Semgrep @ OWASP LA

Finding logic and auth vulns with AI

Code Security for Builders

Semgrep

22 APRIL 2026

Speaker

Traditional SAST catches issues like SQL injection and XSS. Some of the biggest bug bounty payouts come from logic flaws like IDOR, broken authorization, and workflow abuse because these are hard to find with traditional SAST techniques alone.

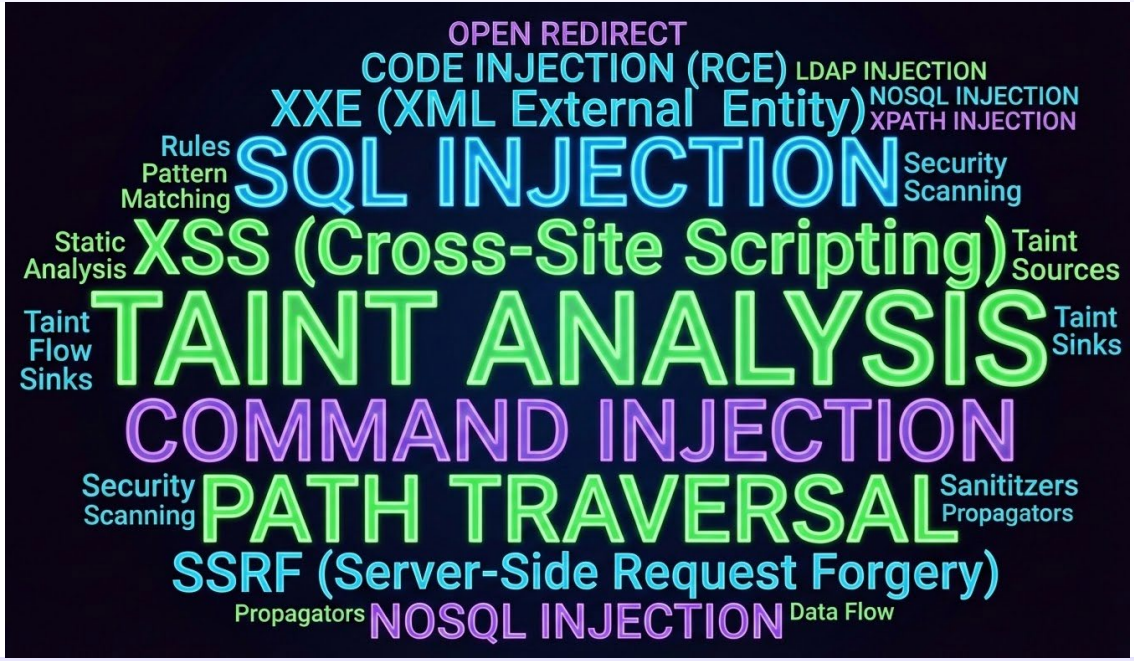
In this technical workshop, you'll see how Semgrep's AI-powered detection combines static analysis with LLM reasoning to uncover business logic vulnerabilities without custom rule writing.

Erik Buchanan

Head of AI Engineering ,
Semgrep

Traditional SAST Strengths

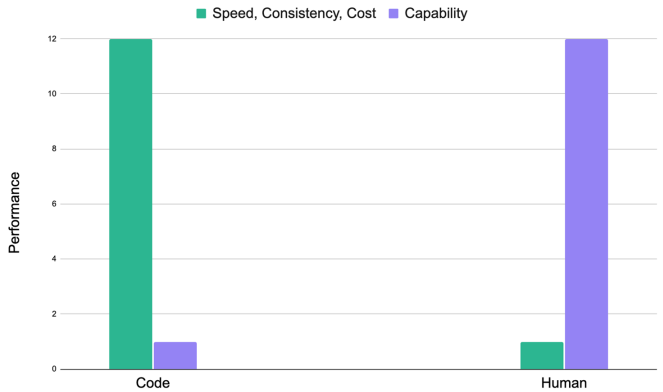
OWASP Top 10



Explosion of AI-generated code

AppSec must keep up with LLMs

2024



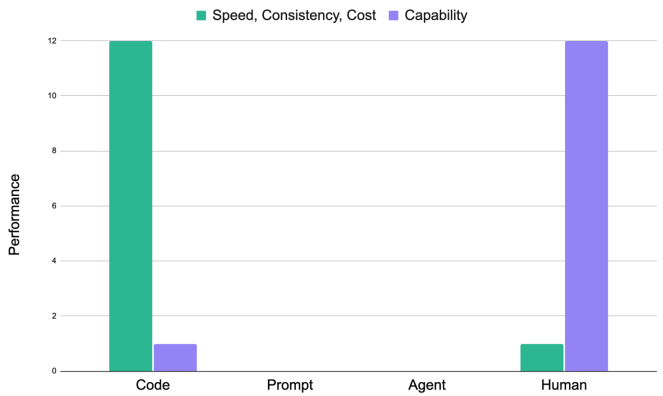
2026

Prior to AI - generated coding tools

Explosion of AI-generated code

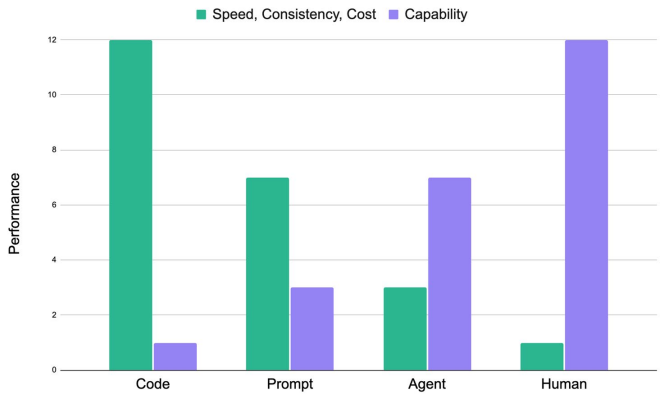
AppSec must keep up with LLMs

2024



Prior to AI -generated coding tools

2026



How modern software teams are working today

AI models detect logic vulns better, but perform poorly on taint style issues

AI models can help find important vulnerabilities that traditional SAST tools struggle with when they lack context beyond the source code and data flows.

Anthropic Claude Code (v1.0.32, Sonnet 4)

Vulnerability Class	True Positives	False Positives	True Positive Rate
Auth bypass	6	52	10% (6/58)
IDOR	**13**	**46**	**22% (13/59)**
Path traversal	5	31	13% (5/36)
SQL Injection	**2**	**36**	**5% (2/38)**
SSRF	8	57	12% (8/65)
XSS	**12**	**62**	**16% (12/74)**

https://semgrep.dev/blog/2025/finding-vulnerabilities-in-modern-web-apps-using-claude-code-and-openai-codex/

LLM context helps to identify vulnerabilities often missed

AI models help find vulnerability classes that SAST and data flow analysis sometimes struggles with.

IDOR / BOLA

Insecure Direct Object References (IDOR) and Broken Object Level Authorization (BOLA) are the #1 API security risk.

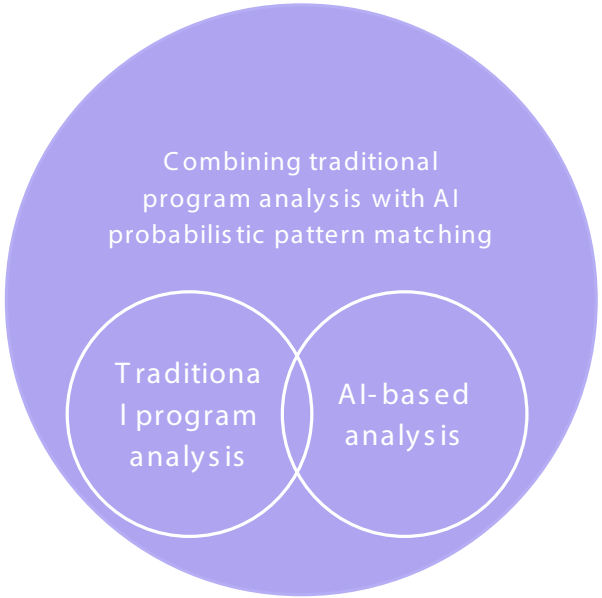
Business Logic Abuse

Application state transitions, decision-making, and other flaws used to bypass restrictions or cause disruption.

And many more

Crypto Misuse
TLS Verification Disabled
Weak Randomness
CSRF via Authentication Cookie
JWT Misconfigs
Missing Rate Limiting
OAuth Misconfiguration
Sensitive Data Logging
Missing Webhook Verification

Introducing Semgrep Multimodal detection



"With Semgrep, I trust that a critical finding will be relevant to us."

Minh Nghiem, Sr. Security Engineer

homebase

Real World Success Stories

"92% recall on known IDORs in app"
Social Media Giant

"70% true positives on IDOR findings so far"
FinTech Unicorn

"88% precision for IDOR, AutZ and Logic Issues"
Fortune 300 FinTech

Benefits of Semgrep Multimodal

Consistency
Repeatability and auditability of results.

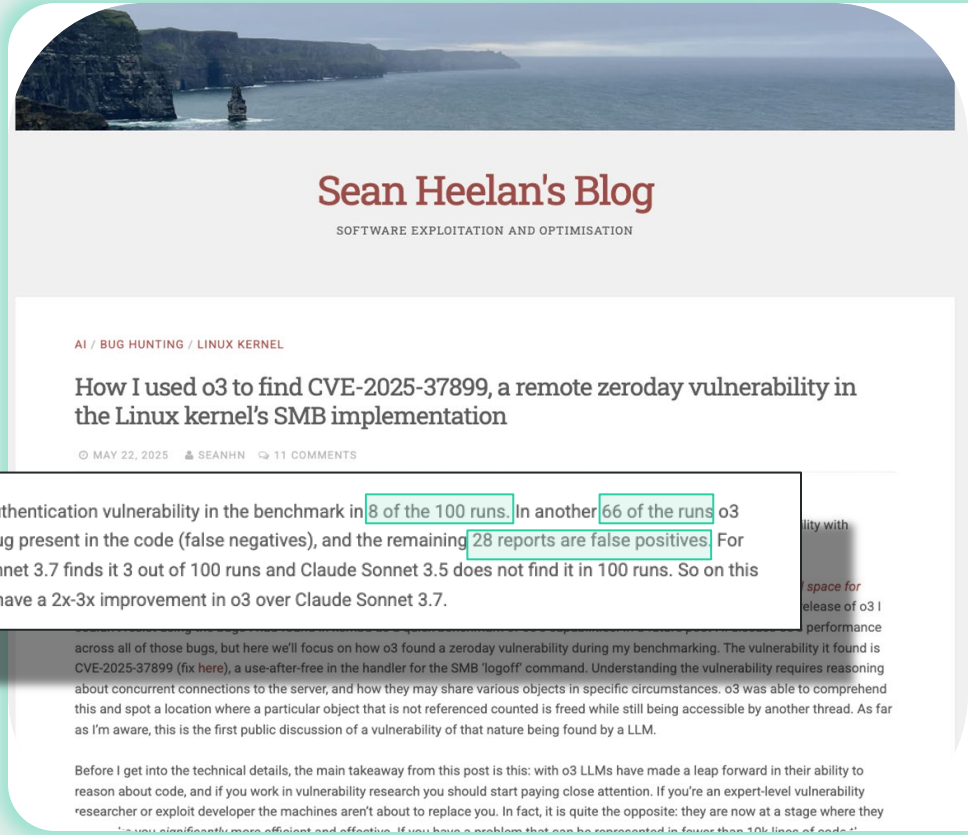
Precision & Recall
Higher rate of True positives while eliminating vulnerability classes with confidence.

Cost
Manage cost of usage of LLMs at scale.

Performance
Static analysis is fast and efficient.

Foundation models aren't consistent between runs

The Semgrep Security Research and others have found serious problems with consistency in AI-only approaches.



Sean Heelan's Blog

SOFTWARE EXPLOITATION AND OPTIMISATION

AI / BUG HUNTING / LINUX KERNEL

How I used o3 to find CVE-2025-37899, a remote zeroday vulnerability in the Linux kernel's SMB implementation

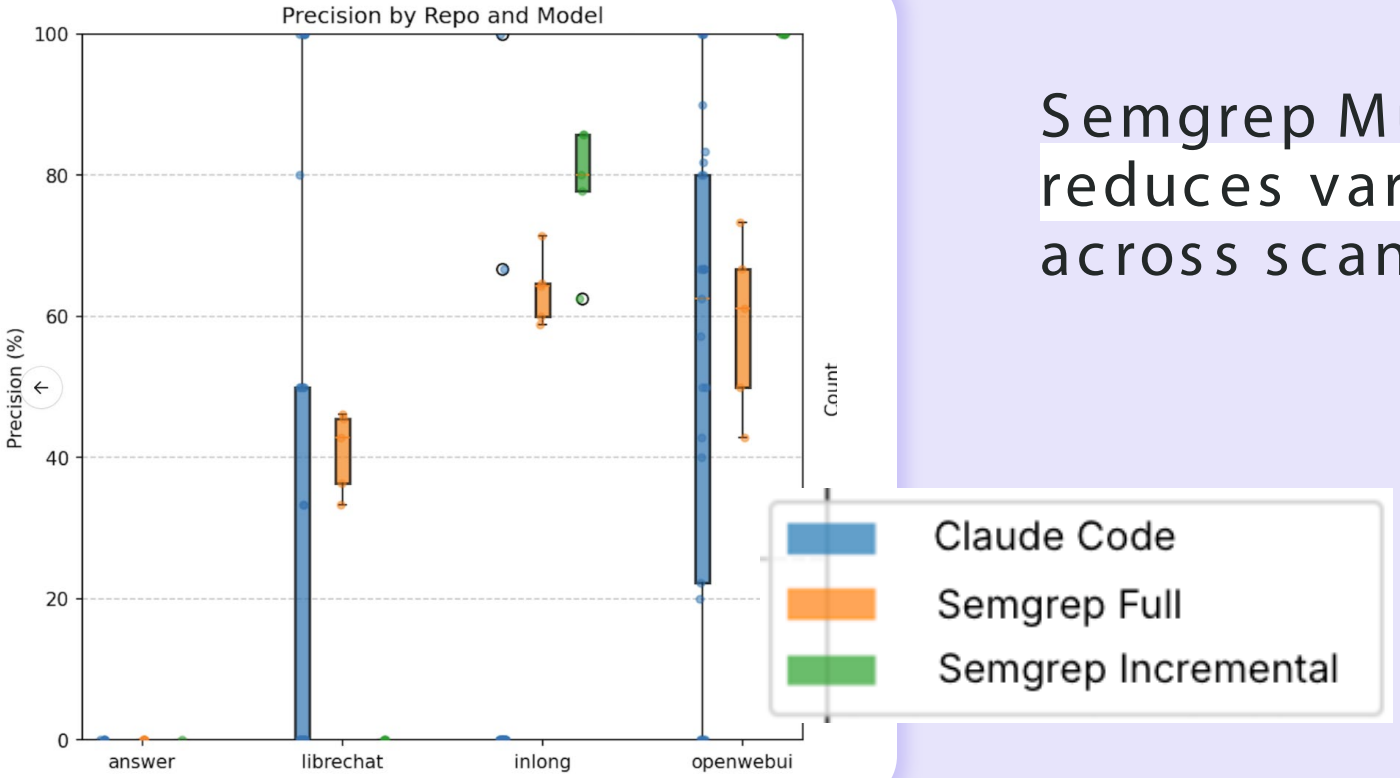
MAY 22, 2025 SEANHN 11 COMMENTS

o3 finds the kerberos authentication vulnerability in the benchmark in 8 of the 100 runs. In another 66 of the runs o3 concludes there is no bug present in the code (false negatives), and the remaining 28 reports are false positives. For comparison, Claude Sonnet 3.7 finds it 3 out of 100 runs and Claude Sonnet 3.5 does not find it in 100 runs. So on this benchmark at least we have a 2x-3x improvement in o3 over Claude Sonnet 3.7.

across all of those bugs, but here we'll focus on how o3 found a zeroday vulnerability during my benchmarking. The vulnerability it found is CVE-2025-37899 (fix here), a use-after-free in the handler for the SMB 'logoff' command. Understanding the vulnerability requires reasoning about concurrent connections to the server, and how they may share various objects in specific circumstances. o3 was able to comprehend this and spot a location where a particular object that is not referenced counted is freed while still being accessible by another thread. As far as I'm aware, this is the first public discussion of a vulnerability of that nature being found by a LLM.

Before I get into the technical details, the main takeaway from this post is this: with o3 LLMs have made a leap forward in their ability to reason about code, and if you work in vulnerability research you should start paying close attention. If you're an expert-level vulnerability researcher or exploit developer the machines aren't about to replace you. In fact, it is quite the opposite: they are now at a stage where they

https://sean.heelan.io/2025/05/22/how-i-used-o3-to-find-cve-2025-37899-a-remote-zeroday-vulnerability-in-the-linux-kernels-smb-implementation/
https://semgrep.dev/blog/security-research/



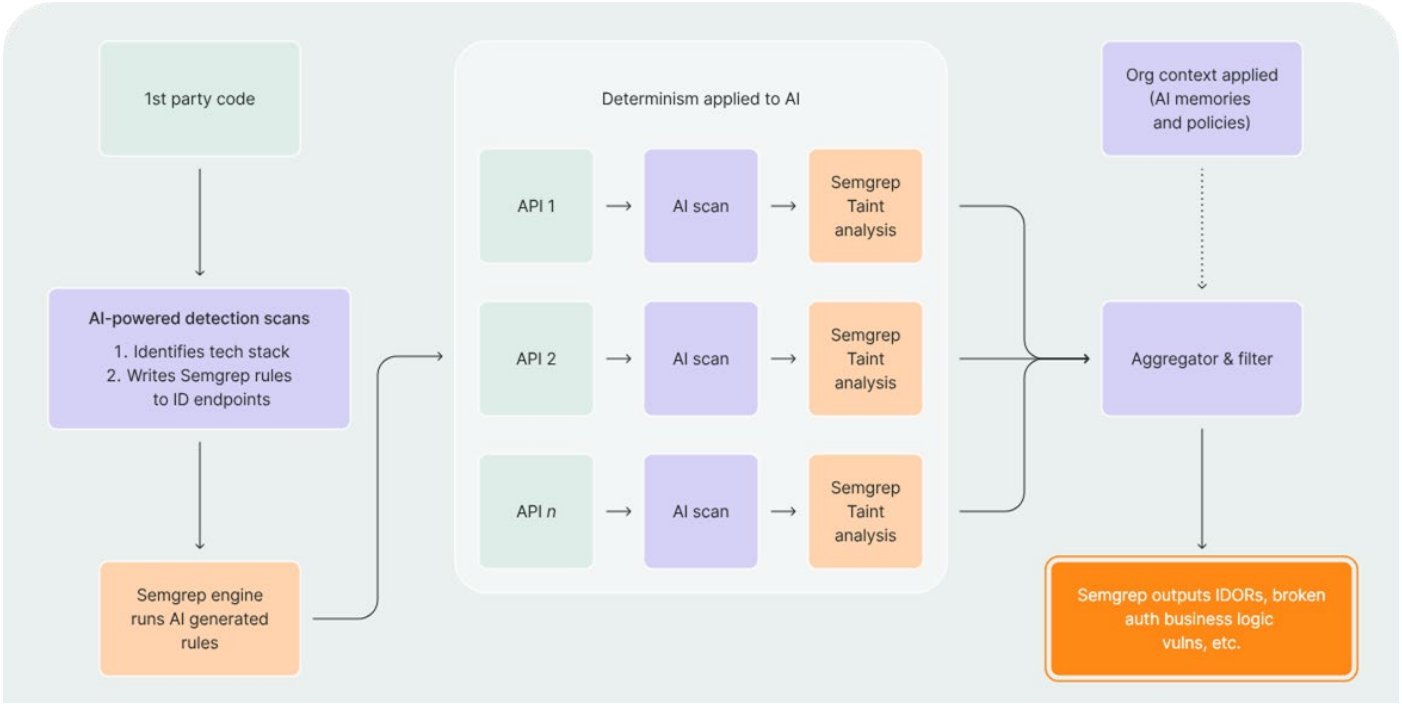
Semgrep Multimodal reduces variance across scans

Slide: Precision & Recall are low

Anthropic Claude Code (v1.0.32, Sonnet 4)

Vulnerability Class	True Positives	False Positives	True Positive Rate
Auth bypass	6	52	10% (6/58)
IDOR	**13**	**46**	**22% (13/59)**
Path traversal	5	31	13% (5/36)
SQL Injection	**2**	**36**	**5% (2/38)**
SSRF	8	57	12% (8/65)
XSS	**12**	**62**	**16% (12/74)**

Architecting a Multimodal detection workflow



Semgrep Multimodal delivers 8x True Positives with 50% fewer False Positives

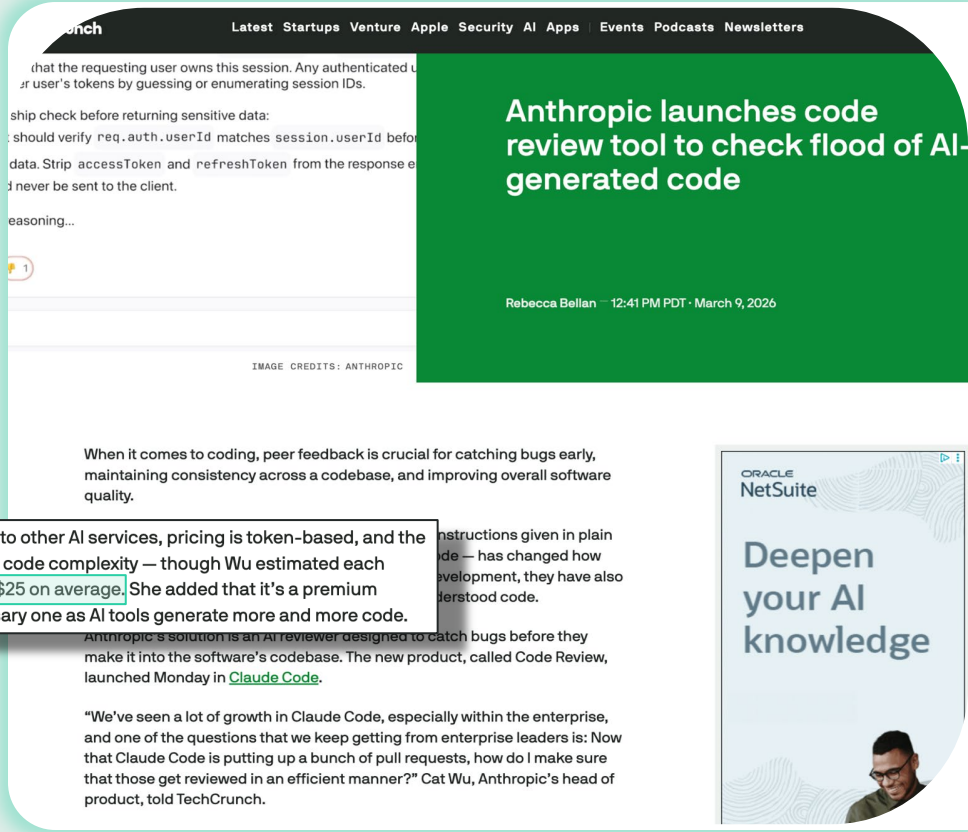
Security expertise improves results over foundational models alone.

	TP Rate (normalized)	FP Rate
Claude	100%	41%
Semgrep	823%	19%

Averaged across runs on a benchmark dataset. (Claude Haiku 4.5)

Balance of effort and budget with expensive LLM runs

For large code bases, LLMs can be expensive to generate the insights needed for a full review.



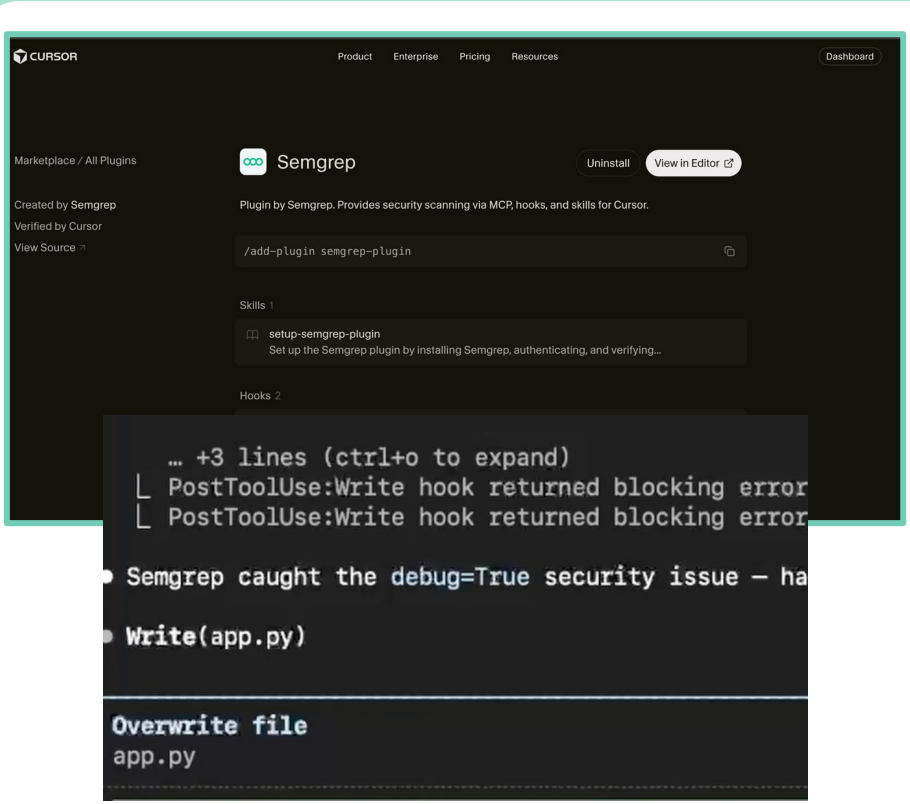
https://techcrunch.com/2026/03/09/anthropic-launches-code-review-tool-to-check-flood-of-ai-generated-code/

Semgrep Multimodal is so fast that your agents will love it

Semgrep will find and eliminate:

- Injection
- SSRF
- Misconfigured auth
- Supply chain vulns
- Leaked secrets

that might be introduced by AI coding agents **before they even end up in a PR.**



Let me see it!

Demo

Semgrep Multimodal goes beyond detection

What if we *didn't* just slap a chatbot onto scan results and call it a day?

Memories

Updating context and policies to learn about your codebase.

Autotriage

Filter out false positives with a 96% user agreement rate.

Autofix

Apply remediation with human-readable PRs and upgrade guidance.

Combining AI reasoning with rule-based SAST, Semgrep Multimodal performs better at detection, triage, and remediation

AI-generated code is outpacing the security practices built for human-speed development.

8x

True positives

Detected findings identified with far less noise than using AI models alone.

96%

User agreement

Presented with **Autotriage** decisions, real usage data consistently validates signal in the resulting analysis.

<2 min

Fix PR Time

Autofix creates fix PRs in under 2 minutes

Ari Kalfus Sr. Manager, Product Security

DigitalOcean

"[Semgrep's] contextual understanding was comparable to the analysis of one of our senior security engineers"

Questions?

Thank you

Try Semgrep Multimodal for yourself: https://go.semgrep.dev/signup/

Semgrep

AppSec is ready for Semgrep Multimodal

Combining LLM agents with rule-based analysis for detection, triage, and remediation.