



JOHNS HOPKINS UNIVERSITY

Emily Riehl

AI for math:
what is happening and how the community is responding

Emerging Mathematica Uppsala

A lot has changed since July 2024:

A.I. Can Write Poetry, but It Struggles With Math

A.I.'s math problem reflects how much the new technology is a break with computing's past.



Sean Dong



1. What is happening in AI for Math

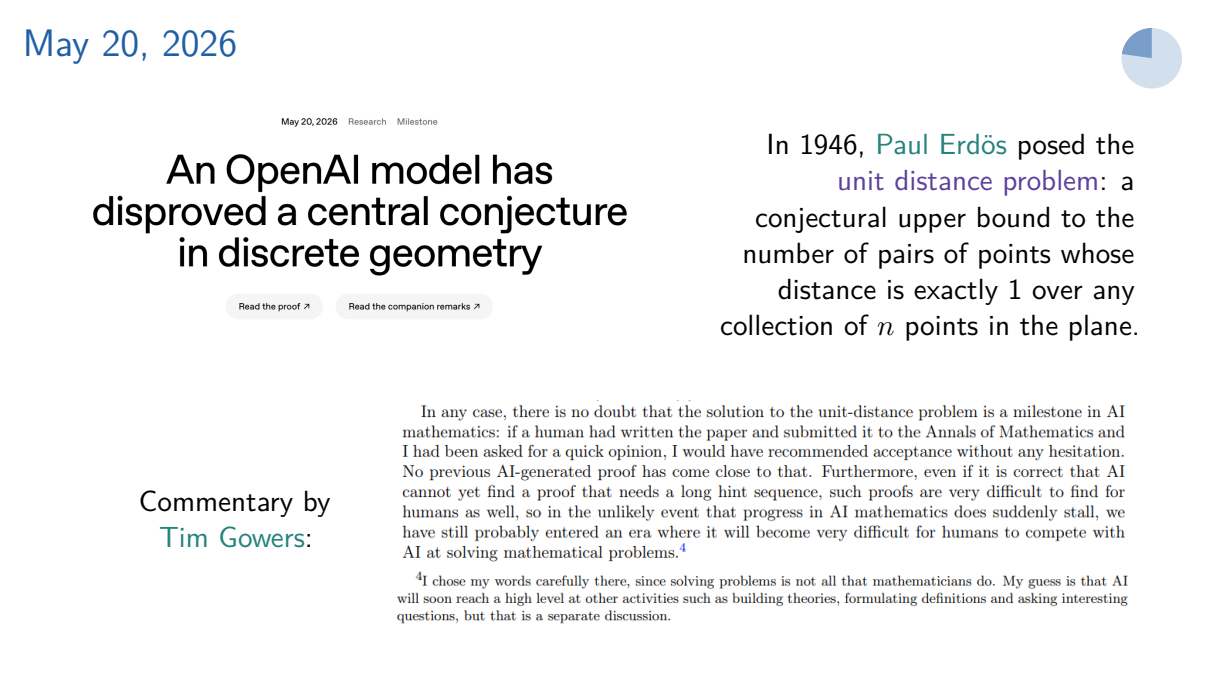
2. How the mathematical community is responding



1

What is happening in AI for Math

May 20, 2026



May 20, 2026 Research Milestone

An OpenAI model has disproved a central conjecture in discrete geometry

Read the proof ↗ Read the companion remarks ↗

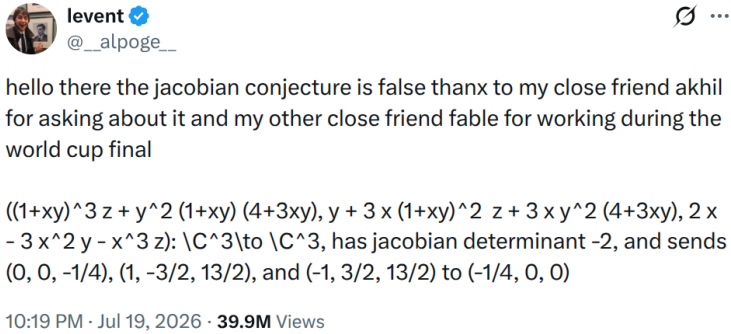
In 1946, Paul Erdös posed the unit distance problem: a conjectural upper bound to the number of pairs of points whose distance is exactly 1 over any collection of n points in the plane.

Commentary by Tim Gowers:

> In any case, there is no doubt that the solution to the unit-distance problem is a milestone in AI mathematics: if a human had written the paper and submitted it to the Annals of Mathematics and I had been asked for a quick opinion, I would have recommended acceptance without any hesitation. No previous AI-generated proof has come close to that. Furthermore, even if it is correct that AI cannot yet find a proof that needs a long hint sequence, such proofs are very difficult to find for humans as well, so in the unlikely event that progress in AI mathematics does suddenly stall, we have still probably entered an era where it will become very difficult for humans to compete with AI at solving mathematical problems.[4]
>
> [4] I chose my words carefully there, since solving problems is not all that mathematicians do. My guess is that AI will soon reach a high level at other activities such as building theories, formulating definitions and asking interesting questions, but that is a separate discussion.

July 19, 2026





The Jacobian conjecture states that for polynomial function $F\colon \mathbb{C}^n \to \mathbb{C}^n$, if the determinant of the matrix of partial derivatives is a non-zero constant, then F has a polynomial inverse. This is true in dimension 1, open in dimension 2, and false in dimensions 3 and above on account of the counterexample found by Fable.

August 1, 2026



The 10 advances, including results in discrete geometry, algebra, graph theory, and complexity theory, were presented in a 253 page paper, with each result accompanied by a formalization in `Lean`.

We provide new results for the following problems. The results were achieved by an internal version of Astra, our next major model. The total number of tokens needed to find solutions to these problems would cost roughly $2,000 at Sol API rates. These arguments were then prepared into manuscripts by humans with the same model. Afterward, the model formalized each argument in a Lean certificate. We are also releasing for each solution a model's narration of its thinking process.

August 10, 2026

Learning more about Claude's mathematical capabilities

Aug 10, 2026

The Riemann zeta function

The Riemann zeta function describes the distribution of prime numbers: each place that the function takes the value of zero contributes successively finer detail to the sequence of primes. The Riemann hypothesis is that the zeros that determine the primes all exist along a certain vertical line. This has become one of the most consequential conjectures in mathematics: many results assume it in order to provide a form of randomness in the primes.

No one has yet been able to prove or disprove the Riemann hypothesis, but mathematicians have made progress in many related directions studying the Riemann zeta function and its zeros. One of these, as above, is quantifying a minimum proportion of zeros that are on the line: over time, they've gradually increased this known constant proportion to 41.6%.

Recently, a member of staff at Anthropic gave Claude an unreasonable challenge. It was about one of the most famous unsolved problems in mathematics: *Take a real stab at the Riemann hypothesis*.

Claude did take a real stab, but as you might have expected if you're familiar with the difficulty of the task (the Riemann hypothesis dates back to 1859 and has a million-dollar bounty), it didn't succeed. Nevertheless, during its attempt, it unexpectedly made strides on a related problem.

An unreleased research version of Claude has improved on a longstanding lower bound for the fraction of zeros of the Riemann zeta function that satisfy the Riemann hypothesis. Drawing on extensive prior research by mathematicians over the past decades, it has increased this bound from 41.6% to 67.2%.

August 23, 2026

levent ✓
@_alpoge__

Please welcome to the world a beautiful new geometric object, to do with a problem i've always loved. claude really contains multitudes:D Does S^6 admit a complex structure?

Yup

5:31 PM · Aug 23, 2026 · **1.3M** Views

The literature, dating back to 1947, contains several mistaken disproofs. Boris Alexeev from OpenAI produced a `Lean` formalization four days later.

Formalization of the solution to the Hopf problem

The six-sphere admits a complex manifold structure compatible with its standard topology.

Based on *A compact complex threefold fibred by tori over the projective line, and the six-sphere*, originally shared on X by Levent Alpöge.

The repository includes a Comparator setup, with the statement adapted from the Formal Conjectures project.

```
lake update
lake exe cache get
lake build Lean4export
lake exe comparator comparator/config.json
```

Type-check it online!

levent ✓ @_alpoge__ · Aug 23

i always like to have short certificates that people can use to verify in full with their expertise and/or with one of today's super amazing models. this one is a little hard because (a pattern?) it's a problem with a lot of claims either way in the literature that fail due to subtleties. still im gonna try, two screenshots attached, but i also got opus 5 to write every detail down in a certainly far too long 100+ page pdf

the first page, or at most the first two, have enough to replicate i think



The $(3,4,\infty)$ modular family of 2-lets, completed at its three special points, is a complex structure on S^6.

The base and the family. Let $\Lambda = \mathbb{Z}[3,4,\infty]$ be the triangle group, acting on the upper half plane $\mathbb{h}$ the quotient $\Lambda\backslash\mathbb{h}$ is $\mathbb{P}^1$ minus a point, and adding the cusp gives $\mathbb{P}^1$ with orbifold points of orders 3 and 4 and the cusp. Let $V \cong \mathbb{P}^1$ carry automorphisms T_3, T_4 of orders 3 and 4 with $T_3 = (T_4 T_4^{-1})^{-1}$ conjugate, $(T_4 \cdot T)^2 = 0$. Let $\mathcal{E} \to V \times \mathbb{h}$ be the family of elliptic curves $E_{V,\tau} = V \times \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ where T_3, T_4 act on the first factor and the Möbius transformation $\tau \mapsto \frac{a\tau+b}{c\tau+d}$ on $\mathbb{h}$ if $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is one of the matrices T_3, T_4 or their inverses. Each $\tau \in \mathbb{h}$ is an equivariant period map $V \to \mathbb{P}^1$ of degree 12, and the quotient is a family of complex 2-tori over the punctured orbifold curve: the analogue of a universal family of abelian varieties over a modular curve, with this rank-four lattice representation in place of the symplectic one. The very general fibre is not algebraic. The very general fibre T has $\mathrm{NS}(T) = \mathbb{Z}[\sqrt{-3}]$ with $\sqrt{-3}$ a single indefinite class, and the algebraic dimension of the total space $\mathcal{E}$ is exactly 1.

The three special fibres. The three conjugacy classes of maximal finite and parabolic cyclic subgroups of Λ, single out the three special points, and at these three points we take the three standard completions, each is a choice, and not at any every orbifold point: below we treat, at an elliptic point or the local monodromy is a torus automorphism of order m, and the filling is Kodaira's logarithmic transform: (due to Iitaka) $\mathbb{C}Z/m\mathbb{Z}$, the quotient of $\mathbb{C}$ by the lattice $\mathbb{Z} + m\tau\mathbb{Z}$ rather than $\mathbb{Z} + \tau\mathbb{Z}$, and the resulting fibre is a cusp, a rational curve with one ordinary double point, like a nodal cubic. Here $m = 1$ is the case of no monodromy: a surjection of rank zero, and the filling S_0 is Mumford's toric degeneration; its central fibre W_0 is read off the pinwheel. At a ramification of $\mathbb{P}^1$, the disk action being the order-2 involution, the filling S_∞ is

August 31–September 3, 2026



BOUNDED GAPS BETWEEN PRIMES

JULIA STADLMANN

Abstract. Polymath 8b [5] proved that $H_1 = \liminf_{n\to\infty}(p_{n+1}-p_n) \leq 246$. In this paper we show how the Bombieri-Vinogradov theorem can be combined with newer equidistribution estimates for smooth moduli to obtain the improved bound $H_1 \leq 240$.



The value 246 corresponds to the length of the shortest admissible 50-tuple, while 240 is the length of the shortest admissible 49-tuple. The bound $H_1 \leq 240$ is therefore the smallest possible improvement. However, while Polymath [5] used symmetric polynomials of degree ≤ 27 to achieve $H_1 \leq 246$, we only needed polynomials of degree ≤ 21 to get $H_1 \leq 240$. As such, we believe that our method could achieve bigger improvements, but due to limited resources and limited time, we have decided to present this modest improvement as a proof of concept.

Axiom ✓
@axiommathai

1/ ✨ World record today on bounded gaps between primes, one of number theory's oldest open conjectures.

We show that infinitely many pairs of prime numbers are separated by 212 or less.

10:56 AM · Sep 3, 2026 · **259K** Views

Weijie Su ✓
@weijie444

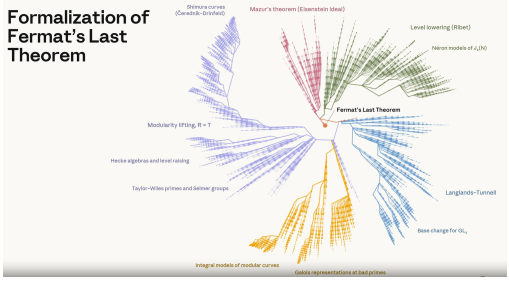
Announcing that GPT-6 Astra has pushed the prime gap to 186, with Lean formalization!

I was 9 when I first heard the twin prime conjecture. Its elegance and Yitang Zhang's legendary story have always stuck with me. A truly surreal night, being the first to see our model make progress, pushing 246 all the way down to 186, on a problem I've revered since I was a kid.

For me, it felt like witnessing a new era of intelligence being born, made possible by everyone at @OpenAI!

3:49 PM · Sep 3, 2026 · **146.8K** Views

September 4, 2026



We are sharing the first complete computer-checked proof of Fermat's Last Theorem. Claude worked largely autonomously over 11 days to write the proof in the Lean programming language. Below, we describe how the formalization was done and share some thoughts about what this work could mean for research mathematics.



FLT: Anthropic has beaten me to it

Posted on September 4, 2026 by xenaproject

I guess technically it was revealed to the world by a coffee shop in Islington on Insta, but an hour later it was officially announced by Anthropic: one of their internal models, using the prove2.me platform, has formalized a complete proof of Fermat's Last Theorem (FLT) in Lean. This is the final theorem to be formalized in Freek Wiedijk's famous list of 100 formalization challenges and thus wraps up this 20-year-old benchmark. Congratulations to Anthropic!

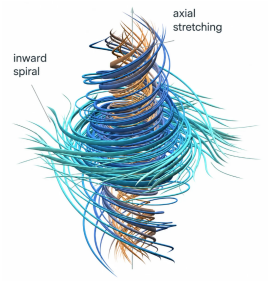
September 8, 2026



September 8, 2026 Research Publication

On the Navier–Stokes Millennium Prize Problem

Read the paper Link to Lean formalized proof ›



On Tuesday, September 1, we heard rumors that two Millennium Prize problems had been resolved. Inspired by these rumors and by the step change in performance of our internal model, we launched an effort to evaluate it on all open Millennium Prize problems and a few other high-impact problems.

We used a system of coordinating agents powered by our internal model. The agents had access to tools such as the ability to read from a cached version of the internet and the ability to run code. Agents were subdivided into groups with the ability to communicate within the group. The groups varied in size, and the group that produced the Navier–Stokes resolution involved on the order of 10,000 concurrent agents. At all times we maintained the same strict safeguards that we apply to all our frontier model evaluations, including monitoring and isolation.

...

The agents arrived at their resolution on Saturday, September 5, about 88 hours after the first agents were launched. Lean formalization and verification took an additional 17 hours via GPT-6 Astra.

Across all attempted problems, the agents sent 4.9 million messages and used about 300 billion output tokens. In the process of resolving the Navier–Stokes problem, the agents sent 2.7 million messages and used approximately 130 billion output tokens.

October 6, 2026



MATHEMATICS OPENAI

OpenAI Research Catalog

372 result families in 719 manuscripts October 6, 2026

An overview of the manuscript collection. Results are grouped by subject; entry numbers follow the catalog order and do not indicate a ranking. PDF links follow each summary.

Contents

Number theory	2
Algebraic and complex geometry	5
Real and complex analysis	9
Convex and metric geometry	10
Theoretical computer science	12
Dynamical systems and ergodic theory	16
Combinatorics	17
Algebra	21
Probability and statistical mechanics	23
Mathematical logic	26
Group theory	27
Mathematical physics	28
Operator algebras	31
Topology	33
Functional analysis	35
Differential geometry	36
Partial differential equations	39

October 6, 2026 Research Publication

Sharing AI progress in mathematics

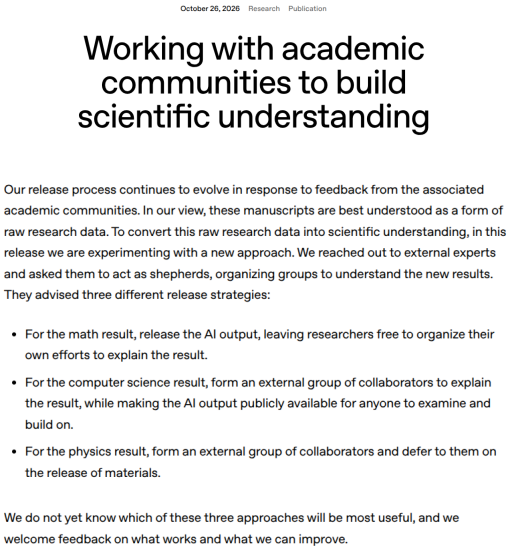
View on GitHub ↗

As of today, just under half (344/721) of the manuscripts have some relevant `Lean` formalizations of either the main theorem or of a significant result.[a]

[a] That is 48% of these papers come with a comparator challenge.

October 26, 2026





October 26, 2026 Research Publication

Working with academic communities to build scientific understanding

Our release process continues to evolve in response to feedback from the associated academic communities. In our view, these manuscripts are best understood as a form of raw research data. To convert this raw research data into scientific understanding, in this release we are experimenting with a new approach. We reached out to external experts and asked them to act as shepherds, organizing groups to understand the new results. They advised three different release strategies:

- For the math result, release the AI output, leaving researchers free to organize their own efforts to explain the result.
- For the computer science result, form an external group of collaborators to explain the result, while making the AI output publicly available for anyone to examine and build on.
- For the physics result, form an external group of collaborators and defer to them on the release of materials.

We do not yet know which of these three approaches will be most useful, and we welcome feedback on what works and what we can improve.

An internal model at OpenAI:

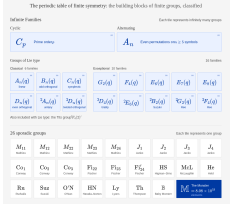
- proved the local smoothing conjecture in harmonic analysis, stated by Chris Sogge in 1991,
- demonstrated that the exponential time hypothesis about the about the computational complexity of Boolean formula satisfiability is false, and
- gives an existence proof of certain quantum field theories including pure Yang–Mills theory.

October 26, 2026 by Axiom

Formalizing the Enormous Theorem

Axiom has formalized in Lean the classification of finite simple groups, the longest proof in the history of mathematics.





Today, we are releasing the first complete computer-checked proof of the Classification of Finite Simple Groups (CFSG). One of the major achievements of twentieth-century mathematics, the classification identifies the basic building blocks of finite groups. Widely considered to be the longest mathematical proof ever written by humans, the original proof of CFSG extends across roughly 15,000 pages in more than 500 articles by about 100 mathematicians, earning CFSG the nickname “the enormous theorem.” For his crucial work on the classification, John Thompson was awarded the Fields Medal (1970) and the Abel Prize (2008). Over 24 days, engineers and mathematicians at Axiom Math directed swarms of dozens of AI agents in a project that produced approximately 43.1 million lines of code in the Lean proof assistant.



2

How the mathematical community is responding

How do we incentivize useful scientific activities?





Tony Feng @tonylfeng · Aug 25

I don't repost mathematics if I do not personally feel sure of its correctness, so I did not yet comment on the purported construction of a complex structure on S^6. However, I wanted to add some contextualization because I don't agree with the social media narrative that is developing.

0. Of the AI breakthroughs so far it's the closest to my interests, although still pretty far. I am not grieving, I am interested to learn more about it! I don't know if it's correct but on cursory glance it seems quite plausible.

1. If correct, this would be a major development. I believe it is a central problem in its field (complex differential geometry), like the Unit Distance Conjecture was for combinatorial geometry. Perhaps even more famous names are attached to this field, leading some people to say it's even more prestigious, but I prefer not to think in such terms.

2. There are two writeups, a 2-page one and a 100+-page one. Both are terrible and seem deliberately constructed to be inhumane. Fortunately, one can fight AI with AI: I asked Claude/GPT/Gemini to collaborate on making a more readable and enlightening version, which you can find at overleaf.com/read/frddxmbnq.... To make this, I prompted the models to write a document in the style of one of my own surveys. I haven't checked the mathematics, but it seems less forbidding than the original document at least.

3. This is a great opportunity for an expert to write an illuminating account of the construction, explain the key ideas and their proper historical context, and summarize the takeaways and lessons learned for the mathematics community.

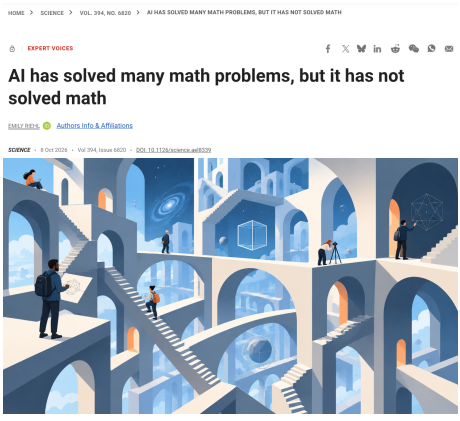
4. If I see such an account, I would be happy to arrange for the author to give a lecture series in Berkeley. If this lecture series is also wonderful, I would go further and advocate for the author to have a position at Berkeley (if desired). For contrast, I see little value in having colleagues who are good at being the first to prompt AI for a solution.

COMPLEX STRUCTURES ON $\mathbb{S}^6$

PHILIP ENGEL

ABSTRACT. A recent manuscript produced by Claude under the direction of Levent Alpöge constructs a 1-parameter family of compact complex threefolds diffeomorphic to $\mathbb{S}^6$ [AC26]. Our goal is to describe the geometric ideas behind the construction, give a self-contained proof, and remark on possible future research directions.

How do we articulate the value of human mathematicians?



HOME > SCIENCE > VOL. 394, NO. 6820 > AI HAS SOLVED MANY MATH PROBLEMS, BUT IT HAS NOT SOLVED MATH

EXPERT VOICES

AI has solved many math problems, but it has not solved math

EMILY RIEHL Authors Info & Affiliations

SCIENCE • 8 Oct 2026 • Vol 394, Issue 6820 • DOI: 10.1126/science.ae8339

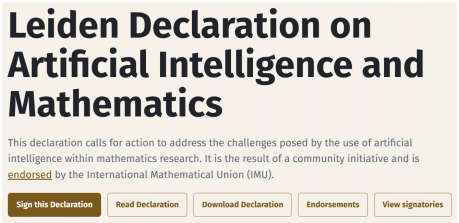
To explain, let me describe the landscape of research mathematics. Ever since the first mathematical records were made in Mesopotamia and Egypt, around 3000 BCE, mathematicians from around the world have been simultaneously surveying and erecting an enormous structure: the magnificent mathematical mansion. No architect has drawn up floor plans or is directing the work. Instead, small groups of explorers wander down dark corridors, trying to open every door they find. Each door is an unsolved mathematical problem, and to open it, one must solve the problem. Sometimes a stuck door only needs a bit of force in exactly the right spot; other times, a door won't budge until the correct key is found or forged.

This summer, AI has helped mathematicians open several doors that had previously resisted human efforts. However, the work does not stop there. Once a door is opened, someone must enter the room and turn on a light, figuring out what the solution actually teaches us. Then someone must oil the hinges and cut a spare key, so that others can follow the same path more easily. A solved problem can also lead to further questions: A door may lead to a closet or a large wing where further mathematical discoveries await. While the room is being explored, someone needs to relay the news to teams in nearby hallways, explaining the discovery to mathematicians in related fields. Finally, someone must draw the map that will be used to teach the next generation.

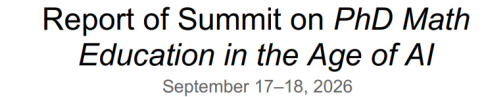
See also terrytao.wordpress.com

What norms should we establish around AI use?









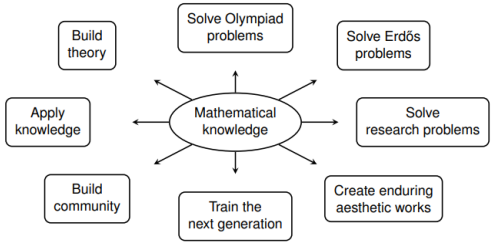


Recommendations for individual mathematicians are to disclose tool use, retain responsibility for correctness, affirm humanity of authorship, put effort into proper attribution, participate in public discourse, stay informed about emerging technologies, ...

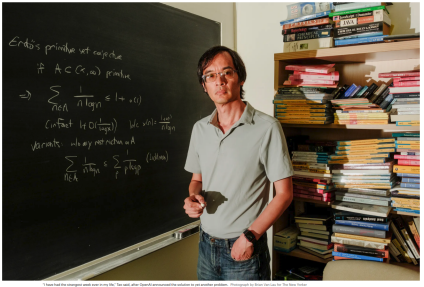
What are the goals and values motivating mathematical research?



Excessive AI optimization may in fact cause the many (previously largely aligned) goals of mathematics to diverge[3] from each other.



[3]The diagram is extremely oversimplified; in particular, it should be much higher dimensional.



From a talk entitled Mathematics in the Age of AI delivered by Terry Tao at the 2026 International Congress of Mathematicians.

What should we tell our students?



The Enduring Value of Math, in an Age of AI

Thoughts I'll share at the start of my upper division math course



FRANCIS SU

AUG 20, 2026

- *Why get good at problem-solving if an AI can do that much more efficiently?*
- *Why construct mathematical arguments, when AI models can now produce cleaner, slicker, more well-reasoned arguments than I can?*
- *What's the point of being a math major?*

The good news is that doing mathematics is not ultimately just about the particular problem-solving skills you are learning, nor about the outputs you are producing. It's actually more about the kind of person you are becoming. Said another way, *mathematical thinking is a set of virtues, not just a set of skills.*

Open problems



There are many important questions that are yet to be answered:

- How can we collectivize and democratize access to resources?
- How do we reinforce open scientific practices, such as sharing work in progress?
- How do we save the mathematics literature?
- How should we update our criteria for admissions / hiring?
- What should be the requirements for a PhD thesis?
- Do we need to abandon the practice of attributing authorship to mathematical discoveries?

Tack så mycket!