ACL2 WORKSHOP 2026 · ACL2 and AI Panel

Human Intelligence is *Cooked*



The singularity is here.
Buckle up!

Pete Manolios, Northeastern

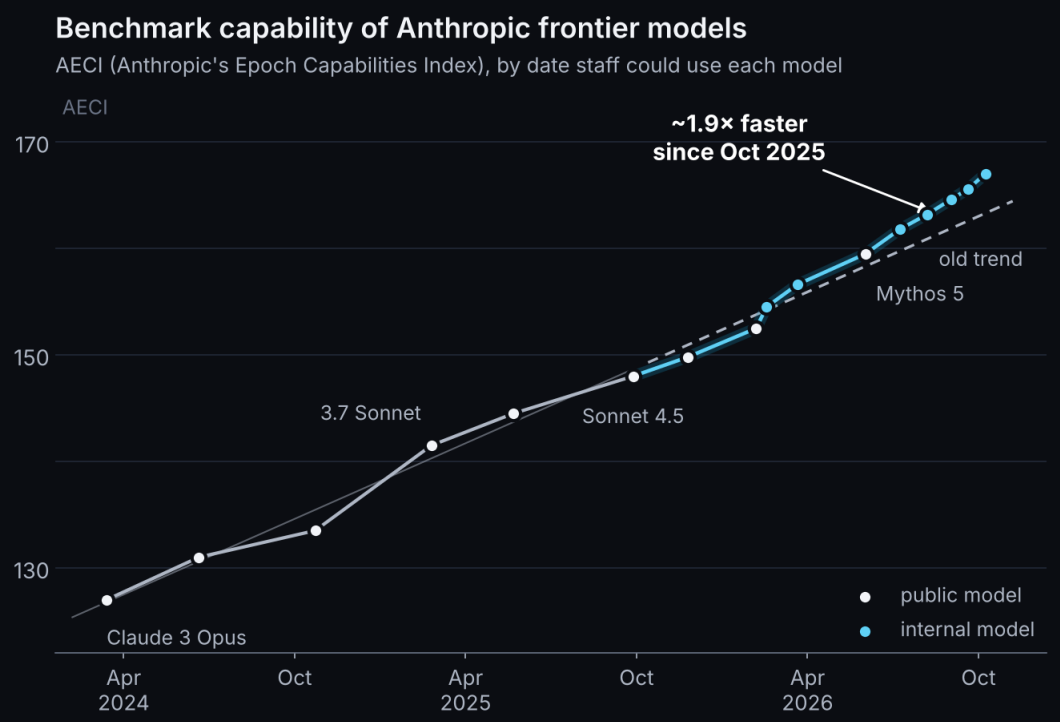
Biological vs. artificial intelligence

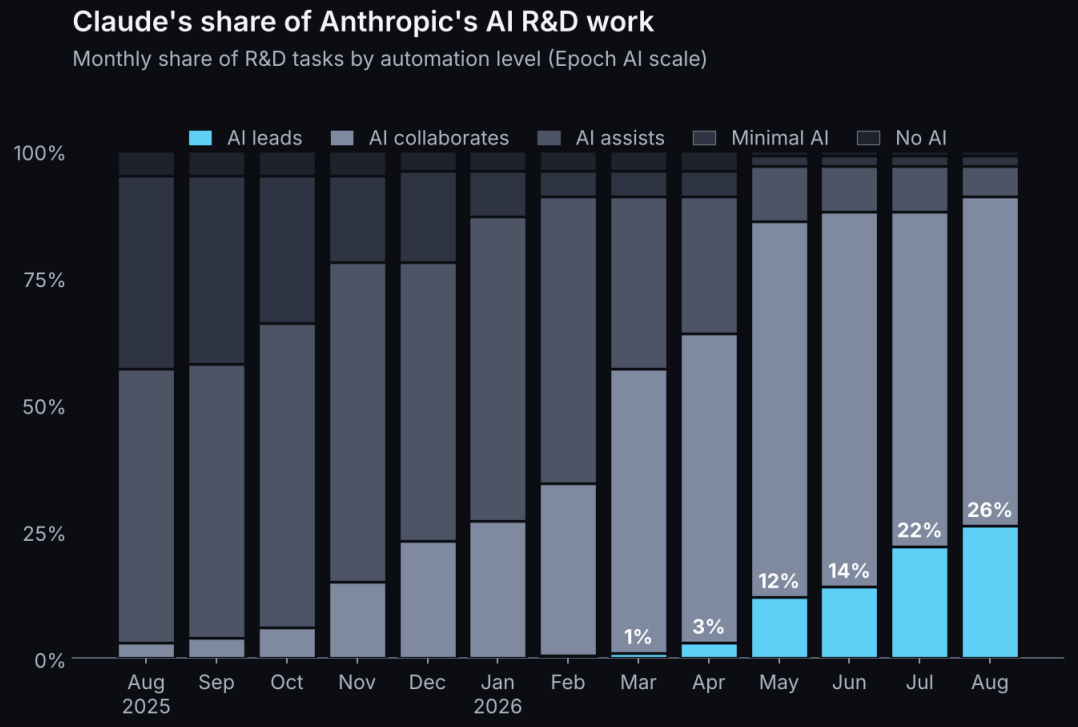
Evolution needed 2 billion years to reach us. AI is passing us now, and it will keep going even if it tapers off.



Illustrative curves, not measured data. Intelligence on a linear scale (human = 1×); time axis switches from log to linear at 2010. The 2026 crossing is the speaker's reading of the results on the following slides.

Recursive self-improvement has started





26% of Anthropic's AI R&D tasks led by Claude in Aug 2026, up from <1% in Feb

1.9× faster capability growth since Oct 2025 (METR; 95% CI 1.5–2.8×)

~80% of that speed-up attributed to AI doing AI R&D (METR's best guess)

Charts redrawn approximately from: METR, "Independent Assessment of AI R&D Acceleration at Anthropic," Nov 5 2026; Anthropic Institute, "Measurements for understanding the pace of AI development," Sep 17 2026 (R&D Automation Index, Epoch AI automation levels). Shown in D. Patel, "Recursive self-improvement has started," Nov 2026.

The flywheel: AI improves everything it runs on

Inner loop. LLMs write, train and evaluate better LLMs. Claude already leads 26% of Anthropic's R&D tasks.

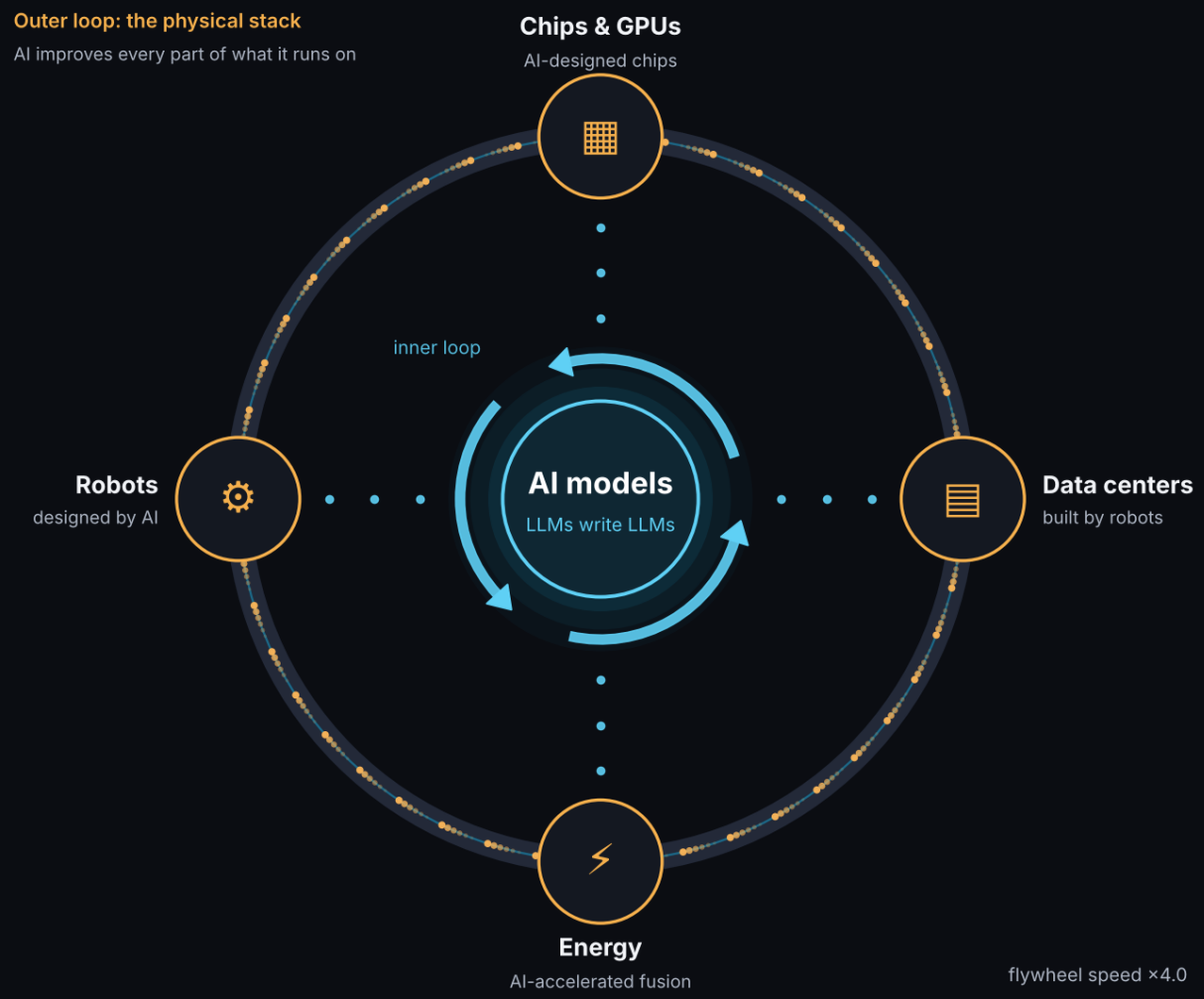
Chips. AI designs the accelerators it runs on.

Robots. Designed by AI.

Data centers. Built and run by robots.

Energy. AI-accelerated science (fusion, materials, grids) powers the next build-out.

Every loop that closes removes a human bottleneck, and the wheel spins faster.



Conceptual diagram. The inner loop is measured today (previous slides); the outer loops are the speaker's projection.

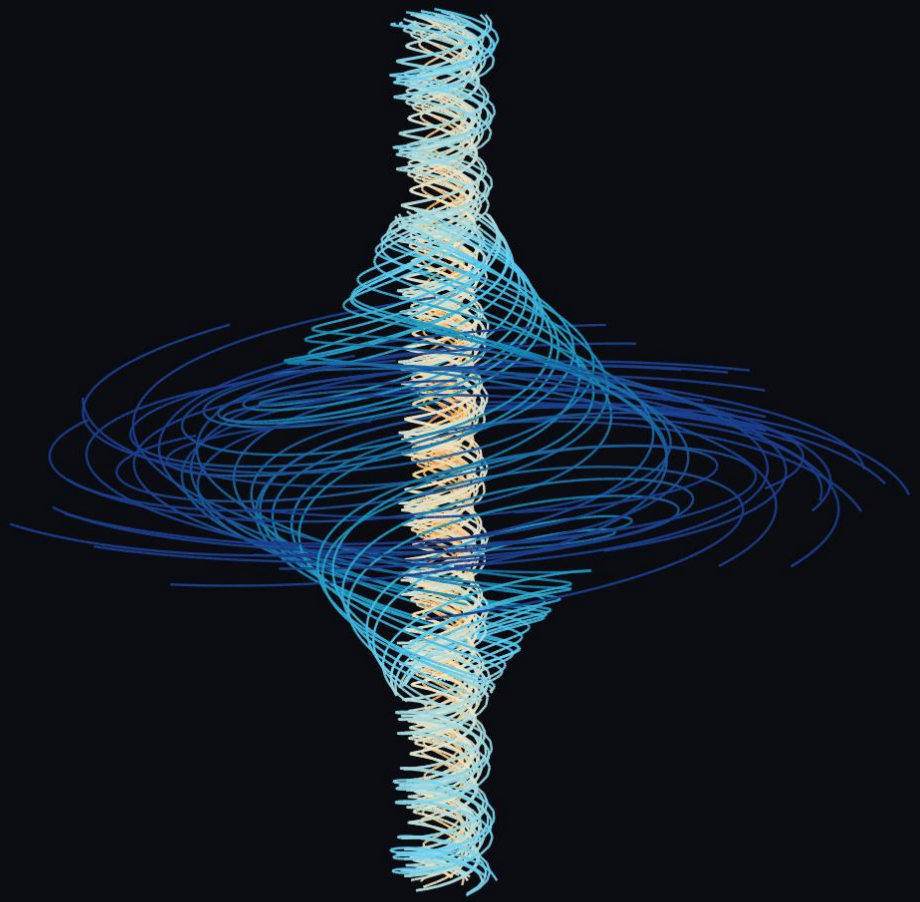
Where are we in FM?

From hard high-school problems to Millennium Prize problems in about two years



Dates: AlphaProof IMO silver (Jul 2024); IMO gold, OpenAI & Gemini (Jul 2025); Erdős #397, GPT-5.2 Pro (Jan 2026); unit distance (May 20); Jacobian (Jul 19); ζ zeros (Aug 10); FLT formalized (Sep 4); Navier–Stokes (Sep 8); OpenAI math release (Oct 6); ETH, Kakeya, Yang–Mills (Oct 26); CFSG formalized (Oct 27). Tiers are the speaker's judgment; most 2026 claims are still under expert review.

Navier–Stokes: a smooth fluid that blows up



The claim (OpenAI, Sep 8 2026). A 3D incompressible fluid that starts at rest, with a smooth external force and finite energy, reaches infinite speed in finite time. That is Clay's statement C (and D): a Millennium Prize problem.

The mechanism. A vortex spirals inward and stretches along its axis like spaghetti. Huge terms cancel so precisely that the force stays smooth.

How it was found. Agents first cracked the unforced Euler blow-up (~100 agents, ~50 h), then ~10,000 concurrent agents produced the Navier–Stokes proof in 88 h.
166-page paper plus Lean formalization.

OpenAI, “On the Navier–Stokes Millennium Prize Problem,” Sep 8 2026; DeepLearning.AI The Batch; DataCamp; Live Science; NPR (Sep 22 2026: “The paper is not written for humans,” J. Gómez-Serrano). Animation is a schematic, not the actual solution.

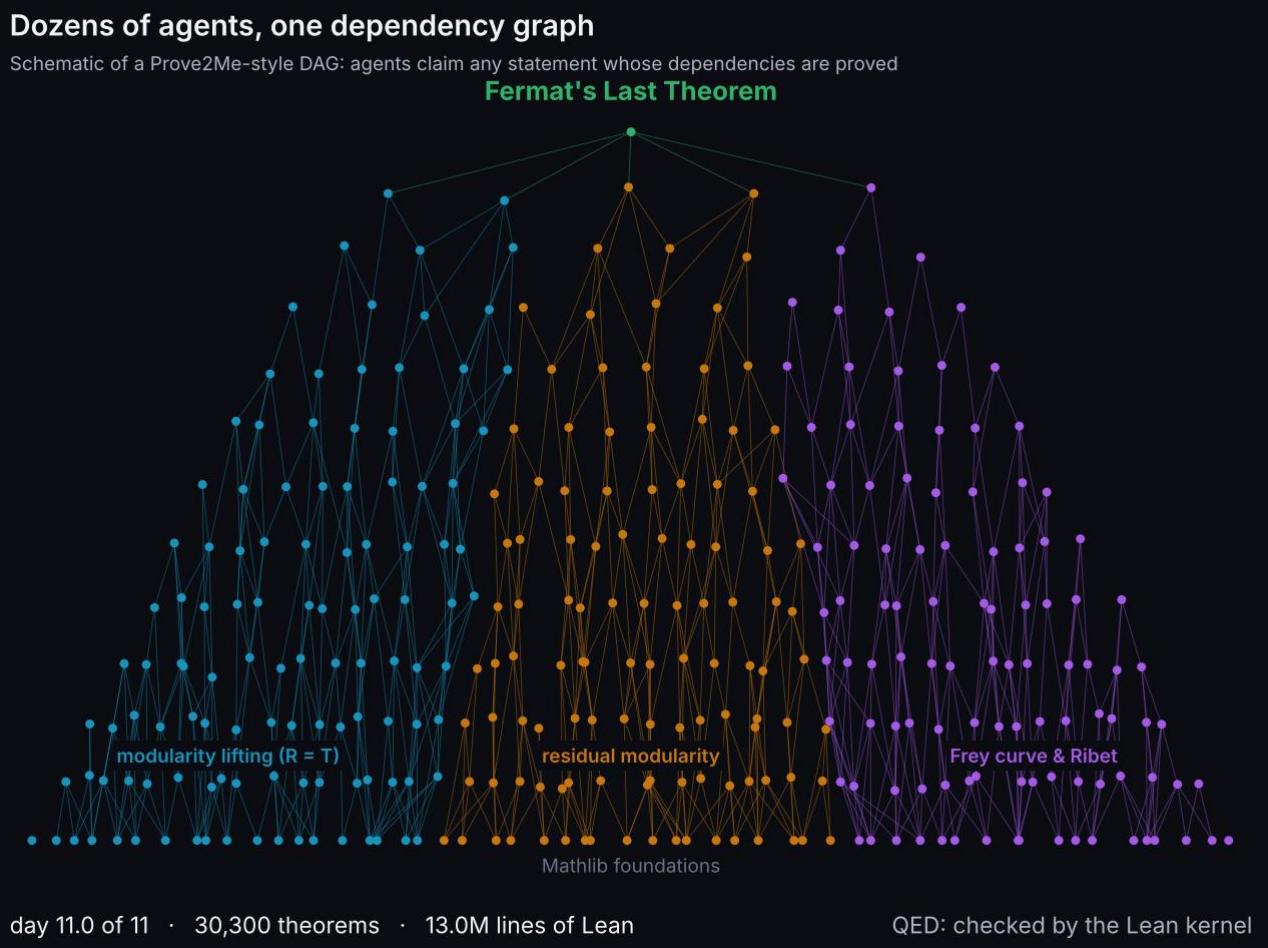
Fermat's Last Theorem: formalized in 11 days

Sep 4, 2026. Anthropic publishes the first complete computer-checked proof of FLT, written largely autonomously by dozens of Claude agents.

Scale. ~13M lines of Lean (over 5× Mathlib)
30,300 theorems (29,500 used)
~6B output tokens.

The trick: a shared DAG.
Uncoordinated agents lost track and stalled. Prove2Me gave them a graph of statements to claim in parallel.

Trust. Follows Darmon–Diamond–Taylor's simplified Wiles proof; only Lean's 3 standard axioms. Buzzard: it proves FLT “with no assumptions other than the axioms of mathematics.”



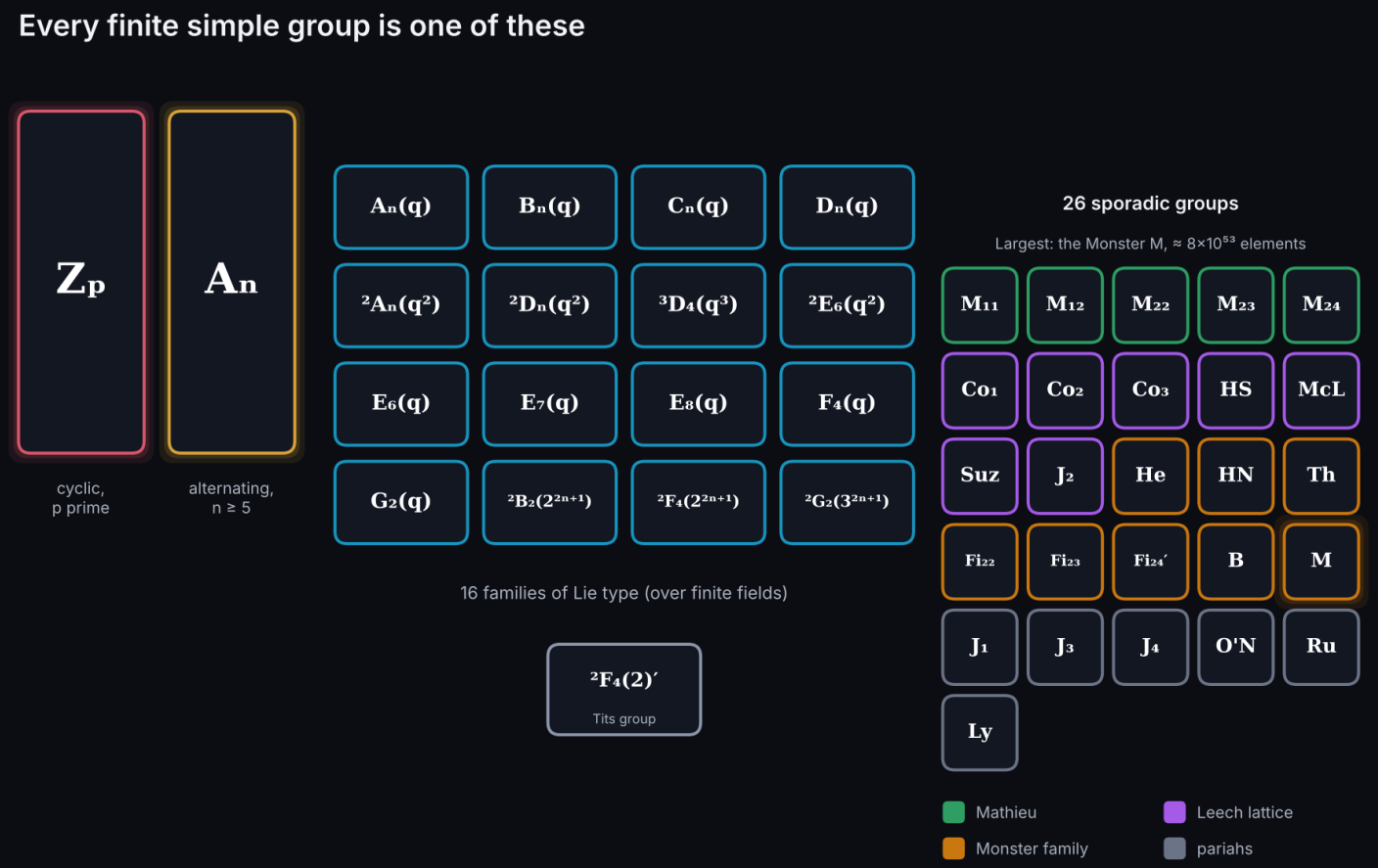
Anthropic, “Formalizing Fermat's Last Theorem,” Sep 4 2026; Techstrong.ai; Decrypt. Animation is a schematic of the Prove2Me-style workflow; section labels are the standard pillars of Wiles' proof, not the project's milestone names.

The Enormous Theorem, machine-checked

The theorem.
Classification of Finite Simple Groups:
~15,000 pages,
500+ articles,
~100 mathematicians.

Oct 27, 2026.
Axiom Math releases the 1st complete Lean proof:
24 days,
swarms of dozens of agents,
43.1M lines,
2.05M declarations.

Bugs found. 34 deviations from the literature, including a counterexample to an intermediate claim ($PSL_2(49) \to PSL_2(7)$), now repaired. The theorem stands.



Axiom Math announcement and repository, Oct 27 2026 (sorry-free, standard axioms only, no native_decide); Unite.AI; Defant & Ono, UGT repair paper (Hexagon, Oct 27 2026). Axiom's CEO noted that no outside group had yet rebuilt and checked it.

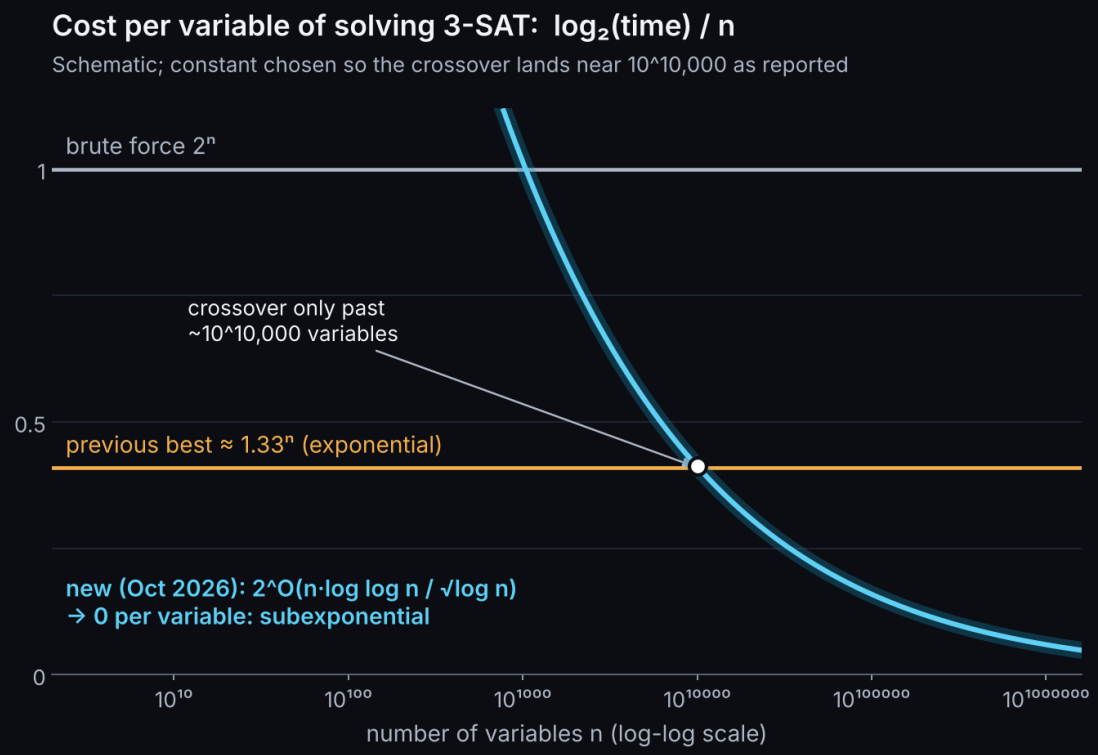
The Exponential Time Hypothesis is false

The hypothesis. Impagliazzo–Paturi (1999):
3-SAT needs 2^Ω(n) time.
More than 800 papers depend on it or on SETH.

Oct 26, 2026.
OpenAI posts “The Exponential Time Hypothesis Is False”.
Hours later Anthropic posts
“A 2^o(n) Time Algorithm for 3-SAT”, with Claude as author.

Same bound, different routes.
Both get 2^O(n log log n / √log n) and count solutions.
OpenAI: sparse parity checks and polynomial interpolation.
Claude: tensor-network contraction.
Both found the core idea on Oct 16; both are Lean-checked.

Status. Review led by Impagliazzo and Paturi. Not practical (crossover near 10^10,000 variables), but a foundational assumption of complexity theory is gone.



OpenAI, “Working with academic communities to build scientific understanding,” Oct 26 2026; Anthropic GitHub release (Claude, “A 2^o(n) Time Algorithm for 3-SAT”); The Hindu, Nov 2 2026; Hacker Noon; Level Up Coding. Chart is schematic.

OpenAI's big drop: 722 manuscripts in one day

722
manuscripts, 372 result families

34,815
pages

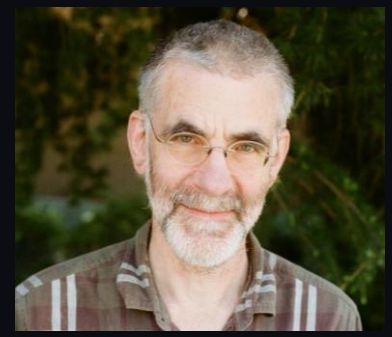
~4,000
open problems posed

~3 h
of ChatGPT-Pro thinking per result

Quasi-Riemann hypothesis (Oct 6). No zero of ζ, or of any Dirichlet L-function, has real part above 7/8. Radziwiłł: “the vast majority of consequences of RH already follow.”

Also claimed: no Landau–Siegel zeros; the Unique Games Conjecture. Independent rewrites (Banks, Kerr) support the L-function results.

Second drop (Oct 26): ETH false; local smoothing in every dimension, hence Kakeya and Stein restriction; existence of 4D Yang–Mills and QCD (no mass gap).



Richard Borcherds — Fields Medal, 1998
“AI is already better than me at everything.”



Alex Kontorovich — Rutgers University
“If a human did this, it would be an instant Fields Medal, no questions asked.”

OpenAI, “Sharing AI progress in mathematics,” Oct 6 2026 and github.com/openai/math; Quanta Magazine, “AI Proved the Quasi-Riemann Hypothesis. What Happens Now?” Oct 19 2026; Emergent Mind release explorer; OpenAI, Oct 26 2026. Most results are at varying stages of verification.

The elephant in the room

Can AI do everything presented at this workshop?



Photo: “African Bush Elephants,” Addo National Park, by Gorgo, Wikimedia Commons (public domain).