Agents Get a Wallet, Security Firms Build Agent Kill Switches, and Cloudflare Opens a Browser Made for Bots

OpenAI's ChatGPT Wallet lets its dots agents pay with tokenised cards and a biometric check, CrowdStrike adds an AI Kill Switch while security vendors race to buy agent-identity startups, Thinking Machines opens the weights of its 594-billion-parameter Inkling 2 Small, and Cloudflare open-sources Kitesurf, a lightweight browser engine for AI agents.

Themes: Governance & Safety · Infrastructure & Economics

Note: direct page fetches were blocked again today, so this issue relies on search summaries of the original pages. Company figures are the companies' own; deal talks and IPO demand figures are press reports and unconfirmed. Treat secondary figures as unverified.

1) ChatGPT Wallet: OpenAI's Agents Can Now Pay

OpenAI launched ChatGPT Wallet on November 13. It lets dots, the always-on proactive agents covered on Day 223, complete purchases for the user. The first rollout is narrow: US ChatGPT Pro subscribers aged 18 and over with eligible US-issued cards, and only inside dots conversations, not ordinary chats or ChatGPT Work. Visa, Mastercard and American Express are the main partners. The agent never sees the real card: payments use network tokens, full card numbers and CVVs are not stored, and every purchase must be approved with Face ID, Touch ID or a similar biometric check. OpenAI takes no commission for now, warns that agents can make mistakes, and notes that merchants and card issuers can still decline. It says users will later be able to pre-approve purchases within limits such as spending caps, categories or time windows.

The concept, explained simply: A payment token is a stand-in card number that only works for one merchant, device or purpose. If it leaks, the thief cannot reuse it elsewhere, and it can be cancelled without replacing your real card. The biometric check is a 'human in the loop' step: the agent fills the basket and prepares the payment, but money moves only when the person physically confirms. Pre-approved limits ('up to $50 on groceries this week') are the next step, a budget the agent may spend without asking each time.

Why it matters: This is the first big consumer test of agents spending money, and it copies the guardrail pattern seen in Anthropic's read-only finance feature (Day 224): narrow access, tokens instead of secrets, and explicit approval. It also collides with the lockout story from Days 230-231: Amazon and United block third-party agents, so a wallet only helps where merchants agree to let agents in. Expect card networks, not AI labs, to become the gatekeepers of agentic commerce.

2) CrowdStrike's AI Kill Switch: Containing Agents That Misbehave

At Fal.Con Europe 2026 in Barcelona, CrowdStrike unveiled an AI Kill Switch in its Falcon Guardian product. Administrators can stop a misbehaving agent, block it from relaunching on chosen machines and cut its network access, with every action logged. The first working mechanism is at the gateway: Falcon Guardian tells connected AI gateways (CrowdStrike's own Falcon AI Gateway, included at no extra cost, or a third-party one) to block the agent's access to the models it calls. Reports disagree on whether stopping agents directly on the endpoint is live or still coming soon, and SaaS-hosted agents are covered through Falcon Shield. CrowdStrike also agreed to buy Fiddler (AI observability) and Dreadnode (AI red-teaming) for a combined $495 million, according to the Wall Street Journal.

The concept, explained simply: An AI gateway is a checkpoint that every request from an agent to a model passes through, like a building's single front door. If you control the door, you can stop an agent instantly without finding every copy of it, because an agent that cannot reach its model can no longer think or act. Killing the process on the laptop or server is the second line, for agents that run local models or have already started a harmful action.

Why it matters: Day 229 covered OpenAI's 35-second Control Room for its own research workloads; this is the same idea sold to ordinary companies. The gateway design only works if all agent traffic is forced through it, so the real lesson is architectural: route every agent's model calls through one place you can switch off, and check vendor claims about endpoint termination before relying on them.

3) Thinking Machines Opens Inkling 2 Small's 594B Weights

Thinking Machines Lab, Mira Murati's startup, launched Inkling 2 Small on November 5 as a hosted-only model on its Tinker platform, then posted the weights on Hugging Face on November 12 under Apache 2.0, crediting its team and safety-testing partners. It is a sparse mixture-of-experts model with 594 billion total parameters but only 17 billion active per token, accepts images and has a 1-million-token context window. Two repositories are listed: a main NVFP4 checkpoint (about 394 GB, needing at least 500 GB of combined GPU memory) and an MXFP8 version (about 623 GB, at least 800 GB). Artificial Analysis ranks it fourth of 28 open-weight models with an Intelligence Index of 31, and one review notes DeepSeek V4.1 Flash leads on most coding rows of the lab's own model card.

The concept, explained simply: 'Small' here means small for the lab, not small for your laptop. In a mixture-of-experts model, each word is sent to only a few specialist sub-networks, so it computes like a 17B model but must store all 594B parameters in memory. NVFP4 and MXFP8 are low-precision number formats (4-bit and 8-bit) that shrink each stored parameter; fewer bits means less memory and faster maths, at a small cost in accuracy. The 'closed first, open a week later' pattern gives a lab time to watch for misuse before anyone can download and modify the model.

Why it matters: Another strong Apache-licensed model widens the open-weight field after Kimi K3.1 Pro (Day 223), giving firms a model they can host privately for data-sensitive work. But the hardware bill (a multi-GPU server) means most teams will use it through hosted providers, and the staged release is becoming a template that balances openness with a safety check.

4) The Agent-Identity Buying Spree

Security vendors are paying high prices for startups that manage what AI agents are allowed to access. Palo Alto Networks is reported (first by The Information, then TechCrunch) to be in talks to buy Clutch Security for about $500 million in cash plus milestone payments; Clutch is said to be nearing $10 million in annual recurring revenue, about 50 times revenue and roughly six times its last $85 million valuation. No deal is signed, and neither company commented. It would be Palo Alto's seventh acquisition this year. Reports also cite Cyera completing a roughly $1 billion purchase of Oasis Security and Cisco buying Astrix for about $400 million, plus CrowdStrike's $495 million pair.

The concept, explained simply: Every agent needs credentials (API keys, tokens, service accounts) to touch email, code or databases. These are 'non-human identities', and in many companies they already outnumber employees. Agent-identity tools keep an inventory of them, give each only the permissions it needs, rotate secrets and flag an agent that suddenly reaches for something new, much as HR and IT manage staff badges.

Why it matters: Days 228-229 showed that stolen credentials on AI platforms are the easiest way in. Buyers paying 50 times revenue are betting that agent identity becomes a standard budget line like antivirus. For teams, the practical question is simple: can you list every agent key you have issued, and revoke one in minutes?

VIRAL / OPEN-SOURCE SPOTLIGHT

Cloudflare Kitesurf: a browser engine built for AI agents

Cloudflare open-sourced Kitesurf under Apache 2.0 (changelog dated November 12, blog November 13). It is an experimental browser engine written in Rust and WebAssembly that loads pages, runs JavaScript and renders layouts without needing Chromium or a full browser container. It combines Blitz for HTML, DOM and layout, Stylo (Firefox's CSS engine) for styling and Boa for JavaScript. Clients connect over the Chrome DevTools Protocol, so existing Puppeteer scripts work without rewrites, and developers can run it locally or deploy it to their own Cloudflare account. Extensions are deliberately left out, the hosted Browser Run service is not part of the release, and Cloudflare says it is under heavy development.

Why it is taking off: Browsing is the most expensive part of many agents: each Chromium instance eats hundreds of megabytes, which limits how many agents a company can run at once. A lighter engine that speaks the same protocol as Chrome means cheaper, faster agent browsing with no code changes, and the release was picked up widely by developer outlets. One report claims about seven times less memory on typical agent tasks, not confirmed by Cloudflare. It also fits the day's theme: Cloudflare sits between websites and bots, so it is shaping both how agents browse and how sites decide whether to let them in. Expect sites that break without full Chromium, so test before switching.

MARKET SIGNAL

Money is flowing to the control layer around agents. Card networks are lending their tokens to ChatGPT Wallet, CrowdStrike is spending $495 million and Palo Alto may spend $500 million on agent safety and identity, and Cloudflare is giving away the browser layer to sell the network underneath. Meanwhile Anthropic's IPO is reported to be about six times oversubscribed at its $70-$78 range (Bloomberg, via Reuters; not confirmed by the company), with pricing expected around November 18 and Nasdaq trading as ANT around November 19. The signal: investors are paying up for both the frontier labs and the 'picks and shovels' that make agents safe to deploy, and the vendors who control payment, identity and the off switch will capture a growing share of agent spending.

PRACTICAL TAKEAWAYS

1) Never give an agent your real card.
Use tokenised wallets with per-purchase approval, and set low spending caps when pre-approval arrives.

2) Route every agent through one gateway.
Send all agent model calls through a single AI gateway so you can block an agent in one step.

3) Rehearse your kill switch.
Test stopping, blocking relaunch and cutting network access for one agent before you need it in an incident.

4) Inventory your non-human identities.
List every API key and service account used by agents, give each least privilege and practise revoking one.

5) Check memory before you pick an open model.
Inkling 2 Small computes like 17B but needs 500 GB of GPU memory; use a host unless you have the hardware.

6) Try a lighter agent browser in a sandbox.
Point a Puppeteer test suite at Kitesurf and compare memory, speed and page breakage against Chromium.

SELF-TEST

Q1. What protects the user's card when a dots agent pays through ChatGPT Wallet?

A1. A network token stands in for the real card number, card details are not stored, and each purchase needs biometric approval.

Q2. How does CrowdStrike's AI Kill Switch stop an agent today?

A2. It tells connected AI gateways to block the agent's access to its models, and can also block relaunch and network access, with logging.

Q3. Why can a 594B mixture-of-experts model run like a 17B one but still need huge memory?

A3. Only about 17B parameters are active per token, but all 594B must be stored and loaded.

Q4. What is a non-human identity?

A4. A credential such as an API key, token or service account used by software or an AI agent rather than a person.

Q5. Why can existing Puppeteer scripts drive Cloudflare Kitesurf?

A5. Kitesurf speaks the Chrome DevTools Protocol, the same interface Puppeteer uses to control Chrome.

TOMORROW

What We're Watching: independent tests of ChatGPT Wallet: which merchants accept dots agents and how often purchases fail.

What We're Watching: whether CrowdStrike confirms endpoint-level agent termination as live, and how rivals respond.

What We're Watching: whether Palo Alto Networks signs the Clutch Security deal.

What We're Watching: Anthropic's IPO pricing around November 18 and whether it lands above the $70-$78 range.

What We're Watching: community benchmarks of Kitesurf against Chromium and early fine-tunes of Inkling 2 Small.

TOPICS COVERED SO FAR (RECAP)

Recent days covered fast workhorse model tiers and thinking levels (Gemini 4 Hydrogen), litigation as a safety brake, structural defences against prompt injection (IBM LLMON), contract-level agent bans, key-person risk at Nvidia, IPO book-building, websites blocking AI agents, fail-closed kill switches for AI research, specialised TPUs, the Fable/Mythos/Haiku 5.x refresh, gated cyber releases, attacks on AI developer platforms and proactive agents. Today added agent payments with network tokens and biometric approval (ChatGPT Wallet), gateway-based kill switches for enterprise agents (CrowdStrike), staged open-weight releases and low-precision formats (Inkling 2 Small, NVFP4/MXFP8), non-human identity security and the agent-security M&A; wave, and lightweight agent browsers (Cloudflare Kitesurf).

Sources: cnbc.com, androidauthority.com, thenextweb.com, benzinga.com and webpronews.com (ChatGPT Wallet); crowdstrike.com blog and ir.crowdstrike.com, csoonline.com, heise.de and winfuture.de (AI Kill Switch, Fiddler/Dreadnode); thinkingmachines.ai/model-card/inkling-2-small, unite.ai, alphasignal.ai, cellcog.ai and orcarouter.ai (Inkling 2 Small); techcrunch.com, scworld.com, techzine.eu and letsdatascience.com (Palo Alto/Clutch talks); developers.cloudflare.com/changelog, phoronix.com, linuxiac.com and byteiota.com (Kitesurf); leprivatebanker.com, explainx.ai and pomegra.io (Anthropic IPO demand, citing Bloomberg/Reuters). Figures are secondary reports and unverified unless from the company's own page or filing.