NEWS RELEASE

Cisco Unveils Frontier AI Red Teaming and Agent Observability to Secure the Agentic Enterprise

2026-11-09

AI Defense, Splunk, and Cisco IQ innovations give security and operations teams one loop to find risk, measure value, and prove trust.

News Summary:

- Cisco AI Defense's adaptive AI red teaming framework is adding Anthropic's Claude Mythos, harnessing advanced adversarial reasoning to simulate real-world attacks on AI agents.
- With Cisco's enhanced Tokenomics, organizations can optimize AI spend to defend what matters most by connecting agent cost directly with ROI in real time.
- Cisco IQ, which leverages decades of vulnerability and troubleshooting data, now correlates live threat detection data with potentially exposed network infrastructure at machine speed to accelerate remediation.
- Together, these innovations offer enterprises an always-on trust loop to secure, govern, and observe increasingly capable AI agents and complex enterprise AI deployments.

MELBOURNE, Australia, Nov. 9, 2026 /PRNewswire/ -- **CISCO LIVE APJC** -- Cisco (NASDAQ: CSCO) today announced it is adding Claude Mythos to AI Defense for red teaming AI applications and agents, empowering defenders to stress-test AI systems with frontier adversary reasoning. Coupled with new Splunk Agent Observability and Cisco IQ advancements, Cisco is preparing enterprises for a new era of trust that requires securing, observing, and governing agents at machine speed.

"You cannot scale what you cannot secure, govern, or observe," said Cisco President and Chief Product Officer Jeetu Patel. "By uniting frontier-level AI runtime protections with machine-speed agent quality controls, we are building a continuous trust loop: protecting agents from compromise while ensuring every token spent delivers measurable business value."

As AI agents proliferate and connect to complex tools and sensitive data, 81% of ITOps and engineering professionals cite increased cooperation between security and observability teams, according to a

global Splunk research report last month. Cisco's Splunk and AI Defense innovations equip security leaders to account for token spend and effective tool use while conducting tailored security testing against highly capable AI agents.

Since Cisco launched AI Defense in January 2025, it has helped hundreds of customers identify critical flaws in AI models, applications, and agents before they reach production. Its adaptive red teaming allows security teams to quickly build customized, multi-step attacks on target systems to uncover specific risks. Claude Mythos will augment this framework with advanced reasoning that plans tests that matter and translates the results into attack ideas for each system.

Features include:

- **Realistic Adversarial Ideation:** generates multi-step, contextual attack paths designed to help identify subtle vulnerabilities across complex architectures.
- **Secure, Contained Access:** By design, customers interact only through AI Defense's structured interface: Customers have no open-ended prompt access.
- **Audit Trails and Repeatable Results:** AI Defense records each testing action for audit, and results can be carried over to inform follow-up attack loops.

"Enterprises are increasingly building AI agents into their core operations and want to test those agents thoroughly before putting them into production," said Michael Moore, Cybersecurity Lead at Anthropic. "By bringing Claude Mythos to Cisco AI Defense, Cisco can help security teams red team their agents with advanced reasoning that plans tests and generates attack ideas for each system."

Cisco engineered AI Defense to use Claude Mythos with multi-layered safeguards:

- **Registered, Authorized Targets:** Operates exclusively against targets explicitly registered, owned, and authorized for testing by the enterprise customer; it cannot scan arbitrary internet infrastructure.
- **Named Operator Attribution:** Every red teaming engagement requires authentication and is cryptographically tied to a named enterprise operator.
- **Revocable Execution:** Engagements run with a fixed job and objectives defined by the human user, with time-bound sessions that can be halted on demand.
- **Dual-Layer Oversight:** Product-level controls in AI Defense operate alongside program-level safeguards and monitoring.

Closing the Loop: From Test Findings to Measurable Agent ROI

Frontier AI models can excel at complex security tasks such as crafting AI red teaming strategies. That capability cuts both ways. AI is lowering the barrier to entry for attackers just as many enterprises are working to balance security, governance, and ROI considerations across growing AI portfolios.

For organizations to confidently deploy AI agents at scale, they need unified, real-time visibility into AI agent cost, business performance, and quality. Cisco is bridging the trust gap with new Splunk Agent Observability enhancements:

- **Tokenomics:** cost governance, soon to be available in AWS Marketplace, that ties AI spend to output quality. Using usage data from Amazon Bedrock and Amazon CloudWatch, it pairs each agent's cost with evaluations that grade the quality and impact of its work, giving organizations a

single view of AI ROI. This extends the token-spend attribution launched at .conf26, which tracks costs across AI agents and coding tools such as Claude Code, Codex, and Cursor.

- **Eval-G1 Evaluation Models:** Splunk Agent Observability extends our support for turning agentic traces and expert feedback into custom small language models that execute evals and guardrails at up to 98% lower cost than a commercial LLM. These next generation Eval-G1 models deliver evaluation response times 2.5x faster than traditional LLMs and enable real-time agent control with high accuracy.
- **The Evaluation Tuner:** Teams can turn their agentic traces and subject matter expert feedback into fine-tuned evaluator models to enforce quality standards and steer agents at machine speed.

Adaptive red teaming with Claude Mythos is coming soon for select enterprises. Tokenomics will be available in private preview in December, with general availability in AWS Marketplace early 2027. The evaluation models and the evaluation tuner are generally available starting this month.

Connecting Infrastructure Exposure to Security Action with Cisco IQ and Cisco Data Fabric Powered by the Splunk Platform

While understanding vulnerability exposure is critical, defenders must also accelerate how they remediate risk across complex enterprise estates. Cisco IQ identifies critical exposures, and the powerful ecosystem of Cisco Partners and Professional Services turns insight into action. Partners help deliver remediation at scale and modernize infrastructure, and Cisco Professional Services bring deep expertise to put new Cisco capabilities into practice. Together, Cisco and its partners can help streamline operations and future-proof your IT infrastructure for the AI era.

As AI-accelerated threats compress the window between vulnerability disclosure and exploitation from weeks to hours, point-in-time audits leave organizations exposed. To help security and network operations teams act at machine speed, Cisco is extending real-time intelligence from Cisco IQ—which has close to 17,000 customers across more than 22 million devices—directly into enterprise security environments, open agentic protocols, and automated lifecycle workflows.

Key updates include:

- **Cisco Data Fabric Integration:** Ingests Cisco IQ vulnerability, lifecycle, and configuration data directly into Splunk Platform via open APIs, enabling SecOps teams and AI agents to correlate live threat detections with underlying network exposure in Splunk Enterprise Security.
- **Model Context Protocol (MCP) & ServiceNow Connectors:** Provides structured infrastructure context to supported AI agents and enterprise IT Service Management platforms, automatically translating exposure findings into asset-linked ServiceNow problem records to coordinate remediation without manual data reconciliation.
- **AI-Powered Resilient Infrastructure Playbook:** Surfaces critical lifecycle vulnerabilities, including personalized alerts for known defects and quantum-safe cryptographic readiness**,** allowing teams and certified partners to scope and prioritize modernization roadmaps.

"AI is shrinking the time organizations have to respond to threats, and knowing where you're exposed only matters if you can act on it," said Liz Centoni, Cisco EVP and Chief Customer Experience Officer. "Bringing Cisco IQ vulnerability insights into Cisco Data Fabric powered by the Splunk Platform gives analysts that context in the middle of an investigation, so they can decide what to do next. With our partners and Professional Services teams, we're helping customers move from identifying risk to actually reducing it."

To learn more about securing and observing the agentic enterprise, visit cisco.com/go/ai-defense.

Read more in our blogs:

- Cisco AI Defense to Add Anthropic's Claude Mythos for AI Red Teaming
- Splunk by Cisco Brings Agent ROI to AWS Marketplace and Real-time Evaluations to Everyone
- From Risk to Resilience: Close the Gap Between Knowing and Fixing
- Moving Fast and Staying in Control in the Agentic AI Era

About Cisco
Cisco (NASDAQ: CSCO) is the worldwide technology leader that is revolutionizing the way organizations connect and protect in the AI era. For more than 40 years, Cisco has securely connected the world. With its industry leading AI-powered solutions and services, Cisco enables its customers, partners and communities to unlock innovation, enhance productivity and strengthen digital resilience. With purpose at its core, Cisco remains committed to creating a more connected and inclusive future for all. Discover more on The Newsroom and follow us on X at @Cisco.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at http://www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word 'partner' does not imply a partnership relationship between Cisco and any other company.

Disclaimer: The timeline for the release of some products, features and integrations is subject to change given ongoing evolution in development and innovation.

View original content to download multimedia:https://www.prnewswire.com/news-releases/cisco-unveils-frontier-ai-red-teaming-and-agent-observability-to-secure-the-agentic-enterprise-302933708.html

SOURCE Cisco