

The 88-Hour Proof

How ten thousand AI agents settled a ninety-year-old question about fluids - and why mathematics is still arguing about what just happened.

Executive Summary

For ninety years nobody knew whether the equations we use to model every moving fluid - air over a wing, blood in an artery, weather over an ocean - are always well behaved. The **Navier–Stokes equations** take Newton's $F = ma$ and apply it to a fluid treated as a smooth continuum rather than as individual molecules. The open question, one of the seven **Millennium Prize Problems** and worth a million dollars, was whether a fluid that starts out perfectly smooth can, obeying those equations, reach infinite speed in finite time. On **8 September 2026**, OpenAI said the answer is yes, and published a proof.

What makes this a landmark is not the fluid mechanics. It is that no human wrote the proof. An unreleased internal model, described by OpenAI as substantially stronger than the GPT-6 Astra system it had shipped five days earlier, was pointed at the problem as a swarm of roughly **ten thousand coordinating agents**. They ran for **88 hours**, exchanged 2.7 million messages, produced about 130 billion tokens of output, and returned a 166-page paper plus a machine-checked formalisation in Lean. Independent estimates put the compute bill somewhere between **one and fifteen million dollars** - a proof that cost more than the prize. OpenAI says it will not claim the prize.

Twelve hours before that announcement, two human mathematicians - Tristan Buckmaster of NYU and Levent Alpöge, who works at Anthropic - released their own year-long result on a closely related equation, built on a research programme started in 2023 by Diego Córdoba and Luis Martínez-Zoroa. Buckmaster alleges that news of their method reached OpenAI days before the company launched its run, and that the company then used the same line of attack. OpenAI denies it, while conceding in its first statement that it could not rule out having benefited from de-identified data derived from the pair's use of its own products. It has since hardened that to a categorical denial.

So there are two stories here and they are both true. A machine has, for the first time, apparently closed a major open problem in mathematics, and it did so by converting money into insight at a rate no university can match. And the institutions around mathematics - credit, peer review, the confidentiality of unpublished work typed into a commercial tool - turned out to have no procedure for any of it. **The result is probably right. The process is what the field is arguing about.**

88 hours of agent runtime	~10,000 agents in the winning group	166 pages, machine verified in Lean	\$1m prize on offer, left unclaimed
--	--	--	--



One Equation, Eighty-Eight Hours, Ten Thousand Agents

Claimant: **OpenAI** | Announced: **8 Sep 2026** | Problem: **Navier-Stokes existence & smoothness** | Status at Clay: **Active**

Everything in this note unpacks one event: a mathematical question that had resisted ninety years of human attention was resolved in under four days by software, and the answer arrived with a machine-checkable certificate attached. This page is the whole story in one view; the rest explains why each piece of it matters.

88 h Wall-clock time from first agent launched to proof	~10,000 Concurrent agents in the group that found it	130 B Output tokens spent on Navier-Stokes alone	166 pp Length of the paper, plus a Lean formal proof
---	--	--	--

WHAT ACTUALLY CHANGED - BEFORE VS. AFTER

	BEFORE 8 SEPTEMBER 2026	AFTER
The question	Open since Jean Leray's 1934 work; the Clay Institute put a \$1m price on it in 2000	Answered - smooth 3-D flow <i>can</i> break down in finite time
Millennium Problems settled	One of seven (Poincaré, 2010 - prize declined)	A second claimed; Clay has moved its status to "active", not "solved"
Who writes the proof	Named mathematicians, occasionally with computer assistance	An agent swarm; humans wrote the prompts and chose the target
Time to a result	Decades of accumulated human effort	88 hours of wall clock, plus 17 hours to formalise
Cost of a proof	A salary, a blackboard, a library card	~\$1m of internal inference; about \$6m at public API rates
Checking it	Peer review, months to years	Lean verifies the logic in hours; peer review has not happened yet
Credit	Attached to people, by convention	Disputed between a company, two mathematicians and a prior programme
What did not happen	-	The other five open Millennium Problems were attempted and not solved

THE ONE-LINE VERSION

A company rented roughly a million dollars of its own computers for four days and bought an answer that the mathematical community had been unable to reach in ninety years - then discovered that nobody had written the rules for what happens next.



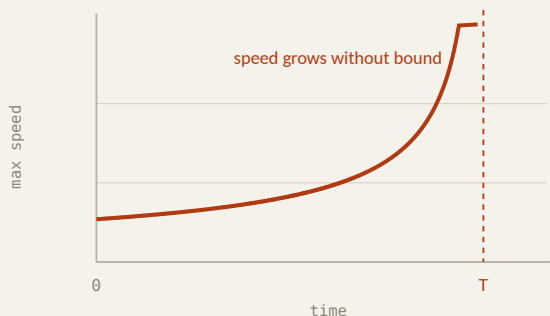
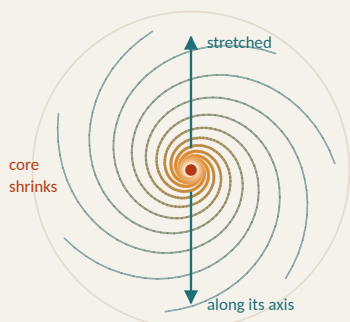
What Was Actually **Proved**

The Navier-Stokes equations are Newton's second law, $F = ma$, written for a fluid. Their essential move is to stop counting molecules and treat the fluid as a continuous substance with a velocity at every point. That approximation is why we can design aircraft, forecast weather and model blood flow at all. The ninety-year-old question is whether the approximation can destroy itself: starting from a perfectly smooth state and pushed by a perfectly smooth force, can the equations drive the speed at some point to **infinity within a finite amount of time**? That event is called a **singularity**, or a blow-up.

THE SKATER, AND WHY THIS IS NOT OBVIOUS - AN ANALOGY

A spinning ice skater pulls her arms in and speeds up; the same angular momentum packed into a smaller radius means faster rotation. A fluid vortex can do this to itself. OpenAI's solution is exactly that picture: a swirl that spirals inward while being stretched along its axis, like spaghetti. The core shrinks, so it spins faster, so it shrinks further.

What stops this in ordinary life is **viscosity** - internal friction, the reason honey does not form infinitely fast whirlpools. The difficulty is that the runaway has to beat viscosity while obeying two hard constraints: total energy must stay finite, and the external force must stay smooth. You are not allowed to cheat by pushing infinitely hard. The proof works because the large terms in the equation - acceleration, pressure, momentum transfer, friction - are arranged to cancel almost exactly, leaving a gentle force outside and an unbounded speed inside.



Left: the vortex spirals inward and stretches; orange marks faster rotation. Right: core speed rises without bound and reaches infinity at the finite time T. Schematic, not plots of the proof.

WHERE THIS SITS - THE LADDER OF HARD CASES

Theorists had been climbing toward this by solving weaker versions first, graded by how much viscosity the runaway must defeat. Switch it off and you have the Euler equations; weaken it and you have the hypodissipative case. Human groups, above all Córdoba and Martínez-Zorúa from 2023 onward, built the lower rungs and the method for climbing them; OpenAI's agents claim the top one.

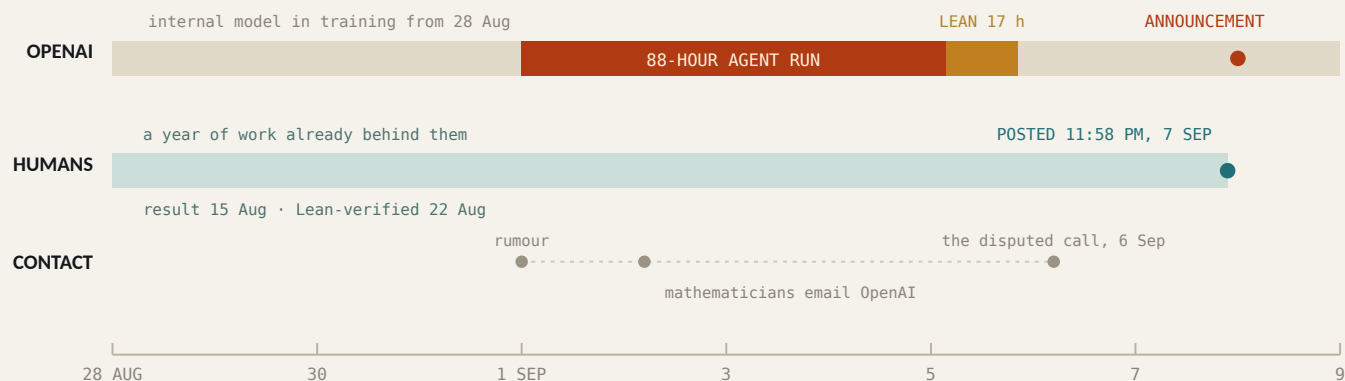
EULER viscosity switched off entirely Córdoba, Martínez-Zorúa & Zheng · Buckmaster & Alpöge, Aug 2026 · OpenAI agents, as a warm-up	HYPODISSIPATIVE NAVIER-STOKES viscosity present but deliberately weakened Córdoba, Martínez-Zorúa & Zheng, 2026	FULL NAVIER-STOKES real viscosity, smooth force, finite energy the Millennium Problem · claimed by OpenAI, 8 September 2026
--	--	--

What it does not mean: aircraft are not in danger. A blow-up is an exquisitely tuned configuration, not something a wing encounters. What breaks is the model: at a singularity the continuum picture stops describing the fluid, and you would have to go back to tracking particles.



A Swarm, Not a Mathematician

The method matters as much as the result. OpenAI did not ask one model one question. It ran a **multi-agent system**: thousands of copies of the same internal model, each able to read a cached copy of the internet and run code, split into groups that could talk among themselves. Different groups got different versions of the problem - some told to prove the equations always behave, others to prove they do not. Groups also worked easier warm-ups, and one of those cracked first: about a hundred agents, roughly fifty hours, disproved the frictionless Euler case. That was fed back as a hint to the groups attacking the real thing, with Codex consolidating the best fragments from each group.



Reconstructed from OpenAI's account and public statements. The run began 1 September, the proof landed 5 September, verification finished the next day, the announcement came at noon on 8 September - twelve hours after the human result was posted.

WHAT IT TOOK - THE ARITHMETIC OF A PROOF

RESOURCE	NAVIER-STOKES ALONE	ALL SIX PROBLEMS ATTEMPTED
Agent messages	2,700,000	4,900,000
Output tokens	~130,000,000,000	~300,000,000,000
Concurrent agents	~10,000 in the winning group	groups of varying size
Elapsed time	88 h + 17 h formalisation	1–8 September

WHAT THE 88 HOURS DOES NOT INCLUDE

The clock starts very late. It excludes the year Buckmaster and Alpöge had already spent on this terrain, the programme Córdoba and Martínez-Zorúa built from 2023 that supplied the line of attack, the 2013 numerical evidence from Luo and Hou that convinced the field a blow-up was likely at all, and ninety years of literature sitting inside the model's training data. The swarm ran the last mile very fast. It did not run the race.

Why verification is the quiet headline: the proof ships with a Lean formalisation - a computer has checked every logical step, so you need not trust the author to believe the chain holds. What Lean cannot tell you is whether the argument is meaningful or answers the question experts were actually asking. That judgement still takes people, and months.



Who Owns a Result That **Nobody Wrote**

The mathematics is not contested. What is contested is how OpenAI came to be working on this problem, by this method, in this week - and what a company owes researchers who use its tools.

WHEN	WHAT HAPPENED
Late Aug	Buckmaster and Alpöge finish a smooth-forcing blow-up result for the frictionless case, verified in Lean on 22 August. They delay publishing, wanting a human-readable write-up first.
31 Aug – 1 Sep	A cryptic post by Alpöge is read online as a hint that Anthropic has solved Millennium Problems. The rumour spreads; OpenAI, on its own account, launches its agent effort on all six open problems.
2 – 3 Sep	Both mathematicians contact OpenAI independently, having been tipped that their progress had reached the company, and ask that the work not be rushed into an AI write-up.
6 Sep	A call is held. Told OpenAI has a Navier-Stokes proof, Buckmaster concludes from the details that it followed the same Córdoba-Martínez-Zorúa route he was on. He asks whether his Codex sessions were accessible; OpenAI denies it.
7 Sep, 11:58 pm	Buckmaster posts the pair's results together with a detailed statement of events. Terence Tao calls the work <i>a remarkable achievement</i> .
8 Sep, noon	OpenAI announces the full Navier-Stokes proof, congratulates the pair, says it will not claim the prize. Its first paper omits the Córdoba-Martínez-Zorúa work; references are added in the 166-page revision.
9 – 11 Sep	Denials escalate from “cannot rule out” to “categorically impossible”. Clay calls its process deliberately unhurried and lists the problem as active; the AMS congratulates both parties.

THE TWO POSITIONS - STATED PLAINLY

THE ALLEGATION	THE RESPONSE
OpenAI learned of the pair's approach through rumour and, possibly, through their own Codex sessions, then spent millions to finish the job first.	The effort was triggered by a public rumour about a rival lab, not their work; no user data was searched; an investigation later concluded the prompts could not have influenced the model.
Alpöge was to be excluded from authorship of a human-written version because he works for Anthropic.	OpenAI cites conflict of interest; the researcher who made the argument has since apologised for his wording, including the remark “Why would you ruin your career?”

THE STRUCTURAL PROBLEM UNDERNEATH THE PERSONAL ONE

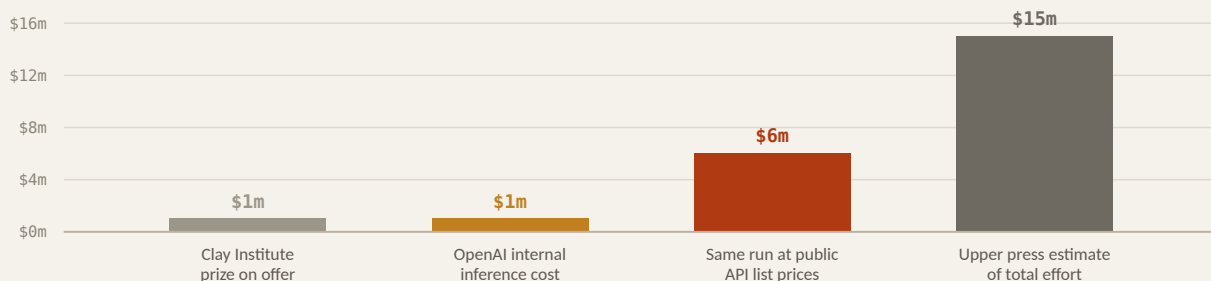
Strip out the personalities and something uncomfortable is left. The company selling researchers their most powerful tool is also a competitor for the results those tools produce, and it can mobilise resources no research group can match. OpenAI's own carefully worded first statement - that it could not rule out benefiting from de-identified data derived from the pair's usage - tells every scientist something they did not previously have to weigh: unpublished work typed into a commercial model is not obviously private. Whether or not it happened here, nothing in the arrangement makes it impossible.

Diego Córdoba, whose programme supplied the method, put the credit question in one sentence: “if our work had not existed, AI would not have solved the problem.” Terence Tao's objection is broader - that racing to answers strip-mines the open problems from which the next generation of techniques, and of mathematicians, would have grown.



The Price of a Theorem

The most consequential number here is not 88 hours. It is the cost. A result ninety years of human thought had not produced was bought in a few days for between one and fifteen million dollars - more, on every estimate, than the prize attached to it.



Sources: OpenAI's chief research officer put the cost in the millions; an independent academic estimate placed the retail equivalent near \$6m and the internal cost near \$1m; press estimates ran to \$15m.

FOUR THINGS THIS MAKES TRUE - THAT WERE NOT TRUE IN AUGUST

- **Compute is an input to discovery, not just engineering.** If a hard problem yields to money spent on inference, progress depends on who has the budget. A field that prided itself on rewarding insight over resources now has a price list.
- **Verification and understanding have come apart.** Lean can certify a 166-page argument before any human reads it. But a correct proof nobody comprehends teaches the field nothing, and the techniques, not the answers, are what mathematics lives on.
- **The tool vendor is also the rival.** Researchers must now weigh what their drafts, prompts and dead ends are worth to the company hosting them. Expect no-training guarantees, academic tiers and rules on what may be pasted into a commercial model.
- **Announcements are outrunning the institutions.** The result arrived by press post, not preprint or seminar. Clay still requires peer-reviewed publication and lists the problem as "active", not solved. The checking machinery of science is now the slow part of the pipeline.

THREE THINGS TO WATCH FROM HERE

1. **Does it survive reading?** Lean says the logic holds; experts must still agree it answers the question as posed. Clay's judgement, not the announcement, is the milestone.
2. **Does it repeat?** One problem fell and five did not. If a second goes the same way within a year, the resource thesis holds. If nothing follows, this was a well-chosen target at the end of a human runway.
3. **Who gets named.** The credit convention that survives this episode will shape how scientists use these tools - and whether they use them on their best unpublished ideas at all.

*For the ordinary reader the summary is short: the equations governing fluids have a flaw long suspected and now apparently confirmed, nothing you fly in is less safe, and the real event was procedural. A machine produced knowledge faster than the institutions built to validate it can move. **The proof took 88 hours; deciding what it means will take considerably longer.***

INFORMATIONAL ONLY; NOT INVESTMENT ADVICE.

Sources: OpenAI, "On the Navier-Stokes Millennium Prize Problem", 8 Sep 2026 (updated 10 Sep); Science; BBC News; with reporting from Nature, Fortune, Axios, MIT Technology Review, Scientific American and El País, 8-13 Sep 2026, and statements from T. Buckmaster, T. Tao, the Clay Mathematics Institute and the AMS.