

# The Next AI Revolution

## Robotics And The World It Will Remake

In November 2022, a large language model (LLM) fine-tuned for chat and wrapped in a clean user interface became the fastest product ever to reach one million users. You probably know it as ChatGPT.

An enormous capital investment in AI and data centers followed as companies sought to cash in, with the large hyperscalers spending \$470 billion last year alone and pledging a nearly 60% increase in spending for 2026 (see Figure 1).

AI skeptics remain convinced that the chatbot phenomenon will soon flame out, with revenues never materializing to justify the excessive capex spent, let alone enable hyperscalers to eke out profits.

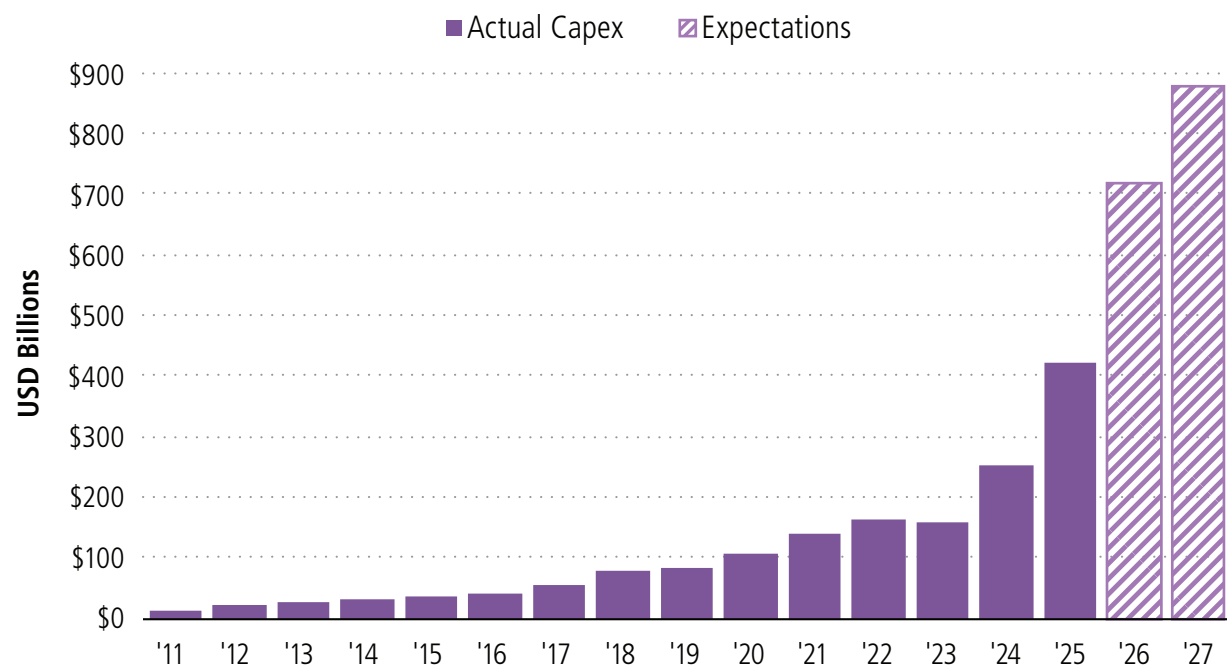
**«AI CHATBOTS AND AUTONOMOUS AGENTS ARE JUST THE BEGINNING.»**

We’ve got bad news for the bears: AI chatbots and autonomous agents are just the beginning.

We’re preparing for the next wave—this time in “physical AI”—enabled by the data and compute flywheel that developed chatbots. Autonomous robot manipulation models, the emerging class of systems capable of perceiving, reasoning, and acting in the real world, are the key technology for deploying AI in the physical world.

While many obstacles remain, recent breakthroughs in autonomous robots make

fig 1. CAPEX CREATION: ANNUAL HYPERSCALER\* CAPITAL EXPENDITURES



Sources: Bloomberg, Payden Calculations

\*Top five largest hyperscalers, which includes Alphabet, Amazon, Apple, Meta, and Microsoft

it worth imagining how the next wave of AI in robotics could unfold—and how it might reshape the economy.

If the physical AI path materializes, the risk is not a compute-capex glut but rather an insufficient compute capacity.

### FROM WORDS TO MOTION

The transition from words to motion is moving at an impressive pace.

Before the ubiquity of LLMs like ChatGPT and Claude, robotic systems could perform only repetitive, short-horizon tasks in choreographed settings, such as placing an item on a table in a pre-set location. Any change in the object’s size or the table’s height could easily “break” them.

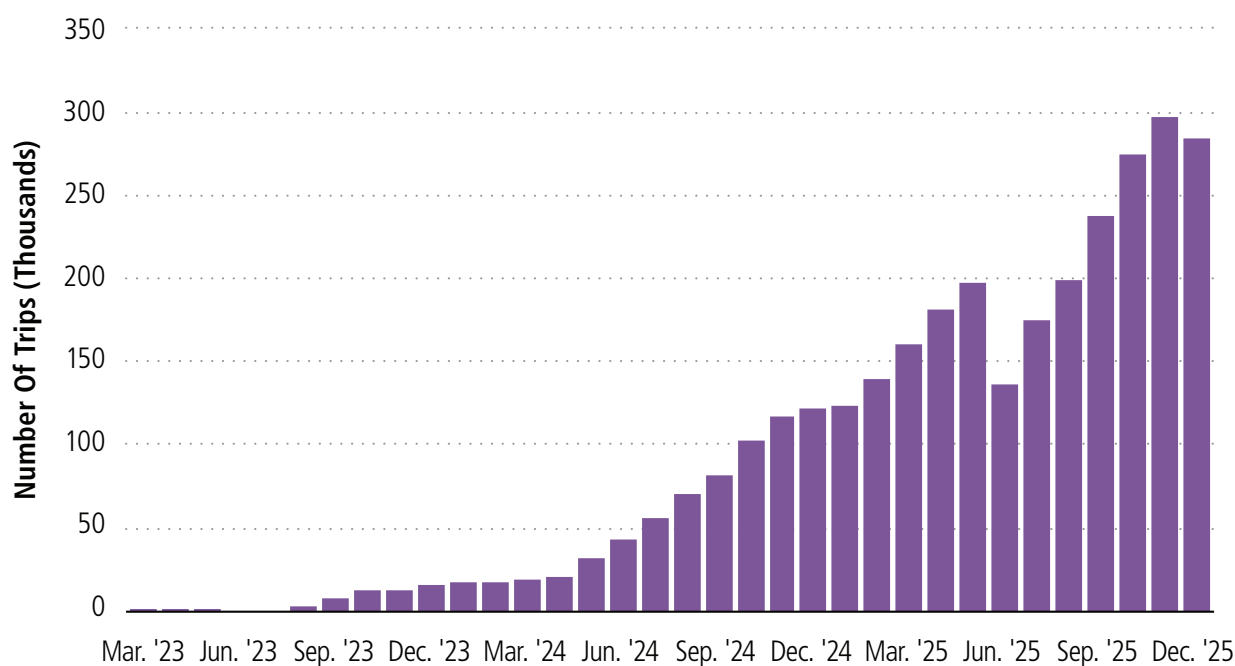
But applying LLMs to robotic models enables *planning*, a key feature that we humans do unconsciously. For robots, planning enables them to perform long-horizon, more dynamic tasks, such as making a sandwich or cleaning a room. Sounds simple, but these tasks were nearly impossible for robots to do without intensive human intervention before 2022.

**«ROBOTS, WITHOUT BEING TRAINED ON WHO TAYLOR SWIFT IS, CAN IDENTIFY HER PICTURES USING PUBLICLY AVAILABLE WEB IMAGES.»**

Then, in 2023, Google DeepMind developed a model that can connect robots to web-scale knowledge, the first step of integrating AI into physical motion.<sup>1</sup> For example, Google DeepMind’s robots, without being trained on who Taylor Swift is, can identify her pictures using publicly available web images.

In late 2024, Physical Intelligence developed a robot “brain” that can control eight different types of robots, a key step toward creating

fig 2. DELAYED DEPLOYMENT: WAYMO AUTONOMOUS VEHICLES AVERAGE WEEKLY PAID TRIPS IN CALIFORNIA



Sources: California Public Utilities Commission, 2026 AI Index Report

“generalist” robotic models that can be applied to any robot design.<sup>2</sup>

Fast forward a year, and Figure AI created the first autonomous robotic model capable of simultaneously controlling the full upper body of a humanoid, from the head, torso, and wrists to individual fingers, which is the closest to replicating how the human body works—one brain for all actions.<sup>3</sup> Before this, autonomous robots used at least two models: one for “brain” planning and one for physical action, although most had multiple systems controlling separate body parts. In turn, the more systems an autonomous robot relies on, the more training and fine-tuning it requires to learn new tasks.

And the key to Figure AI’s breakthrough? Over 500 hours of teleoperation—human operators guiding the robot through tasks to create training data. With teleoperation enabling robots to learn and record how the human body moves, it’s now easier to generate high-quality training data for all possible task combinations.

Figure AI’s new model can even use one brain for *two* bodies, controlling two humanoids simultaneously to perform collaborative group tasks—sounds like it could do better teamwork than two humans?

In early 2026, Nvidia’s robotic model can now expand beyond teleoperation. Rather than requiring a human operator, their model can extract joint-level data from watching videos of humans doing various tasks.<sup>4</sup> Although it is still early in the process, the ability to scale training data means that one day robots will learn how to move from any public web source with human data (like the thousands of vlogs out there).

**“THESE AREN’T THE DROIDS YOU’RE LOOKING FOR”**

We can already hear the objections, “Are you serious?! Autonomous robots? We’ve been hearing autonomous vehicles are only a year or two away since 2015; now it’s 2026 and we still don’t see that many!”

Indeed, while experts predicted full vehicle autonomy by 2020, the leading autonomous vehicle company, Waymo, didn’t reach mass-scale deployment until 2025 and remains limited to only a few U.S. cities (see Figure 2 on page 11).<sup>5</sup>

For autonomous vehicles, the challenge proved harder than anticipated, not because the core capabilities were absent, but because we are deploying them on roads exposed to all kinds of tricky, unpredictable edge cases.

Similarly, robots today can execute various tasks in labs, but actual “real-world” deployment faces an analogous problem. Adding to the complication, uncontrolled consumer environments like private homes and hospitals are even messier than roads, since they don’t have pre-specified lanes or structured traffic rules.

High-quality data for autonomous humanoids is also much harder to obtain than for cars.

Autonomous vehicles can be trained on thousands of miles of recorded driving data from cameras, GPS, steering wheel, acceleration, and braking. Just by driving around, a Tesla or Waymo is providing training data. But capturing joint and muscle-level data from human movement is much harder. For example, before Nvidia’s invention in 2026, online cooking videos could help robotic models learn the steps involved in making a sandwich, but couldn’t teach robots how to move their joint actuators to physically grab a slice of bread.<sup>6</sup>

Another obstacle that still stands in the way is that the “intelligence” bar for replicating humans is very high. For example, frontier autonomous robotic models are reaching 80%–90% accuracy on small tasks, but that’s far from sufficient to replace human workers, especially at scale. Imagine hiring a humanoid home cleaner with 90% accuracy at handling glassware—you would be out of glasses in no time!

Beyond technical difficulties, humanoids are still too costly—McKinsey estimates that per-unit prices must fall by 80%–90% from the current \$150,000–\$500,000 range before large-scale industrial adoption becomes viable.<sup>7</sup>

As a result, the autonomous humanoid market was smaller than \$250 million in 2025—compared with over \$40 billion spent on generative AI startups during the same year—with most end users still being universities and research institutions.<sup>8</sup>

**THE WRONG BLUEPRINT?**

But do we need to wait for humanoids to fully replicate humans to deploy physical AI at scale?

History suggests no. After all, we did not build mechanical horses to travel faster. Instead, we built automobiles.

What’s more, when the automobile became viable, the response was not to teach cars to navigate cobblestone streets designed for horses and pedestrians, but to build highways, reconfigure cities, relocate commerce, and redesign roughly a third of urban land surface to accommodate the car’s physical constraints (see *Did You Know? The Phoenix Effect*).

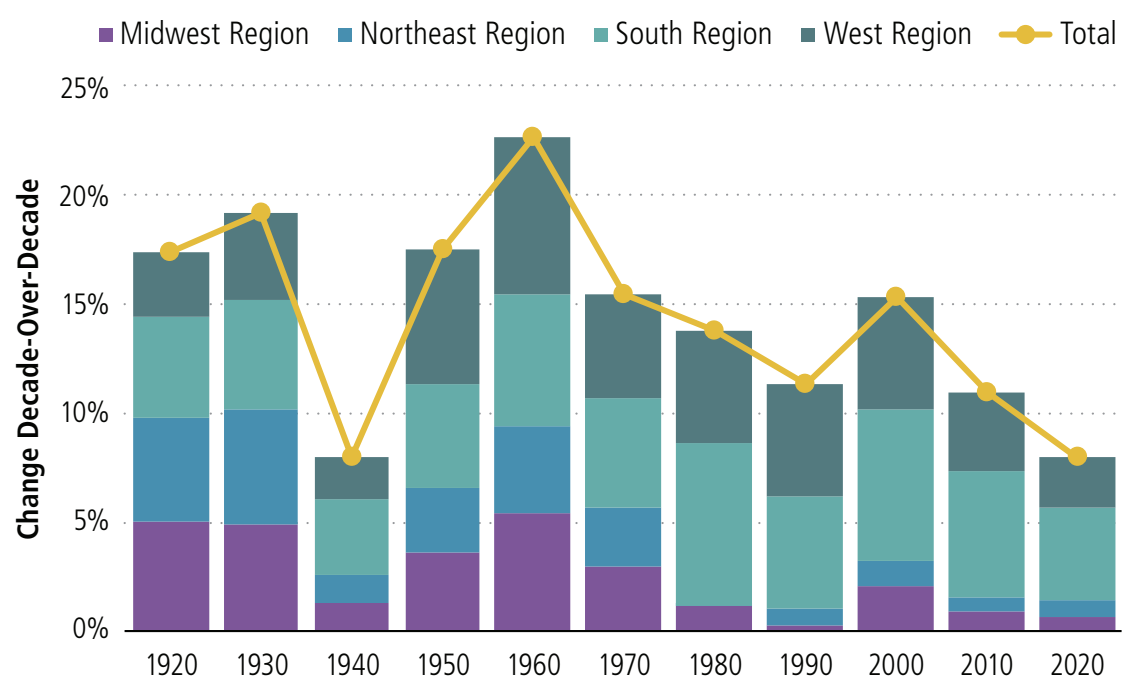
Similarly, for autonomous robots, the key to mass deployment may be engineering environments to reduce intelligence requirements rather than maximize them.

**DID YOU KNOW?**

**THE PHOENIX EFFECT**

The American cities that grew the fastest in the 20th century were the areas not beholden to urban path dependencies. Between 1940 and 1980, the Sunbelt region—built from the ground up around wide arterials, low-density lots, and single-use zoning—grew in population by 112%. Air conditioning helped, too. For example, the Phoenix population expanded by an astonishing 1,138%. In 1940, 60% of the national population growth came from the South and West regions.<sup>9</sup> By 1980, it accounted for 90% of the nation’s population growth (see Figure 3). Meanwhile, the older Northeast and Midwest cities whose street networks were laid out before the private car are still walkable and historic, but they accounted for less than 10% of the national population growth. The lesson: regions with fewer constraints may experience the robot boom more directly.

fig 3. HIGHWAY HUBS:  
U.S. POPULATION GROWTH, CONTRIBUTION BY REGION



Sources: U.S. Census Bureau, Payden Calculations

Warehouses and factories, which are more controlled than homes or offices, can be redesigned to achieve fully autonomous automation relatively easily, especially since most manufacturers have the capital and operational sophistication to adopt early-stage technologies.

For example, in 2012, when Amazon acquired robotics company Kiva Systems, their robots didn't look "human" at all—they were squat and rectangular. But Amazon didn't ask the robots to adapt to its warehouses. It built new warehouses around the robots—specific floor tolerances, standardized shelf heights, and inventory logic redesigned from scratch. Today, Amazon operates over a million Kiva-derived robots across its facilities, which are engineered for machine operation.<sup>10</sup>

And it's not just warehouses—we are already seeing deployments of autonomous industrial robots in manufacturing, though most remain in the pilot stage, especially in the electronics and automotive sectors (see Figures 4 and 5).

In 2023, Foxconn developed fully automated "lights-out" factories in China, designed from the ground up for machine operation. The factories are dark because they don't need lights—autonomous robots assemble, inspect, and polish electronic products 24/7, 365 days a year. Human intervention is required only for maintenance, material replenishment, and monitoring of system health.

In the U.S., autonomous robots are beginning to unload auto parts, move metal sheet components across the welding line, do sub-assembly work, and run inspections at automobile manufacturing plants.<sup>11</sup>

Even for environments that can't be redesigned, such as homes, hospitals, and care settings, it's possible that, instead of a single general-purpose humanoid housekeeper or caretaker, autonomous robots can be designed to specialize in cooking, skimming the pool, and washing the windows.

**COMPUTE AGE**

Regardless of which path the physical AI revolution takes—whether through better models or by redesigning existing environments to enable mass-scale deployment—if autonomous robots reach mass-scale deployment, the demand for computing power will accelerate exponentially.

Current estimates suggest that the compute demand for training and for the use (inference) of chatbots and AI agents will increase by 166% over the next five years.<sup>12</sup>

The rollout of autonomous robots will further amplify compute demand, especially since autonomous robotic model training is still in its early stages, with current autonomous robotic models using only about one-hundredth as much training compute as frontier LLMs (see Figure 6 on page 14). Future improvements in autonomous robotic models will require significantly more compute.

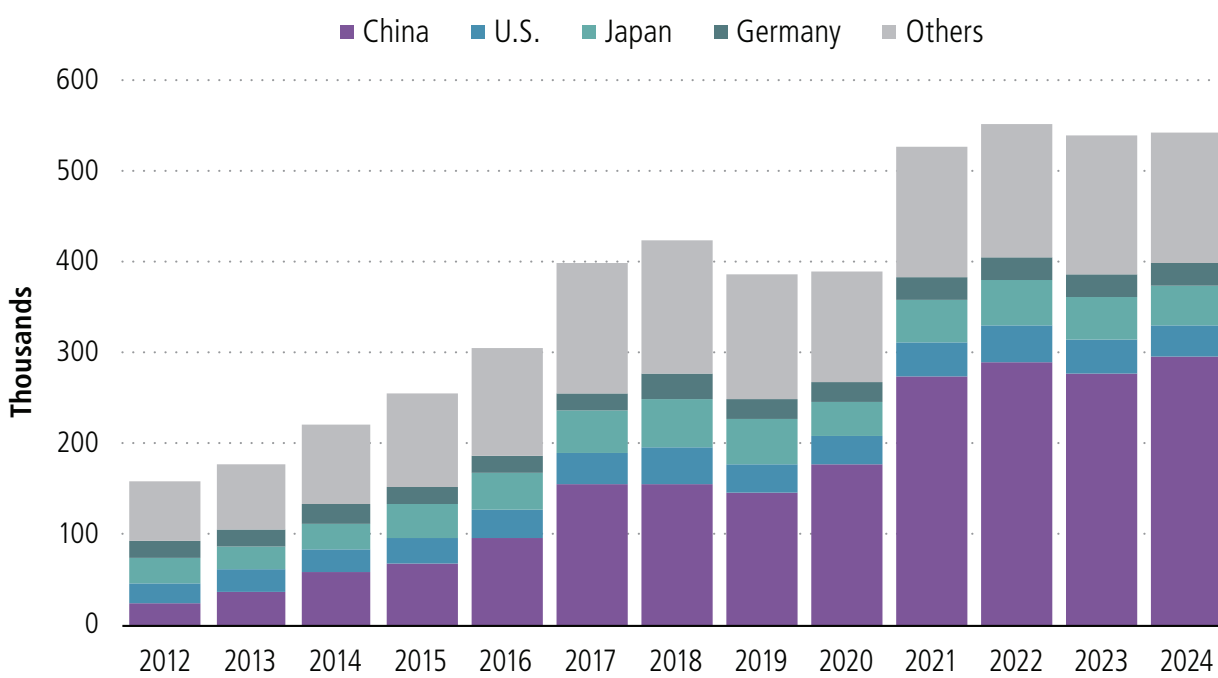
More importantly, the inference demand in mass-scale autonomous robot deployment will be enormous.

For example, for a chatbot or AI agent, compute is required per user query. But an autonomous robot needs compute continuously to perceive from sight, sounds, and sensors, and to "plan" its actions, and execute commands.

The computational intensity of operating autonomous robots is evident in the fact that advanced systems, such as the humanoid mentioned earlier, require dedicated GPUs inside the robots. Furthermore, unlike online chatbots, autonomous robots require the lowest possible latency. For example, if a humanoid is about to fall down the stairs, it has just a few milliseconds to "think" and reposition itself, while online chatbots have a few more seconds to generate a response to a message.

fig 4. BOT BUILDOUT:

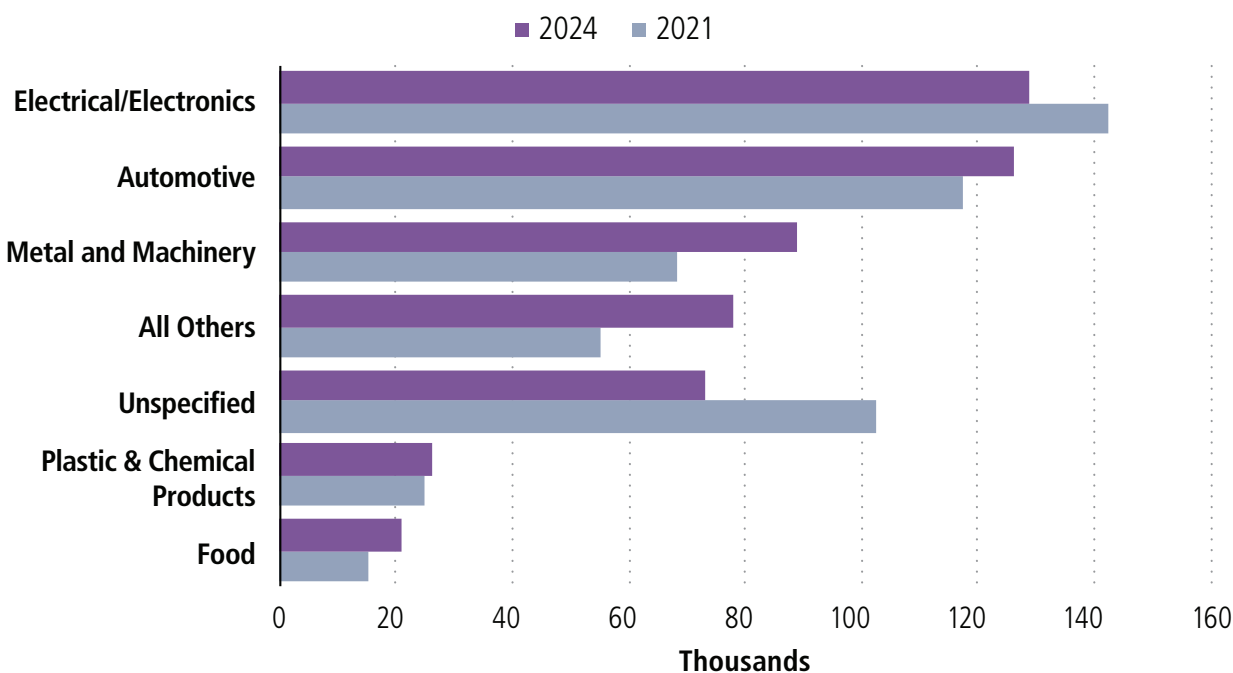
NUMBER OF NEW INDUSTRIAL ROBOTS INSTALLED IN THE WORLD PER YEAR



Sources: International Federation of Robotics, AI Index Report 2025

fig 5. AI INDUSTRIALIZATION:

ANNUAL INSTALLATIONS OF INDUSTRIAL ROBOTS IN THE WORLD BY CUSTOMER INDUSTRY



Source: World Robotics

**«AN AUTONOMOUS ROBOT ON THE LATEST NVIDIA GPUS FOR ONE HOUR COULD REQUIRE THE COMPUTE EQUIVALENT TO 100 CHATGPT QUERIES.»**

It's hard to estimate the actual scale of compute required for autonomous robots, since we haven't seen enough commercial use cases. However, a very rough estimate suggests that running an autonomous robot on the latest Nvidia GPUs for one hour could require the compute equivalent to 1000 ChatGPT queries.<sup>13</sup> If every goods-producing sector worker gets replaced by an

autonomous robot running on 24-hour shifts, that would require more than 100 times the inference compute of ChatGPT per day!<sup>14</sup>

**THE SECOND WAVE**

While AI skeptics question the massive capital expenditures driven by the ChatGPT moment, chatbots are just the beginning. The second wave of the AI revolution is already underway in the physical world.

But unlike chatbots that made AI “legible” in words, the physical AI moment is probably not going to be creating humanoids that will completely replicate human interaction with the world. Rather, it’s more likely that there will be a gradual redesign of the physical world around what machines can do—the same process that turned horse paths into highways and eventually open fields into lights-out factories.

Still think ongoing investments in training models, designing better hardware, and building of more data centers is all just hyperscalers overinvesting in a computing glut?


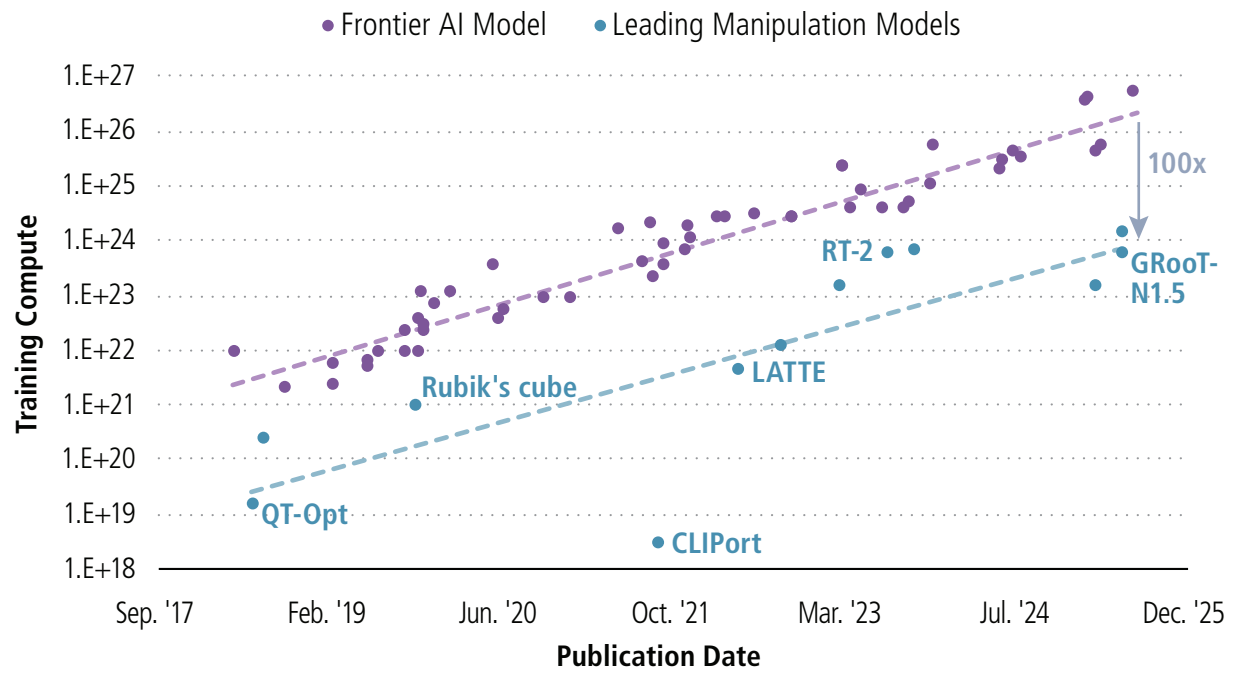
If the future of AI is mere chatbots, you might be right. But what if chatbots are only scratching the surface of intelligence? 

fig 6. TRAINING GAP:  
TRAINING COMPUTE OF FRONTIER AI AND ROBOTICS  
MANIPULATION MODELS



Source: EPOCH AI

ENDNOTES

1. Brohan, A., et al. (2023). RT-2: Vision-Language-Action Model Transfer Web Knowledge to Robotic Control. Google DeepMind.
2. Black, K., et al. (2024).  $\pi_0$ : A Vision-Language-Action Flow Model for General Robot Control. Physical Intelligence.
3. Figure AI. (2025, February). Helix: A Vision-Language-Action Model for Humanoid Whole-Body Control.
4. Sequoia Capital. (2026, May 3). *Robotics' end game: Nvidia's Jim Fan* [Video]. YouTube. [https://www.youtube.com/watch?v=3Y8aq\\_ofEVs](https://www.youtube.com/watch?v=3Y8aq_ofEVs)
5. Business Insider. (2015, October 10). *Google, Apple, and Tesla are racing to develop self-driving cars by 2020*. Business Insider. <https://www.businessinsider.com/google-apple-tesla-race-to-develop-self-driving-cars-by-2020-2015-10>.
6. With Nvidia's robotic model being able to train on human videos, now robots can learn at least the hand movements from cooking videos, although the technology is still early in the process.
7. Yee, L., et al. (2025, November 25). *Agents, robots, and us: Skill partnerships in the age of AI*. McKinsey Global Institute.
8. LaBerge, J. (2026, March 26). Humanoid robots: Timeline, impact, and investability. The Bank Credit Analyst Special Report. *BCA Research*, 77(10). The \$200–250M estimate adjusts headline figures by excluding wheeled social robots and bundled software revenues.
9. U.S. Census Bureau. (January, 2013). Data visualization gallery. U.S. Department of Commerce. <https://www.census.gov/dataviz/visualizations/049/508.php>
10. Greenawalt, T. (2025, October 22). *Amazon robotics: Meet the robots inside fulfillment centers*. About Amazon. <https://www.aboutamazon.com/news/operations/amazon-robotics-robots-fulfillment-center>
11. LaBerge, J. (2026), table II-1. Figure AI's Figure 02 spent ten months on a welding line at BMW's Spartanburg, South Carolina facility, moving over 90,000 sheet metal components across more than 30,000 vehicles. Aptronik's Apollo is running inspection and sub-assembly work at Jabil. Boston Dynamics' Atlas is being developed at Hyundai's Robotics Metaplant Application Center, with parts sequencing targeted for 2028. Agility Robotics' Digit completed a year-long pilot unloading totes on Toyota's RAV4 production line in Canada earlier this year.
12. McKinsey & Company. (2025, January 28). *The next big shifts in AI workloads and hyperscaler strategies*. McKinsey & Company. <https://www.mckinsey.com/industries/technology-media-and-telecommunications/our-insights/the-next-big-shifts-in-ai-workloads-and-hyperscaler-strategies>.
13. Nvidia's Jetson Orin GPU specializes in edge computing for robots, which has a maximum capacity of 275 TOPs. Assuming a 10% utilization rate, running an autonomous bot on Jetson Orin GPU can require (275 x 10% x 3600 seconds) 100,000 TFLOPS/hour. Meanwhile, EPOCH AI estimates that a typical ChatGPT-4o query of 500 tokens uses 100 TFLOPS.
14. Open AI outputs 20 trillion tokens every day, and each token requires 0.2 TFLOPs, so total inference compute is 4E12 TFLOPs, whereas 1,000,000 autonomous robots running on the Jetson Orin at a 10% utilization rate require 2.4E12 TFLOPs from calculations in footnote 13. Nearly 20 million workers are employed in the production, transportation, and material moving sector as of March 2026.