

Cecuro Reaches 91.45% Detection Rate on EVMBench Security Benchmark

Cecuro reports 91.45% vulnerability detection on EVMBench, the independent benchmark from OpenAI and Paradigm, up from 87.17% and more than double the best general-purpose frontier model.

San Francisco, United States, June 23, 2026 [Cecuro](#), an agentic smart contract auditing platform, reported a vulnerability detection rate of 91.45% on EVMBench, the independent benchmark built by OpenAI and Paradigm. The figure rises from a previously published 87.17% on the same scored set and keeps Cecuro ranked first among all systems tested. The result lands as automated attacks accelerate and on-chain losses climb.

EVMBench measures whether an AI system can detect, patch, and exploit real smart contract vulnerabilities drawn from competitive audits. Cecuro's 91.45% covers a set of 117 vulnerabilities sourced from 40 audits, including findings from public competitions and a purpose-built stablecoin Layer 1. The benchmark scores three distinct tasks: finding a flaw, fixing it without breaking the contract, and writing an end-to-end exploit. Among general frontier models, the strongest detection rate recorded by the benchmark authors was 45.6%. Cecuro's 91.45% reflects how the work is structured around a model, not a larger model. The 4.28 point gain over the prior figure came from refinements to that orchestration rather than a change in the underlying frontier model.

The timing reflects a sharpening threat environment. The UK AI Security Institute estimates that frontier AI cyber-offense capability is now doubling roughly every 4.7 months, down from an 8-month estimate in late 2025. Anthropic reported that the share of tracked threat actors using AI rose from 33% to 56% over a single year. The losses are concrete: Chainalysis reported about \$3.4 billion stolen in crypto during 2025, including the roughly \$1.5 billion Bybit hack in February. As offensive tooling compounds, defenses that improve at the same pace carry the advantage.

"Detection is the bottleneck, not exploitation," said Daniel Delouya, CEO at Cecuro. "Once a flaw is identified, writing the exploit is often mechanical, and attackers already automate that step. The hard part is reasoning across a whole protocol to find the flaw in the first place. This jump from 87.17% to 91.45% did not come from a bigger model. It came from how we coordinate many specialized agents pursuing separate hypotheses, then confirm each finding with a working proof-of-concept. That is why the same model performs so differently inside our system, and why the gap over general-purpose tools keeps widening."

Cecuro runs roughly 180 specialized AI agents over an average of about eight hours per audit, with findings confirmed by proof-of-concept exploits and symbolic execution before they reach a report. The platform supports all chains and smart contract

languages, with reports starting at \$799. Full EVMBench methodology is published by the benchmark authors at <https://openai.com/index/introducing-evmbench/>, and teams can start a free review at [Cecuro's site](#).

About Cecuro

Cecuro is an agentic smart-contract auditing platform that ranks #1 on OpenAI's exploit benchmark, EVMBench, and has a consistent track record of finding vulnerabilities that top audit firms miss. Each security review is done in hours, not weeks, at roughly 90 percent lower cost than traditional audits. Its proprietary agentic system deploys hundreds of specialized agents, reasons about contracts from multiple perspectives, and validates each finding with proof-of-concept exploits before reporting. Cecuro supports all blockchain networks and smart-contract languages, with a free trial included. To learn more or start a review, visit <https://cecuro.ai>.

Media Contact

Daniel Delouya CEO, Cecuro daniel@cecuro.ai

###