

Cheaper Frontier AI Meets a VC Shift Toward Infrastructure

VC Tech Radar

2026-09-23

Cheaper Frontier AI Meets a VC Shift Toward Infrastructure

By VC Tech Radar • September 23, 2026

OpenAI's claimed 50% API-price cut and Xiaomi's open-weight MiMo release coincide with a VC panel more bullish on AI infrastructure than new foundation-model entrants. The brief also covers World Labs' Atlas, Perplexity's agent-reliability result, Andy's enterprise network, and a16z's \$42M AI-education bet.

1. Funding & Deals

The Horowitz Andreessen Academy disclosed \$42M in total funding led by a16z, without naming a financing stage. It is a selective San Francisco school for young builders; its first offering is a tuition-free, one-year fellowship scheduled to begin in Fall 2027. Ten founding partners—including Anduril, Anthropic, Google, NVIDIA, and OpenAI—are expected to contribute compute or hardware, curriculum input, and co-op opportunities; the launch also lists 50+ hiring partners. The launch interview describes the incoming CEO as a decade-plus entrepreneur who co-founded Udemy and Maven. This is best read as a bet on AI-era talent formation and employer access, rather than a conventional AI-product financing. [1, 2, 3]

2. Emerging Teams

Andy has come out of stealth with a network-first wedge into fragmented corporate entertainment. Founder Loit Sarma co-founded an ADP-backed HR venture and spent 2.5 years at ADP overseeing money movement and payments. He says Andy has contracts with 60 large companies, more than 1,600 hospitality groups, and 93,000 venues; case studies claim 12–15% savings, while the roughly \$400M of platform volume projected by year-end remains a forecast. Its AI is positioned as a concierge and response layer: the company

says humans remain involved and agents are not yet negotiating on customers' behalf. Diligence should focus on money movement, tax jurisdictions, venue onboarding, and activating additional cities. [4]

3. AI & Tech Breakthroughs

Xiaomi's MiMo-V2.6-Pro and Flash extend the open-weight contest into natively omnimodal models. Artificial Analysis, as relayed by AINews, ranked Pro first among open-weight models on its Intelligence Index (46), at \$0.13 per index task; the report gives 1.02T total and 42B active parameters. The reported asynchronous RL stack spans coding, agents, visual work, and cyber tasks. Xiaomi says it will open-source environment code and training recipes, but not the complete 7,000+ task dataset. AINews cites \$2.6M as a conditional estimate for a 130-hour, 75B-token RL run—not an all-in training-cost disclosure. The diligence questions are whether the open toolchain can produce repeatable gains and what costs sit outside that RL run. [5]

World Labs is preparing Atlas for product release as a controllable 3D world model. Fei-Fei Li describes combining camera control and pixel consistency with image generation and reconstruction of intricate 3D structure. She says Atlas outperforms specialized state-of-the-art models, but offers no quantitative comparison in the interview. The application she emphasizes is creating realistic simulations to train and evaluate robots in data-scarce settings such as pharmaceutical manufacturing and laboratory science; she also argues that world-model benchmarking needs independent academic and public-sector participation. [6]

Agent reliability is improving through post-training, but automated evaluation still misses product judgment. Perplexity says its Computer agent learned from real user sessions by imitating good trajectories and correcting avoidable tool-call errors; a live A/B test reported 21.2% fewer tool-call failures than an earlier checkpoint. In a separate analysis of 100 production traces from an apartment-leasing assistant, automated evaluators caught obvious contradictions but missed context-dependent product failures and sometimes flagged good responses. The authors recommend human review and active sampling, making the feedback and adjudication loop—not just the model—the diligence target. [7, 8, 9]

4. Market Signals

A curated VC panel favors AI infrastructure over new foundation-model entrants. Newcomer's first sentiment report draws on interviews with 25 investors whose firms collectively manage more than \$100B. Its published excerpt describes respondents as bullish on AI infrastructure but soured on new model companies beyond the leading labs; they point to billion-dollar entry valuations, incremental Transformer variants, and the need for model vendors

to keep delivering strong new releases. Treat this as a qualitative signal from a selected panel, not a representative measure of capital allocation. [10]

The price/performance bar is moving down, and token prices alone are a poor underwriting metric. OpenAI says GPT-6 Sol and Luna bring many of Astra’s strengths into faster, more affordable models, with API prices 50% below GPT-5.6 promotional pricing; Sam Altman says they cost half as much per token and less per task. Separately, Epoch AI estimates that cost at a given performance level has fallen about 47% per quarter since 2023. These are company and research-body claims, not verified margin data. For diligence, compare realized cost per successful task and the usage response; Altman himself argues that per-task pricing is the relevant comparison. [11, 12, 13, 14]

5. Worth Your Time

- **Watch — Fei-Fei Li on Atlas.** The most useful passage connects world models to realistic robotics training and evaluation environments for data-scarce lab and pharmaceutical work, then makes the case for independent



benchmarking. [6]

Fei-Fei Li: AI's Future Is 'About Humans' (2:49)

- **Read — Lenny’s “Advanced evals.”** Practical on why teams should inspect production failures before writing metrics, and where automated trace review still needs human judgment. [9]

Sources

1. X article by @gaganbiyani
2. X post by @gaganbiyani
3. Why Education Has to Change
4. Why Corporate Entertainment Needs Its Own Operating System
5. [AINews] Xiaomi MiMo-V2.6-Pro 1T-A42B: the new top Open Weights model, trained for \$3M
6. Fei-Fei Li: AI's Future Is 'About Humans'
7. X post by @AravSrinivas
8. X post by @perplexity_ai
9. Advanced evals: How to find (and fix) hidden AI failures in your product
10. Introducing Newcomer's VC Sentiment Report: Unfiltered Insights from Silicon Valley's Top Venture Capitalists
11. X post by @OpenAI
12. X post by @sama
13. X post by @sama
14. X post by @EpochAIResearch