

Outcome-Driven Agents Win Seed Capital as AI Inference and Seed Valuations Split

VC Tech Radar

2026-07-30

Outcome-Driven Agents Win Seed Capital as AI Inference and Seed Valuations Split

By VC Tech Radar • July 30, 2026

This brief tracks a \$15M seed round for outcome-driven customer agents, a high-pedigree team pursuing post-Transformer architectures, and the widening split between frontier AI, local inference, and seed-market valuations.

1. Funding & Deals

telli raises a \$15M seed for B2C customer operations

telli announced a \$15M seed led by redalpine, with participation from strategic investors, Mutschler, angels including Whirlpool CEO Marc Bitzer, and existing backers Cherry Ventures and Y Combinator. [1] The company sells AI agents that handle calls, lead qualification, appointment booking, follow-ups, service, and support; it says the agents already manage millions of conversations for customers including Sky, Enpal, and Vaillant. [1]

The traction claim makes this more consequential than a generic “agent for X” launch: the underwriting question is whether telli can turn deployed conversations into a durable operating advantage, while maintaining quality and margins across channels. The company is hiring across more than 15 engineering, go-to-market, and customer-operations roles. [1]

2. Emerging Teams

Core Automation is betting on a post-Transformer research company

Core Automation was founded by Jerry, a former OpenAI VP who worked on the Strawberry and reasoning teams, and Rohan, a former Gemini pre-training lead who also led fundamental research at Google Brain and worked at Anthropic. [2] The team argues that current Transformers cannot support the

kind of continual learning needed for models to learn from deployment, and is pursuing a replacement architecture through an unusually automated research lab. [2]

The company’s wedge is research throughput: it is rebuilding the deep-learning stack to increase architectural experiments from roughly one per day toward 10 and eventually 100. [2] A concrete bottleneck is kernel generation. A human-plus-search loop produced a QR kernel reported to be $60\times$ faster after roughly \$100,000 and four weeks of work, while current commercial models were described as still unable to solve the problem. [2] For investors, the relevant question is whether this team can convert exceptional research pedigree and tooling into a repeatable architecture-discovery advantage before larger labs regain appetite for non-consensus paths. The founders’ stated reason for starting outside the largest labs is that competitive pressure is pushing those labs toward scaling Transformers and coding agents rather than exploring alternatives. [2]

3. AI & Tech Breakthroughs

AI-for-science claims are moving from assistance toward discovery

Nathan Benaich reports that DeepMind’s AI co-scientist reproduced, as its top hypothesis, the answer to a bacterial gene-transfer problem that a laboratory had spent roughly 10 years solving—in two days. [3] The result is a high-value signal for scientific-agent diligence, but it is still a reported case study rather than an independently inspected benchmark; the important question is whether the system can generate experimentally useful hypotheses repeatedly across domains.

Local inference is becoming an economic alternative, not just a privacy feature

A YC Paper Club presentation reported that up to 88.7% of queries in its study could be routed to local accelerators running open models; it also reported a roughly $3\times$ improvement in intelligence per watt over two years and an $18\times$ improvement in intelligence per joule over about 16 months. [4] The presenters estimated that even imperfect routing could save 50–70% of energy, compute, and dollar cost. [4] That shifts the infrastructure opportunity toward routers, quantization, local accelerators, and heterogeneous serving—not only larger frontier-model clusters. The study’s collaborators included Nvidia, Google, Apple, AMD, OpenRouter, and SambaNova. [4]

4. Market Signals

Seed investing is splitting between high-legibility AI and overlooked companies

Carta data presented by 500 Global shows the top 5% of seed-stage companies separating sharply from the median in 2025. The presenters said AI startup

valuations are now well above the 2021 peak, while some companies in the top valuation tier have little or no revenue and are being funded on team quality and a future product thesis. [5] At the other end of the funnel, fewer than 20% of companies that raised seed in late 2023 had reached Series A after two years, against an estimated healthy benchmark of roughly 40%. [5] The practical implication is a sharper choice between paying for legible frontier-lab spinouts and hunting lower-priced, non-consensus teams; the data no longer supports treating “seed AI” as one homogeneous market.

The market is moving from copilots toward action, while model demand bifurcates

Andrew Chen characterizes the startup shift as “Copilot for X” becoming “Agent for X”: users want systems to take actions and generate outcomes rather than create more work for a human to review. [6] At the model layer, he argues that local and open-weight models are improving rapidly and may cover more than 90% of consumer and prosumer use cases, while coding, science, mathematics, and robotics remain the higher-value frontier battleground. [7] The economics are not frictionless: Martin Casado says the reported commercial agreement for Kimi includes a 30% take, because hosting a model of that size requires a sophisticated provider. [8]

Frontier labs are buying distribution in scientific workflows

OpenAI said it will provide scientists, mathematicians, and engineers free access to its frontier models, starting with 10,000 researchers and expanding to 100,000 through 2027. [9] Combined with the DeepMind result above, this points to scientific workflows becoming a strategic distribution channel for model providers, not merely a research-demo category.

5. Worth Your Time

- **Core Automation on continual learning and automated research.** Jerry and Rohan explain why they believe the current path will not remove humans from the loop, why Transformers need a replacement, and how they plan to automate architecture search. [2]



Building the Automated AGI Lab: Core Automation's Jerry Tworek and Rohan Anil (44:46)

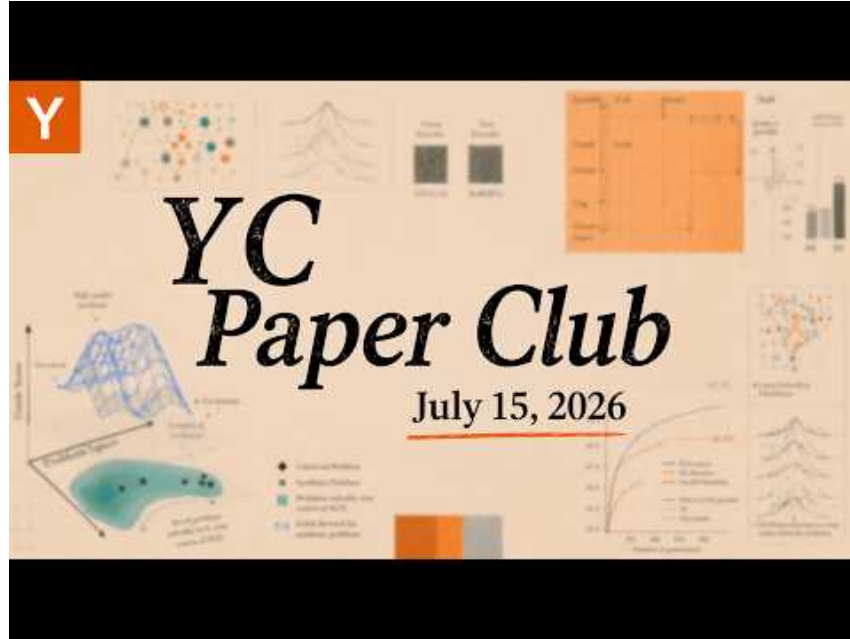
- **500 Global's early-stage VC charts.** The discussion is useful for calibrating the distance between AI seed valuations, revenue, and actual Seed-



to-Series-A conversion. [5]

Aha Charts: Counterintuitive Insights for VCs (5:46)

- **YC Paper Club on intelligence per watt and inference specialization.** The presentation connects model improvement, local accelerators, routing, and the emerging case for heterogeneous inference infrastructure.



[4]
Multi-GPU Kernels, Intelligence per Watt, Heterogeneous Inference, and More | YC Paper Club (26:31)

Sources

1. telli Raises \$18.5M in Total Funding to Build AI for Customer-Facing Operations
2. Building the Automated AGI Lab: Core Automation's Jerry Tworek and Rohan Anil
3. X post by @nathanbenaich
4. Multi-GPU Kernels, Intelligence per Watt, Heterogeneous Inference, and More | YC Paper Club
5. Aha Charts: Counterintuitive Insights for VCs
6. X post by @andrewchen
7. X post by @andrewchen
8. X post by @martin_casado
9. X post by @OpenAI