OpenAI's $110B raise and AWS compute deal; classified-network deployment agreement with DoW safety terms

AI News Digest

2026-02-28

OpenAI's $110B raise and AWS compute deal; classified-network deployment agreement with DoW safety terms

By AI News Digest • February 28, 2026

OpenAI disclosed a $110B funding round and a multi-year AWS partnership centered on Trainium and a "stateful runtime environment," while also announcing an agreement to deploy models in the DoW's classified network with explicit safety terms. Elsewhere, Codex adoption and agent tooling kept accelerating, alongside notable research on RL training, tool-calling inference speedups, and privacy-preserving "unlinkable inference."

OpenAI's $110B raise + Amazon partnership (compute, chips, and "stateful runtime")

OpenAI confirms a $110B funding round backed by Amazon, NVIDIA, and SoftBank

OpenAI leaders said the company raised a **$110B** round from **Amazon, NVIDIA, and SoftBank** [1][2]. In a CNBC "Squawk Pod" segment, hosts and guests also described the round as valuing OpenAI at **$730B**, with Amazon as the largest investor committing **$50B** (structured as **$15B upfront** plus

---

[1] post by @sama

[2] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

$35B tied to milestones) [3][4][5][6].

Why it matters: the scale and structure (two tranches, milestone-based) is a major signal about how capital is being deployed to secure **compute supply and strategic distribution** for frontier AI.

Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only (28:55)

AWS becomes a multi-year strategic partner, with Trainium and a "stateful runtime environment" in Bedrock

OpenAI and Amazon announced a **multi-year strategic partnership** [7]. Reporting in the same segment said OpenAI will consume **2 gigawatts of training capacity** through AWS infrastructure—some of it described as **exclusive**

---

[3]Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[4]Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[5]Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[6]Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[7]Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

to Amazon [8].

The partnership centers on a **"stateful runtime environment"** powered by OpenAI GPT models that will be available in **Amazon Bedrock** [9]. Amazon CEO Andy Jassy described it as enabling developers to access state (e.g., memory, identity) and call tools/compute "in a stateful way," claiming "there's nothing else like that today" [10].

OpenAI CEO Sam Altman also highlighted Amazon's **Trainium** as part of the relationship [11], while Jassy referenced **30–40% better price** from leveraging Trainium for training [12] and said Amazon now has "the two largest AI labs… significantly betting on Trainium" [13].

OpenAI's cloud positioning: AWS expansion alongside continued Azure exclusivity for the stateless API

Altman said OpenAI "continue[s] to have a great relationship with Microsoft," and that its **stateless API will remain exclusive to Azure**, while OpenAI will "build out much more capacity" with Microsoft [14].

Why it matters: OpenAI is explicitly describing a **split deployment model**: Azure exclusivity for one interface, while simultaneously scaling via AWS for other advanced workloads.

Pentagon / "Department of War" flashpoint: OpenAI reaches a classified-network deployment agreement

OpenAI says it reached an agreement to deploy models on the DoW's classified network—with explicit safety principles

Altman said OpenAI reached an agreement with the **Department of War (DoW)** to deploy models in its **classified network** [15][16][17]. He said the agreement includes two core safety principles: **prohibitions on domestic mass**

---

[8] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

[9] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

[10] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

[11] post by @sama

[12] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

[13] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only

[14] post by @sama

[15] post by @sama

[16] post by @sama

[17] post by @sama

surveillance and **human responsibility for the use of force**, including autonomous weapon systems—principles he said the DoW agrees with and reflects in law/policy [18][19][20].

Altman also said OpenAI will implement technical safeguards, deploy **FDEs**, and deploy "on cloud networks only" [21][22][23]. He added OpenAI is asking the DoW to offer "these same terms to all AI companies" and said OpenAI wants de-escalation "away from legal and governmental actions" toward "reasonable agreements" [24][25][26].

Context: Anthropic dispute + DPA pressure, and diverging interpretations

Commentary across sources framed OpenAI's agreement against the backdrop of the Pentagon/DoW dispute with Anthropic and reported talk of using the **Defense Production Act (DPA)** to compel access [27][28][29]. Matt Wolfe summarized the dispute as the Pentagon demanding Claude for "all lawful purposes," while Anthropic refused to lift safeguards around **mass domestic surveillance** and **fully autonomous weapons** [30].

Some reactions praised the stance as aligned with Anthropic's position (Ilya Sutskever: "It's extremely good that Anthropic has not backed down, and it's significant that OpenAI has taken a similar stance.") [31]. Others described the outcome as confusing or politically fraught, with Nathan Lambert warning it "stands to fracture the AI community" [32] and separately calling Altman's move a "weird scooping up of the DoW contract" [33].

An account labeled @UnderSecretaryF argued the DoW contract approach is grounded in "**all lawful use**," references legal authorities, and includes mutually agreed safety mechanisms—calling this a compromise that Anthropic "was offered, and rejected" [34].

Critics (including Gary Marcus) questioned whether the government effectively

---

[18] post by @sama
[19] post by @sama
[20] post by @sama
[21] post by @sama
[22] post by @sama
[23] post by @sama
[24] post by @sama
[25] post by @sama
[26] post by @sama
[27] AI News: AI's Biggest Stand Just Happened
[28] post by @Polymarket
[29] post by @nabeelqu
[30] AI News: AI's Biggest Stand Just Happened
[31] post by @ilyasut
[32] post by @natolambert
[33] post by @natolambert
[34] post by @UnderSecretaryF

offered OpenAI terms similar to what it publicly rejected from Anthropic, and raised concerns about unequal treatment and political influence [35][36][37].

Agents & coding: Codex traction, "agentic" tooling, and enterprise demand signals

Codex usage and momentum: demand spike + weekly growth signals

In the Squawk Pod discussion, Altman said demand is "rapidly growing" and cited **Codex** as having grown "like **30 something percent** in the last week," as an indicator of enterprise readiness [38]. Separately, @swyx reported OpenAI Codex added **600k users in the last 3 weeks**, reaching **1.6M WAU** on Feb 27 (up from 1M WAU on Feb 4), and said it is **>3×** up from Jan 1 (including a Feb 2 app launch) [39].

Why it matters: OpenAI is pointing to Codex as a near-term adoption barometer, while external tracking highlights unusually fast growth over a short window.

Brockman spotlights Codex 5.3's software-engineering capability on a complex task

Greg Brockman highlighted Codex 5.3 for "complicated software engineering" [40][41], pointing to a description of the model "one-shotting" a task that involved bypassing HuggingFace's KV cache abstraction, monkey-patching attention at the module level, handling M-RoPE, coordinating prompt-level memory state with KV cache state, and performing "granular surgical eviction with span tracking" [42].

Agent products continue to converge on "computer use" and long-running workflows

Matt Wolfe described new agent features across vendors, including **Cursor agents** that can control virtual computers and record videos of their actions [43], and **Microsoft Copilot Tasks** (waitlist) that aims to take plans/drafts into "completed tasks" such as building slide decks and booking appointments [44].

----

[35] post by @GaryMarcus
[36] post by @GaryMarcus
[37] post by @GaryMarcus
[38] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[39] post by @swyx
[40] post by @gdb
[41] post by @gdb
[42] post by @eigenron
[43] AI News: AI's Biggest Stand Just Happened
[44] AI News: AI's Biggest Stand Just Happened

Andrej Karpathy, reacting to Cursor usage trends, described a shifting progression of workflows ("None → Tab → Agent → Parallel agents → Agent Teams (?) → ???") and argued the practical balance is **80%** work in reliable setups and **20%** exploration of what's next [45][46].

Research & infrastructure: RL training shift, faster tool-calling inference, and privacy architecture

Scale AI: reinforcement learning becomes the majority of training work (for agents)

Scale AI's Chetan Rane said **more than half** of Scale's training work now involves **reinforcement learning (RL)**—up from **less than a quarter** six months ago [47]. The piece describes RL as goal-oriented training meant to push models from "responses" into doing things online, but notes limits around generalization and higher compute needs [48][49][50].

ContextCache: persistent KV caching for tool schemas claims large TTFT gains without quality loss

A paper and code release introduced **ContextCache**, a persistent KV cache system for tool-calling LLMs that caches prefill states for tool schemas indexed by a content hash and restores them on subsequent requests [51]. The authors report ~**200ms** cached TTFT from 5 to 50 tools versus full prefill growing from **466ms to 5,625ms**, with a **29× speedup** at 50 tools (skipping **99%** of prompt tokens per request) and "zero quality degradation" on listed metrics [52].

Paper/code: https://zenodo.org/records/18795189 [53] and https://github.com/spranab/contextcache [54]

Open Anonymity: "unlinkable inference" as a VPN-like layer for AI prompts

The Open Anonymity project described an "**unlinkable inference**" layer— framed as a "VPN for AI inference"—to reduce longitudinal user tracking by routing requests through decentralized proxies and using blind authentication

---

[45] post by @karpathy

[46] post by @karpathy

[47] Most AI Training Is Moving To Reinforcement Learning, Scale AI Says

[48] Most AI Training Is Moving To Reinforcement Learning, Scale AI Says

[49] Most AI Training Is Moving To Reinforcement Learning, Scale AI Says

[50] Most AI Training Is Moving To Reinforcement Learning, Scale AI Says

[51] r/MachineLearning post by u/PlayfulLingonberry73

[52] r/MachineLearning post by u/PlayfulLingonberry73

[53] r/MachineLearning post by u/PlayfulLingonberry73

[54] r/MachineLearning post by u/PlayfulLingonberry73

plus ephemeral keys [55]. The project also released **oa-chat**, storing chat history locally and sending each query to OpenAI under a temporary key intended to be unlinkable to other queries; Percy Liang said he switched to it as a "convenient drop-in replacement" [56].

Project links: https://openanonymity.ai/ [57] and blog: https://openanonymity.ai/blog/unlinkable-inference/ [58]

Labor + markets: layoffs attributed to AI, and competing interpretations

A Squawk Pod segment described Block's plan to reduce headcount from 10,000+ to under 6,000, and quoted Jack Dorsey saying "AI tools have changed what it means to run a company" while predicting many companies will make similar structural changes [59]. In contrast, Elad Gil argued many "layoffs due to AI" narratives are near-term corrections for **2020-era over-hiring** rather than direct AI effects, adding that most AI productivity impact is "still around [the] bend" [60][61].

Why it matters: "AI-driven layoffs" is becoming a public storyline, but even AI investors and executives are split on how much is signal versus post-hiring-cycle adjustment.

Sources

1. post by @sama
2. Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
3. post by @sama
4. post by @sama
5. post by @sama
6. post by @sama
7. post by @sama
8. AI News: AI's Biggest Stand Just Happened
9. post by @Polymarket
10. post by @nabeelqu
11. post by @ilyasut

---

[55] post by @kenziyuliu
[56] post by @percyliang
[57] post by @kenziyuliu
[58] post by @kenziyuliu
[59] Squawk Pod: OpenAI CEO Sam Altman & Amazon CEO Andy Jassy 02/27/26 |Audio Only
[60] post by @eladgil
[61] post by @eladgil

12. post by @natolambert
13. post by @natolambert
14. post by @UnderSecretaryF
15. post by @GaryMarcus
16. post by @GaryMarcus
17. post by @GaryMarcus
18. post by @swyx
19. post by @gdb
20. post by @eigenron
21. post by @karpathy
22. Most AI Training Is Moving To Reinforcement Learning, Scale AI Says
23. r/MachineLearning post by u/PlayfulLingonberry73
24. post by @kenziyuliu
25. post by @percyliang
26. post by @eladgil