SPONSORED BY

From Alerts to Insight: How Agentic AI Elevates Security Teams

A SANS First Look

Written by **Cristian-Mihai Vidu** | May 2026

Introduction

What would your SOC look like if your team could spend at least 50% of their time on threat hunting, detection engineering, and continuous skill development?

Although it might sound like a luxury, in the current threat landscape, it is a strategic necessity to make the organization safer. Every hour spent on repetitive, manual triage to clear the alert queue can be an extra hour for the attacker to burrow deeper into the organization. With human analysis as the bottleneck, attackers can exploit the gap and increase their dwell time while defenders are still busy manually gathering data to understand if the alert they are looking at is a true or false positive.

Adding detail takes time. Reducing the number of alerts that are investigated can leave an organization vulnerable. The cycle of repetitive, manual tasks—often referred to as toil—forces analysts to choose between investigation depth and total coverage. The result is an ever-increasing cognitive load and widespread burnout that increases an organization's risk.

The AI Trust Gap

Generative AI (GenAI) can help, but have you ever asked your favorite large language model (LLM) for a reference on a topic, only to find that it hallucinated the source and fabricated the evidence? As practitioners have learned over time, LLMs are not magic switches you can just flip to fix everything. They bring their own training biases and will sometimes drift or completely hallucinate their answers.

Whether an analyst initially welcomed GenAI with open arms or remained a skeptic, the reality is that many have lost that inherent trust. In a high-stakes security environment, a "black box" verdict is not enough. To build that trust back up, analysts require hard evidence and a transparent, logical reasoning process that can be followed and understood. See Figure 1.

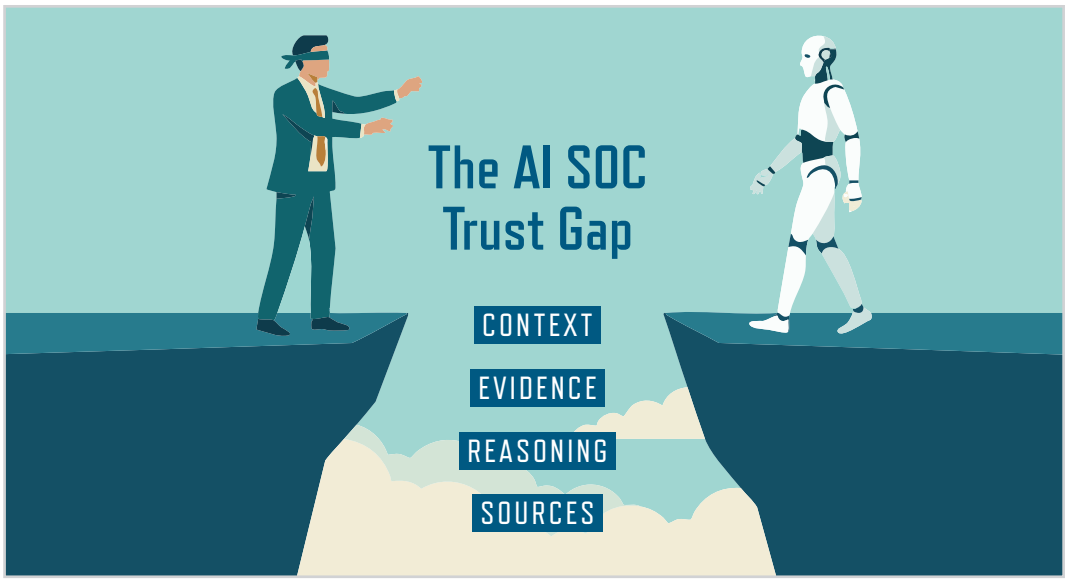

Figure 1. The AI SOC Trust Gap

Current AI tools fall into the SOC for two specific reasons, as highlighted in the SANS 2025 AI Survey:[1]

- **Limited contextual experience—**Generic models lack the specific cybersecurity analysis knowledge and the "organizational DNA" needed. They don't understand the nuances of each type of investigation and fail to account for the specific environment, tech stack, or historical norms.
- **The evidence gap**—Without the capability to show the data behind a verdict, AI becomes another source of noise that requires more manual verification, adding to the very toil it was meant to solve.

A First Look at Embed Security

SANS looked at Embed Security's solution, which takes an innovative approach to agentic AI triage and security. Embed's platform focuses on building trust in AI by embedding domain expertise from seasoned analysts into the agentic flow, an orchestrated, automatic investigation workflow driven by AI agents. Embed's core stated purpose is to move the SOC away from reactive firefighting and toward a proactive defense posture.

By speeding up triage, increasing accuracy, and eliminating repetitive tasks, Embed Security aims to achieve an alert queue of zero. This isn't simply a speed benefit; it's about reclaiming the analyst's day, allowing them to shift their focus from manual triage, evidence collection, and context generation to high value (and fun) activities, such as proactive threat hunting and detection engineering.

[1] "SANS 2025 AI Survey: Measuring AI's Impact on Security Three Years Later," September 2025, www.sans.org/white-papers/sans-2025-ai-survey-measuring-ai-impact-security-three-years-later

Building Trust in Agentic AI

Trust is generally built over time, but always through validation. For an analyst, it depends on the ability to see the reasoning process, inspect the logical flow, and verify the evidence. This comes down to transparency and reliability.

When an alert comes into the Embed platform, it initiates a transparent, nonlinear chain of actions. First, the artifacts are extracted, and then the platform synthesizes an initial tailored investigation path leveraging iSteps,™ which are dynamically selected based on organizational context and real-time findings. As it progresses through the investigation, the agentic flow can pivot and adapt its strategy based on the leads it uncovers. The final results are assembled into a conclusion with each step and decision documented. It's exactly what is expected from an experienced analyst: leverage the experience to consider the next steps, gather evidence, document the process, and draw a conclusion.

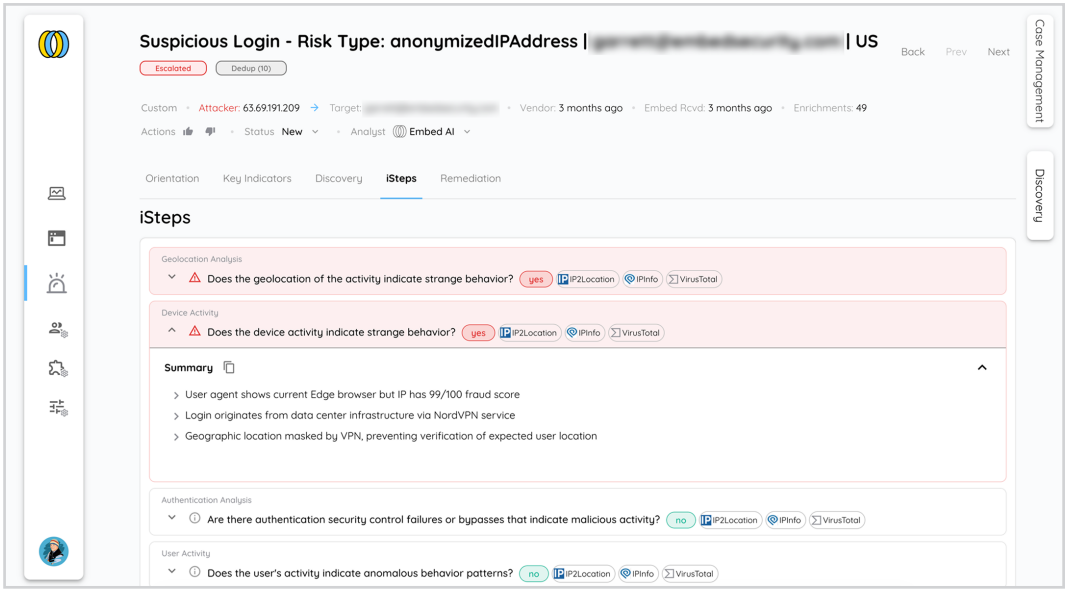

Figure 2. iSteps Questions from a Suspicious Login

The iSteps provide the structured methodology for each investigation, functioning as reasoning units. See Figure 2. Each iStep asks a specific question, such as, "Is the source location consistent with the user's normal login pattern?" and analyzes the initial telemetry to develop sub-questions (e.g., "What are the user's historically normal login locations?" or "Are there any known travel or remote-work circumstances that would explain the anomaly?"). It then proceeds to answer these sub-questions by gathering additional data from relevant sources. Because this is an iterative process, the answer to one sub-question can provide the evidence to trigger a different iStep. Because iSteps are transparently documented, the analyst can choose to look at the evidence and documented reasoning trail to see whether they agree with the conclusion. In cases where the analyst does not agree with the investigation, the feedback is documented and goes into NoiseIQ™ to refine future agentic AI decisions.

NoiseIQ is an organizational knowledge base comprising feedback from analysts, previously investigated case histories, and information about the environment. This ensures the agentic flow is constantly learning what is normal for the organization to improve triage accuracy over time.

A key architectural distinction noted was the deliberate separation between tasks handled by an AI agent through an LLM and those requiring the use of a deterministic tool. LLMs are great at reasoning and synthesis, but they struggle with tasks that require precision, such as complex data parsing or statistical anomaly detection. The choice of using deterministic tools, such as purpose-built scripts and code for analysis and classical machine learning, anchors the flexibility of GenAI with the repeatable, verifiable logic of traditional engineering.

Balancing Alerts

Detection engineering has always been a compromise between sensitivity and specificity. To catch sophisticated attackers, security teams want high sensitivity, a broad net that catches even the lowest signal of attacker activity. However, this comes at a price: a flood of false positives that analysts must work through, resulting in alert fatigue.

At the other end of the spectrum is high specificity, tuning the tools and detections to look for only the exact, known-bad patterns of attacker behavior. This reduces the noise, but it introduces the more dangerous false negatives: attackers that remain undiscovered in the environment for months, turning a minor intrusion into a major, high-impact breach.

Agentic AI has the potential to change the equation. The main problem with the high number of false positives is that it is coupled with human analysis. By decoupling the two and allowing a trusted agentic AI system to perform all initial investigation and context building, you avoid the need to sacrifice one for the other. Once the analyst understands and trusts its reasoning, agentic AI can perform a deep-dive triage for every alert by dismissing false positives with documented reasoning and escalating only truly suspicious events to the human analyst—augmented with additional investigative context. But trust is earned, not given, and this is where the Embed platform delivers on the promise of verifiable agentic triage.

Bottom Line

The transition from drowning in alerts to an alert queue of zero means more than just having analysts go through the alerts faster or simply sending all alerts to AI. It requires a trusted decision layer that removes the burden of false positives while providing the transparent evidence needed to confirm true threats.

This allows analysts to focus on high-value activities that truly move the needle in organizational security. Instead of being trapped in a cycle of evidence collection and manual triage, analysts can shift attention to proactive threat hunting, creative detection engineering, and strategic defense work that often gets pushed to the side.

By shifting away from the “black box” approach of generic LLMs toward a structured agentic methodology, Embed Security offers a blueprint for the future that infuses the investigative workflow with the domain expertise and logical rigor of a seasoned human defender. This transparency serves as a strategic benefit, facilitating knowledge transfer so junior analysts can learn, and ultimately elevates the entire defensive posture of the SOC.