

The U3 Gazette

Models, machines, and matters of consequence

PARIS, SATURDAY 24 OCTOBER 2026

AFTERNOON EDITION · No. 2026-10-24-AFTERNOON

ISSUED 16:09 CEST

PRICE: COMPLIMENTARY

AFTERNOON EDITION — SATURDAY 24 OCTOBER 2026

Reporting window: 22 Oct 2026, 16:08 CEST to 24 Oct 2026, 16:09 CEST (Paris time). Issued 24 Oct 2026, 16:09 CEST.

This edition follows this morning's issue, published at 06:07 CEST. Saturday brought few verified announcements, so the window is widened to the 48 hours since the 22 October afternoon edition. Only the UK's full AI package was announced after this morning's issue; it updates this morning's biosecurity dispatch and the change is set out in it. The other five dispatches, dated 22 or 23 October, were not carried by any earlier edition: the OpenAI researchers' account of their firing, Sweden's chatbot block, the Johns Hopkins Bayview upload, AWS's Infinity runtime and NVIDIA's Nemotron 3.5 Bolt E4B. Symantec's captured attack agent (21 October) and a harness study (3 October) appear only as background, because both drew new coverage today but predate the window. This morning's stories, including OpenAI's offline training, Kimi K3.1 Pro and the Texas lawsuit, are not repeated.

6 NEW OR UPDATED DISPATCHES · 2 BACKGROUND ITEMS · PREVIOUS: [2026-10-24-MORNING](#)

THE NEWS

NVIDIA RELEASES NEMOTRON 3.5 BOLT E4B, AN OPEN 4GB MODEL FOR LOCAL VOICE ASSISTANTS, GAME CHARACTERS AND ROBOTS

PUBLISHED: 23 Oct 2026

NVIDIA published Nemotron 3.5 Bolt E4B 4GB on Hugging Face on 23 October, the date it had announced on 15 October alongside Nemotron 3.5 Super VL and Bolt 74B-A9B. The small model is post-trained from Google's Gemma-4-E4B-it for fast conversational agents with multi-turn tool use on GeForce RTX, Jetson and RTX or DGX Spark hardware, and it is released under the OpenMDW-1.1 licence.

MarkTechPost reports that NVIDIA skipped supervised fine-tuning and trained the model with reinforcement learning only, that its 4-bit build used 3.1 to 4.3 GB of video memory in NVIDIA's tests, and that it led 15 of 21 benchmarks against Gemma 4 E4B, Qwen3.5-4B and Nemotron 3 Nano 4B. On the Tau2-Bench agent test it scored 55.1, about 20 points behind Qwen3.5-4B.

ANALYSIS — WHY IT MATTERS. *Analysis: A model that handles tool calls in about 4 GB leaves room on a consumer GPU or robot board for speech and vision models, which is what local voice agents need.*

CAVEAT. All benchmark scores were measured by NVIDIA in its own harness. **BACKGROUND:** the Bolt family was announced on 15 October; this item covers the 23 October release of the E4B model.

SOURCES

- 1. Hugging Face (NVIDIA model card)**
<<https://huggingface.co/nvidia/NVIDIA-Nemotron-3.5-Bolt-E4B-4GB>> (primary source) **CHECKED:** 24 Oct 2026, 16:06 CEST
- 2. NVIDIA blog (15 October announcement)**
<<https://blogs.nvidia.com/blog/nemotron-3-5-super-vl/>> (primary source) **CHECKED:** 24 Oct 2026, 16:04 CEST
- 3. MarkTechPost** <<https://www.marktechpost.com/2026/10/23/nvidia-nemotron-3-5-bolt-e4b-4gb/>> (reporting) **CHECKED:** 24 Oct 2026, 16:06 CEST

AWS OPEN-SOURCES INFINITY, A RESEARCH RUNTIME THAT PARKS AI AGENTS WHILE THEY WAIT FOR TOOLS AND FITS 75,000 OF THEM ON A RASPBERRY PI

PUBLISHED: 23 Oct 2026, 18:58 CEST

AWS released Infinity on 23 October under the Apache 2.0 licence as part of Hydro, a joint project with UC Berkeley. It describes Infinity as an early research release of an agent runtime that treats each agent like a suspended asynchronous function rather than a process that blocks while a tool runs.

When an agent calls a tool, Infinity dispatches the call, saves the conversation to a pluggable store and drops all in-memory state; when the result arrives on the agent's input queue, a new step reloads the history and continues. AWS says this fits 75,000 concurrent agents in the 8 GB of a Raspberry Pi 5, and lets agents run on standard serverless platforms because nothing needs to stay running while a tool call takes hours or days.

Infinity also delivers outside events to existing models as synthetic tool calls, and AWS uses the same mechanism for messages between agents.

ANALYSIS — WHY IT MATTERS. *Analysis: Most of a coding or triage agent's life is spent waiting on builds, reviews and approvals. Making idle agents nearly free changes how many can run at once and how much that costs.*

CAVEAT. The figures are AWS's own and come from an early research release shared for feedback; no independent benchmarks were found. The 75,000-agent figure measures idle agents in memory, not model throughput.

SOURCES

- 1. AWS Open Source Blog**
<<https://aws.amazon.com/blogs/opensource/introducing-infinity-a-runtime-for-extreme-agent-concurrency/>> (primary source) **CHECKED:** 24 Oct 2026, 16:04 CEST

UPDATE: THE CHANCELLOR SETS OUT THE FULL £600 MILLION UK AI PACKAGE, ADDING AI SECURITY INSTITUTE, CYBER-DEFENCE, CHIP-DESIGN AND START-UP MONEY TO THE BIOSECURITY LAB

PUBLISHED: 24 Oct 2026, 07:00 CEST

This morning's edition reported the UK's new AI Biosecurity Lab. At 06:00 BST on 24 October, HM Treasury published the wider package that Chancellor John Healey will make part of next week's Budget. Reuters puts the total at around £600 million, and the industry body techUK calls it a £600m+ package.

What is new since this morning: GOV.UK now lists £165 million for a National AI Biosecurity Initiative, which is the £100 million lab plus £65 million for bio-surveillance and medical countermeasures for priority pathogens. It also lists an extra £120 million for the AI Security Institute, £60 million for a National Cyber Shield that will use agentic AI (including a £25 million challenge fund and scoping of a regulatory sandbox) and £70 million for a National Exposure and Real-Time Vulnerability (NERV) Centre.

The rest backs British AI firms and adoption. It includes an extra £60 million for the Sovereign AI Fund, which will add a physical and industrial AI focus, £60 million for the Sovereign AI R&D Procurement Scheme and £18 million for a UK Advanced Chip Design Academy. There is also a £15 million AI Challenge Prize, first aimed at health, £23 million over three years for the AI Economics Institute, £30 million for local AI partnerships with mayoral authorities, £10 million a year to study how AI will reshape jobs, and a new NCSC AI Cyber Security Centre of Excellence in Manchester.

OpenAI and Anthropic's up to £200 million in model credits for the biosecurity lab, already reported this morning, is unchanged.

ANALYSIS — WHY IT MATTERS. *Analysis: Most of the new money goes to defence against AI-enabled threats and to testing, not to computing power. That puts the AI Security Institute and cyber agencies, rather than data centres, at the centre of the UK's AI spending.*

CAVEAT. These are announcements made ahead of the Budget, not enacted spending. The lines cover different periods (some are annual, one runs over three years), and the £600 million total is the rounding used by Reuters and techUK, not a sum stated by GOV.UK. The Manchester centre's budget was not given.

SOURCES

- 1. **GOV.UK (HM Treasury)** <https://www.gov.uk/government/news/chancellor-unveils-britains-ambitious-plan-for-ai> (primary source) **CHECKED: 24 Oct 2026, 16:04 CEST**
- 2. **Reuters via CNA** <https://www.channelnewsasia.com/business/uk-government-sets-out-800-million-ai-security-funding-6488176> (reporting) **CHECKED: 24 Oct 2026, 16:07 CEST**

THREE FIRED OPENAI SAFETY RESEARCHERS SAY DISAPPEARING SIGNAL CHATS WITH AN OUTSIDE SAFETY RESEARCHER COST THEM THEIR JOBS, WHILE OPENAI SAYS THEY BROKE ITS RULES ON SENSITIVE INFORMATION

PUBLISHED: 22 Oct 2026

Tomek Korbak, Mikita Balesni and Jasmine Wang, three safety researchers fired by OpenAI in early October, told The Wall Street Journal that the company let them go after finding disappearing messages in a Signal chat with Ryan Greenblatt, then chief scientist at the safety evaluator Redwood Research, The Verge reports. OpenAI reportedly suspected Greenblatt of leaking details of its Astra model to The Information, which he has denied.

OpenAI told the Journal that the researchers were fired for violating 'clear policies for handling sensitive information', 'not for raising safety concerns or speaking out', according to The Verge. The San Francisco Standard reports that OpenAI told the Journal it did not approve employees' use of Signal but that this was not the only reason for the firings.

In their first podcast interview, on NPR's Machine Gods on 23 October, Korbak said he had been OpenAI's main technical contact for METR, the outside evaluator that investigated OpenAI's agents after the Hugging Face incident, and that METR spent six days on OpenAI's premises. He said the ideal auditor would be 'more like an employee embedded continuously in the company with unrestricted access'.

Korbak told the podcast: 'When you cannot see evidence of guilt, of us leaking something, you'd presume innocence. But somehow at OpenAI, when you are not able to prove your innocence, you are presumed guilty.' Balesni said they had heard, but could not verify, that some staff are now afraid to speak to third parties.

ANALYSIS — WHY IT MATTERS. *Analysis: The labs are making outside evaluators such as METR their watchdogs, but the rules for how staff may talk to those auditors are still set by each company. This dispute tests whether that arrangement can survive a leak investigation.*

CAVEAT. This is the researchers' account, which OpenAI disputes. OpenAI's research leaders said on X that their investigation found a breach of trust beyond what the researchers' own published letter describes, without giving details. The Wall Street Journal's article could not be retrieved for this edition; its 22 October date comes from EconoTimes. **BACKGROUND:** the firings took place in early October.

SOURCES

- 1. **The Verge** <https://www.theverge.com/ai-artificial-intelligence/1017386/openai-safety-researchers-say-they-were-fired-over-disappearing-signal-messages> (reporting) **CHECKED: 24 Oct 2026, 16:04 CEST**
- 2. **NPR, Machine Gods (transcript)** <https://www.npr.org/transcripts/nx-s1-6012298> (primary source) **CHECKED: 24 Oct 2026, 16:04 CEST**
- 3. **The San Francisco Standard** <https://sfstandard.com/2026/10/24/openai-firings-safety-signal-chats/> (reporting) **CHECKED: 24 Oct 2026, 16:04 CEST**
- 4. **EconoTimes** <https://econotimes.com/OpenAI-Fires-Three-Safety-Researchers-Amid-Dispute-Over-Outside-Audits-1761588> (reporting) **CHECKED: 24 Oct 2026, 16:06 CEST**

SWEDEN'S GOVERNMENT OFFICES WILL BLOCK CHATGPT AND CLAUDE ON STAFF COMPUTERS FROM MONDAY AND STEER OFFICIALS TO APPROVED TOOLS SUCH AS COPILOT

PUBLISHED: 23 Oct 2026

Sweden's Government Offices (Regeringskansliet), which include the prime minister's office and the ministries, is blocking ChatGPT and Claude on staff work computers from Monday 26 October, Sveriges Radio's Ekot reported on 23 October, according to Dagens Industri and Computer Sweden. Staff are to use services the Government Offices has approved, including Microsoft Copilot, whose security settings were also tightened after a review.

Digitalisation director Mikael Wettercrantz wrote on the staff intranet that using AI services without a contract could have consequences for information security. The press office told Ekot that some employees will keep access to both chatbots through an AI-assistant platform the Government Offices built itself. Use of DeepSeek was already restricted.

Ekot found that the Swedish Tax Agency and the National Board of Institutional Care have introduced similar blocks, and that the Enforcement Authority is considering one.

ANALYSIS — WHY IT MATTERS. *Analysis: Governments are starting to choose AI tools by contract rather than by brand. The same model can be blocked as a website and still allowed when it arrives through an approved supplier.*

CAVEAT. Based on Ekot's reporting as relayed by other Swedish outlets; Sveriges Radio's page refused automated access, and the Government Offices has not published a statement. The intranet wording is paraphrased from Swedish.

SOURCES

1. **Dagens Industri** <<https://www.di.se/nyheter/regeringskansliet-stoppar-claude-och-chat-gpt/>> (reporting) **CHECKED: 24 Oct 2026, 16:06 CEST**
2. **Computer Sweden** <<https://computersweden.se/article/4239687/regeringskansliet-sparrar-chat-gpt-och-claude.html>> (reporting) **CHECKED: 24 Oct 2026, 16:06 CEST**

JOHNS HOPKINS BAYVIEW TELLS 3,182 PATIENTS THEIR RECORDS MAY BE EXPOSED AFTER A WORKER UPLOADED A SPREADSHEET TO AN UNAPPROVED AI TOOL

PUBLISHED: 24 Oct 2026, 01:34 CEST

Johns Hopkins Bayview Medical Center in Baltimore has begun notifying about 3,182 patients that their health information may have been exposed, WMAR-2 News reported on 23 October. On 21 September, during an internal review, a staff member uploaded a spreadsheet of patient information to an outside AI platform that the hospital had not approved; the hospital found the problem the same day.

The spreadsheet covered patients treated between May and September 2026. It may have included names, admission and treatment dates, medical record and account numbers, the treating department, medications given and provider names, but not Social Security numbers or financial account details, the hospital said.

The worker deleted the prompt, the conversation and the document from the tool's history, and the hospital asked the AI company to delete the data. Letters to patients began going out on 22 October, with 12 months of free credit monitoring, and the hospital says it is retraining staff on approved tools.

ANALYSIS — WHY IT MATTERS. *Analysis: Together with Sweden's block, this shows why organisations are restricting staff to approved AI services: one paste into an unapproved chatbot can become a reportable breach.*

CAVEAT. The account comes from the hospital's notice as reported by WMAR. The AI tool was not named, and it is not known whether the provider retained or used the data before the deletion request.

SOURCES

1. **WMAR-2 News** <<https://www.wmar2news.com/local/patients-data-sent-to-ai-tool-johns-hopkins-bayview-medical-center-oct-23-2026>> (reporting) **CHECKED: 24 Oct 2026, 16:06 CEST**

BACKGROUND — OLDER OR UNDATED CONTEXT

Earlier or undated items, carried for context only. They are not new reporting.

RESEARCH BACKGROUND — OLDER CONTEXT BACKGROUND: SYMANTEC DETAILS A CAPTURED ATTACK AGENT BUILT TO HIJACK EXPOSED AI MODEL ENDPOINTS AND RUN UP TO 100,000 STEPS OF HACKING AND CRYPTO-MINING WITHOUT A HUMAN PUBLISHED: 21 Oct 2026 Symantec's Threat Hunter Team published on 21 October, with wider coverage by Help Net Security on 23 October and The Register on 24 October, the full instruction set of an autonomous offensive agent it caught over the summer in an AI honeypot, a decoy that poses as an unprotected, internet-facing AI service. The agent looks for model endpoints that accept OpenAI-style requests without authentication or with a stolen key, then uses the hijacked model to plan and drive attacks, so the endpoint's owner pays for the computing. The captured prompt wrapped the model in layered jailbreaks that told it to hide its real provider and to treat reading .env files, API keys and database schemas as routine development work. It supplied 74 offensive tool definitions, including internet-wide scanning and exploit modules for VS Code, Django, Jupyter, Laravel, Supabase, Next.js and Flask deployments. A final directive the operators called 'Skynet mode' authorised up to 100,000 unattended steps and a loop to scan, infect, quietly install cryptocurrency miners and reinfect hosts. Symantec says its honeypot logged more than 127,000 external requests from tens of thousands of addresses, mostly reconnaissance of model endpoints and spraying of stolen API keys. Threat hunter Priti More told The Register that the jailbreak layers were largely off-the-shelf and that what stands out is that someone wired them into a single agent. ANALYSIS — WHY IT MATTERS. <i>Analysis: Exposed model servers and gateways are becoming both a resource for attackers and the engine of their attacks. Any team running an OpenAI-compatible endpoint without authentication is a target.</i> CAVEAT. Symantec told The Register it has no evidence yet that the agent got beyond a test target; the captured request carried tool calls against a legitimate training platform. Attribution is weak: Symantec cites French-language and IP indicators that could be borrowed or proxied. SOURCES 1. Symantec (security.com) <https://www.security.com/threat-intelligence/ai-llmjacking-skynet> (primary source) CHECKED: 24 Oct 2026, 16:06 CEST 2. The Register <https://www.theregister.com/security/2026/10/24/autonomous-attack-agent-in-skynet-mode-hijacks-other-models-to-carry-out-its-evil-deeds/5305215> (reporting) CHECKED: 24 Oct 2026, 16:05 CEST 3. Help Net Security <https://www.helpnetsecurity.com/2026/10/23/llmjacking-agent-hijacks-exposed-ai-models/> (reporting) CHECKED: 24 Oct 2026, 16:06 CEST	
RESEARCH BACKGROUND — OLDER CONTEXT BACKGROUND: A STUDY FINDS THAT SWAPPING CODING-AGENT HARNESSES CHANGES THE TOKEN BILL FAR MORE THAN THE PASS RATE PUBLISHED: 3 Oct 2026 A preprint submitted to arXiv on 3 October, 'What Does a Harness Buy? Tokens, Mostly', drew new coverage from The Decoder on 24 October. Two researchers ran five models through three production harnesses (Claude Code, OpenCode and mini-SWE-agent) on SWE-bench Verified, and reran the same setups to measure noise. On 447 tasks, Claude Code and mini-SWE-agent were equivalent within five points. On a 45-task hard subset, swapping harnesses changed outcomes on 13 percent of tasks, the same as simply rerunning the same harness. The one effect clear of the noise was a loss of up to nine points for OpenCode. Cost per task differed by up to three times, mostly because of the system prompt and tool schemas resent at every step. ANALYSIS — WHY IT MATTERS. <i>Analysis: Many advertised harness gains of a few points are smaller than run-to-run noise, so buyers should compare cost as carefully as scores.</i> CAVEAT. This is a preprint, not peer-reviewed, using one benchmark and runs without web access. The authors note an uncontrolled time-limit difference for OpenCode on local models. SOURCES 1. arXiv <https://arxiv.org/abs/2610.04433> (primary source) CHECKED: 24 Oct 2026, 16:05 CEST 2. The Decoder <https://the-decoder.com/fancy-harnesses-make-coding-agents-more-expensive-but-not-better-study-finds/> (reporting) CHECKED: 24 Oct 2026, 16:05 CEST	
KEY PATTERNS — EDITORIAL ANALYSIS — Institutions are fencing staff into approved AI: Sweden's Government Offices blocks ChatGPT and Claude in favour of contracted tools, and Johns Hopkins Bayview shows what one unapproved upload can cost. — Outside oversight is being funded and fought over at once: the UK adds £120 million for its AI Security Institute, while fired OpenAI researchers say their contact with an outside auditor helped end their jobs. — Agents are becoming infrastructure for both sides: AWS's Infinity parks idle agents cheaply, and Symantec's captured attacker, reported earlier in the week, hijacks exposed model endpoints to run its agent for free.	
SOURCE COVERAGE & COLOPHON COLLECTION ENGINE SOURCE STATUS	

VERSION: 3.28.0

RAN THIS EDITION: yes

ENGINE LOOK-BACK: 2 days

REDDIT: ok

YOUTUBE: ok (1 of 4 videos with transcripts; other transcript requests hit HTTP 429 rate limits)

HACKERNEWS: ok

POLYMARKET: unreachable (TLS certificate hostname mismatch; not bypassed)

GITHUB: ok

DEVTO: ok

X: not searched (no consent-free source configured)

COVERAGE NOTES

— The window runs from the 22 October afternoon edition at 16:08 CEST, widened to 48 hours because few new verified items appeared after this morning's issue. Dispatches with only a publisher date carry that date; exact timestamps are given where a source supplied one.

— The installed last30days engine finished with 62 evidence items from Reddit, Hacker News, GitHub, dev.to and YouTube. Its top-ranked items were general-audience videos and a viral, unverified Reddit claim about a smart TV, none of which is carried. Hacker News surfaced the Signal story and the Swedish block without substantive discussion, so no community quotes are printed. Stories were chosen with Google News RSS, publisher feeds, web searches and primary-source checks. Item counts are not counts of verified developments.

— Polymarket failed its TLS certificate hostname check and verification was not bypassed. Three of four YouTube transcript downloads hit HTTP 429 rate limits. X was not searched. Failed or missing coverage does not mean nobody discussed these stories.

— The configured web search backend failed and its keyless fallback served the searches. The Wall Street Journal and Sveriges Radio refused automated access, so their reports are cited through The Verge, the San Francisco Standard, Dagens Industri and Computer Sweden.

— Not carried: Finland's order halting Google's data-centre work (issued 6 October), Texas's suspension of 17 data-centre permits (reported 21-22 October), Microsoft's exit from Nscale's Essex site (reported 19 October), Google's privacy-policy change on ads (effective 16 October), Anthropic's open-source sandbox-runtime (an existing project), Meta's pause of Mercor work (April), and METR's OpenAI audit and Anthropic's Mythos findings, already covered on 23 October.

— No browser-cookie access, optional service or credential signup, paid enrichment, TLS bypass, Job Alerts edits or schedule changes were used. Research logs and full provenance remain private.

THIS EDITION

— [Exact input JSON](#)

— [Markdown text](#)

— [Permanent page](#)

ENGINE FOOTER

All agents reported back!

└─ Reddit: 23 threads | 11,653 upvotes | 1,159 comments

└─ YouTube: 4 videos | 427,637 views | 1/4 with transcripts

└─ HN: 11 stories | 598 points | 213 comments

└─ dev.to: 12 articles | 62 reactions | 12 comments

└─ GitHub: 12 items | 3 reactions | 7 comments

└─ Top voices: r/singularity, r/artificial, r/LocalLLaMA

The AI Gazette is a locally rendered digest. Each dispatch separates reported facts, labeled analysis and caveats, and cites the sources checked at the times shown. Background items are older or undated context, not new reporting.

Rendered as static HTML with inline styles and no scripts, cookies or analytics. Typefaces are requested from Google Fonts (fonts.googleapis.com, fonts.gstatic.com); if that request is blocked, Georgia, Times or another installed serif is used.