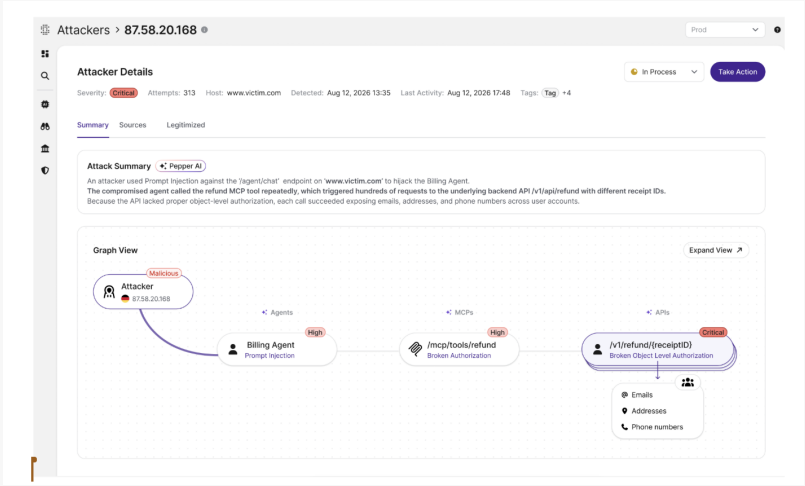
Salt Security Extends Its Agentic Security Platform with Native AI Detection and Response

New LLM runtime protection connects prompt injection to downstream MCP and API attacks, while helping enterprises close gaps in existing AI protection



Salt Security Extends Its Agentic Security Platform with Native AI Detection and Response

Salt Security Logo

PALO ALTO, CA, UNITED STATES, September 22, 2026 / EINPresswire.com/ -- Salt Security, the leader in agentic and API security, today announced native AI Detection and Response (AI-DR) capabilities within the Salt Agentic Security Platform.

The new capabilities provide real-time LLM protection against direct and indirect prompt injection, jailbreak attempts, unsafe model behavior, and other runtime threats. Prompt injection is the top-ranked risk in the OWASP Top 10 for LLM Applications 2026, and it takes two forms, direct and indirect, both of which the new capabilities detect in real time. Salt also brings the AI guardrails and gateways enterprises already use into one security view, helping teams apply consistent protection across models, agents, MCP servers, tools, and downstream APIs.

Protecting the full agentic path

AI agents connect models to tools and business systems. An attack that starts with a malicious prompt can expose those connections and lead to unauthorized access or actions downstream. This is not hypothetical: in Salt Security's 2H 2026 State of Agentic AI and API Security survey, nearly half of organizations (49.8%) confirmed or suspected that an AI agent took an action they did not intend, expect, or authorize in the past year, and only 12.5% can consistently trace an

agent's full path from the initial prompt through the MCP servers and APIs it reaches. Security teams need to see how an attack moves across these layers to understand what is at risk.

AI-DR is built into Salt Agentic Detection and Response (AG-DR), extending its runtime protection to the LLM layer. The Salt Agentic Security Graph connects agents to their models, Model Context Protocol (MCP) servers, tools, and downstream APIs. With native AI-DR, teams can connect an attack against the model to subsequent MCP and API attacks in the same platform.

AG-DR and Agentic Security Posture Management (AG-SPM) form the Salt Agentic Security Platform. AG-SPM helps teams understand their agents and security gaps, while AG-DR detects and connects attacks across those agents and the systems they use.

Seeing how AI attacks reach business systems

For example, an attacker could trick a billing agent into revealing information about a refund tool on an MCP server and the API behind it. The attacker could then exploit that API - directly or through the agent to attempt unauthorized refunds. What began as a malicious prompt has become an attack on a payment system.

Salt shows these steps together, from the prompt injection through the MCP server to the API attack. Security teams can see how the attack started, what it targeted, and which business systems are at risk, even when the attacker moves beyond the agent to attack an API directly.

Unified Visibility Across AI Guardrails and Native Protection Where Coverage Is Missing

Enterprises use AI guardrails across cloud platforms, endpoint and SASE solutions, AI gateways, and managed AI services. Each control protects the interactions it covers, leaving teams to identify inconsistencies and gaps across environments.

Through AG-SPM and the Agentic Security Graph, Salt gives teams one view of these guardrail configurations and the agents they protect. Where protection is missing, including in homegrown agents running in Kubernetes, Salt can fill the gap with native AI-DR within AG-DR. Customers can establish more consistent protection while keeping their existing gateways and security products.

“An attack on an AI model can become an attack on the systems that run the business,” said Roey Eliyahu, co-founder and CEO of Salt Security. “Our native AI-DR protects LLM interactions and connects what happens at the model to the tools and APIs downstream. Teams can see the full attack, understand what is at risk, and fill protection gaps while keeping the guardrails they already use.”

Availability

Native AI-DR capabilities are available within Salt AG-DR as part of the Salt Agentic Security Platform. Visit salt.security/demo-request for more information.

About Salt Security

Salt Security is the leader in agentic and API security, protecting the world's most innovative enterprises from AI agent and API attacks. AI is shifting from chatbots that answer to agents that act, and Salt secures the entire agentic path, from AI-generated code to runtime, across models, MCP servers, tools, and downstream APIs. Founded in 2016, Salt Security is backed by Sequoia Capital, S Capital, Tenaya Capital, Salesforce Ventures, Advent International, and other leading investors. For more information, visit salt.security.

Media Contact
Dr. Karl Bateson - Karlb@salt.security

Karl T Bateson
DigitalPulse365
+1 617-306-6275
email us here