

MirendilAI's \$200M Seed, Coval's Voice-Agent Momentum, and Ngram's Continual-Learning Thesis

VC Tech Radar

2026-06-25

MirendilAI's \$200M Seed, Coval's Voice-Agent Momentum, and Ngram's Continual-Learning Thesis

By VC Tech Radar • June 25, 2026

This brief tracks the cycle's clearest funding signals in self-improving AI systems, voice-agent infrastructure, and enterprise AI control layers. It also highlights standout founding teams, technical bets on continual learning and low-latency inference, and market signals around tighter Series A standards and the diligence gap in AI-built products.

1) Funding & Deals

- **MirendilAI:** announced a **\$200M seed** led by **a16z** and **Kleiner Perkins**, followed by a major investment from **NVIDIA** [1, 2]. The company says it is focused on **self-accelerating AI R&D** to speed progress across science and technology while democratizing frontier capabilities beyond a small number of labs [1]. a16z describes the product as a system that trains frontier models for AI R&D and then loops over research and engineering problems with its own GPU control [2].
- **Coval:** raised a **\$28.2M Series A** for a simulation and observability platform that helps enterprises **test, monitor, and evaluate AI-powered voice agents at scale** [3]. The operating signal is stronger than a pure tooling pitch: YC says Coval processes **tens of millions of calls per month** for customers including **Perplexity** and **Deepgram** [3].
- **Runlayer:** announced a **\$30M round** from **Felicis** and **Khosla Ventures** [4]. The pitch is an enterprise AI control plane combining **enable-**

ment, security, and control in one platform [4], and Vinod Khosla explicitly framed it as “an important new category” [5].

2) Emerging Teams

- **MirendilAI’s founding bench is unusually concentrated around frontier labs.** The company says its founding team includes **20 researchers and engineers** from **Anthropic, xAI, Google DeepMind, and OpenAI** [1]. The founder announced the company alongside co-founders **Harsh Mehta, Shayan Salehian, and Tara Rezaei** [1], and a16z separately highlighted the group as one of the few teams with the experience to make an end-to-end self-accelerating system work [2].
- **Ngram combines a distinct technical thesis with a named early partner set.** In a Sequoia interview, the founders described Ngram as a company focused on **memory and continual learning** [6]. They said they are already working with **Notion, Microsoft, and Harvey** to train **per-team models** on documents and interactions over time [6].
- **Suhail’s new AI startup is still sparse on product details, but the early buildout is visible.** He said the effort started with **two 8xB200s** [7], that he had been working on **image models** [8], and that he is now letting an “**autonomous ai scientist**” work on new optimizations [9]. He also said the **seed round is done**, the **domain/name is acquired**, and he is **hiring employee #1** [10, 11, 12].
- **VentureLync is an early but notable vertical-agent bet for VC workflows.** The founder describes it as an AI operating system for funds with **three agents—Analyst, Associate, and Operations—running on a persistent memory layer** for sourcing, diligence, portfolio monitoring, and LP reporting [13]. The product is already live with funds using it, with design partners signed and more funds in active conversations [13].

3) AI & Tech Breakthroughs

- **MirendilAI is one of the clearest current bets on AI systems improving AI systems.** a16z says Mirendil is building frontier models specialized for **AI R&D** and wrapping them in a product that can make progress on research and engineering problems without human intervention [2]. Martin Casado framed the broader shift as “**AI-to-accelerate-AI-development**” becoming more broadly available [14].

“It’s like a coding agent built for AI research that controls its own GPUs.” [2]

- **Ngram’s thesis is that company context should be learned into weights, not just retrieved at inference time.** The founders said their models are “**always training**” and use adapter fine-tuning methods

such as **LoRAs** to internalize team and workspace context [6]. They also said the approach needs **white-box access to weights**, making **open-source models** the easiest fit today [6].

“It can be 100x fewer tokens.” [6]

- **Kog’s open-source Laneformer release is a clean latency signal.** Clement Delangue highlighted that Kog open-sourced the **2B Laneformer model** it used to demonstrate **3,000+ tokens per second** inference speed on Hugging Face [15].

4) Market Signals

- **Voice is emerging as one of the first productionized autonomous-agent categories.** YC said Coval’s founder discussed why **voice** is becoming the first productionized use case for autonomous agents [3]. Coval says it processes **tens of millions of calls per month** for customers including **Perplexity** and **Deepgram** [3].
- **Enterprise AI control layers are getting category-level framing from investors.** Runlayer is being pitched as a single platform for **enablement, security, and control** [4], and Khosla described that wedge as “**an important new category**” [5].
- **The Series A bar remains high.** Harry Stebbings said that a company finishing this year at **\$1.5M ARR** and next year at **\$5M ARR** is, in his view, **not enough to raise a good Series A** in the current market [16]. His framing was blunt: “**Opportunity cost of cash is real.**” [16]
- **AI-built products can reach revenue quickly and still fail diligence on code quality.** One r/SaaS example described a non-technical solo founder who used **Cursor and Claude** to get to **\$8k MRR** with real users in roughly four months [17], but technical diligence exposed **three auth implementations, 17 database tables, contradictory relationships, and no tests** [17]. After a roughly **\$28k rebuild**, the founder closed a **\$1M round** [17], reinforcing the gap between fast AI-built PMF and investor-ready software [17].
- **In creator tooling, some builders are moving from generation to research.** One early-stage founder said creators were spending hours across **TikTok, Reddit, X, podcasts, newsletters, and news sites** looking for topics [18], and concluded that **idea discovery** may be a bigger bottleneck than content production [18].

5) Worth Your Time

- **Ngram on memory and continual learning** — the best primary source here on the “always training” thesis, per-team fine-tuning, open-

weight requirements, and the claim that internalizing context can materially cut inference tokens [6].



Memory and Continual Learning: Engram's Dan Biderman and Jessy Lin (6:51)

- **Mirendil founder announcement and a16z's companion thread** — the cleanest source pair on the financing, founding team, and the product thesis around self-accelerating AI R&D [1, 2].
- **YC on Coval** — worth reading for the strongest concise case in this batch that enterprise voice agents are already a scaled evals market, plus Brooke Hopkins' Waymo-to-voice transfer story [3].
- **Kog's Laneforner 2B blog post** — a direct look at the **3,000+ tokens/second** latency claim and why latency-first model design is attracting attention [15].
- **r/SaaS on AI-built codebases and diligence** — a useful operator essay on how an AI-built MVP reached revenue fast, failed technical review, and then closed financing only after a backend rebuild [17].

Sources

1. X post by @bneyshabur

2. X post by @a16z
3. X post by @ycombinator
4. X post by @berman66
5. X post by @vkhosla
6. Memory and Continual Learning: Engram's Dan Biderman and Jessy Lin
7. X post by @Suhail
8. X post by @Suhail
9. X post by @Suhail
10. X post by @Suhail
11. X post by @Suhail
12. X post by @Suhail
13. r/artificial post by u/GlitteringEditor6671
14. X post by @martin_casado
15. X post by @ClementDelangue
16. X post by @HarryStebbing
17. r/SaaS post by u/Warm-Reaction-456
18. r/SaaS post by u/Candid_Remove_6922