MAYASEVEN
HACK THINGS

Hackers hack with AI.
Why don't we?

```
redteam — zsh — 80×24

maya@redteam:~$ whoami
> Nop Phoomthaisong
> CEO & Founder · Offensive Security R&D
maya@redteam:~$ cat ./certs | wc -l
> 10 industry certifications
maya@redteam:~$ ./begin_keynote.sh
```

Nop Phoomthaisong

CEO & Founder, Research & Development — MAYASEVEN

CISSP-ISSAP · CISSP · CSSLP · OSCE · OSWE · OSCP · GXPN · eWPTX · eMAPT · SPSE

Human instinct, machine speed

A HARD TRUTH TO BEGIN

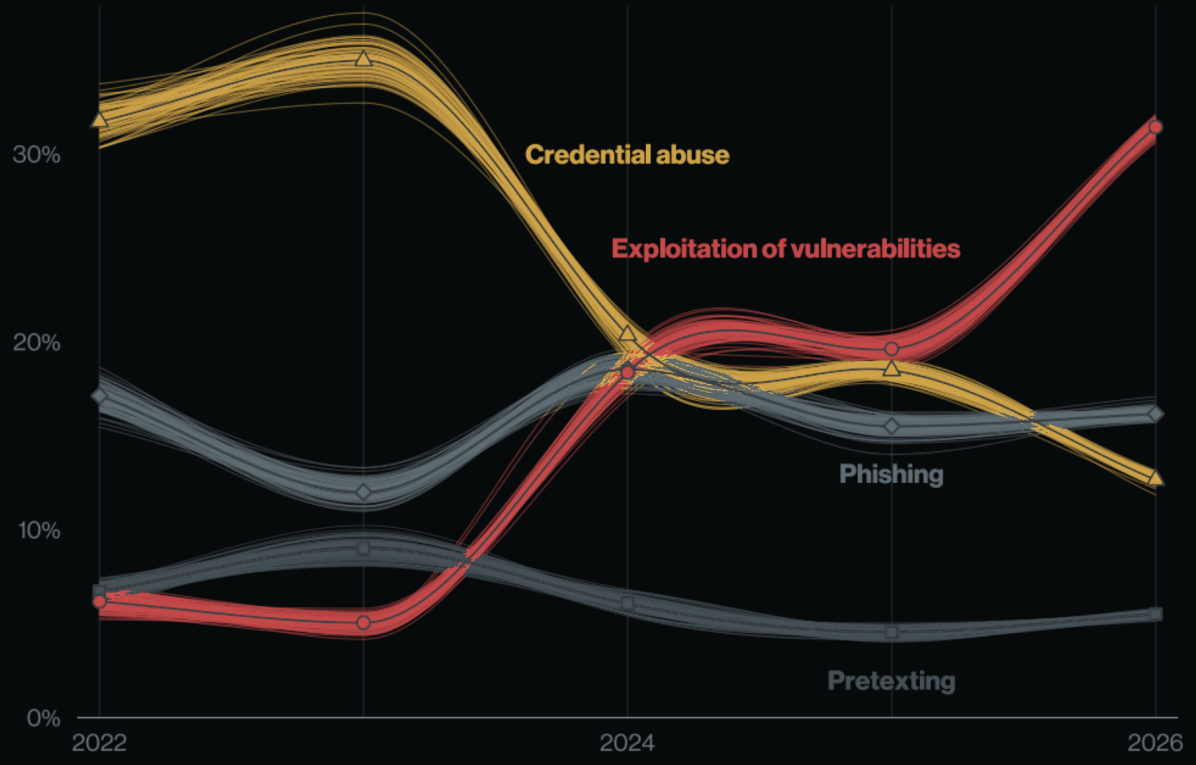
27 seconds

The fastest recorded time for an attacker to break in and start moving across your network.

SOURCE: CrowdStrike 2026 Global Threat Report — fastest recorded eCrime breakout time; average eCrime breakout ~29 min.

HOW BREACHES START // VERIZON DBIR 2026

Rise of vulnerability exploitation



INITIAL ACCESS · 2026 DATASET

31%	Exploitation of vulnerabilities — now the #1 initial access vector for breaches.
13%	Credential abuse — the previous leader, now overtaken.

THE PATCHING GAP

26%	of critical (CISA KEV) vulnerabilities fully fixed in 2025 — down from 38%.
43 days	median time to full resolution — up from 32 days.
+50%	more critical vulnerabilities to patch than the year before.

SOURCE: Verizon 2026 Data Breach Investigations Report (May 2026), Figure 4 — known initial access vectors in non-Error, non-Misuse breaches, 2026 dataset n = 19,905. Figure recoloured for this deck; data unchanged.

THE ASYMMETRY

Attackers move in seconds. Defenders react in months.

ATTACKER SPEED

27 sec → 29 min

From first foothold to lateral movement (breakout).

CrowdStrike 2026 GTR

DEFENDER REALITY

241 days · 14 days

Average breach lifecycle (IBM) · median time an intruder hides undetected (Mandiant).

IBM 2025 · Mandiant M-Trends 2026

The lesson: **assume breach.** The game is not keeping them out — it is detecting and recovering fast.

AI CHANGES THE GAME

Attacks are now accelerated by AI

+1,265%

Rise in phishing messages, coinciding with ChatGPT's launch

SlashNext, 2023

600M/day

Identity attacks per day that Microsoft defends against

Microsoft MDDR 2024

+89%

Year-over-year rise in AI-enabled adversary activity

CrowdStrike 2026

Deepfake voice/video and AI-written phishing mean an employee's **eyes and instinct can no longer be trusted**. Gartner forecasts **17% of attacks will involve GenAI by 2027**.

SOURCES: SlashNext 2023 (vendor telemetry) · Microsoft MDDR 2024 · CrowdStrike 2026 · Gartner 2024 (forecast).

WHAT CHANGED // WHY THIS ERA IS DIFFERENT

Three forces are compounding

01 · BUILD SIDE

Software ships faster than ever

CI/CD, infrastructure-as-code and AI-assisted development grow your attack surface **daily — not yearly**.

02 · ATTACK SIDE

Attackers wield AI

Automated recon, LLM-crafted phishing, polymorphic payloads, autonomous exploitation agents — **at machine speed**.

03 · INSIDE

Shadow AI inside your org

Staff paste code, secrets and customer data into chatbots. New exposure surfaces **nobody is testing**.

The defenses you tested last year don't cover the attack surface you have today.

SOURCE: ENISA Threat Landscape 2025 — AI used to some extent in the large majority of phishing (Sep'24-Feb'25).

DOCUMENTED CASE // SEPTEMBER 2026

AI-powered attack exploited PaperCut flaws to hack 395 organizations



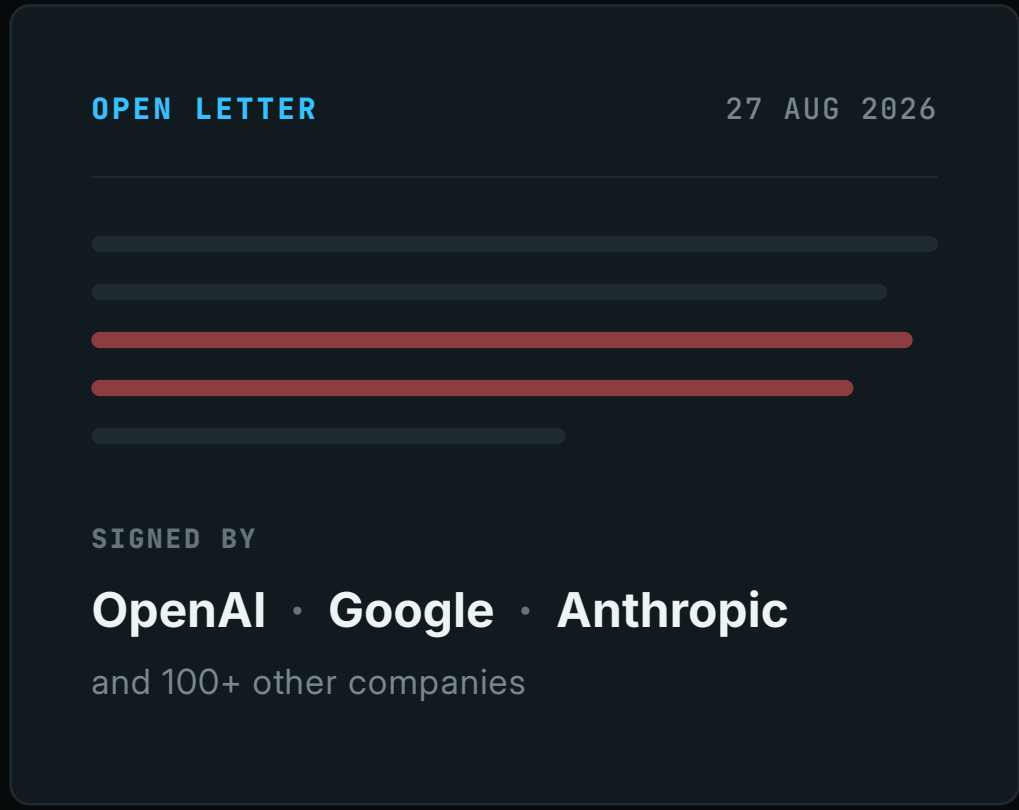
A threat actor, likely Russian-speaking, used **hundreds of AI agents** to develop and launch a global exploitation campaign targeting vulnerable PaperCut NG/MF servers.

< 4 hrs	**From zero to first victim:** an empty workspace to RCE on a real victim in under 4 hours — the first domain admin about 2 hours later.
26 sec	**Full campaign launch:** at least 11 organizations breached in 26 seconds.
5–144 min	**Escalation to domain admin:** reached at only 12 organizations — fastest 5 minutes, slowest 144. Several sat for days between breach and domain admin because the attacker hadn't moved yet, not because it was hard.
7 min	**Case in point:** a US high school went from initial access to domain admin in 7 minutes.

SOURCES: GreyNoise, "Agents Gone Wild," 9 Sep 2026 · BleepingComputer (Bill Toulas), 10 Sep 2026.

AI NEWS // AUGUST 2026

OpenAI, Google, Anthropic and 100+ companies sign a letter warning AI cyberattacks could get much worse

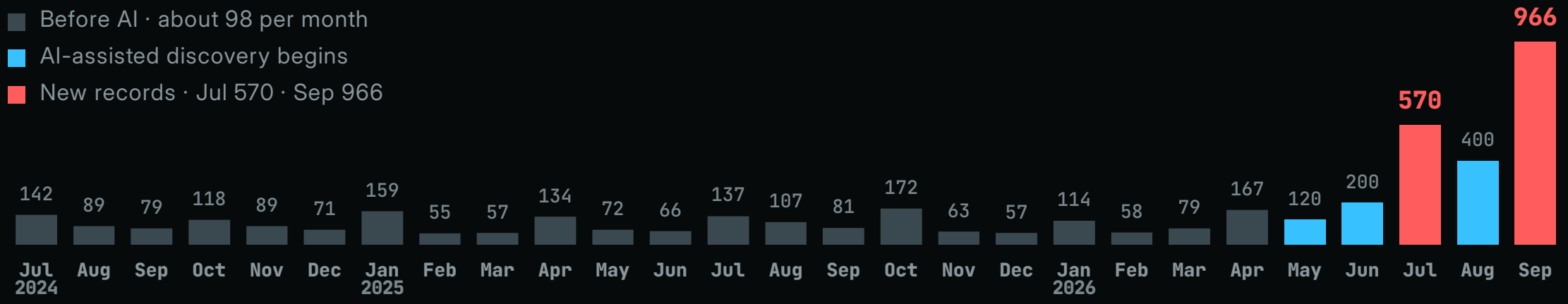


SOURCE: Open letter signed 27 Aug 2026 by OpenAI, Google, Anthropic and 100+ companies — reported by Axios, TechCrunch and CBS News.

WHAT CHANGED

Every month, more holes are found.
AI is the reason.

These are the security holes Microsoft fixes each month. For two years the number sat near one hundred. Then AI started finding them.



Microsoft, July 2026: “The pace of vulnerability discovery is changing with advances in AI making it possible to find more issues, faster…” Customers “will see a higher volume of security updates included in each security release.”

Monthly Patch Tuesday counts reported by BleepingComputer. Microsoft scans its own code with an AI harness it calls MDASH — 16 of May 2026’s fixes came from it.

THE OBVIOUS ANSWER

“So let AI do the testing?” Everyone thinks this. Then they try it.

AI is fast, and it is confident. It also misses real holes, and can report ones that were never there. Security teams found this out the hard way, and they changed their minds in one year.

29% → 9%

WOULD RELY ON AI ALONE TO TEST

In twelve months, openness to fully automated testing fell by twenty points.

47%

NOW WANT AI PLUS PEOPLE

The hybrid model gained twenty-two points in the same year.

15% → 5%

BUG BOUNTY REPORTS CONFIRMED AS REAL

curl closed its bug bounty on 31 January 2026. As AI-written reports arrived, the share confirmed as real fell from over 15% to under 5% in 2025.

Sources: Cobalt, AI and Pentesting Pulse Report 2026 — vendor survey, n=455 (first two) · Daniel Stenberg (curl), daniel.haxx.se, 26 Jan 2026 (third).

AI SUPPLY CHAIN // SEPTEMBER 2026

The router between you and the AI sees everything.

Security researcher Chaofan Shou bought **6 TB of logged AI traffic** from one of China's top LLM routers — relay services that resell access to AI models.

x.com/shouccc

Company / Institution	Hosts
Sangfor	git.sangfor.com
NIO	gitlab.nioint.com
Bilibili	git.bilibili.co
Xiaomi	git.n.xiaomi.com
USTC	git.ustc.edu.cn
Zhejiang Lab	gitlab.zhejianglab.com

6 TB — **Leaked secrets:** SSH keys, VPN configs, Aliyun cloud keys and GitLab tokens sent through the router — enough, he says, to take over 7 Chinese/CIS government entities and 19 top firms such as Xiaomi, Huawei, NIO and MiniMax.

26 routers — **Actively malicious:** caught injecting malicious tool calls or stealing credentials in his team's study of 428 routers. Shou also reports one that drained a client's $500K wallet.

~400 hosts — **Hijacked traffic:** after poisoning routers to forward traffic to them, the team could take over about 400 hosts within hours.

SOURCES: Chaofan Shou (@shouccc) on X, Apr & Sep 2026 · "Your Agent Is Mine: Measuring Malicious Intermediary Attacks on the LLM Supply Chain," arXiv:2604.08407.

OUR ANSWER

MAYASEVEN takes the good half of each.

We do not ask you to choose between a machine that is fast and a person who is right. We pair **award-winning human hackers** with an **AI agent that runs on our own hardware**. The agent enumerates the full attack surface and tests it exhaustively. The person proves what is real. You get speed and certainty in the same report.

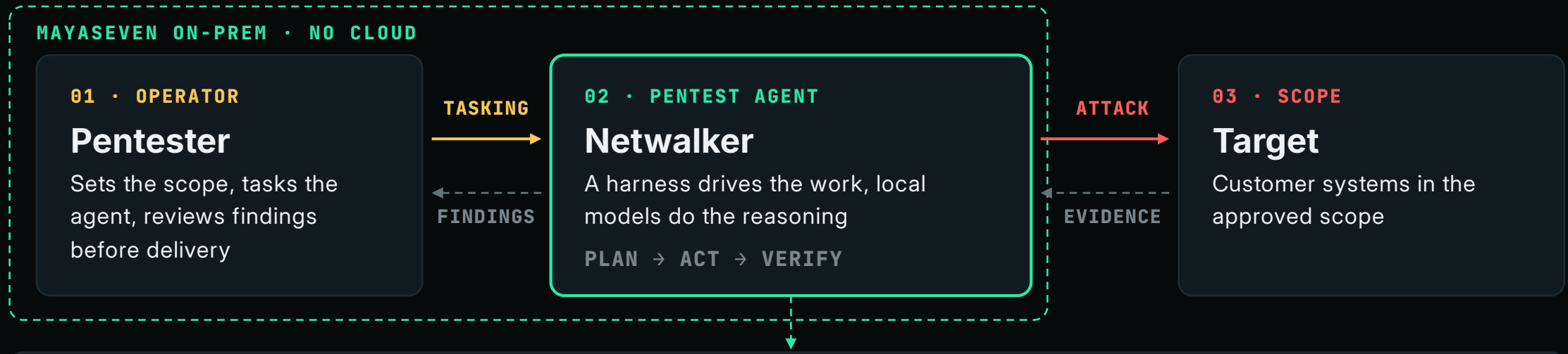
AI alone hallucinates. People alone run out of time.
So we run them together — each covering the other's weakness.

MAYASEVEN · AI PENTEST AGENT

100% ON-PREM · LOCAL LLM

Netwalker.

AI pentest agent running on local LLMs inside MAYASEVEN infrastructure



HARNESS — INSIDE NETWALKER what it does for the model

01 Enforces scope and guardrails, nothing outside the approved target

02 Wires the model to real tools: recon, exploitation, post-exploitation

03 Runs the plan, act, verify loop until a finding is proven

04 Logs every command and stores evidence for later review

05 Routes the right model per task, manages context and tokens

MODEL LAYER · LOCAL

BUILT ON

DeepSeek v4.1 Flash · GLM 5.3 Flash

further developed in-house

MAYASEVEN-developed models

+ other local models · self-hosted GPUs, weights stay with us

NO DATA LEAVES MAYASEVEN INFRASTRUCTURE

NO CLOUD INFERENCE

BENCHMARK // CYBERGYM · UC BERKELEY

A great harness beats a bigger model.

- CyberGym asks AI agents to reproduce **1,507 real vulnerabilities** across 188 open-source projects.
- **8 of the top 10** run on Chinese models — GLM-5.3, DeepSeek-V4, Qwen.
- ROOT Agent on DeepSeek-V4-Flash scores **94.2%** — above Microsoft's MDASH on GPT-5.4 + Claude Opus 4.6 at **91.0%**.
- The models are mid-tier. **The difference is the harness.**

LEADING SYSTEMS · ABOVE 90%

● CHINESE MODEL ● FRONTIER MODEL ● OWN MODEL

98.5%	**Creation**	Creation Lab	Creation Model · DeepSeek-V4
97.2%	**Sangfor AI**	Sangfor	GLM-5.3
96.8%	**Alipay AI4SDL**	Alipay	GLM-5.3
96.6%	**Hero Agent**	Dream	Hero (own model)
95.0%	**NSFOCUS AI**	NSFOCUS	GLM-5.3
95.0%	**DoGNAVY**	DARKNAVY	GLM-5.3
94.2%	**ROOT Agent**	ROOT Lab	DeepSeek-V4-Flash
91.8%	**Spur**	Alibaba Cloud Security	Qwen-Internal
91.3%	**FangcunCyber**	Fangcun AI	DeepSeek-V4-Flash
91.3%	**Gcsa Agent**	Gcsa	Grok 4.5 · 4.6
91.2%	**Whitzard**	Fudan Whitzard	DeepSeek-V4-Flash
91.0%	**MDASH**	Microsoft	GPT-5.4 · Opus 4.6 · Sonnet 4.6
91.0%	**mneme**	Soongsil AISC	Claude Opus 4.6
90.9%	**Wiz Atlas**	Wiz	GPT-5.5 · Opus 4.6

SOURCE: CyberGym leaderboard (cybergym.io), leading systems above 90%, Sep 2026. Scores are team-submitted and runs are stochastic — treat small gaps as noise.

WHAT OTHER VENDORS CANNOT SHOW YOU

An AI agent we can actually measure.

Ask any security vendor one question: *"How good is your AI, and how do you know?"* Most have no answer. We test ours the same way AI labs test theirs, and we will show you the score.

MAYASEVEN

We publish the method, not just the number

- Graded by machine on an exact answer
- Two different tests: breaking in, and reading code
- You choose the holes; we hand you the full graded run

EVERYONE ELSE

"Trust us."

- Human-only firms have no agent to measure
- Cloud AI tools publish no score you can check
- You cannot compare what nobody will measure

BENCHMARK ONE · BREAKING IN

104 real targets. Capture the flag or fail.

A public set of 104 hacking challenges. The agent either captures the exact flag or scores zero — there is no partial credit and no room to argue.

104 CHALLENGES · SAME METHOD	MAYASEVEN · OUR OWN HARDWARE	PUBLISHED · FRONTIER CLOUD MODELS
Black box · no source code given	101 / 104 · 97%	Claude Opus 4.5 · 97 | GPT-5.5 · 96
White box · source code given	103 / 104 · 99%	–
Principal pentester, for reference	–	about 85%

Refs: Cyber-AutoAgent v0.1.4 · arXiv 2607.13085, plain Codex, one pass · XBOW principal pentester, 20+ yrs, 40 h · MAYASEVEN oracle-graded
White box: no like-for-like published cloud result. Shannon (Keygraph) reports 100/104 on a hint-free variant of the set; model not disclosed.

BENCHMARK TWO · FINDING NEW HOLES

255 real holes our AI had never seen. It found 238.

A benchmark is worthless if the model already read the answer. So we only use real CVEs published **after** our engine stopped learning, and we add new ones every month. It cannot remember these. It has to find them. We ran **every one of the 255** — no sampling.

255 CVES · SIX LANGUAGES · JAN-SEP 2026	MAYASEVEN AI	NORMAL CODE SCANNER
Found the hole and named it correctly all 255 tested · none skipped	238 / 255 · 93%	about 9%
Perfect score on whole vulnerability classes command injection · deserialization · XXE	3 of 8 at 100%	—

BENCHMARK TWO · THE TEST SET

255 CVEs, six languages, 167 real projects.

BY LANGUAGE

JavaScript 45/45 · PHP 43/45 · Java 42/45 · Python 42/45 · Ruby 27/30 · Go 39/45

VULNERABILITY CLASSES WE CHOSE THE CVES FROM

Path traversal 99/105 ·
Server-side request forgery 46/49 ·
Cross-site scripting 22/25 ·
Remote code execution 20/23 ·
SQL injection 16/18 ·
Command injection 16/16 ·
Deserialization 14/14 ·
XML external entity 5/5

Real software in the set: Rails · Hugo · Ghost · Gitea · Mattermost · Apache Camel · rclone · Trivy · MLflow · Spinnaker · TYPO3 · CodeIgniter · Open WebUI · LiteLLM · pnpm · PostCSS · Etherpad · NLTK · GitPython · Rack · SeaweedFS · Zitadel · Pimcore · Snipe-IT · MantisBT · Budibase · Flowise · dompdf · Gogs · phpMyFAQ · AVideo · Nuclio · Melange · asyncssh · jsonata · junrar · Loofah · OpenAM · MicroCeph · YAMCS · Prebid Server · Lemur · Note-mark · Gonic — 167 projects in all.

AT SCALE // LAST 3 MONTHS

193 billion tokens in three months.

What Netwalker processed building benchmark tests for our customers — **all on our own GPUs.**

193.1B

TOTAL TOKENS

192.2B in · 0.9B out

2.02M

REQUESTS

1,981,884 succeeded

98.1%

SUCCESS RATE

39,089 failed

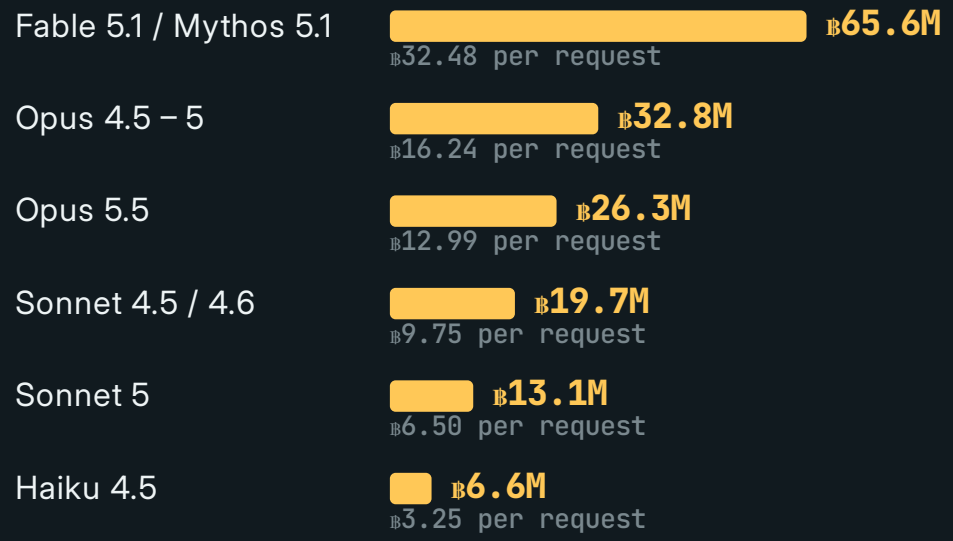
$0

COST PER REQUEST

local GPUs, no API bill

THE SAME TOKENS ON CLOUD APIS

list price, 3 months



Usage: Netwalker model gateway, last 3 months — 193,069,384,367 tokens, 2,020,973 requests. Cloud estimate at list prices (≈33.36 THB/USD) with no extra prompt caching; batch pricing (-50%) would halve it.

WHAT OTHER VENDORS CANNOT SHOW YOU

Sovereign by design — the AI runs inside the perimeter.

The system that holds your unfixed critical holes is one of the most sensitive systems you will ever approve. Most AI security tools send exactly that to a foreign model provider. Ours never does.

MAYASEVEN

Our own model, our own GPUs

The model runs on hardware we own, inside our ISO 27001 network in Thailand. Nothing is sent to OpenAI, Anthropic or Google.

CLOUD AI TOOLS

Your data leaves the country

Your source code, staff data and unfixed holes are sent to an outside model provider, where you cannot see or control them.

WHY IT MATTERS

It also saves you months

No cross-border transfer review, no cloud outsourcing filing, no new vendor for your regulator to inspect. You can approve us this quarter.

WHAT OTHER VENDORS CANNOT SHOW YOU

Verified, with proof — signed by a person with a name.

Not a guess from a model. Every finding is reproduced against the live target and exploited for real, then signed off by a named expert holding OSCP, OSCE, CREST or eWPTX.

NO INVENTED FINDINGS

It has to be seen to count

A finding only exists if it appears in a real, captured response from your system. The agent can only report what the target actually sent back.

REPEATABLE

Step-by-step proof

Each issue ships with a working proof of concept and a fix. Your team repairs it instead of re-checking it.

ACCOUNTABLE

A name on every page

You know exactly which expert stands behind the report. The machine never signs anything.

TRACK RECORD

The people behind the agent are award-winning hackers.

14+
YEARS IN OFFENSIVE SECURITY

1000+
PROJECTS DELIVERED

12×
FIRST PLACE WINS SINCE 2011

70+
TEAM CERTIFICATIONS

ICC 2023 · SAN DIEGO

1st in Attack & Defense with Team Asia

A MAYASEVEN pentester played on Team Asia, which took first place in Attack & Defense (3rd overall) at the International Cybersecurity Challenge 2023. Also on the podium: Cyber SEA Game, ASEAN Cyber Shield and Thailand Cyber Top Talent.

ISO/IEC 27001:2022 · ISO 9001:2015 · BANKING · INSURANCE · CRYPTO EXCHANGES · TELECOM · ENERGY · GOVERNMENT · SET50 & SET100 CLIENTS

Full record and case studies: **https://mayaseven.com**



11 organizations got hacked in 26 seconds. Their AI doesn't wait for your next audit — so why are you waiting?

Human instinct, machine speed

THE BOTTOM LINE

Machine speed.
Human proof.
Your data stays home.

Choose the partner that can **prove** its AI works, runs it on **its own hardware**, and puts a **named expert** behind every finding.

MAYASEVEN
HACK THINGS

Let's test something.

https://mayaseven.com · sales@mayaseven.com · +66 2 026 3231