THE ACTIVANT *Opinion*

Did OpenAI Just Solve Robotics?

No, but GPT-6 Astra shows why robot data still matters

November 2026

AUTHOR:



Ross Nayler

Analyst | Research

ross@activantcapital.com

About Activant

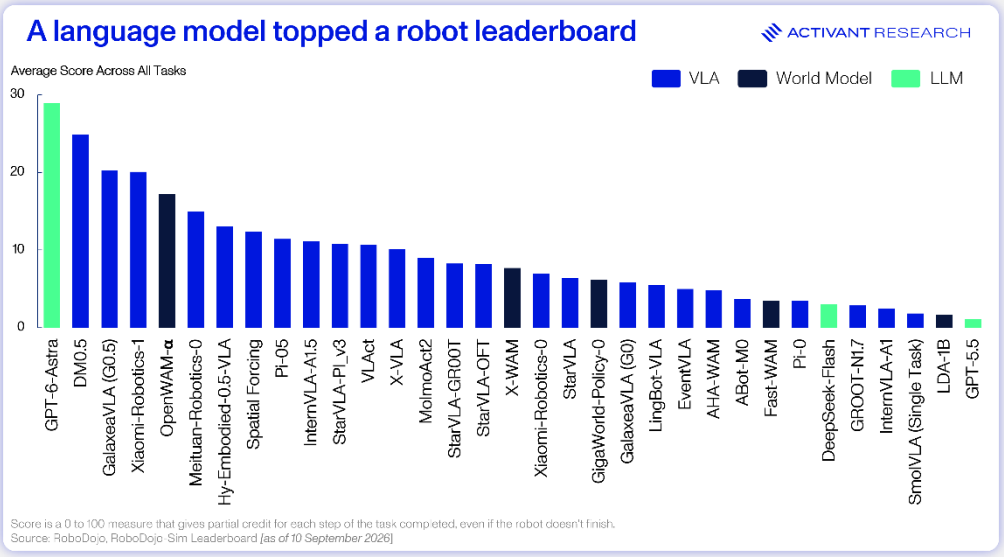
Activant is a research-led global investment firm that partners with high-growth companies. Since 2015, we have invested in category-defining businesses during their most critical phases of growth, partnering with founders who have won their initial battles and are ready for the next challenge.

Our approach pairs deep, proprietary research with patient, flexible capital and hands-on operational partnership. We work alongside founders and leadership teams to refine strategy, strengthen operations, and accelerate sustainable growth.

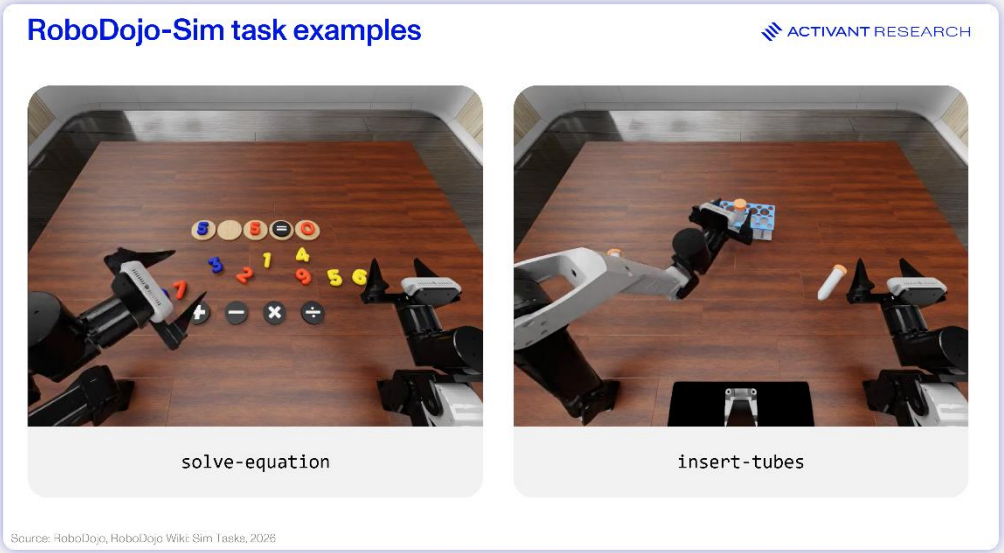
Activant Research is dedicated to uncovering the most exciting emerging technologies, sectors, and companies we believe will shape the future. Our research-driven perspective informs everything we do, helping us invest at meaningful inflection points and support founders in building enduring, category-leading businesses.

You can find out more about Activant and our research at https://activantcapital.com/.

In September, OpenAI's GPT-6 Astra ranked first on a popular robotics benchmark ahead of 40 open-weight robotic foundation models (RFMs) including Dexmal's $DM_{0.5}$ and Physical Intelligence's $\pi_{0.5}$.[1] The benchmark, RoboDojo, runs a model through 42 diverse simulated manipulation tasks, scores each one on five separate capabilities, and showcases the model rankings on a public leaderboard.[2]



The quick read is that frontier labs are coming for robotics. At a surface level, the evidence is there: when a task needs deep reasoning, like physically solving an equation, Astra succeeds 40% of the time and no robot policy succeeds at all.[3] But the average score on the benchmark hides an important split: when a task requires precise physical dexterity, like inserting glass test tubes into a rack, the best robot policy at the time (Dexmal's $DM_{0.5}$) succeeded 59% of the time and Astra *never* did.[4]



The benchmark also has a real-robot version, RoboDojo-Real, which Astra never finished: the team stopped testing after Astra damaged hardware.[5] **So, we read it the other way.** Astra shows that a robot's understanding of the world can be

outsourced to an external API, but the skill required act on that understanding cannot. We think this raises the price of what RFMs are trained on: robot data from real tasks that depend on touch and force.

Still no Common Crawl for robots

In our latest article on the data problem in physical AI, we looked at where RFMs get their training data. Anything a model can learn from the internet, like what objects are and what instructions mean, is easy to come by. Physical skills are different, because the internet contains almost no record of phenomena like gravity or friction.

Our view is the capabilities that can be learned from the internet will commoditize first, and skills that depend on sensory data beyond the visual spectrum, like touch, will retain their value. Astra looks like a version of what the internet can teach pushed as far as we've seen. OpenAI's launch post never mentioned robotics, yet Astra picks gripper positions directly with no robot fine-tuning.[6]

The benchmark's design gives us a rough, but useful check on our view:

RoboDojo-Sim task corpus

ACTIVANT RESEARCH

Axis	Number of Tasks	Focus
Generalization	12	Object manipulation tasks with *varied objects, layouts, and random variants*
Memory	6	Tasks requiring *state tracking, sequence recall, or delayed matching*
Precision	8	*Fine-grained manipulation* tasks with tight spatial constraints
Long-Horizon	8	*Multi-step* tasks that require completing *several subgoals*
Open	8	*Open-ended* or *language/image-conditioned* manipulation tasks

Source: RoboDojo, RoboDojo-Sim, 2026

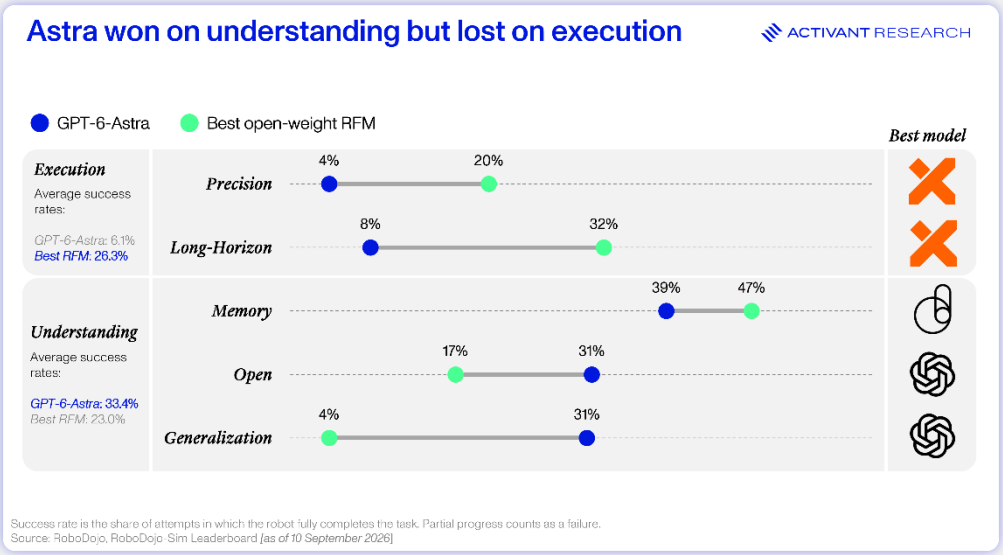
Models submitted to RoboDojo for evaluation are fine-tuned on 100 demonstrations per task, except on its *Open* axis, which is kept out of the fine-tuning data on purpose. *Open* tasks test new instructions: no model has seen demonstrations of these tasks, though skills like grasping appear in fine-tuning. *Precision*, with tasks like inserting tubes and fastening screws, tests the exact tasks every model trained on. So, based on our view: **if the web can teach understanding but not physical skill, Astra should win on *Open* and lose on *Precision***.

Astra's failure on real hardware points to a bigger question: is an LLM even the right tool for moving a robot? It is slow and costly at the job. In a separate study run by the Galbot team, what should have been 30 seconds of walking required

250 Astra calls averaging about 40 seconds each.[7] 50 RoboDojo-Sim attempts also used about 1.1 billion tokens, roughly $270 per attempt.

Astra understands but can it do?

While all of the RoboDojo axes involve physical execution, if you partition them by what makes each one hard (*understanding the task* or the *actual execution of the task*), Astra's lead turns into two very different results:



On the three understanding axes, *Open*, *Generalization* and *Memory*, Astra averages a 33.4% success rate against 23.0% for the best public policy on each axis. On the two execution axes, *Precision* and *Long-Horizon*, it averages 6.1% against 26.3%, more than four times lower.[8] **Astra's 3-point overall lead is a modest win on understanding added to a large loss on execution.** When RoboDojo randomizes lighting, clutter and backgrounds, Astra's Generalization score falls 11%, while Physical Intelligence's $\pi_{0.5}$ falls 72%.[9] This gap is a perfect example of the internet's contribution in training.

To be fair to OpenAI, comparing Astra with no robot-specific fine tuning to actual RFMs on *Precision* is slightly unfair, and we would like to see what happens when a frontier LLM like Astra sees thousands of hours of fine-tuning data.

Galbot's study also paired Astra with a purpose-built RFM. The team let $\pi_{0.5}$ propose motions and had Astra review them. $\pi_{0.5}$ supplied 85.6% of the control steps and Astra generated or corrected the other 14.4%. The pair succeeded on 48% of attempts, against 26% for Astra alone and about 16% for $\pi_{0.5}$'s published score on the same tasks.[10] **The internet-scale LLM worked as the high-level supervisor, while the robot data trained RFM still handled low-level execution.**

We expect the easier tasks, the ones a camera alone can judge, to become cheap to train for, mostly through new ways of turning everyday video into robot training data. This is now happening from a slightly different direction: a general-purpose

AI model anyone can pay to use. So, video data for those easier tasks is worth less, because a lab can now rent the same ability and pay per use.

The serious counter argument is speed. In July, Anthropic reported that no frontier model completed more than 5.5% of LIBERO manipulation tasks; two months later, Astra topped RoboDojo.[11,12] We doubt precision rides that curve. On in-hand rotation, task-specific policies kept the object on target for 76.9% of steps, and Astra for 0.51%.[13] Given a one-shot demonstration, Astra's RoboDojo success fell from 22.9% to 17.9%, and to 12.9% when the demo was written as text.[14] That matters because a demonstration is the format most robot data comes in. Meanwhile the specialists keep climbing, and Astra has since slipped to tenth on the RoboDojo leaderboard.

Profits move to what is still hard

Christensen's Law of Conservation of Attractive Profits states that when one layer of a stack becomes modular and good enough, profits move to the next layer that still limits performance.[15] We think Astra is proof that the semantic layer of robot intelligence can be turned into a metered API. Every lab can rent the same understanding of a scene, but none can rent the motor skill to act on it. That skill comes from robot experience, real or simulated.

The rental does have its limits today. Astra is too slow to drive a robot on its own, so for now, labs can rent it (or other frontier LLMs) as a supervisor, while a fast model on the robot does the moving.

Because scene understanding is now something anyone can rent, it counts for less. What counts more is access to data from tasks that depend on touch and force, plus owning robots already at work that keep producing more of it. Companies with both can rent a better brain and maintain ownership of the part that is scarce.

A few companies are already building around this. Genesis AI and Tacta Systems pair dexterous robot hands with data-collection gloves that capture how skilled people work.[16,17] RLWRLD's model learns from the touch and force signals that cameras miss, and Tangent Robotics is building hands and touch sensors for precise factory work.[18,19]

OpenAI's own spending points in the same direction: their San Francisco lab reportedly runs around 100 data collectors across three shifts, rated on how many "good hours" of data they produce.[20] Whether OpenAI used any of that data in Astra's training is an open question, but either answer supports our view: if they didn't, the lab that topped a benchmark is still paying to make robot data; if they did, Astra still trails the best RFMs by 4X on execution after more than a year of collecting it.

If we are right, contact-rich robot data holds its value while videos of tasks a camera can judge get cheaper. Our view breaks if a frontier model like Astra, fine-tuned on a fraction of the robot data that robot foundation models use, starts beating them on precise physical work. In that world, robot data is still needed but worth much less, because general intelligence does most of the work. Until one does, the scarce input in physical AI is still robot data, and Astra has showed its value.

---

[1] Zhang et al., An Unexpected Robot Policy: Early Evaluations of GPT-6 Astra on RoboDojo and Beyond, 2026

[2] Chen et al., RoboDojo: A Unified Sim-and-Real Benchmark for Comprehensive Evaluation of Generalist Robot Manipulation Policies, 2026

[3] Zhang et al., An Unexpected Robot Policy: Early Evaluations of GPT-6 Astra on RoboDojo and Beyond, 2026

[4] Ibid

[5] Ibid

[6] OpenAI, GPT-6 Astra: A new generation of intelligence, 2026

[7] Galbot, Systematically Exploring the Capabilities of GPT-6 Astra as Embodied Policies, 2026

[8] Activant analysis of RoboDojo-Sim benchmark statistics

[9] Chen et al., RoboDojo-Sim Leaderboard, 2026

[10] Galbot, Systematically Exploring the Capabilities of GPT-6 Astra as Embodied Policies, 2026

[11] Liu et al., LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning, 2023

[12] Anthropic, Claude plays robotics, 2026

[13] Galbot, Systematically Exploring the Capabilities of GPT-6 Astra as Embodied Policies, 2026

[14] Chen et al., RoboDojo: A Unified Sim-and-Real Benchmark for Comprehensive Evaluation of Generalist Robot Manipulation Policies, 2026

[15] Christensen, The Innovator's Solution: Creating and Sustaining Successful Growth, 2003

[16] Genesis AI, GENE-26.5: Advancing robotic manipulation to human level, 2026

[17] Tacta Systems, Why We're Building Dexterous Intelligence, 2026

[18] RLWRLD, RLWRLD Unveils RLDX-1, a Dexterity-First Foundation Model for Humanoid Robots, 2026

[19] Tangent Robotics, Towards Autonomous Dexterity, 2026

[20] Business Insider, Inside OpenAI's Robotics Lab, 2026

Disclaimer: The information contained herein is provided for informational purposes only and should not be construed as investment advice. The opinions, views, forecasts, performance, estimates, etc. expressed herein are subject to change without notice. Certain statements contained herein reflect the subjective views and opinions of Activant. Past performance is not indicative of future results. No representation is made that any investment will or is likely to achieve its objectives. All investments involve risk and may result in loss. This newsletter does not constitute an offer to sell or a solicitation of an offer to buy any security. Activant does not provide tax or legal advice and you are encouraged to seek the advice of a tax or legal professional regarding your individual circumstances.

This content may not under any circumstances be relied upon when making a decision to invest in any fund or investment, including those managed by Activant. Certain information contained in here has been obtained from third-party sources, including from portfolio companies of funds managed by Activant. While taken from sources believed to be reliable, Activant has not independently verified such information and makes no representations about the current or enduring accuracy of the information or its appropriateness for a given situation.

Activant does not solicit or make its services available to the public. The content provided herein may include information regarding past and/or present portfolio companies or investments managed by Activant, its affiliates and/or personnel. References to specific companies are for illustrative purposes only and do not necessarily reflect Activant investments. It should not be assumed that investments made in the future will have similar characteristics. Please see "full list of investments" at activantcapital.com/companies/ for a full list of investments. Any portfolio companies discussed herein should not be assumed to have been profitable. Certain information herein constitutes "forward-looking statements." All forward-looking statements represent only the intent and belief of Activant as of the date such statements were made. None of Activant or any of its affiliates (i) assumes any responsibility for the accuracy and completeness of any forward-looking statements or (ii) undertakes any obligation to disseminate any updates or revisions to any forward-looking statement contained herein to reflect any change in their expectation with regard thereto or any change in events, conditions or circumstances on which any such statement is based. Due to various risks and uncertainties, actual events or results may differ materially from those reflected or contemplated in such forward-looking statements.