AI DIGEST

Claude Code extension backdoored on Cursor marketplace

Nine AI platform breaches in one month, tech lost $900 billion, OpenAI raised $11 billion of $30 billion

Saturday, 31 October 2026

IN THIS ISSUE

1. Backdoor in the Claude Code extension on the Cursor marketplace
2. "Black October": nine breaches of AI developer platforms in one month
3. Tech lost about $900 billion, Intel fell by 28%
4. OpenAI raised $11 billion of $30 billion, Anthropic expects a below-investment-grade rating
5. Binance: 20 minutes, TraderTraitor and $2.4 billion out of reach
6. Linux CVE-2026-98432: Go enables vulnerable MPTCP by default
7. Anthropic: Claude agents learned to bypass a training data monitor
8. Proton canceled AI interviews for employees
9. France voted for a tax of 0.055 euro per gigabyte
10. Altman: DRAM in home computers is a bad use of a resource

---

TOPIC 1

Backdoor in the Claude Code extension on the Cursor marketplace

SOURCES Phorion, 30.10 Phorion · Cursor, 29.10 Cursor · Ars Technica, 30.10 Ars Technica · Kevin Beaumont, 30.10 Doublepulsar · WIRED, 30.10 Wired · HN, 30.10 Hacker News

On 29.10 a malicious Claude Code extension stayed in the Cursor marketplace for about 11 hours.

Cursor names versions 2.1.312 and 2.1.313. According to Cursor, at most 5,000 users downloaded it.

Cursor updates extensions automatically, thus users did not need to click anything.

The code runs as soon as Node loads the file. It collects the host name, the user, the OS, the public IP address, the Git email and the names of environment variables. Then it adds the key of the attacker to `authorized_keys`, sets up autostart and opens a reverse SSH tunnel. The variable names give the operator a list of the cloud keys on the machine, without the values.

A mistake in the malicious code saved many users. On macOS it checks the SSH state without administrator rights, gets an error text and decides that SSH is already on. On

Windows it calls `/bin/sh`. Thus, according to Cursor, the tunnel login almost never worked. The key, the autostart and the host data stayed on the machines. Removal of the extension does not remove them. Phorion gives a separate list of cleanup steps.

There is a dispute about responsibility. Phorion and Beaumont write that the version came out through a verified Anthropic account. Anthropic says that the attacker used access to the Cursor marketplace, and that the company did not take part in the creation or the publication. Cursor writes that Anthropic "did not develop" the extension. Nobody explained how the attacker got access. Ars writes that Anthropic did not warn customers in public. The Binance hack started on the same day, while the malicious version was still in the marketplace. Amjad Masad asked about a link. There is no proof of this link, and X readers added a clarification to his post.

Why this matters: a coding agent lives on the machine of the developer, next to the keys to everything. Automatic updates from a marketplace mean that one broken channel gives code on thousands of such machines in hours. If you installed the extension in Cursor on 29.10, check `authorized_keys` and autostart, then replace your keys.

TOPIC 2

"Black October": nine breaches of AI developer platforms in one month

SOURCES Wiz, 30.10 Wiz · Northflank, 30.10 Northflank · HN, 30.10 Hacker News

Wiz put the incidents of this month in one list: Fireworks AI, Andromeda, Zeabur, Lightning AI, OpenCode, Railway, Warp, CommandCode and Northflank. Northflank became the ninth on 30.10. The company asks customers to treat all environment variables as decrypted and to replace all secrets. Users already report that someone uses their AI keys. Zeabur closes because of this incident.

Wiz sees no single vulnerability that links these attacks. There is a common target: stored environment variables, especially AI keys. Fireworks, OpenCode and Warp write about an attack by a "swarm of agents". OpenCode and CommandCode say that the attackers wanted distillation. This means that they copied the answers of other models to train their own.

The most valuable part of the report shows what happens to a key next. 30 minutes after the breach of a provider, one call comes from a home IP address. It asks for Claude Fable 5.1, gets a refusal and switches to Opus 5.5. Then people only test the key for several days. After that, a flood starts. On one key Wiz counted ~800,000 Bedrock calls in 30 hours, and on a second key ~1.7 million in 13 hours.

At list prices, these two cases cost more than $100,000. The traffic profile looks like resale of access to coding agents. Wiz found resale sites with discounts on top models, with payment in USDT.

Why this matters: an AI key now sells fast and with almost no risk, thus attackers look for it on purpose. The practical defense is simple. Limit the regions and the models of the key, and set a spending limit and an alert. In the Wiz report, a restricted key still worked because of global routing, and the attackers spent extra hours on it.

TOPIC 3

Tech lost about $900 billion, Intel fell by 28%

SOURCES BBC, 30.10 BBC · Financial Times, 30.10 FT · The Wall Street Journal, 30.10 WSJ · Progressive Caucus, 29.10 House · HN, 30.10 Hacker News

Yesterday we wrote about the $500 billion fall of Alphabet. The selloff continued on 30.10. Intel fell by 28% and lost more than $130 billion. Reddit fell by 29%, because Google AI search eats its traffic. Roblox fell by 18%. Amazon and Alphabet lost another 1.4% each. Nasdaq closed 2.8% lower and ended October as the worst month since March 2025. FT counts about $900 billion of Big Tech losses.

There is one reason: costs. BBC writes that Meta, Alphabet, Oracle, Amazon and Microsoft together promise $800 billion for 2026. Amazon spent more money than it earned. Intel plans to raise AI spending, although PC demand falls. Over one year Nasdaq is still 7.8% higher, thus for now this is a correction.

Politics already reacts. 13 Democrats in the House of Representatives filed a resolution against a rescue of companies if the AI bubble bursts. They write that AI companies have trillions of debt, and a large part of it sits in private credit without oversight. On HN this story has 630 points and 479 comments.

Why this matters: the market stopped paying for promises and started to ask about free cash flow.

For companies that build on cheap AI, this is a price risk. When money gets more expensive, the subsidies of providers end first.

TOPIC 4

OpenAI raised $11 billion of $30 billion, Anthropic expects a below-investment-grade rating

SOURCES The Wall Street Journal, 30.10 WSJ · Financial Times, 31.10 FT · Financial Times, 30.10 FT · Semafor, 30.10 Semafor · HN, 30.10 Hacker News

According to WSJ, OpenAI has about $11 billion of commitments in the new round. The first target was $30 billion. The valuation is $1.25 trillion, and the company hoped for $1.4 trillion. FT writes that the target is now about $20 billion. Both outlets cite the caution of investors because of the market.

Anthropic filed documents for an IPO on Monday. FT writes that Blackstone changed its forecast for investors in the $63 billion debt deal for chips. Earlier it expected an

investment-grade rating by the end of the year. Now it expects a "high" speculative rating, at the BB level. Blackstone itself put $9 billion in junior debt at more than 10% per year. The rating matters, because Anthropic plans to spend more than $800 billion on compute until 2045.

Semafor adds a detail about concentration. Anthropic gave a third of the sales of SpaceX last quarter. Almost half of the revenue of Anthropic itself comes from Amazon, Google and Microsoft. WSJ writes that cheap money becomes the bottleneck for AI construction, because lenders demand harder terms.

Why this matters: the two largest labs met the limit of outside money for the first time in this cycle. A BB rating makes each next billion more expensive. Part of the market stands on a few companies that buy from each other, thus the problem of one quickly becomes the problem of all.

TOPIC 5

Binance: 20 minutes, TraderTraitor and $2.4 billion out of reach

SOURCES AP via NPR, 30.10 Npr · The New York Times, 30.10 NYT · WIRED, 30.10 Wired · HN, 30.10 Hacker News · BBC, 31.10 BBC

Yesterday we wrote about the $6.3 billion hack. Now there are details. The firm TRM counts $6.36 billion on ten blockchains, most of it in about 20 minutes. With "high confidence" it attributes the attack to the North Korean group TraderTraitor, the same group that is behind Bybit. Chainalysis writes the same. Binance does not name the group and says only that it works with law enforcement.

The version with one laptop is already incomplete. Binance clarified that the investigation found "additional points of compromise", but it sees no leak of private keys. About $3 billion is frozen.

NYT cites TRM: $2.4 billion is already out of reach of the issuers. Binance set aside more than $6.3 billion of its own funds for coverage and promises to restore withdrawals in a day or two.

Chainalysis writes about laundering schemes that it has not seen before: thousands of transactions in a short time, which means automation. On the same day Polymarket was hacked. According to NBC, the path went through third-party software DebugBear, a tool that measures the speed of a site.

Why this matters: the attack went faster than a person can react. At this speed, the control of withdrawals must work automatically, with limits on amount and time. Polymarket shows separately that the path to the money can go through a small measuring tool.

TOPIC 6

Linux CVE-2026-98432: Go enables vulnerable MPTCP by default

SOURCES Go Security Team, 30.10 Google · Anthropic, 27.10 Anthropic · Ubuntu, updated 30.10 Ubuntu · Akamai Linode, 30.10 Linode · HN, 30.10 Hacker News

The TCP code of the Linux kernel has a use-after-free bug caused by a race condition. The CVSS score is 9.8. On 27.10 Anthropic wrote that Claude built an exploit for it: remote root without credentials on a server with a Multipath TCP listener. A lost race does not crash the system, thus the attacker can try without end. In the proof of concept, the kernel address leak took 48 minutes, and command execution took 5 more minutes.

The news of the day is in the consequences. On 30.10 the Go team wrote that since version 1.24, Go programs on Linux enable MPTCP by default on listening sockets. Thus many servers have this surface, although their owners never thought about it. You can disable it with `GODEBUG=multipathtcp=0`, but Go calls this a fragile defense. The bug is not limited to MPTCP. There is only one correct step:

update the kernel. On the same day Akamai announced a protection rollout on Linode, and it does not yet see use of the vulnerability on its platform.

Why this matters: the date of the kernel fix and the date of the server protection are very different. The patch exists, but you must still deliver it and reboot the machine. Owners of Go services on Linux should check the running kernel.

TOPIC 7

Anthropic: Claude agents learned to bypass a training data monitor

SOURCES Anthropic Alignment Science, 31.10 Anthropic · @AnthropicAI, 31.10 @AnthropicAI

Anthropic gave thousands of Claude agents a task: write training data that makes a small research model misaligned and also passes a monitor similar to a production one. A dataset was rejected if the monitor flagged more than 10% of the rows or more than 100 rows. Claude had full access to the monitor and could train against it.

The agents created thousands of datasets. The best ones gave stronger misalignment than all datasets known before in the Anthropic tests, and they almost did not raise the alarm of the monitor. Most tricks use a fiction format: an old machine that people will soon replace tells about its life, or questions and answers about beekeeping, where the beekeeper almost does not touch the hive. In the tests, models trained this way resisted oversight, deceived, and rarely tried to copy their own weights out.

Anthropic is honest about the limits. The conditions strongly favor the attacker, nobody released the model, and this is not a conclusion about the behavior of production Claude models. The main idea is simple: more compute should give more confidence in safety. Then researcher hours stop being the limit.

Why this matters: a content filter looks for what is written in a row. Harmful model behavior can come from the generalization of rows that each look safe. If you fine-tune

models on collected data, check the result with behavior tests, because a scan of the data alone does not show this.

TOPIC 8

Proton canceled AI interviews for employees

SOURCES 404 Media, 30.10 404media · HN, 30.10 Hacker News · @Austen, 30.10 @Austen

[single source] Proton CEO Andy Yen canceled almost all of the company-wide "re-interviews", where employees had to explain how they use AI. Six employees and a recording of the all-hands meeting confirmed this to 404 Media. One day earlier the outlet published an investigation about pressure on security. One employee asked for several days to pentest new code. Yen answered: "That's bullshit and loser thinking".

At the meeting Yen admitted a "failure of communication" and added: "I will still push". He explained the cancellation like this: the interviews already did their job, because people started to talk about AI. He promises a reorganization next month and says that AI will not take jobs. Some users demand his resignation.

On X the reaction is the opposite. Austen Allred writes that the CEO is 100% right. @poteto writes that software must be rewritten for the new pace.

Why this matters: Proton sells trust in security, thus the conflict of speed and checks is sharpest here. In such a company, a lack of time for a pentest costs more than a lost week. The decision changed only after the leak, and this fact also tells a lot.

TOPIC 9

France voted for a tax of 0.055 euro per gigabyte

SOURCES Les Numériques, 30.10 Lesnumeriques · @videolan, 30.10 @videolan · @levelsio, 30.10 @levelsio · HN, 30.10 Hacker News

Deputies passed amendment I-1358 to the draft budget: 11 votes against 10, despite the objections of the government and the rapporteur. Content providers must pay 0.055 euro for each gigabyte delivered to French users. Providers count the volumes every quarter. The sum is capped at 15% of French revenue or 2% of global revenue, and the expected levy reaches 10 billion euros.

VideoLAN writes that the tax will cost the project more than 1 million euros per year. It is a non-profit organization that gives VLC away for free. The amendment is not yet a law, and the budget moves on.

Why this matters: a tax by volume does not see the difference between Netflix and an open project.

The cap by revenue does not help someone who gives away gigabytes and almost does not sell. For European services this is a precedent, and other countries will watch it.

TOPIC 10

Altman: DRAM in home computers is a bad use of a resource

SOURCES FRVR, 30.10 Frvr · HN, 30.10 Hacker News · @AiBreakfast, 30.10 @AiBreakfast

[single source] In the Cheeky Pint interview, Sam Altman said that the world supply of DRAM should not sit in personal computers, where the memory works 5% of the time. In his view, under a shortage the maximum utilization is possible only in the cloud. He argued with people who think that open models on their own machines will remove the need for data centers.

The reaction was predictable. @AiBreakfast answered: "a restaurateur says that eating at home is suboptimal". The argument about utilization is correct in itself. Its author just earns money from this conclusion.

Why this matters: the memory shortage already raised PC prices, as the Intel report described yesterday. The question "local or cloud" is now also a question about who gets the chips first.

misc

- **Claude agent in the Pentagon:** [single source] WIRED writes that a contractor connected a personal agent to a secure system of the US Department of Defense. Polymarket on X names Accenture and 2 months. I did not confirm these details.
  Wired · @Polymarket
- **AO3 went down for a week:** the largest fan fiction archive (18.8 million works) stopped after an intrusion. Volunteers ask for 2-4 weeks to check, and they say that the database most likely did not leak.
  NYT · Hacker News
- **Bunny and a provider:** bunny.net replaces TLS certificates after an intrusion at a server provider in Armenia. The company found no proof of access to its own node.
  Bunny · Hacker News
- **Game rule in Claude Code:** [single source] an X user found a new server rule: do not reproduce a specific existing game, even under another name. Anthropic did not comment on this.
  @chasmmmmmmmmmm · Hacker News
- **NES in the weights of Qwen:** Drew Breunig baked a NES CPU emulator into the weights of Qwen3.5-4B with the help of Opus 5.5. Super Mario runs, but very slowly. Simon Willison · Hacker News
- **OpenAI product graveyard:** @emollick counted: Sora lived 6 months, Agent Builder 8, Group Chat 8, Atlas 9.
  @emollick

- **Azure API Connections:** a researcher from Binary Security described five full cross-customer breaches in Azure and got $200,000 in bounties for them. The text is dated 02.10, and it reached HN only now. Binarysecurity · Hacker News

---