

OpenAI's math release meets organized pushback as Texas freezes new data-center permits

VC Tech Radar

2026-10-08

OpenAI's math release meets organized pushback as Texas freezes new data-center permits

By VC Tech Radar • October 8, 2026

Mathematicians push back on OpenAI's math results, cheaper models launch, Texas halts new data-center permits, and healthcare AI and RL-infrastructure startups raise money.

The math release gets organized pushback

The day after OpenAI published results from an unreleased frontier model, the response turned from awe to opposition. Exponential View put the release at 722 manuscripts in 372 families, averaging the equivalent of three hours of ChatGPT Pro thinking each. It noted that many results have been verified in Lean, but not all [1]. The same post floated a split into “machine mathematics,” verified but mostly read by AIs, and a smaller human “effective theory” of what people can understand [1].

Terence Tao reposted a statement from the Association for Human Mathematics [2]. It urges mathematicians to “discontinue their work with OpenAI and to return to a vision of science that centers human understanding” [3]. One commenter pointed out that the repost does not necessarily reflect Tao's own view [4].

The investor angle comes from Leo Polovets. He asks whether organizations will start “tokenmaxxing” research, for example a seed fund spending \$50K or UPS spending \$500M on tokens to find better routing algorithms [5]. His reasoning: AI's R&D impact so far has been mostly on development rather than research. If compute can produce genuinely novel results, the right AI budget for some organizations could be much larger than they now assume [6].

A related data point: P0 Research reports that frontier models' “experimental

research taste” has doubled roughly every three months since December 2025. It says Opus 5.5 now beats its expert human baseline, though most of those experts have not worked at a frontier lab [7].

Models: GPT-6 for everyone, cheaper Haiku

OpenAI is rolling out GPT-6 and “Intelligent UI” to all ChatGPT users. The new UI turns answers into interactive visuals and tools [8].

Anthropic released Claude Haiku 5.5, saying it costs about 75% less to run than Haiku 4.5 [9]. Early reviews are mixed: - **Jerry Liu:** for document parsing it costs about \$1.20 per 1,000 pages and handles tables and reading order well for that price. It is weaker on charts, semantic formatting and bounding boxes [10]. - **Bindu Reddy:** calls it “MUCH worse than DeepSeek” and worse than Luna, but gives no benchmarks [11].

Separately, Anthropic’s startup program offers up to \$7,000 in Claude credits for a year [12]. A Reddit commenter flags terms worth reading before applying. These reportedly allow Anthropic to develop competing technology, bar using the services to build competing products, and let Anthropic change the terms without notice. This is one commenter’s reading of the terms, not verified here [13].

Power is the bottleneck: Texas freezes data-center permits

a16z reports that Texas has frozen new data-center permits. The large-load queue went from 63 GW to 474 GW in 18 months, more than 5x record peak demand. Only 9.5 GW is approved and about 4.3 GW is running [14]. ERCOT has also paused approvals for data centers of 75 MW or more to switch on, including 17 that had finished every other step. The timeline is “TBD,” and full grid connection can take 5–10 years [15].

Behind-the-meter power is a bridge rather than a replacement. SemiAnalysis tracks 75 GW of equipment orders for on-site power, but sites still plan to connect to the grid because on-site power almost always costs more [15]. a16z also argues that data centers willing to cut usage at peak hours could add 100 GW to US grids without new plants [16]. For scale, US grid additions planned this year total 86 GW: 43.4 GW of that is solar and none is new nuclear [17].

a16z’s Ryan McEntush sees two ways for startups to break in: winning on integration or business model (Base Power), or bringing better technology (Heron Power’s solid-state transformers) [15]. His diligence point is that big firm orders rarely go to equipment without thousands of operating hours. So getting designed into a real project “matters far more than early deposits or revenue” [15].

The same day, a16z said Base Power raised \$2.5B. Three years in, the company says it built a battery factory in 8 months, installs 100 batteries a day and

powers more than 30,000 homes [18].

Agents move to the desktop, and gatekeepers resist

Sriram Krishnan described Microsoft’s new Windows agent features [19]: - **MXC containers:** a local sandbox where agents run. - **Hybrid routing:** actions can run on local models or be sent to cloud models. - **Copilot agents:** one demo did taxes with no remote models. - **Hardware:** a new generation starting with Surface Laptop Ultra and Nvidia RTX Spark.

Replit launched a Windows desktop preview that builds apps locally, with each build sandboxed using Microsoft Execution Containers and Nvidia OpenShell [20]. Amjad Masad cites supply-chain attacks and catastrophic agent mistakes as the risks this design addresses [21].

On consumer agents, a guest on an Alexandr Wang–hosted interview described Muse’s plan to bring agents to non-coders, whom the industry largely overlooked [22]. Muse is free with subscription tiers and is exploring small fees on purchases or savings it facilitates rather than ads [22].

Paul Graham argues that Amazon banning agents is the first opening he has seen for a startup to build an Amazon competitor [23]. More generally, he says any business that bans agents shows there is demand for a competitor that allows them [24].

Deals and funds

- **HealthLeap:** raised \$38M across Seed and Series A from Sequoia, First Round and Hummingbird. It reports 39% more appropriate diagnoses with the same staff at Cedars-Sinai, and growth from 3 to 50+ hospitals in a year [25].
- **Vitalize Care (YC W23):** raised a \$31M Series A for hospital staffing software. It says it manages millions of shifts a week across 15+ health systems, and that St. Luke’s cut overtime and agency spend 54% in 12 weeks [26].
- **Preference Model:** a16z invested. The company builds RL environments for AI research and ML engineering for leading labs, and is open-sourcing Karotte, a framework it says was hardened over more than 1M evaluation runs [27].
- **Sriram Krishnan:** raising a \$500M VC fund, per Axios via Dan Primack [28].
- **a16z Japan:** a16z opened a Japan office led by Hitoshi Yoshida, former president of Microsoft Japan and HP Japan. It is targeting AI, cybersecurity, robotics and defense [29].

Early-stage market read: Hustle Fund’s September numbers

- **Volume:** Hustle Fund passed on 1,901 of 1,922 pitches, with 21 still in progress [30].
- **Top pass reason:** incorporation, at 47%. The firm wants Delaware C-corps or Canadian or Singaporean equivalents [31].
- **Valuations:** about 70% of founders sought valuations under \$10M [32].
- **Categories:** 38% self-identified as AI/ML [33].

Tooling and science

- **Perplexity:** open-sourced pplx-embed-v2-late, multimodal embeddings in 9B and 0.6B sizes that share one embedding space. It enables OCR-free PDF search and querying on the device [34].
- **OpenDocRouter:** LlamaIndex launched a unified document-parsing API that serves models at cost plus a small transaction cut [35].
- **OCR copying allegation:** Vik Paruchuri alleged that Interfaze’s new open-weight model copied and relicensed Chandra OCR 2 [36]. This is an allegation, not established.
- **AlphaGenome:** DeepMind says it matched or beat the strongest comparison model on 25 of 26 variant-effect benchmarks [37]. Its Atlas precomputes effects for about 9 billion single-nucleotide variants and is 30x the size of the AlphaFold database [37]. Linking these molecular predictions to disease risk still needs more research [37].

Sources

1. What is left to do
2. r/artificial post by u/Eliv_nurotic
3. r/artificial comment by u/kitten_orchestra
4. r/artificial comment by u/Phytor_c
5. X post by @lplovets
6. X post by @lplovets
7. X post by @pzeroresearch
8. X post by @OpenAI
9. X post by @claudeai
10. X post by @jerryjliu0
11. X post by @bindureddy
12. r/SideProject post by u/Responsible-Fan-5442
13. r/SideProject comment by u/throwawayaccount931A
14. X post by @a16z
15. X article by @rmcentush
16. X post by @a16z
17. X post by @a16z

18. X post by @a16z
19. X post by @sriramk
20. X post by @Replit
21. X post by @amasad
22. The Biggest Launch Since ChatGPT (Muse)
23. X post by @paulg
24. X post by @paulg
25. X post by @josiahbuild
26. X post by @ycombinator
27. X post by @a16z
28. X post by @danprimack
29. X post by @RaghuRaghuram
30. X post by @dunkhippo33
31. X post by @dunkhippo33
32. X post by @dunkhippo33
33. X post by @dunkhippo33
34. X post by @AravSrinivas
35. X post by @jerryjliu0
36. X post by @VikParuchuri
37. DeepMind's New AI Just Cracked The Code Of Life