A Stolen Key Poisons a Coding-Agent Extension, Nine Nations Sign a Pact to Guard AI Supply Chains, and IBM and Microsoft Ship 6,000 Agents

Cursor traces a trojanised Claude Code extension to a leaked Fireworks AI key, the US and eight partners sign the Pax Silica supply-chain accord, IBM and Microsoft launch Imagine Enterprise, Fractile raises $720 million for memory-first inference chips, and AiSOC is the open-source AI security centre developers are starring.

Themes: Governance & Safety · Infrastructure & Economics

Note: this issue relies on search summaries of the original pages. Company figures are the companies' own; the Pax Silica details come from Financial Times reporting relayed by other outlets, and no official text had been published at the time of writing. Treat secondary figures as unverified.

1) One Leaked Key Poisons Cursor's Claude Code Extension

Cursor disclosed on November 13 that a credential it had issued to Fireworks AI, the inference provider whose breach was covered on Day 228, was used to attack its extension marketplace. Attackers first copied data from a read-only storage bucket Cursor shared with Fireworks. Then, on October 29, they used a stolen credential to place trojanised copies of Anthropic's Claude Code extension in Cursor's marketplace cache. The bad build was live for about 11 hours and tried to plant an SSH key and set up background persistence on Windows and macOS. Cursor estimates that at most 5,000 users who requested the cached version may have received it. It deactivated the Fireworks key, rotated credentials, moved inference off Fireworks and rebuilt its marketplace.

The concept, explained simply: This is a software supply-chain attack. Instead of hacking you directly, the attacker poisons something you trust and install automatically, here a coding-agent extension. The weak point was a cache: a local copy of popular packages kept for speed. If someone can write to the cache, every user who downloads from it gets their version, even though the original source (Open VSX, according to one report) was never infected. A coding-agent extension is a very valuable target because it runs with your permissions, can read your code and secrets, and can run shell commands.

Why it matters: Days 228-229 warned that AI platform credentials are the easiest way in; this is the follow-on harm, showing how one leaked partner key spread across companies. Coding-agent extensions now sit in the same risk class as browser extensions and npm packages. Expect signed extensions, pinned versions and per-partner, least-privilege keys to become the norm for AI developer tools.

2) Pax Silica: The US and Eight Partners Sign an AI Supply-Chain Pact

The Financial Times reported on November 15 that Japan, Australia, Israel, the UK, South Korea, India, Qatar and the United Arab Emirates signed an accord with the United States to protect AI supply chains, under a US-led initiative called Pax Silica. Germany, Sweden, the Netherlands and the European Commission were expected to sign the following week. The pact covers semiconductors, critical minerals, energy and industrial capacity, and sets up working groups to find weak points and build economic hubs. Reports say it does not create new export controls or chip-tracking rules and names no companies, and it comes ahead of a planned Trump-Xi meeting.

The concept, explained simply: An AI supply chain is everything needed to build and run AI: rare minerals, chip factories, lithography machines, power plants and data centres. A pact like this is a 'friend-shoring' agreement: countries agree to source key parts from each other and plan together, so no single rival can cut off supply. It is a coordination framework, not a law; the details come later through working groups.

Why it matters: Earlier issues covered the compute race, and Days 229-230 covered Google's specialised TPUs and Nvidia's leadership change. Pax Silica is the political layer underneath: where chips and power come from is now a matter of national security. For companies, it means AI infrastructure choices (which cloud region, which chip, which country) will increasingly carry geopolitical risk, and Gulf states are becoming central partners.

3) IBM and Microsoft Launch Imagine Enterprise with 6,000+ Agents and Assets

IBM and Microsoft announced Imagine Enterprise on November 10, a joint platform for building, deploying and governing AI agents in large companies. IBM Consulting contributes more than 6,000 prebuilt agents and assets; Microsoft supplies the technology through Microsoft Foundry, GitHub Copilot and Copilot Studio, plus its cloud, data and security services. The first industry frameworks target financial services, insurance and retail, and IBM forward-deployed engineers help clients set it up. Pricing, general availability and named production customers have not been disclosed, and the 6,000 figure counts agents and other assets together. In the same week Perplexity launched Teammates, persistent named agents for Max subscribers that keep their own memory and act on triggers from tools like Slack and GitHub.

The concept, explained simply: Think of an agent catalogue as an app store of ready-made workers: a claims-triage agent for an insurer, or a returns agent for a retailer, each pre-wired to common systems. A forward-deployed engineer is a vendor engineer who works inside the customer's company to fit those parts to real data and processes. The hard part of enterprise agents is rarely the model; it is connecting to old systems, permissions and audit, which is what consultants sell.

Why it matters: Earlier issues covered enterprise agent platforms; this deal shows the market dividing between labs that sell models and integrators that sell finished, governed agents. When evaluating such offers, ask how many of the 6,000 items are actually deployable agents, which run in production today, and who owns the logs and kill switch (Day 232). Persistent agents like Teammates show the same direction for smaller teams.

4) Fractile Raises $720 Million to Attack the Memory Wall

London-based Fractile raised a $720 million Series C at a reported $7 billion post-money valuation, according to Reuters and others, led by Thrive Capital, Founders Fund, Lightspeed and Redpoint, with backing from the UK's Sovereign AI Fund. That is roughly six times its valuation from the $220 million Series B in May. Anthropic reportedly signed an initial deal in August to buy about $250 million of Fractile's inference chips. Its first commercial chip is not due until 2027, and no independently verified performance numbers have been published.

The concept, explained simply: When a model answers a question, most of the time is spent not calculating but moving model weights from memory to the processor. This bottleneck is called the memory wall. GPUs keep weights in separate DRAM chips and fetch them over a bus, which is like a chef walking to a distant pantry for every ingredient. Fractile puts the maths units right next to fast on-chip memory (SRAM) on the same die, so the ingredients are on the counter. The trade-off: SRAM is small and expensive, so very large models must be split across many chips.

Why it matters: Inference, not training, is becoming the largest AI cost as agents run around the clock. A big valuation before any chip ships shows how badly labs want alternatives to Nvidia and to scarce, expensive high-bandwidth memory. For buyers, the lesson is to compare cost per token and latency on your own workload, not headline speed claims.

VIRAL / OPEN-SOURCE SPOTLIGHT

AiSOC: an open-source AI security operations centre

AiSOC (github.com/beenuar/AiSOC) is an MIT-licensed, self-hostable security operations centre in which AI agents investigate alerts. It ingests security telemetry, correlates related alerts into incidents, and lets agents research each one, but any response they propose (such as isolating a machine) must be approved by a human. Every prompt, tool call and line of reasoning is logged and can be replayed. It starts with a clone and 'make up' on Docker (one guide suggests 8 GB of free memory), and threat intelligence works without any API key by using CISA's public Known Exploited Vulnerabilities catalogue. One report says the security firm Cyble funds the work; one tracker put it at about 5,500 stars in late October, and it has kept appearing on trending lists in November.

Why it is taking off: Security teams are overwhelmed by alerts, and commercial 'AI SOC' products are expensive. AiSOC gives small teams a free, inspectable version they can run on their own hardware, so sensitive logs never leave the building. Its design also matches the safety patterns this series keeps returning to: agents propose, humans approve, and everything is replayable. It fits today's theme too, since an incident like Cursor's poisoned extension is exactly what a SOC must spot. Treat it as a lab tool first; check the code and test it on sample data before connecting production systems.

MARKET SIGNAL

Capital is pricing in AI as strategic infrastructure. Fractile's $7 billion valuation before a single chip ships, a government fund in its round, and a nine-nation pact on chips, minerals and power all point the same way: compute supply is now treated like energy security. Meanwhile, the Cursor incident shows the hidden cost side: every partner key and extension cache is a liability that can spread across companies. Anthropic's IPO, reported as heavily oversubscribed at its $70-$78 range, is expected to price around November 18 and begin Nasdaq trading as ANT around November 19 (press reports, not confirmed by the company). The signal: investors reward owners of scarce compute and trusted distribution, while security failures in the AI toolchain are becoming a material business risk.

PRACTICAL TAKEAWAYS

1) Treat agent extensions like code you ship.

Pin versions, prefer signed extensions and review what your coding agent's plugins can run on your machine.

2) Give each partner its own narrow key.

Issue a separate, least-privilege, rotatable credential to every vendor so one leak cannot reach your other systems.

3) Check your machine after an extension scare.

Look for unknown SSH keys in authorized_keys and new startup items, then rotate any secrets the IDE could read.

4) Map where your AI compute comes from.

List the cloud regions, chips and countries your AI workloads depend on and note any single points of geopolitical risk.

5) Question agent-catalogue numbers.

Ask vendors how many catalogue items are deployable agents, which run in production, and who controls logs and shutdown.

6) Benchmark cost per token on your own workload.

Compare new inference hardware or providers by cost and latency on your real prompts, not by headline speed claims.

7) Try an AI SOC in a sandbox.

Run AiSOC on sample logs to see how agent-led triage with human approval could fit your security process.

SELF-TEST

Q1. How did attackers deliver a malicious Claude Code extension to Cursor users?

A1. They used a stolen credential linked to the Fireworks AI breach to place trojanised copies in Cursor's marketplace cache, which served them for about 11 hours.

Q2. What is a software supply-chain attack?

A2. An attack that poisons something users trust and install automatically, such as a package, extension or update, instead of attacking each user directly.

Q3. What does the Pax Silica accord cover?

A3. Coordinated protection of AI supply chains: semiconductors, critical minerals, energy and industrial capacity, through working groups rather than new export controls.

Q4. What is the memory wall, and how does Fractile try to beat it?

A4. Inference is slowed by moving weights from off-chip DRAM to the processor; Fractile places compute next to on-chip SRAM on the same die.

Q5. Why does AiSOC require human approval for responses?

A5. Its agents only propose actions such as isolating a machine; a person approves them, and every step is logged and replayable.

TOMORROW

What We're Watching: whether Cursor publishes a full post-mortem and whether other AI tools find similar cache or partner-key exposure.

What We're Watching: Germany, Sweden, the Netherlands and the European Commission signing Pax Silica, and any official text from the State Department.

What We're Watching: Anthropic's IPO pricing around November 18 and its first day of trading as ANT.

What We're Watching: the first named production customers and pricing for IBM and Microsoft's Imagine Enterprise.

What We're Watching: independent benchmarks of Fractile's chip and whether other labs follow Anthropic as customers.

TOPICS COVERED SO FAR (RECAP)

Recent days covered agent payments with network tokens (ChatGPT Wallet), gateway kill switches (CrowdStrike), staged open-weight releases (Inkling 2 Small), non-human identity security, lightweight agent browsers (Kitesurf), fast workhorse model tiers (Gemini 4 Hydrogen), litigation as a safety brake, prompt-injection defences (IBM LLMON), websites blocking AI agents, IPO book-building, fail-closed research kill switches, specialised TPUs, the Fable/Mythos/Haiku 5.x refresh, gated cyber releases and attacks on AI developer platforms. Today added software supply-chain attacks through coding-agent extensions and caches (Cursor), friend-shoring of AI supply chains (Pax Silica), enterprise agent catalogues and forward-deployed engineers (Imagine Enterprise, Perplexity Teammates), the memory wall and SRAM inference chips (Fractile), and open-source AI security operations (AiSOC).

Sources: bleepingcomputer.com, thehackernews.com, securityonline.info and dev.to (Cursor / Fireworks extension incident); bloomberg.com and seekingalpha.com (citing the Financial Times), thenews.com.pk, polandinsight.com and mlq.ai (Pax Silica); ciodive.com, analyticsindiamag.com and runtimewire.com (Imagine Enterprise); techrepublic.com and letsdatascience.com (Perplexity Teammates); siliconangle.com, thenextweb.com, tomshardware.com and runtimewire.com (Fractile); github.com/beenuar/AiSOC, beenuar.github.io/AiSOC and explainx.ai (AiSOC). Figures are secondary reports and unverified unless from the company's own page or filing.