working day, and many alerts are never investigated in depth.

Research on SOCs has highlighted this gap. Technology and people are well studied, but the processes that connect them, including how analysts turn data into decisions, have received less attention (Vielberth et al., 2020). The result is a familiar pattern: rich data, overloaded analysts and slow investigations.

Large language models offer a new kind of interface to this data. Brown et al. (2020) showed that large models can perform a wide range of language tasks from instructions and a few examples, without task-specific training. In a SOC, this means an LLM can translate an analyst's question into a SIEM query, summarise thousands of log lines, explain an alert in plain language and draft an incident report. Grounding the model in the organisation's own data through retrieval-augmented generation makes its answers more factual and verifiable (Lewis et al., 2020). Such a model, embedded in the analyst's workflow, is often called a security copilot.

A copilot is only useful if analysts can trust it. Machine learning detections already suffer from opacity (Zhang et al., 2022), and an LLM that summarises those detections without showing evidence would add a second black box. Copilots in enterprise settings also touch sensitive data. In SAP landscapes, they must respect trust boundaries and identity context (Asuri, 2021) as well as privacy and retention obligations (Asuri, 2022a).

This article proposes an integration architecture that makes copilots explainable, grounded and governed. Section 2 reviews the literature. Section 3 presents the architecture. Section 4 describes the investigation workflow. Section 5 covers security, privacy and governance of the copilot itself. Section 6 discusses challenges and future research, and Section 7 concludes.

2. Literature review

The literature behind this article covers four areas: SIEM and SOC operations, explainable detection, large language models, and the enterprise SAP context.

2.1 SIEM and SOC operations

González-Granadillo et al. (2021) review the most widely used SIEM systems, their functionality and their use in critical infrastructures. They describe how SIEMs have grown into comprehensive platforms that provide broad visibility of risk and support faster incident response, and they observe a convergence between SIEMs and big data analytics. Among the enhancements they propose for future SIEMs are tighter integration with security orchestration, automation and response, stronger AI and machine learning capabilities, and more user-friendly interfaces for decision-making. A natural-language copilot directly addresses the last of these.

Vielberth et al. (2020) conducted a systematic study of SOC research and derived the main building blocks of a SOC. Their central finding is that research has concentrated on people and technology but neglected the processes, particularly non-technical ones, that connect them. A copilot is, in effect, a process technology: it changes how analysts move from data to decision. This makes Vielberth et al.'s finding directly relevant to how copilots should be designed and evaluated.

2.2 Explainable detection

Mahbooba et al. (2021) apply explainable AI to trust management in intrusion detection systems. They use decision tree models, whose rules resemble human decision-making, and extract those rules from the widely used KDD benchmark dataset. They compare the approach with other state-of-the-art algorithms and show that interpretable models let experts understand why an intrusion was flagged. This strengthens trust in detection.

Zhang et al. (2022) survey XAI applications in cybersecurity, covering intrusion detection, malware detection and spam filtering. They note that AI-based defences outperform rule- and signature-based approaches but are mostly opaque. This opacity reduces human confidence as attacks become more varied and complex. They argue for XAI as the basis for cyber defences that humans can understand and manage.

2.3 Large language models

Brown et al. (2020) introduced GPT-3, a 175-billion-parameter language model, and showed that scaling language models substantially improves few-shot performance. Tasks can be specified entirely through text, without fine-tuning. This general-purpose capability is what allows one copilot to handle query writing, summarisation, explanation and reporting.

Lewis et al. (2020) introduced retrieval-augmented generation, which combines a language model with a retriever over an external index. RAG produced more specific and factual output than a generation-only model, and its knowledge could be updated by replacing the index. For a SIEM copilot, RAG provides the mechanism for grounding answers in logs, alerts, runbooks and asset data, and for citing that evidence.

2.4 The enterprise SAP context

Asuri (2021) addresses the security of cross-platform SAP ecosystems through trust-boundary enforcement, identity context preservation, reduction of the SAP Cloud Connector's attack surface, and enterprise API security governance. These controls define important detection use cases for a copilot, such as calls that lose identity context or cross trust boundaries unexpectedly. They also define how the copilot itself must access SAP data.

Asuri (2022a) sets out an approach to enterprise data privacy governance across SAP ecosystems based on context-aware masking, lifecycle retention enforcement and cross-jurisdictional compliance. SIEM data often contains personal and business-sensitive information, so the same controls must govern what the copilot retrieves, stores and displays.

Asuri (2022b) proposes an intelligent enterprise framework that combines SAP S/4HANA and SAP Analytics Cloud for real-time decision intelligence. For this article, its relevance lies in connecting security findings to business context. A copilot can draw on real-time business data to explain which processes an incident affects and how urgently to respond.

Together, these studies show that the components of an explainable copilot exist. What remains is to integrate them with the SIEM in a governed way.

3. Copilot–SIEM integration architecture

The architecture places the copilot beside the SIEM rather than inside it. The SIEM remains the system of record for events and alerts, and the copilot reads from it, reasons about it and proposes actions through controlled interfaces. Five components make this work.

1. **SIEM data and detection layer.** The SIEM ingests events from across the enterprise, including SAP S/4HANA security logs, SAP BTP and Cloud Connector activity, API gateways, identity providers and network and endpoint tools. Correlation rules and machine learning models generate alerts (González-Granadillo et al., 2021). Before storage, sensitive fields are masked according to context, and retention rules are applied (Asuri, 2022a).
2. **Explanation service.** Every model-generated alert is accompanied by a structured explanation of the factors that drove it. Interpretable models, such as decision trees, are used where their accuracy is sufficient, because their rules can be shown directly (Mahbooba et al., 2021). For complex models, feature-attribution methods are applied (Zhang et al., 2022). The explanation is stored with the alert so that the copilot and the analyst see the same evidence.
3. **Retrieval layer.** A retrieval index holds runbooks, detection rule documentation, asset and identity inventories, SAP trust-boundary and API configurations (Asuri, 2021), threat intelligence and closed incident reports. When the copilot answers a question, it retrieves the relevant documents and SIEM results and cites them (Lewis et al., 2020).
4. **Copilot reasoning layer.** The LLM (Brown et al., 2020) receives the analyst's request or a new alert, plans the steps needed, and calls tools. It translates questions into SIEM queries, runs them, inspects results, retrieves documents and iterates until it can answer. It then returns a summary, a timeline, an assessment and recommended actions, each linked to the queries and records that support it.
5. **Action gateway.** Any action the copilot proposes, such as disabling an account, revoking an API key or blocking a Cloud Connector route, passes through a gateway. The gateway checks the action against policy, the copilot's permissions and the required approval level before handing it to an orchestration playbook. Business context from real-time enterprise analytics informs the approval decision, for example whether the affected system supports a critical process that day (Asuri, 2022b).

Two principles govern the design. **Evidence first**: the copilot never states a conclusion it cannot link to SIEM data or a retrieved document. **The SIEM stays authoritative**: the copilot may read, query and propose, but the record of what happened always lives in the SIEM.

4. Explainable detection and automated investigation workflow

The workflow runs in five steps. An example from an SAP landscape illustrates each.

1. **Explained alert.** The SIEM raises an alert on an API client that is calling an on-premises S/4HANA system through the Cloud Connector. The explanation shows three drivers: the client was registered two days ago, its calls carry a technical user rather than the originating user's identity, and it is reading customer master data at an unusual rate.

2. **Automated triage.** The copilot retrieves the trust-boundary and API policies for that connection (Asuri, 2021) and confirms that propagating the user's identity is mandatory. It classifies the alert as a likely policy violation with possible data exfiltration and raises its priority.
3. **Investigation.** The copilot writes and runs SIEM queries for the client's full activity, the account that registered it and related network flows. It finds that the registering account logged in from an unfamiliar location shortly before. It builds a timeline and links each entry to the underlying events.
4. **Assessment and recommendation.** The copilot summarises the incident in plain language. It estimates how many customer records were accessed, flags the privacy implications and the applicable retention and notification rules (Asuri, 2022a), and recommends revoking the API client and suspending the registering account. Because customer-facing processes depend on the S/4HANA system (Asuri, 2022b), it marks the suspension for analyst approval rather than automatic execution.
5. **Analyst decision and learning.** The analyst reviews the explanation, timeline and evidence, approves both actions and asks the copilot to draft the incident report. The analyst's confirmation is fed back to improve the detection model and the copilot's retrieval content.

In this example, an investigation that might take an analyst several hours of querying and reading is reduced to a few minutes of review. More importantly, every conclusion can be checked against the SIEM, so the analyst remains accountable and in control.

Analysts can also use the copilot conversationally outside the alert flow. They can ask, for example, "Which API clients lost identity context this week?" or "Summarise all failed logins to finance systems since Monday." The copilot translates each question into SIEM queries and returns the answer with its sources.

5. Security, privacy and governance of the copilot

The copilot is itself a privileged system and must be secured like one. It should run under a dedicated, least-privilege identity and preserve the identity context of the analyst who invoked it (Asuri, 2021). Sensitive data should be masked before it reaches prompts or the retrieval index, with retention rules applied to conversation logs and embeddings (Asuri, 2022a). Inputs drawn from logs must be treated as untrusted to resist prompt injection. Every answer, query and proposed action should be logged for audit.

6. Challenges and future research

Remaining challenges include LLM errors, incorrect query generation, inference cost at SIEM scale and analyst over-reliance. Future work should measure investigation time and accuracy with and without copilots in live SOCs, and evaluate how explanation quality affects analyst trust.

7. Conclusion

LLM security copilots can turn the SIEM from a data store that analysts must interrogate into a partner that explains, investigates and recommends. Their value depends on explainable detections, answers grounded in SIEM evidence, and strict identity, privacy and approval controls, especially in SAP landscapes. Built this way, copilots speed up investigations while keeping analysts accountable.

References

Asuri, R. (2021). Securing cross-platform SAP ecosystems through trust-boundary enforcement, identity context preservation, cloud connector attack-surface reduction, and enterprise API security governance. *SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology, 13*(2), 205–223.

Asuri, R. (2022a). Establishing enterprise data privacy governance across SAP ecosystems through context-aware masking, lifecycle retention enforcement, and cross-jurisdictional regulatory compliance. *International Journal of Humanities and Information Technology, 4*(01-03), 352–375.

Asuri, R. (2022b). Leveraging SAP S/4HANA and SAP Analytics Cloud to build an intelligent enterprise framework for real time decision intelligence digital innovation and sustainable business value creation. *International Journal of Science, Technology & Society, 6*(1), 85–100. https://doi.org/10.18091/ijsts.v6i1-113

Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., ... Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems, 33*, 1877–1901. https://arxiv.org/abs/2005.14165

González-Granadillo, G., González-Zarzosa, S., & Diaz, R. (2021). Security information and event management (SIEM): Analysis, trends, and usage in critical infrastructures. *Sensors, 21*(14), 4759. https://doi.org/10.3390/s21144759

Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W., Rocktäschel, T., Riedel, S., & Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. *Advances in Neural Information Processing Systems, 33*. https://arxiv.org/abs/2005.11401

Mahbooba, B., Timilsina, M., Sahal, R., & Serrano, M. (2021). Explainable artificial intelligence (XAI) to enhance trust management in intrusion detection systems using decision tree model. *Complexity, 2021*, 6634811. https://doi.org/10.1155/2021/6634811

Vielberth, M., Böhm, F., Fichtinger, I., & Pernul, G. (2020). Security operations center: A systematic study and open challenges. *IEEE Access, 8*, 227756–227779. https://doi.org/10.1109/ACCESS.2020.3045514

Zhang, Z., Al Hamadi, H., Damiani, E., Yeun, C. Y., & Taher, F. (2022). Explainable artificial intelligence applications in cyber security: State-of-the-art in research. *IEEE Access, 10*, 93104–93139. https://doi.org/10.1109/ACCESS.2022.3204051