Eight Ways Agentic AI Will Reshape the SOC

Security operations centers (SOCs) face intensifying pressure from high alert volumes, a persistent talent shortage, and adversaries who move faster than ever. Agentic AI is emerging as a practical way to augment SOC performance by accelerating triage, investigations, and governed automation, all while keeping humans in control.

This guide provides a forward-looking evaluation framework for your 2026 investments in agentic AI. It identifies eight critical capabilities that separate durable value from hype:

- **Supervised autonomy:** Start with guardrails and explicit human-in-the-loop controls.
- **Policy-driven governance:** Enforce rules with centralized, role-based permissions.
- **Advanced reasoning:** Use multi-step planning for complex investigations.
- **Domain specialization:** Apply rich security context to reduce noise.
- **Deep integration:** Orchestrate actions across the entire security stack.
- **Full transparency:** Deliver clear explanations and complete auditability.
- **Secure by design:** Build for resilience against adversarial manipulation.
- **Continuous learning:** Adapt intelligently to new threats and environments.

Applied together, these criteria help you reduce analyst fatigue, compress response times, and improve consistency without ceding control. The guide concludes with next steps for piloting agentic AI, measuring its impact, and scaling responsibly.

Introduction

SOCs are the core of enterprise cyber defense, but they are under immense strain. Analyst burnout is a critical issue, with 84% of cyber security professionals reporting it and more than half quitting as a result. Teams are so overwhelmed that they ignore 62% of alerts, creating opportunities for attackers to succeed.

Adversaries use generative AI to create malicious code faster, widening the global cybersecurity skills gap to more than four million professionals. This leaves 67% of organizations with a moderate to critical talent shortage while attackers move too fast for human-led investigation and response alone. CISOs are looking for force multipliers, and they see promise in the advisory, investigation, and response capabilities of agentic AI.

In 2025, Gartner® identified "AI SOC agents" and "cybersecurity AI assistants" as technologies at the "Innovation Trigger" stage of its Security Operations Hype Cycle™, marking a new era of security automation. These systems are designed to perceive, reason, and act in complex SOC workflows, moving beyond static playbooks. At the same time, the Gartner Hype Cycle for AI and Cybersecurity, 2025, highlights the need for robust governance, AI literacy, and security controls, warning that more than half of successful attacks against AI agents through 2029 will target access control weaknesses.

For CISOs, this convergence of opportunity and risk creates a dual mandate: Improve SOC efficiency with next-generation AI while ensuring auditability, trust, and alignment with enterprise risk frameworks. This guide defines the eight capabilities you should demand from any agentic AI solution.

1. Start With Supervised Autonomy

As AI agents become operational in your SOC, you must treat them as privileged users. These systems have wide-reaching access to logs, identity data, and response tools. A misconfigured or compromised agent can cause immediate, large-scale damage. To counter emerging attack vectors like prompt injection and data poisoning, you should pair identity and access controls with AI-specific runtime defenses and governance from day one.

The safest path forward is to deploy AI agents in a phased, supervised manner. Begin with advisory modes where the AI gathers data, enriches it, and triages alerts, but stops short of taking action without human review. Introduce human-in-the-loop approvals for high-risk actions such as quarantining endpoints or revoking credentials. As monitoring confirms the AI behaves as expected, you can gradually expand its autonomy.

Guardrails should include clear policies that define permitted actions, confidence thresholds that trigger human escalation, and continuous monitoring to detect anomalous behavior. Your SOC must maintain audit logs of every AI-driven action.

By combining strong oversight with least-privilege access, you can unlock AI's efficiency benefits while reducing risk.

2. Align AI With Governance and Risk

Effective guardrails depend on the policies and governance structures that define them. Ensuring AI behavior aligns with your risk appetite, compliance needs, and security strategy requires formal governance. Gartner highlights dedicated AI governance platforms and AI Trust, Risk, and Security Management (TriSM) practices as emerging ways to centralize oversight for AI applications.

Effective governance starts by codifying rules into formal policies. These policies should specify approvals for supervised automation, define confidence thresholds, and set escalation paths. Baselining your current state and setting success criteria before pilots is the best way to justify costs and verify that the benefits are real.

Because the SOC does not operate in a silo, your Legal, Compliance, Privacy, and Risk Management teams must help determine how AI is allowed to act. This cross-functional approach ensures AI autonomy does not conflict with data protection rules or corporate risk tolerance. Finally, governance must be iterative. As AI systems mature, policies should evolve to grant more autonomy where it is safe. This disciplined approach keeps you firmly in control of risk while unlocking the efficiency of automation.

3. Move From Alerts to Entire Campaigns

SOC teams spend too much time chasing individual alerts instead of solving entire attack campaigns. Traditional automation, which follows pre-programmed playbooks, breaks when attackers change tactics. You need AI that can reason about new events, draw logical conclusions, and plan investigative steps dynamically, thinking more like an experienced analyst than a script.

In the SOC, advanced reasoning means the AI can connect clues across data sources, spot relationships between seemingly unrelated events, and hypothesize an attacker's objective. For example, rather than just flagging a credential misuse alert, a reasoning-capable AI could connect that alert to suspicious lateral movement and present a narrative that explains the likely breach path. This reduces the cognitive load on your analysts, allowing them to validate and respond faster.

Planning is the second half of this capability. An agent should not simply react to alerts but recommend or orchestrate a sequence of next steps to contain a threat. This might include querying more data to confirm scope or suggesting isolation of a compromised host. Crucially, this planning must still be governed by your organizational policy, with analysts reviewing and approving high-impact actions until trust in the AI is established.

4. Prioritize Threats With Security-Specific AI

A high volume of false positives and irrelevant alerts drives SOC inefficiency. Many anomaly detection systems fail because they cannot separate malicious behavior from benign but unusual activity. You need AI that combines deep cybersecurity knowledge with an understanding of your unique environment to produce relevant, actionable, and trustworthy output.

Agentic AI must be purpose-built for security. It should recognize common attack techniques, correlate events using frameworks such as MITRE ATT&CK®, and present findings in a way that matches how your analysts think. An effective AI agent also considers business context, like the criticality of assets, user roles, and historical activity. By applying this context, the AI can suppress noise and prioritize alerts that represent real risk.

Ultimately, the application of context and domain expertise should translate into fewer false positives, higher-fidelity detections, and faster investigations. When AI delivers measurable improvements, you can justify broader adoption and continued investment.

5. Unify the SOC Workflow

Data is not the problem for most SOCs; it's that data is scattered across dozens of tools. Your analysts are slowed by the manual effort required to stitch it all together. Agentic AI should act as a connective layer that automates this cross-tool work, surfaces relevant findings, and orchestrates the next best action across your security ecosystem.

Rather than duplicating SIEM functions, agentic AI should pull together only the data relevant to an investigation and present it in a single, narrative view. It can then recommend a sequence of actions or automatically execute low-risk steps, such as gathering additional evidence or enriching indicators. The value is in reducing friction and decision latency, not replacing existing telemetry pipelines.

This orchestration should give you control over AI autonomy. You can start with a human-in-the-loop model, progress to a human-on-the-loop model for supervision, and eventually move to a human-out-of-the-loop model for routine tasks. This progression leads to a bigger question: Could this create a fully autonomous, self-healing security ecosystem? Agentic AI is a foundational step in that direction.

6. Build Trust Through Transparency

As you adopt AI in the SOC, trust becomes a prerequisite. You and your analysts must understand why an agent made a particular decision. Without this visibility, teams will hesitate to act on AI-driven recommendations, and regulators may block the use of AI for high-impact security actions.

Agentic AI must make its reasoning clear. Each recommendation should come with supporting evidence, a confidence score, and a list of alternatives considered. This transparency helps build confidence, which is essential for moving from supervised to more autonomous modes of operation.

Auditability is equally important. Every action taken by the AI must be logged with enough detail for after-action reviews and compliance reporting. These logs enable you to measure performance, identify errors, and refine policies. By combining transparency with a robust audit trail, you can create an environment where AI outputs are trusted, measurable, and defensible.

7. Secure the Agent

Even with strong guardrails and governance, the AI agent itself is part of your attack surface. If compromised, it could be manipulated to exfiltrate data or trigger unsafe actions. You should expect controls that monitor for prompt abuse, jailbreaks, and data leakage at runtime.

Securing the AI agent begins with protecting its inputs. Providers should implement strict input validation, filtering, and adversarial testing to ensure the system cannot be tricked. Runtime defenses such as rate limiting and context isolation can reduce the impact of a successful attempt.

Privacy is also a core requirement. AI agents must handle sensitive log data and identity information according to data minimization principle, strong encryption, and regional data residency rules. Finally, the AI system must be resilient. It should be monitored for availability and quality, with fallbacks in case performance degrades.

8. Ensure AI Adapts to New Threats

Threats, techniques, and your business operations evolve daily. Your SOC detections and playbooks must evolve, too. A static AI agent risks becoming outdated. The best systems continuously improve as they are exposed to new data, new attacks, and direct analyst feedback.

Continuous learning involves more than annual model retraining. It requires regular updates that reflect new attacker tactics, techniques, and procedures (TTPs) and incorporate emerging threat intelligence. Each time an analyst approves or rejects a recommendation, that feedback can be used to refine the model, making it more attuned to your organization's environment and risk tolerance.

This adaptability should also extend to changes in your infrastructure or user behavior. The AI should be able to reason over new data sources without extensive re-engineering. This approach to learning must be safe and governed, with regular validation to confirm that each update improves performance without introducing new bias.

Conclusion

The modern SOC is pushed to its breaking point by analyst burnout, a persistent skills gap, and adversaries who operate at machine speed. Agentic AI offers a path forward, but only if you adopt it strategically.

This guide outlined eight critical capabilities for evaluating agentic AI: supervised autonomy, policy alignment, advanced reasoning, domain expertise, workflow integration, transparency, built-in security, and continuous learning. Together, they form a framework for a safe, iterative journey toward an AI-augmented SOC.

The vision of self-healing security operations is both exciting and challenging. It promises faster response, fewer missed threats, and a more empowered security team. Now is the time for to evaluate your readiness for agentic AI. Establish governance, measure your current SOC performance, and begin experimenting with trusted pilots. By investing in these capabilities today, you will be best positioned to move from reactive firefighting to proactive defense, with AI serving as a true force multiplier for human expertise.

See how Exabeam delivers agentic AI for the SOC. Watch our webcast: Agentic AI on the Cybersecurity Battlefield.

About Exabeam

Exabeam is a leader in intelligence and automation that powers security operations for the world's smartest companies. As a global cybersecurity innovator, Exabeam provides industry-proven, security-focused, and flexible solutions for faster, more accurate threat detection, investigation, and response (TDIR).

exabeam

Learn more at www.exabeam.com →