

OpenAI Expands Beyond Azure, David Silver Launches a \$1.1B AI Lab, and Anthropic Updates Opus

AI High Signal Digest

2026-04-28

OpenAI Expands Beyond Azure, David Silver Launches a \$1.1B AI Lab, and Anthropic Updates Opus

By AI High Signal Digest • April 28, 2026

OpenAI broadened its cloud distribution beyond Microsoft exclusivity, David Silver launched Ineffable with a massive reported seed round, and Anthropic updated its flagship Opus model. The brief also covers agent research, Xiaomi's open release, enterprise deployments, and China's intervention in the Meta-Manus deal.

Top Stories

Why it matters: Cloud distribution, frontier funding, and flagship model releases all shifted in meaningful ways.

- **OpenAI widened its distribution options.** OpenAI said Microsoft remains its primary cloud partner, but its products and services can now be offered across all clouds; OpenAI also said it will continue providing Microsoft with models and products until 2032, with revenue share through 2030. AWS CEO Andy Jassy added that OpenAI models will be available directly on Bedrock in coming weeks alongside a Stateful Runtime Environment [1, 2].
- **David Silver's new lab launched at unusual scale.** Ineffable Intelligence said it is building a system that discovers knowledge from its own experience, while multiple posts reported a **\$1.1B** raise at a **\$5.1B** post-money valuation. The company is building in London, and Silver committed **100%** of his Ineffable equity proceeds to Founders Pledge [3, 4, 5].

- **Anthropic shipped Claude Opus 4.7.** The update was described as a coding-focused upgrade over Opus 4.6 with improved vision, a new **xhigh** effort setting, and real-world cyber safeguards. On the live GSO benchmark, Opus 4.7 was listed first at **42.2%**, ahead of Opus 4.6 and GPT-5.5 at **37.3%** [6, 7, 8].

Research & Innovation

Why it matters: The most useful research this cycle focused on making agents more reliable, more structured, and less wasteful.

- **A smaller model beat a much larger one in theorem proving.** Self-Guided Self-Play adds a Guide role that filters synthetic problems and reduces reward hacking; in Lean4, a **7B** model outperformed a **671B** baseline in fewer than 200 rounds [9, 10].
- **A major survey gave world-model research a common vocabulary.** The 40-author *Agentic World Modeling* paper proposes a **levels × laws** framework spanning **L1 predictors**, **L2 simulators**, and **L3 evolvers**, synthesizing **400+ works** and **100+ systems** across RL, web agents, video generation, and scientific discovery [11].
- **A new cost paper challenged common assumptions about coding agents.** On SWE-bench Verified, agentic coding used about **1000x** more tokens than chat or code reasoning, varied by up to **30x** across identical runs, and higher spend did not reliably improve accuracy [12].

Products & Launches

Why it matters: New launches are pushing agents deeper into real developer workflows rather than standalone demos.

- **Xiaomi open-sourced MiMo-V2.5 under MIT.** The release includes **MiMo-V2.5-Pro** for agent and coding tasks and **MiMo-V2.5** as a native omni-modal model; both support a **1M-token** context window. vLLM highlighted long-horizon execution across **1000+ tool calls** for the Pro model [13, 14].
- **Cognition launched Devin for Terminal.** It is a local coding agent that runs in the shell with full access to the codebase, tools, and environment, and can hand work off to the cloud after a laptop is closed [15, 16, 17].
- **OpenAI open-sourced Symphony for Codex.** The orchestration layer connects issue trackers such as Linear to coding agents and turns the workflow into: open issue, assign agent, generate PR, then human review [18].

Industry Moves

Why it matters: Enterprise adoption is increasingly arriving through large partnerships, not just model benchmarks.

- **Cognition partnered with Mercedes-Benz** on what it called one of the most extensive deployments of AI software engineering in the automotive industry so far [19, 20].
- **Google DeepMind expanded in South Korea** with a new AI Campus in Seoul, internships for Korean students, and collaboration with the Korean AI Safety Institute [21, 22].
- **Together AI joined the U.S. DOE's Genesis Mission with 17 national laboratories**, aimed at connecting supercomputers, facilities, and datasets to help double American scientific productivity within a decade [23, 24, 25].

Policy & Regulation

Why it matters: Governments are becoming more willing to shape cross-border AI deals directly.

- **China blocked Meta's \$2B acquisition of Manus**, citing concerns over foreign investment and the transfer of strategic AI technology to the U.S. [26].

Quick Takes

Why it matters: A few smaller updates still stood out for pricing, infrastructure, and developer tooling.

- **GitHub Copilot** moves to **usage-based billing** on **June 1** as it supports more agentic and advanced workflows [27].
- **vLLM v0.20.0** added **DeepSeek V4** support, moved to **CUDA 13 / PyTorch 2.11**, and introduced **TurboQuant 2-bit KV cache** with **4x** capacity [28, 29, 30].
- **OpenAI announced gpt-realtime-1.5** for interactive voice-controlled apps and published an open-source repo developers can fork and extend [31, 32].
- **Moonshot open-sourced Kimi K2.6**, a coding and long-horizon agent model that scales to **300 concurrent sub-agents** across **4,000 coordinated steps** [33].

Sources

1. X post by @sama
2. X post by @ajassy
3. X post by @AlexLaterre

4. X post by @etnshow
5. X post by @dejavucoder
6. X post by @dl_weekly
7. X post by @scaling01
8. X post by @scaling01
9. X post by @TheAITimeline
10. X post by @LukeBailey181
11. X post by @omarsar0
12. X post by @dair_ai
13. X post by @XiaomiMiMo
14. X post by @vllm_project
15. X post by @cognition
16. X post by @cognition
17. X post by @cognition
18. X post by @reach_vb
19. X post by @cognition
20. X post by @cognition
21. X post by @_philschmid
22. X post by @_philschmid
23. X post by @togethercompute
24. X post by @togethercompute
25. X post by @togethercompute
26. X post by @kimmonismus
27. X post by @github
28. X post by @vllm_project
29. X post by @vllm_project
30. X post by @vllm_project
31. X post by @OpenAIDevs
32. X post by @OpenAIDevs
33. X post by @dl_weekly