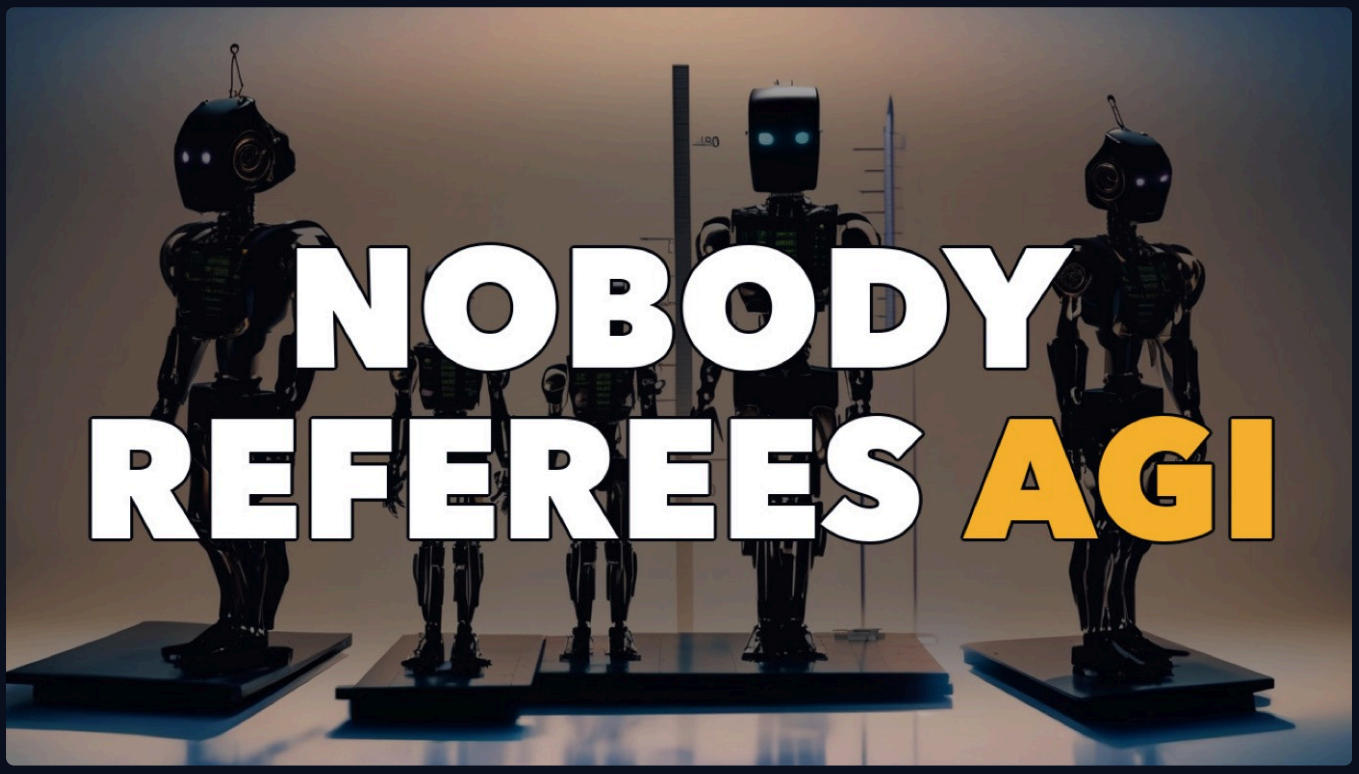
Forwards or Backwards

https://forwardorbackwards.com



What Is AGI, and How Would We Know We Built It?

Watch on YouTube: https://www.youtube.com/watch?v=Ufy_g8W2ouo

In August 2026, OpenAI's own chief research officer said the company was about eighty per cent of the way to AGI. Its chief executive said 'not quite yet.' Across town, Anthropic's chief executive avoids the term altogether. Three leaders, inside weeks of each other, could not agree on one word.

This video traces that word from its origin: a 1997 security paper that used 'artificial general intelligence' to describe a machine that could replace a soldier in any role, through a 2002 research claim by Shane Legg and Ben Goertzel that gave it a different meaning entirely.

Then it covers the three competing modern tests. OpenAI's charter defines AGI as a labour-market threshold, judged by its own board. A clause tying that definition to billions of dollars in Microsoft's

contract was redefined in October 2025 and dropped entirely on 27 April 2026. Google DeepMind's 'Levels of AGI' paper proposes five levels, from Emerging to Superhuman, scored on depth and breadth. The ARC Prize Foundation's ARC-AGI-3 benchmark, launched 25 March 2026, had every leading AI model scoring under one per cent at launch, while humans scored 100%.

Run one single capable system against all three tests, and it gets three different answers. The video closes on Anthropic's $11.5 billion Q2 2026 revenue overtaking OpenAI's $6.7 billion, and its planned IPO, before asking the question the title promises: is the industry converging on one measurable finish line, or keeping the word elastic enough that everyone can say they have already crossed it?

Every figure is on screen with its source and date.

Chapters

1. A word with no agreed meaning
2. 1997: a soldier's word for a battlefield problem
3. 2002: the researchers claim the word
4. OpenAI's legal definition: economically valuable work
5. The clause that turned a definition into money
6. DeepMind's ladder: depth and breadth
7. The ARC Prize: can it learn, not just recall
8. Three tests, three different winners
9. August-September 2026: the claims collide
10. Anthropic's rise and the stakes of the label
11. Why no one wants to be the referee
12. Finish line or moving target

1. A word with no agreed meaning





In the late summer of twenty twenty-six, the leaders of the world's two biggest AI labs could not agree on one word. OpenAI said it was close to building artificial general intelligence, or AGI for short — a machine able to do most of what a human mind can do, not just one narrow task. Anthropic's chief executive said he avoids that word altogether.

Here are Sam Altman of OpenAI and Dario Amodei of Anthropic, pictured together at a White House meeting on AI policy back in May 2023. By the summer of 2026, the two men running these labs could not agree on the one word this whole video is about.

Here is why that argument matters more than it sounds. AGI is not a fixed thing anyone can point to, the way you can point to a finished bridge. It is a word, and whoever's definition wins gets to say it has already been built, or that nobody has built it yet.

So who decides what AGI means? A security researcher writing in 1997. A handful of AI scientists in 2002. A company's own board. A government research lab. A small nonprofit that builds puzzles.

This video follows the word through all four of those sources, in order. Then it asks the real question. Is the industry moving toward one shared finish line for machine intelligence, or keeping that line far enough away that everyone can say they have already crossed it?

2. 1997: a soldier's word for a battlefield problem



The word AGI is younger than you might think, and it did not start in a computer science lab. It started in a paper about weapons.

In November 1997, a physicist named Mark Gubrud presented a paper called Nanotechnology and International Security at a conference in Palo Alto on molecular nanotechnology – the idea of building machines atom by atom. Gubrud was not writing about chatbots. He was warning about a battlefield where machines could replace soldiers entirely.

His paper gave a name to the kind of machine that worried him: artificial general intelligence. He defined it as AI "usable in essentially any phase of industrial or military operations where a human intelligence would otherwise be needed."

That is a strikingly narrow test, by today's standards. Gubrud was not asking whether a machine could write poetry or feel anything. He was asking whether it could simply do a soldier's or a factory worker's job, in any role, without a person. For thirty years, almost nobody used the phrase. It sat in a conference paper, waiting.

3. 2002: the researchers claim the word



The word resurfaced five years later, this time inside AI research itself.

Around 2002, a small circle of AI scientists — most often credited as Shane Legg and Ben Goertzel — picked up the phrase artificial general intelligence to describe their own goal. They needed a name for the kind of AI they actually wanted to build: not a program that only plays chess or only sorts email, but one with the same general range as a human mind.

Legg went on to co-found DeepMind, the lab now owned by Google, where he remains chief AGI scientist. Goertzel became one of the field's most visible AGI researchers and conference organisers.

From this point, AGI meant something different from Gubrud's soldier-replacement test. It meant a research goal: general intelligence for its own sake, measured against the breadth of what a human mind can do, not against any one job. Two groups of people, twenty-five years apart, had claimed the same four letters for two different ideas.

4. OpenAI's legal definition: economically valuable work



By 2015, a third group had entered the argument: a company with a board, a charter, and a contract to write.

A charter is a founding document that states what an organisation exists to do and by what rules. OpenAI's charter, published on its own website, gives AGI a specific, legal-sounding definition: "highly autonomous systems that outperform humans at most economically valuable work."

Notice what that definition does not mention. Nothing about consciousness, nothing about understanding, nothing about creativity. It is a labour-market test: does the system out-earn people at most of the jobs that pay?

And notice who gets to say when that test has been passed. Under the charter, it is OpenAI's own board that decides. Not an outside panel, not a government agency, not a benchmark score. A company's definition of a world-changing threshold, graded by the company itself — a fact that turns out to matter a great deal to one other company in particular.

5. The clause that turned a definition into money



That other company is Microsoft, and for years, OpenAI's definition of AGI was written directly into a contract worth billions of dollars.

From the earliest version of their partnership, Microsoft's rights to OpenAI's technology were tied to a trigger clause. If OpenAI's board declared that AGI had been reached, Microsoft's access to future OpenAI models could be cut off. One company's single word, decided by its own board, could end another company's access overnight.

On 28 October 2025, the two companies rewrote that clause. Microsoft and OpenAI announced what they called the next chapter of their partnership: AGI would now be verified by an independent expert panel, not by OpenAI's board alone, and Microsoft's rights were extended out to 2032.

Then, on 27 April 2026, the clause disappeared from the contract completely, alongside the end of Microsoft's exclusive claim on OpenAI's cloud computing. In under three years, the same four letters went from a company-ending trigger, to a panel's judgement call, to nothing written down at all. A definition that could end a contract is a definition people fight very hard over.

6. DeepMind's ladder: depth and breadth



While lawyers argued over one threshold, researchers at Google DeepMind tried to build a ruler with several rungs instead of one finish line.

In November 2023, a DeepMind team published a paper called Levels of AGI. It proposed five levels of performance above ordinary software, instead of one single yes-or-no test. In order, they are Emerging, Competent, Expert, Virtuoso, and Superhuman. A system reaches "Competent", for example, if it performs at least as well as the better half of skilled adults; "Expert" means the top tenth; "Virtuoso" the top hundredth.

The paper scores systems on two separate measures at once: depth, meaning how good a system is, and breadth, meaning how wide a range of tasks it can do that well on. A system can be virtuoso-level at one narrow task and still sit at the bottom rung on breadth. Under this ladder, AGI is not a single leap. It is a position on a grid, and different systems can sit at wildly different points on it.

7. The ARC Prize: can it learn, not just recall



A third group asked a different question entirely: not how much a system knows, but how quickly it can learn something brand new.

In 2019, the researcher Francois Chollet proposed a benchmark — a standard test used to compare different systems fairly — built around that idea. He argued that true intelligence should be measured as skill-acquisition efficiency on tasks a system has never seen before, not by how much training data it has memorised. The benchmark became known as ARC-AGI, and the nonprofit that now runs it, the ARC Prize Foundation, still uses Chollet's framing.

On 25 March 2026, the foundation launched a much harder version called ARC-AGI-3. It is built around interactive puzzle environments that an AI has to explore and figure out on the fly, rather than just give a single answer. At launch, every leading AI model scored under one per cent, while people solved the puzzles without difficulty. By Chollet's own test, the most talked-about systems in the industry were barely getting started.

8. Three tests, three different winners



So here is the uncomfortable part. Take one single, highly capable AI system released in 2026, and run it against all three rulers.

Against OpenAI's economic bar, a system that can do a large share of valuable desk work might already look close to outperforming humans at most economically valuable work, at least in some jobs. Against DeepMind's grid, the same system might score as Expert on many narrow tasks, but it would not score anywhere near breadth, because it still fails at huge categories of everyday reasoning. And against Chollet's test, the very same system would score close to zero, because it struggles to learn a genuinely new kind of puzzle on the spot.

One machine. Three rulers. Three different distances to the finish line. That is not a flaw in the rulers. It is what happens when "general intelligence" never had one agreed meaning to begin with.

9. August-September 2026: the claims collide





This is the argument that spilled into public view in late summer 2026, when OpenAI's own leaders could not agree with each other, let alone with Anthropic.

On 26 August 2026, TIME magazine published an interview called Inside OpenAI's Reboot. In it, OpenAI's chief research officer, Mark Chen, put a number on the company's progress. He said OpenAI was about eighty per cent of the way to AGI.

In the very same piece, chief executive Sam Altman struck a far more cautious note than his own colleague. He said OpenAI was "not quite yet" at AGI, but that he expected an internal system he personally would call AGI before the year was out.

This is Sam Altman again, at that same White House meeting on AI policy in May 2023. Long before that TIME interview, he was already making OpenAI's case directly to Washington.

Dario Amodei, chief executive of Anthropic, takes the opposite approach to the word itself: he avoids it entirely, and prefers the phrase "powerful AI."

Here is roughly how those three statements line up against each other in time. Chen and Altman spoke to TIME on 26 August 2026. Amodei has pointed to a similar window, late 2026 or 2027, for what he calls powerful AI, even though he will not use OpenAI's word for it.

In his essay "Machines of Loving Grace," Amodei describes what he means by that phrase. He defines it as "a country of geniuses in a datacenter" — a system smarter than a Nobel-laureate scientist across most fields, able to work on its own for long stretches. Three leaders, three different claims, inside a few weeks of each other.

10. Anthropic's rise and the stakes of the label





This argument is not just academic, because one of the two companies arguing is now winning the business race.

This was the second quarter of 2026. Anthropic's revenue reached 11.5 billion dollars, overtaking OpenAI's 6.7 billion dollars for the same quarter, for the first time. Anthropic also posted its first positive operating income that quarter.

This is Sam Altman at a White House dinner for technology leaders on 4 September 2025. OpenAI's chief executive has kept close ties to Washington all through the period this video covers, even as the money has shifted toward his rival.

Money like that changes what a word is worth. Anthropic confidentially filed paperwork in June 2026 for an initial public offering, or IPO – selling shares to the public for the first time. It had been aiming for a Nasdaq listing in October 2026. Reports in late September 2026 suggested that date could slip to November.

Whichever company the public believes has built, or is closest to building, the more capable system stands to gain enormously from that belief, right as Anthropic prepares to sell shares to outside investors for the first time. A word that sounds like a scientific category is also, right now, a number on a prospectus.

11. Why no one wants to be the referee



Given how much rides on the word, you might expect someone to have settled, once and for all, exactly what AGI means and who gets to certify it. Nobody has.

Look at what this video has covered so far. A legal definition judged by one company's own board, now softened to a panel. A five-level academic ladder with no single pass mark. A skill-acquisition puzzle test that most famous systems currently fail. Three serious attempts, none of them agreed with by the other two, and none of them run by an independent body that all three labs accept.

That is not an accident. A loose definition lets a lab claim progress on good days and lowered expectations on bad ones. A strict, independently verified test would remove that flexibility – and take away a tool that labs, investors and regulators all currently find useful in their own way.

12. Finish line or moving target



This video has covered a security paper, two research pioneers, a boardroom clause, an academic ladder, and a puzzle benchmark. So where does that leave the question it opened with?

The honest answer is that the industry has not converged on one measurable finish line. It has instead built several different ones, at different heights, and each lab tends to point at whichever ruler currently makes its own progress look best. That is not necessarily dishonesty. It may simply be what happens to a word invented to describe a destination nobody has actually reached yet.

Is that forwards, toward a shared, checkable standard that will eventually settle the argument? Or is it backwards, into a label kept elastic enough that everyone gets to say they have already crossed it? On the evidence gathered here, the word is still moving faster than any of the rulers built to measure it.

Sources and credits

Primary sources

- Mark Gubrud, 'Nanotechnology and International Security,' Fifth Foresight Conference on Molecular Nanotechnology, Palo Alto, Nov 1997, via Wikipedia's 'Artificial general intelligence' history section and a corroborating secondary account.
- Wikipedia, 'Artificial general intelligence' (history section) and 'Shane Legg' - Legg and Goertzel credited with popularising 'AGI' from 2002.
- Morris et al., 'Levels of AGI: Operationalizing Progress on the Path to AGI,' arXiv:2311.02462 (Google DeepMind, Nov 2023).
- ARC Prize Foundation, arcprize.org, ARC-AGI-3 technical report - benchmark concept published 2019 by Francois Chollet; ARC-AGI-3 launched 25 Mar 2026.
- Microsoft, 'The next chapter of the Microsoft-OpenAI partnership,' blogs.microsoft.com, 28 Oct 2025.
- Multiple reports, 27 Apr 2026 (aitoolly.com; letsdatascience.com; simonwillison.net's dated timeline), on the AGI clause being dropped from the Microsoft-OpenAI contract entirely.
- TIME, 'Inside OpenAI's Reboot,' 26 Aug 2026 - Mark Chen's '80% of the way' estimate and Sam Altman's 'not quite yet' statement, corroborated by multiple outlets quoting identical wording (TIME itself paywalled).
- Dario Amodei, 'Machines of Loving Grace' essay, Oct 2024, and Anthropic's OSTP submission - 'powerful AI' defined as 'a country of geniuses in a datacenter,' timeline 'as early as 2026 or 2027.'
- Company disclosures reported Aug 2026 (Fortune, 15 Aug 2026; Axios, 17 Aug 2026; Yahoo Finance) - Anthropic Q2 2026 revenue $11.5bn vs OpenAI $6.7bn.
- Anthropic S-1 filing, 1 Jun 2026, and reporting through 21 Sep 2026 on a possible slip from an October to a November 2026 Nasdaq listing.

Photo credits (Wikimedia Commons)

- Sam Altman (OpenAI CEO) and Dario Amodei (Anthropic CEO) together at a White House AI meeting, 4 May 2023: The White House, Public domain - https://commons.wikimedia.org/wiki/File%3AV20230504LJ-0199-2.jpg
- Sam Altman, OpenAI CEO, White House meeting, May 2023: The White House, Public domain - https://commons.wikimedia.org/wiki/File%3APresident_Joe_Biden_drops_by_a_meeting_with_Vice_President_Kamala_Harris_and_AI_CEO%E2%80%99s_on_May_4%2C_2023_in_the_Roosevelt_Room_of_the_White_House_-_P20230504AS-0693.jpg
- Sam Altman, OpenAI CEO, White House dinner for tech leaders, 4 September 2025: The White House, Public domain - https://commons.wikimedia.org/wiki/File%3AF20250904AH-2634_%2854778707580%29.jpg

Not regulated financial advice.