Gigascale's Hardware Bet, Nemotron's Open-Model Signal, and New Agentic Teams

VC Tech Radar

2026-06-06

Gigascale's Hardware Bet, Nemotron's Open-Model Signal, and New Agentic Teams

By VC Tech Radar • June 6, 2026

A new $250M hardware fund, YC teams in defense and manufacturing software, and fresh evidence from open models and autonomous agents point to where early AI investing is concentrating. The brief also flags shifts in token economics, legal-tech adoption, and the talent flywheel forming around London neo-labs.

Funding & Deals

- **Gigascale Capital — $250M for AI-era hardware.** Gigascale launched a $250 million first fund targeting early institutional rounds, with typical checks of $1 million to $10 million [1]. Founding partner Mike Schroepfer, former Meta CTO with hardware experience on Ray-Bans and Oculus, says he is looking for companies that are "better, faster, cheaper, but also cleaner" [1]. The firm had already made 25 investments before first close via Schroepfer's family office, including Radiant Nuclear and Fractile, pointing to a thesis around clean power and energy-efficient compute for AI and data-center demand [1].

Emerging Teams

- **9Mothers — low-cost anti-drone defense.** YC says 9Mothers is building AI mission systems for defense, starting with EDDA, a tiny robot meant to protect soldiers and critical assets from Group 1 suicide drones. The design brief is to make it small and cheap enough to deploy on every vehicle, position, and asset. YC named the founding team as rhs, Roman, and Bogdan [2].
- **Walter — legacy ERP automation for factories.** Walter is positioned as an AI employee for the manufacturing back office, logging into

the same ERP systems factories already use and automating manual work the way a human would. YC named founders Nikolas Keller and Lukas Postulka [3].

- **ONYRI Sanitize — privacy tooling before the model call.** The product anonymizes names, passwords, SSNs, health information, and other sensitive data before it is sent to AI systems [4]. The founders say they built it over two months and currently achieve a 95% detection success rate on data from the United States and France, with more language coverage in progress [4].

AI & Tech Breakthroughs

- **Nvidia's Nemotron 3 Ultra is a US open-model release worth tracking.** Perplexity says Nemotron 3 Ultra is Nvidia's new open model built for long-running agents and available to Pro and Max subscribers; Aravind Srinivas described it as "America's leading open-source model" [5, 6]. It arrives as Clouded Judgement argues that the strongest open-source models have largely come from Chinese labs and that open models are the distribution layer for the next generation of AI applications [7].
- **Weco's Aiden posted unusually strong autonomous research results.** In OpenAI's Parameter Golf competition, Aiden produced 7 of the 47 merged leaderboard records—more than 2x the next-best human—and ran autonomously for 22 straight days on a single GPU node using under 4% of visible compute. Its 28% submission acceptance rate was roughly 6x the community rate [8]. Weco says the system was built on the open-source AIDE tree-search loop, and one of the clearest signals was async human-agent collaboration: after a human shipped a new tokenizer, Aiden recombined it with its own prior work to produce the biggest validation jump of the competition [8].
- **AlphaProof Nexus reinforces the importance of the harness around the model.** In a Two Minute Papers breakdown, DeepMind's system is described as having solved 9 of roughly 350 Erdős open problems for a few hundred dollars per problem [9]. The workflow repeatedly generates candidate proofs, uses a cheaper judge model to compare and rank them tournament-style, and iterates from the highest-scoring attempts until a formal proof validates [9]. The same breakdown argues that more of the intelligence now sits in the loop around the model, while also noting that smaller models solved none of the problems [9].
- **Flow Agents push agentic workflows into hardware engineering.** Flow says its agents connect requirements, design, CAD, simulation, code, and testing so teams can surface conflicts early, propagate changes automatically, and iterate faster across aerospace, defense, robotics, energy, and autonomous-systems programs [10]. Roelof Botha's framing is that hardware teams can now access the kind of AI-driven development acceleration that software teams already expect [11].

Market Signals

- **Open-source models matter strategically, but their business model is still unsettled.** Clouded Judgement argues the US lacks a strong open-source model layer even as Chinese labs such as DeepSeek, Qwen, and Kimi lead the category, and says that matters because open models are where developers learn, experiment, and build [7]. The same essay says monetization is hard: inference hosting, fine-tuning, and eval loops all face direct competition from infrastructure vendors and hyperscalers, leaving open questions around long-term margins [7]. One path forward would be either a breakthrough from an open lab or frontier labs open-sourcing older generations of models [7].
- **Token-cost control is becoming its own product category.** Clouded Judgement argues that many workloads do not need frontier-model pricing and may shift to cheaper, "good enough" models if quality is sufficient [7]. Paul Graham said he met a startup that cuts enterprise LLM token costs by about half and splits the savings with customers, implying a TAM equal to a quarter of model companies' corporate revenue; his broader point, echoed by Garry Tan, is that incumbent execution gaps create room for upstarts [12, 13, 14].
- **Investor tolerance still centers on extreme growth.** Sarah Guo's heuristic: if an AI company is growing more than 3x YoY, investors will accept negative to low gross margins and ask fewer durability questions; if growth is only "fast," the company needs a stronger moat in models or enterprise; weak on both is "tough." Her implied scale threshold is roughly $500M+ ARR [15, 16].
- **Legal AI adoption is showing fast tipping behavior.** YC says Legora went from a 2023 founding story around Max Junestrand and two co-founders to $100M ARR in 18 months, with 1,000+ law firms and legal organizations across 50+ markets [17]. Harry Stebbings highlighted one buyer reaction as evidence of the speed of adoption:

  > "We cannot admit to clients that we JUST got Legora. That would be so embarrassing." [18]

- **The startup pipeline still looks active, and talent is recycling into new labs.** Amjad Masad noted that new business creation on Stripe is up 2x YoY [19]. Separately, Andrew Reed highlighted research and product talent leaving DeepMind for verticalized neo-labs in Kings Cross, London—especially around Biology/Chemistry and AI—and argued that keeping large incumbents domestic matters because they train and redistribute talent into the local ecosystem [20, 21].
- **Agent-first and headless software are becoming default assumptions.** Naval's view is that software platforms will be rebuilt for agent-first use cases [22]. Harrison Chase adds that "Every platform will have a headless version," while WitanLabs is explicitly building a headless Office stack for AI agents [23, 24].

Worth Your Time

- **Clouded Judgement: Where Are the American Open Source Models?** — useful for the clearest framing here on the US open-model gap, monetization constraints, and why open models shape future developer ecosystems [7].
- **Weco on Aiden and Parameter Golf** — a concrete case study in autonomous research agents, low-compute performance, and async human-agent collaboration [8].
- **Two Minute Papers on AlphaProof Nexus** — useful for the tournament-and-judge explanation of why the harness around the model is becoming a source of performance [9].

DeepMind's New AI Found A Strange New Way To Think (2:13)

- **Launch pages worth a quick skim:** 9Mothers for low-cost anti-drone defense, and Walter for brownfield ERP automation in manufacturing [2, 3].

Sources

1. AI Pushes Markets into Uncharted Territory as SpaceX & Google Seek $160 Billion Combined with OpenAI & Anthropic in the Wings
2. X post by @ycombinator

3. X post by @ycombinator
4. r/SideProject post by u/No_Computer_1247
5. X post by @perplexity_ai
6. X post by @AravSrinivas
7. Clouded Judgement 6.5.26 - Where Are the American Open Source Models?
8. r/MachineLearning post by u/Educational_Strain_3
9. DeepMind's New AI Found A Strange New Way To Think
10. X post by @buildonfloweng
11. X post by @roelofbotha
12. X post by @paulg
13. X post by @paulg
14. X post by @garrytan
15. X post by @saranormous
16. X post by @saranormous
17. X post by @ycombinator
18. X post by @HarryStebbings
19. X post by @IvanLandabaso
20. X post by @Ronanchamberss
21. X post by @andrew___reed
22. X post by @naval
23. X post by @hwchase17
24. X post by @WitanLabs