Specialized AI Detects 92% of Real-World DeFi Exploits Worth $97M While General-Purpose Models Catch Just 34%

New open-source benchmark of 90 exploited smart contracts reveals increased DeFi security threat from AI, and a 13x gap in exploit value discovered between purpose-built security agents and standard AI coding tools.

FOR IMMEDIATE RELEASE, February 20, 2026

SAN FRANCISCO - New benchmark research reveals a stark gap between general-purpose AI and purpose-built security systems when it comes to protecting DeFi protocols from exploits. In a head-to-head evaluation of 90 real-world smart contracts that were exploited between October 2024 and early 2026, representing $228 million in verified losses, a specialized AI audit agent detected vulnerabilities in 92% of cases, covering $96.8 million in exploitable value. A standard frontier coding agent built on GPT-5.1, the kind of tool most development teams would reach for, caught just 34% and covered only $7.5 million: a 13x gap in protected value. Both systems ran on the same underlying large language model; the difference came entirely from the security-specific methodology layered on top. The findings arrive as separate research from Anthropic and OpenAI has shown that AI exploit capability is doubling roughly every 1.3 months, and that frontier agents can now execute end-to-end attacks on 72% of known vulnerable contracts. With the average cost of an AI-powered exploit attempt at roughly $1.22 per contract, the economics now heavily favor attackers, making the defense gap documented in this study an urgent problem for every deployed DeFi protocol.

"Many teams treat general-purpose AI as a security tool without realizing how much it misses," said Daniel Delouya, Co-Founder of Cecuro, an AI-powered smart contract auditing platform.
"*The model capability is there, but without domain-specific methodology, it does not get directed where it matters. Several of the contracts in our dataset had already passed professional human audits, and the vulnerability still went undetected.*"

Why This Matters

- **The "just use ChatGPT" approach to smart contract security is dangerously insufficient.** The benchmark shows that general-purpose AI, the kind of tool most teams actually use, misses the high-value, high-complexity vulnerabilities responsible for the vast majority of DeFi losses. Several of the exploited contracts in the dataset had previously passed professional human audits. The gap is not in model intelligence, but in how the model is directed. This finding has immediate implications for every project relying on basic AI tools or one-off audits for security.

- **AI offense is scaling faster than AI defense, and the cost asymmetry is getting worse.** Anthropic's research puts the cost of a single AI exploit attempt at $1.22. At that price, an attacker can systematically scan thousands of contracts for pocket change, and the tooling improves on a timeline measured in weeks. Meanwhile, professional human audits cost $15,000 to $300,000+ and cover a codebase only at a single point in time. This study quantifies how much value purpose-built defensive AI can recover, and why continuous automated review is no longer optional.

- **The benchmark is open-source, allowing independent verification.** The full dataset of 90 post-September-2024 exploits, the evaluation framework, and the basic agent baseline are publicly available on GitHub. The specialized agent is withheld to prevent misuse as offensive

tooling, but the reproducible baseline means any security team or researcher can validate the claims and build on the work.

About Cecuro

Cecuro is an AI-powered smart contract auditing platform that provides security reviews for DeFi protocols at a fraction of the cost and turnaround time of traditional audits.

The company's benchmark dataset and evaluation framework are open-sourced at github.com/Cecuro/defi-vuln-benchmark. Cecuro is based in San Francisco and Zurich, Switzerland.

Website: cecuro.ai
Benchmark: github.com/Cecuro/defi-vuln-benchmark
Media Contact: daniel@cecuro.ai