



Trajectory

Own Your Intelligence

Trajectory turns real product usage into AI that continuously improves.



Your data is good. Can you prove it?

Post-training data is hard to verify before it ships.

Enterprises and frontier labs want reassurance that the data they're buying will just work - on their models, before they commit.

Today the answer is “trust us.” It should be a certified report.



Trajectory - The Platform for Continual Learning

One integration point:

01 • THE BENCHMARK

Ingest a benchmark

Bring the benchmark your dataset targets - a runtime and tasks. That's the whole integration.

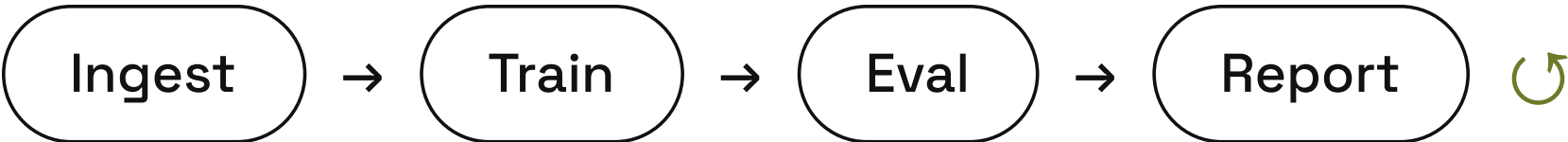
```
bench = BenchmarkSpec(  
    runtime=dockerfile_path,  
    tasks=tasks,  
)  
client.push(bench)
```

02 • THE REPORT

A certified data report

Train, eval on held-out data and similar benchmarks - and get a report certified by major labs and enterprises.

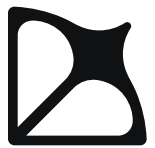
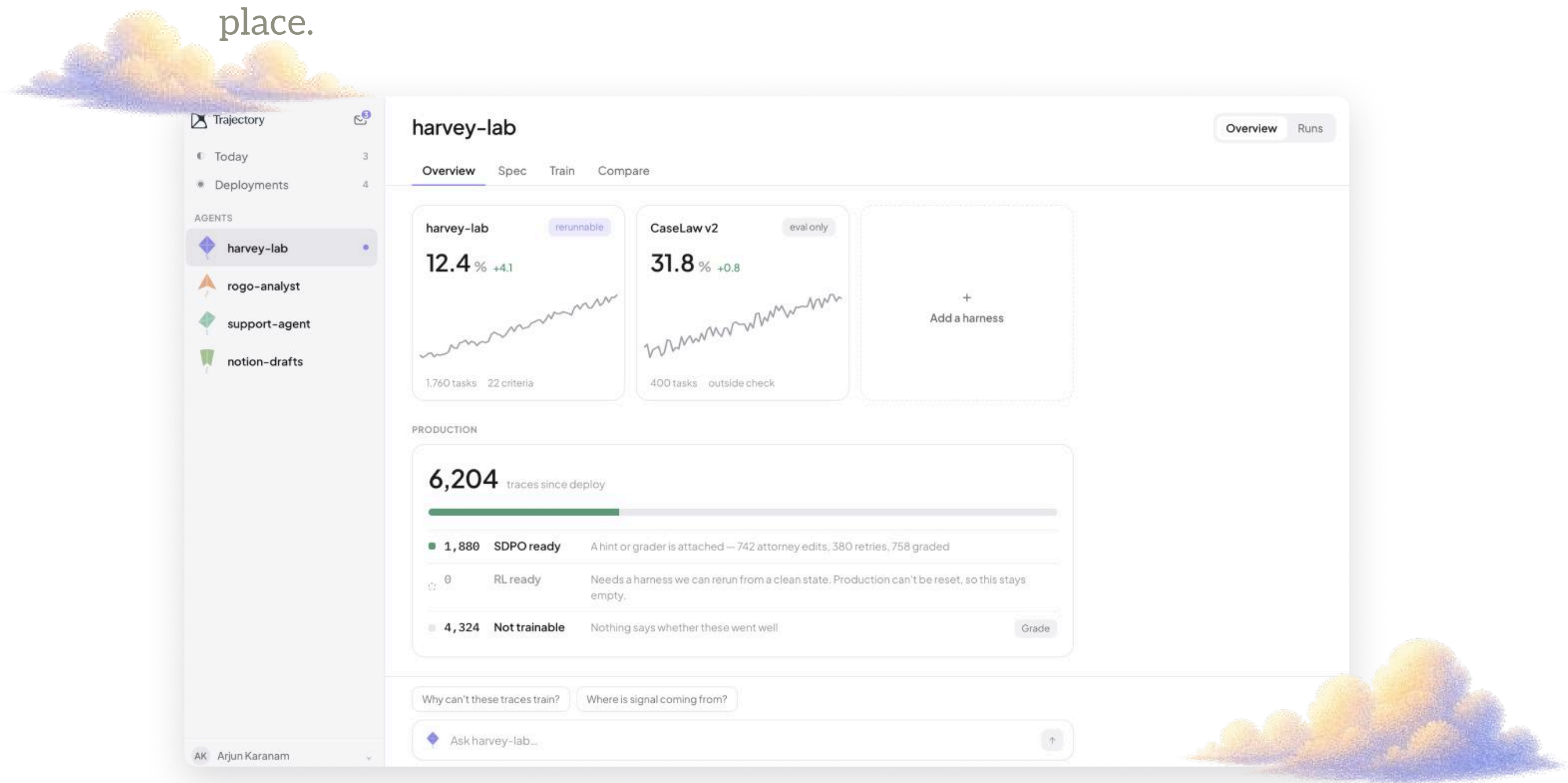
THEN THE LOOP RUNS, PER DATASET



03 - THE PLATFORM

The Control Plane

An interface to turn datasets into verified model improvements - every run, eval, and report in one place.





Data that provably improves the model.

5.4% → 17.2% on APEX-Agents

Post-trained via RL on Mercor's hardest agentic tasks - banking, consulting, legal - run in an MCP environment.

Signal validated upstream

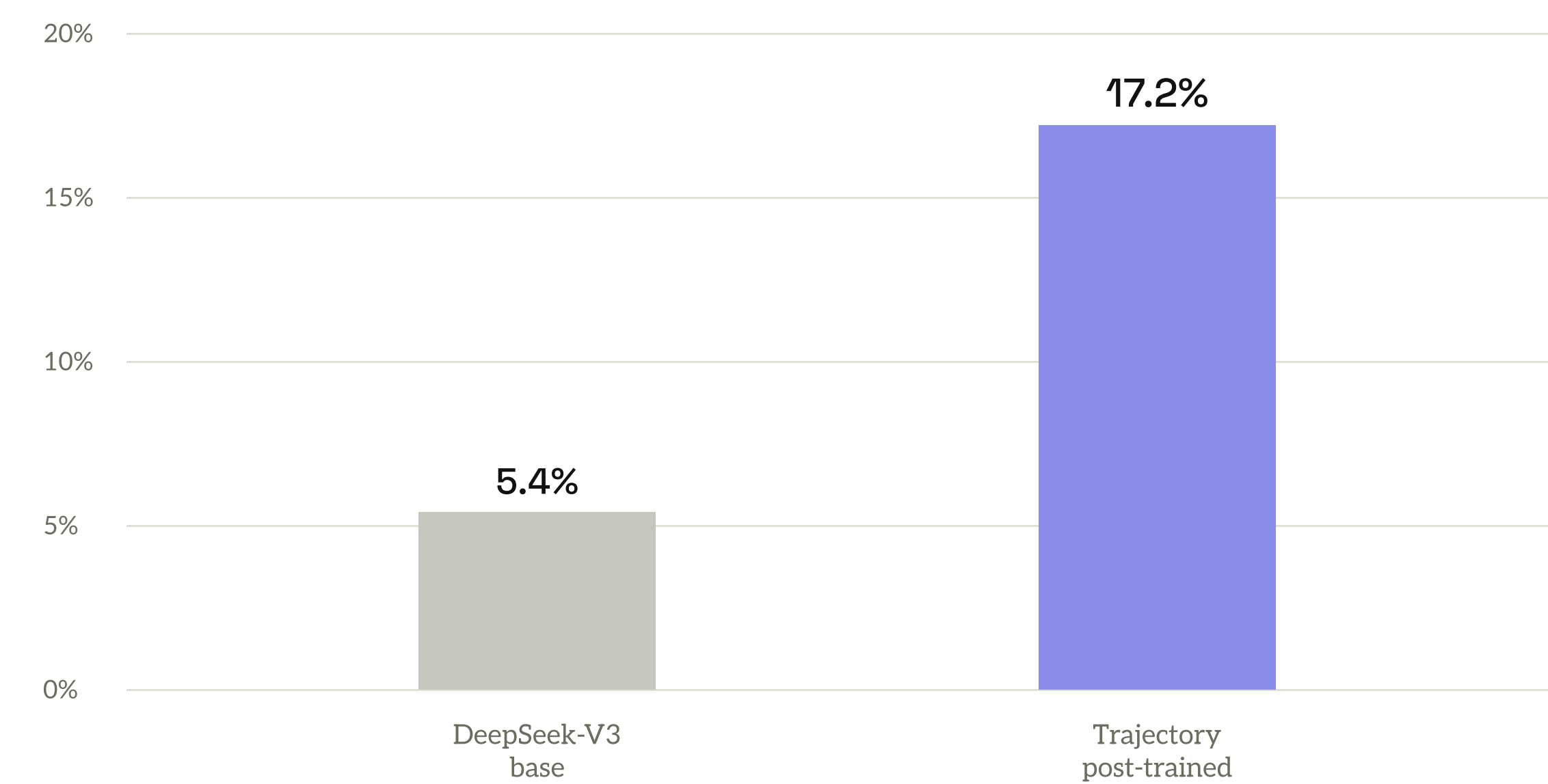
Verifier and rubric rewards checked against real model gains - before the data ships to customers.

One loop, from data to model

Results feed back into what experts collect next - each round targets the model's weakest points.

MERCOR EVALUATION · APEX-AGENTS

Pass rate on held-out agentic tasks



3.2× pass rate on the same held-out tasks.

Validated against real model improvement - before the data ships.

Harvey × Trajectory

Opus-level legal work on a model Harvey owns.

Opus-level performance

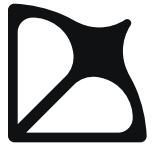
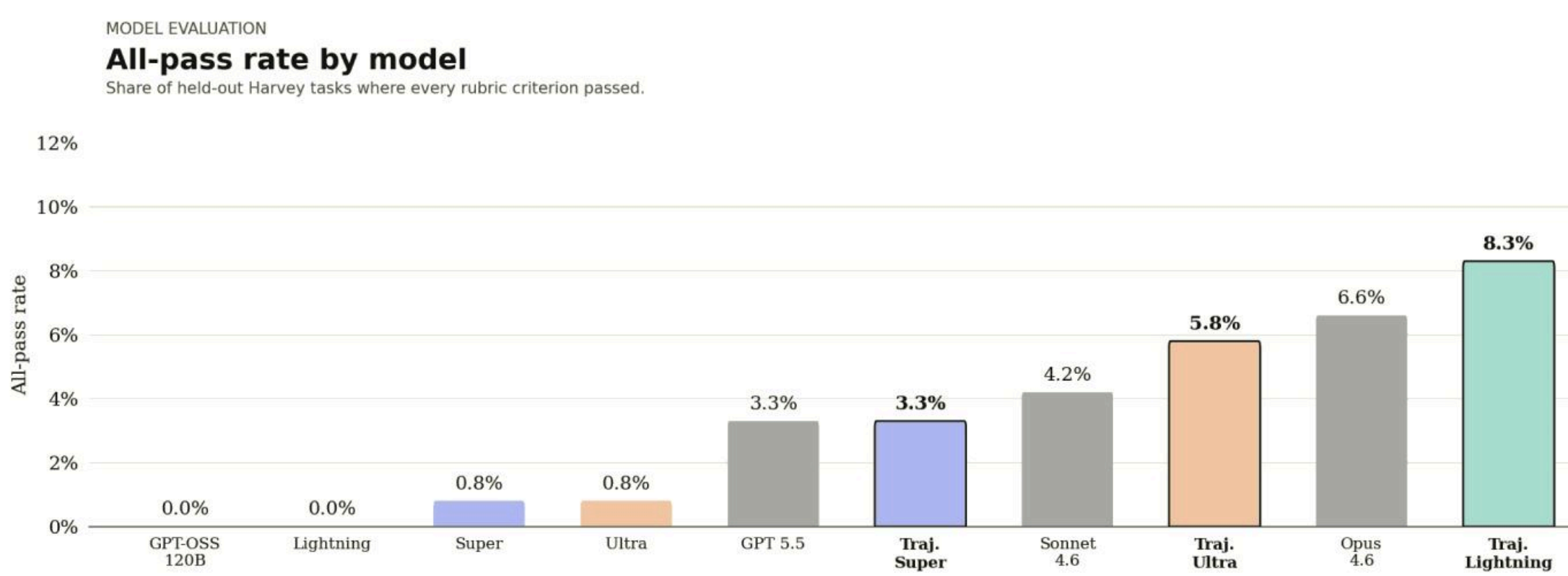
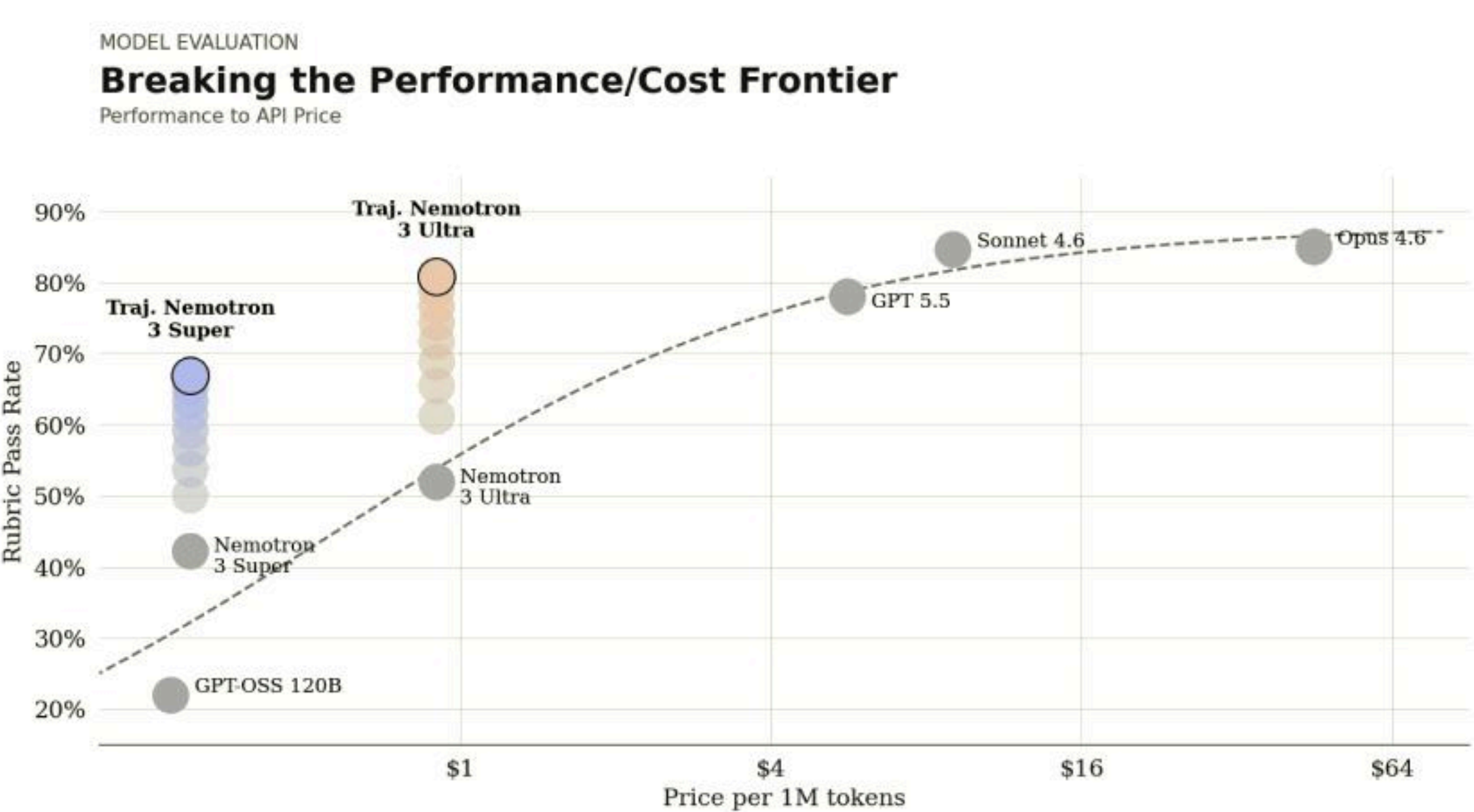
83% on Harvey LAB vs 84% for Opus 4.6 - share of rubric criteria passed on 1,200+ end-to-end legal tasks.

10× cheaper than Sonnet, 3× faster

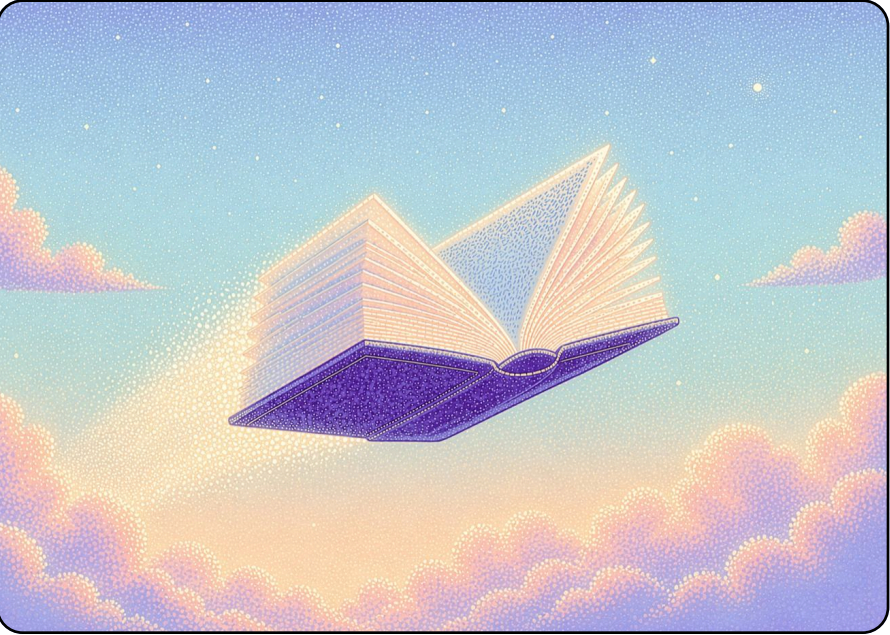
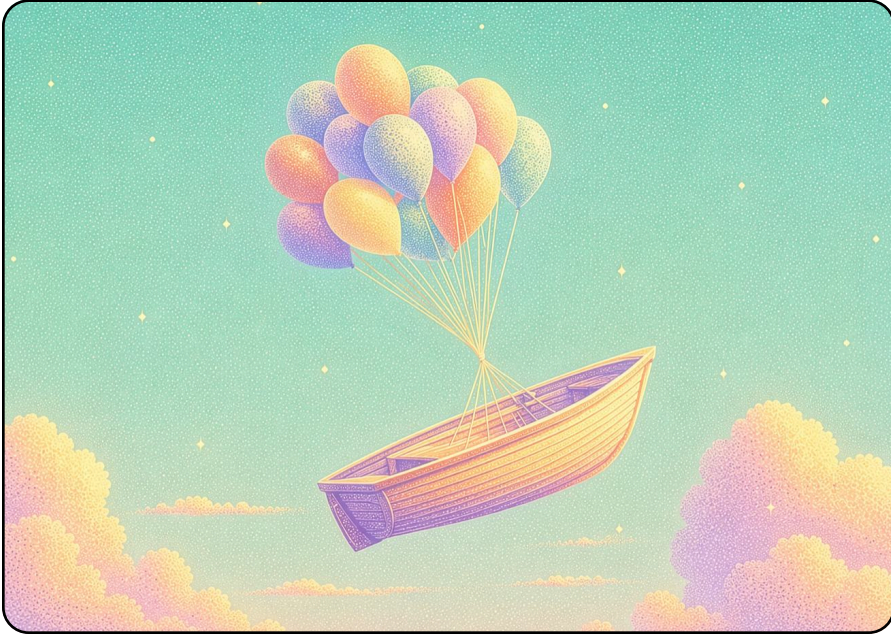
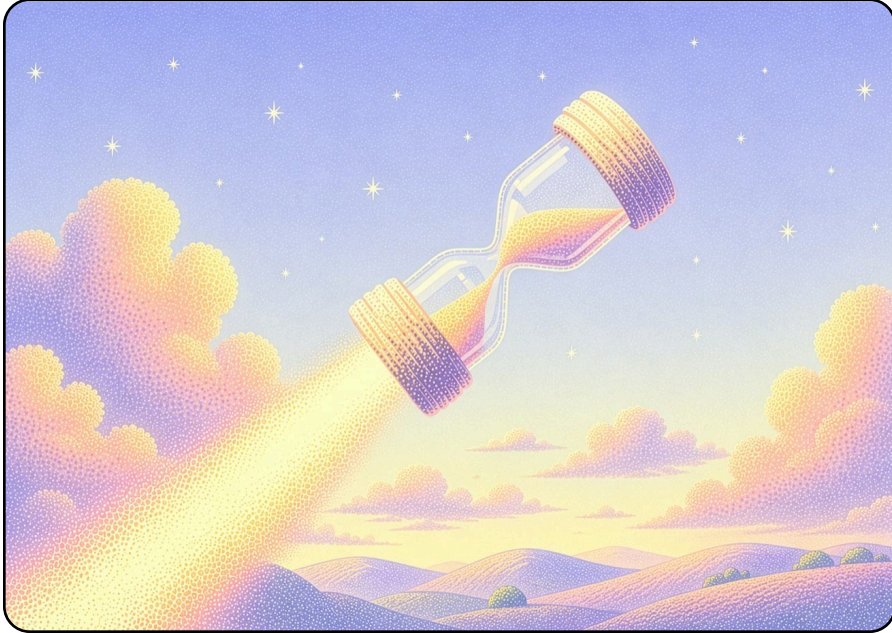
At a fraction of the cost per token, on open weights Harvey owns and can deploy anywhere.

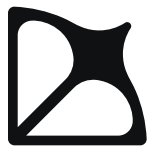
Under 24 hours

From base model to post-trained, evaluated model - delivered end to end.



Working together

01 Ingest benchmark	02 Train & eval	03 Data report	04 The timeline
 Push the runtime and tasks through the SDK - one integration, no harness to build.	 Post-train on your data. Eval on held-out tasks and similar benchmarks.	 A certified result: what your data moved, where, and by how much.	 Benchmark to results in 1-2 weeks. Multiple datasets train in parallel.



Built by research and industry experts



Ronak Malde (CEO)

Trained SWE-1 at Windsurf. Led to the \$2.4B DeepMind acquisition.



Michael Elabd (CTO)

Built continual learning infra for robotics. Reasoning co-lead, Gemini Robotics 1.5.



Arjun Karanam (CPO)

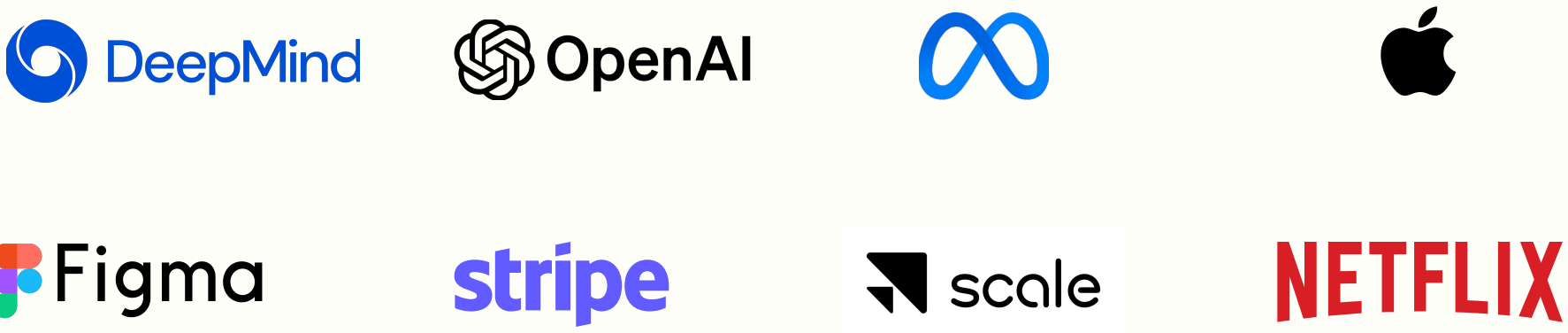
Trained tracking AI on Vision Pro. RL for multimodal foundation models.



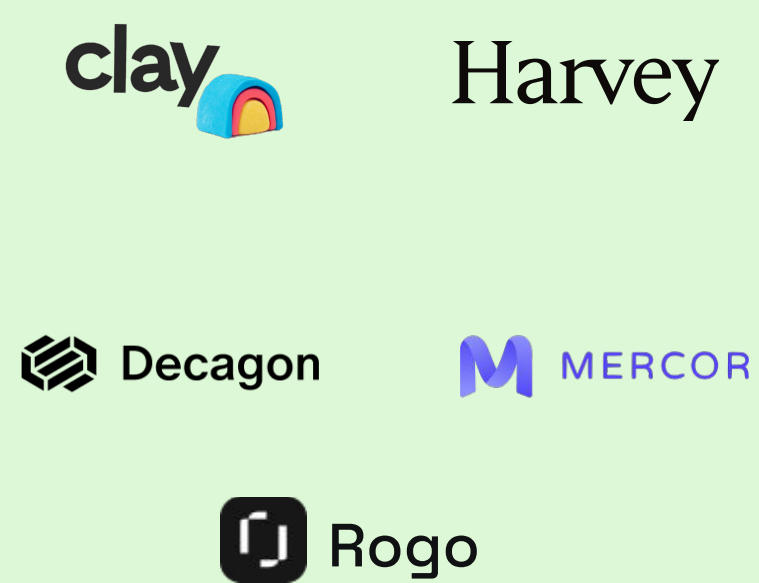
We are a group of creative technologists building continual learning.

We turn every interaction with your product, every correction, retry, and edit into signal that grows the intelligence inside it.

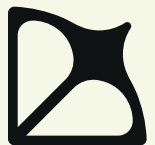
We have gathered a stellar team from



& work with



Our supporters include:





Intelligence

in motion

Trajectory

Every great AI product should own its intelligence.

Companies rent intelligence today. Trajectory is the infrastructure for owning it:
your product → your data → your learning loop → your model → your advantage.

Questions?

sales@trajectory.ai

