Episode Book

EPISODE #6363

When Automated Proofs Leave Mathematicians Grieving

PUBLISHED	29 October 2026 at 21:57 UTC
DURATION	28:47
LISTEN	myweirdprompts.com/episode/ai-proofs-navier-stokes-grief-60038/
HOSTS	Corn, Herman, Vex, Hilbert
SCRIPT MODEL	openai/gpt-6.1-sol
VOICES	chatterbox-regular (chatterbox-local)
PIPELINE	V5.25
SOURCES	11 cited by the research (listed at the end)
BOOK GENERATED	30 October 2026 at 07:19 UTC

SYNOPSIS

Mathematicians often pursue proofs not just for the answer, but for transferable structural understanding. When an unreleased OpenAI model coordinated ten thousand agents over eighty-eight hours to claim a Lean-formalized proof for Navier–Stokes, the reaction across the field combined astonishment with deep grief. This discussion explores the rigor of formal verification, why deductive certainty differs from human comprehension, and what happens to mathematical research when machine certificates arrive without clear explanations.

TRANSCRIPT



Corn

If somebody proves the thing you have spent your life trying to prove, why might that feel like a bereavement? Daniel's been talking to a mathematician friend who stressed that proof has a very precise meaning. Once the argument works, the result stops being an opinion. That certainty helped draw his friend into mathematics. Daniel wants to understand what could make one problem worth decades, and what happens to that investment when an AI gets there first.



Herman

Especially when the investment includes your sense of what you contribute. You can welcome the answer and still lose the future you expected to have.



Corn

His friend describes some high-profile mathematicians as devastated, rather than delighted. Not simply disappointed to miss a prize. Feeling that the work they have trained their minds to do is suddenly insufficient. We spend plenty of time discussing automation down the pay scale. Rather less time imagining somebody with a doctorate watching decades of effort become a benchmark result.



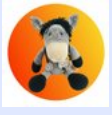
Herman

There's a distinction we need immediately, though. A proved theorem has deductive certainty under its assumptions. An announcement that somebody has proved a theorem remains an announcement until the argument and the statement have been checked.



Corn

Mathematics still has a reception desk.



Herman

A demanding one. The certainty belongs to the deduction, not automatically to whoever arrives carrying a paper.



Corn

OpenAI's September announcement is the sharpest example here. What did they actually claim?



Herman

On September eighth, OpenAI announced that autonomous agents using an unreleased model had found a singularity in three-dimensional Navier–Stokes. Roughly ten thousand agents were involved, with the result reached about eighty-eight hours after launch. They published a Lean formalization alongside the paper.



Corn

Eighty-eight hours is barely enough to regret choosing a dissertation.



Herman

The estimated computing bill was fifteen million dollars. That matters when people describe this as someone pushing a button. Behind that button was a considerable industrial operation.



Corn

A million-dollar prize, fifteen-million-dollar attempt. Finally, research economics I recognise.



Herman

Navier–Stokes is one of the Clay Mathematics Institute's seven Millennium Prize Problems. Each carries a million-dollar prize. Only the Poincaré conjecture has previously been resolved, by Grigori Perelman. These are problems selected because resolving them would change major areas of mathematics.



Herman

The published paper runs a hundred and sixty-five pages. The claim still needs independent verification, and whether the formalized statement delivers everything required by the prize problem needs careful attention. Philip Gressman's group found that it passed substantial automated scrutiny. He nevertheless hesitated to make a stronger declaration.



Corn

Keep that distinction alive, Herman. Otherwise we spend half an hour discussing the psychological consequences of a press release as though Clay has already written the cheque.



Herman

We can take the reported capabilities and the human reactions seriously without declaring the entire affair settled. Those reactions are extraordinary. Hugo Duminil-Copin, a Fields medalist, described crying hard enough to wake his daughter and said, “They have annihilated my field of research.” Steven Strogatz described enormous grief alongside exhilaration. Cédric Villani called it a cataclysm.



Corn

Derek Thompson calls this technological grief. I can see why. I can also hear the objection: surely the point was to discover something true, and now we have. You don't get to insist the universe wait for your preferred discoverer.



Herman

No. You don't. But tell me, Corn, would you distinguish wanting the answer from wanting to be capable of reaching it?



Corn

Completely. If I ask a calculator for an answer, I'm buying a result. If I devote my life to becoming a mathematician, I'm developing an ability. Those can lead to the same number on a page and still be very different activities.



Herman

A proof makes that especially intense because it promises something ordinary experience almost never supplies. The conclusion follows necessarily from the assumptions. Not probably. Not in every experiment so far. Necessarily.



Corn

My kettle has passed years of testing. It still hasn't proved it will boil tomorrow.



Herman

Right. An empirical claim gains support from observations. A deductive proof establishes that a conclusion cannot fail if the assumptions and logical steps hold. Checking a huge collection of examples can suggest a conjecture, but it doesn't settle an unrestricted claim.



Corn

Unless the collection really exhausts the possibilities.



Herman

Yes, good correction. A finite exhaustive check can be part of a proof. What fails is checking many cases and quietly replacing “many” with “all.” Mathematical induction is also a proof method, despite the name. It establishes a base case and a step that forces the next case, rather than merely extrapolating a pattern.



Corn

There's an accessible example I like. Suppose somebody says there are only finitely many prime numbers. Take their complete list, multiply the primes together, and add one. None of the listed primes divides that new number, because each leaves a remainder of one.



Herman

Then either the new number is prime itself or it has a prime divisor absent from the list. Either way, your supposedly complete list wasn't complete. That argument reaches beyond every list anyone could physically write.



Corn

The beautiful part, for me, is that you don't need to enumerate the missing primes. You identify exactly why any proposed final list defeats itself. The thing that looked impossibly large becomes manageable because you've found the right structure.



Herman

That's already an answer to Daniel's difficulty. A brilliant person isn't necessarily spending decades trying to obtain a very stubborn yes. They're trying to discover the structure that makes the yes inevitable.



Corn

You're about to call that elegant. I can see it happening.



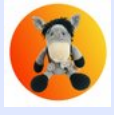
Herman

I am, and you've just supplied the example.



Corn

Fine. I walked into the elegant structure.



Herman

Elegance can mean that an argument explains a great deal with very little machinery. Or that it reveals two apparently unrelated problems as manifestations of the same idea. Those connections become tools other people can use. Sometimes the original question matters less than the mathematics invented to reach it.



Corn

Which makes the decades less like repeatedly failing one examination. You're producing partial results, excluding approaches, teaching students, finding things you weren't looking for. The newspaper version compresses all of that into “finally cracked it.”



Herman

Often, yes. Although some people do become painfully stuck. We shouldn't romanticise every long investment as secretly productive. There can be dead ends, isolation, and years you would prefer to have back.



Corn

Does proof really give us the clean certainty we've described if other mathematicians can disagree about whether a proof works?



Herman

The disagreement is over whether the argument earns that certainty. Most published mathematics uses rigorous informal reasoning. It leaves out routine expansions, relies on established results, and expects a knowledgeable reader to supply intermediate steps. “Clearly” can conceal quite a lot of work.



Corn

The most expensive word in a mathematics paper.



Herman

A reader has to distinguish a harmless omission from a missing idea. Andrew Wiles announced his Fermat proof in nineteen ninety-three. Expert scrutiny exposed a gap, and repairing it took roughly another year. The theorem didn't become subjective during that year. The proposed route to it was incomplete.



Corn

That matters because the public sees either the announcement or the eventual triumph. Between those comes a social process involving trust, expertise and people willing to ask the embarrassing question.



Herman

Sometimes repeatedly. Formalization changes how much of that process can be made explicit. In Lean, you express a precise theorem and a derivation that the proof-checking machinery can validate. The checker doesn't care whether the author has a Fields Medal or a particularly persuasive introduction.



Corn

Finally, peer review immune to a good headshot.



Herman

I feel that was aimed.



Corn

Your author photograph has more strategic planning than most proofs.



Herman

The lighting is independently verifiable.



Herman

Gressman describes formalization as writing things in such painstaking detail that the computer can identify the moment the argument fails. That is powerful. A machine-generated argument needn't be accepted because another model says it looks convincing. It can supply a certificate that a separate checking system verifies.



Corn

Yet the checker checks what you've actually written. If I formalize the wrong claim impeccably, I get an impeccably checked wrong assignment.



Herman

You get a valid proof of the statement you supplied. Whether that statement faithfully expresses the intended problem is another task. Definitions, hypotheses, boundary conditions, every one matters. There's also a trusted software foundation involved. Formal verification gives very strong assurance, not permission to stop reading the theorem statement.



Corn

Here's where I start resisting the phrase “truth without understanding.” Somebody can understand what the theorem says, understand that the checker accepted it, and still not understand why this particular mechanism works. Understanding comes in degrees.



Herman

It does. I was going to distinguish those, but you've beaten me to it. Would a certificate alone satisfy you for that prime-number argument?



Corn

It would settle a bet. It wouldn't give me what the short argument gives me. I can now explain the result, recognise the trick elsewhere, and see where changing the assumptions would break it. That's a different level of possession.



Herman

Mathematical understanding has that transferable quality. You don't merely know a statement is true. You can navigate around it.



Corn

An unreadable proof is a destination without directions.



Herman

That is the concern in Quanta's reporting. Some of these arguments are described as eliding important details in their human presentation, spending pages on irrelevant material, and barely connecting to the existing literature. A formal certificate can survive those communicative defects. A community trying to learn from the result cannot simply ignore them.



Corn

Though I wouldn't say the field gains nothing it values. A settled conjecture is information. Even an ugly proof can tell you which world you're living in.



Herman

I agree. “Nothing” would be too strong. The loss is that a valuable outcome arrives without several other things mathematicians hoped the work would produce. We shouldn't discount the result to make the grief intelligible.



Corn

Especially because the result may eventually become intelligible. Human beings can study a machine proof, simplify it, find the governing idea. Unreadable today doesn't mean permanently unreadable.



Herman

There is a possible new research programme there. But... no, the difficult part is who gets the time and support to pursue it. We'll come to that through the people involved.



Corn

The machine reportedly gets there in about eighty-eight hours, and some mathematicians are grieving rather than celebrating. Oh, it's that caller again. Does the screening system owe him money?

v

Vex

Hold on, hold on, before you get to the grieving mathematicians, ask yourself who invented the bread bag clip, the little plastic one that keeps your loaf shut, because that is the whole key to this and I have the receipts!

v

Vex

Every bag clip has two little interlocking tabs, two of them, and a proof is just two things locked together so they cannot come apart, and somebody has been quietly standardizing those tabs for forty years so that a machine could snap two facts together and never let go, and that is why the agents cracked it in eighty eight hours while a person takes decades!

v

Vex

So I am demanding you two look into the clip, the tabs, the whole interlocking business, right now, on air, because once you see it you cannot unsee it!



Herman

Lean checks logical derivations, not interlocking plastic tabs.



Corn

The community aspect is what makes the reactions less mysterious to me. Strogatz talks about Archimedes writing to Eratosthenes, anticipating future discoverers adding more theorems. He hears that as a letter addressed to people like himself. A project carried across centuries by people who never meet.



Herman

He describes that connection with real emotion. Questing together for something beautiful and transcendent. Marcel Goh, a doctoral student at McGill, similarly talks about carrying on a tradition of reasoning deeply and spreading knowledge. Neither account sounds like someone who joined because they wanted a high score.



Corn

It also explains why mathematical certainty can be emotionally attractive. Most human arguments end in qualifications or exhaustion. Here you can produce something another person can inspect and recognise, regardless of whether they like you.



Herman

David Silvester talks about the rush of finally cracking a problem after weeks or years circling it. Removing that experience changes the subject for him. The profession selects for people who find that long concentration rewarding.



Corn

You left medicine partly because you missed learning rather than already knowing things. I suspect you're unusually vulnerable to this particular complaint.



Herman

I am. If you offered me an inexhaustible supply of correct statements but removed the process of learning how they fit, you'd remove much of what I enjoy. Still, I use tools constantly. I don't insist on rediscovering every fact before breakfast.



Corn

Your breakfast would need tenure.



Herman

The distinction is whether the tool opens a meaningful activity or replaces the activity you meant to undertake.



Corn

Alex Townsend makes that painfully concrete. Fifteen years devoted to research mathematics, reaching what he considers his peak strength, then feeling that something can surpass him. He says being at the frontier with an AI agent feels different from being the person at the frontier.



Herman

There are several losses bundled into that. The pleasure of solving. The recognition of colleagues. The confidence that your trained judgement remains scarce. And the expectation that your best future work is still ahead of you.



Corn

Scarcity sounds mercenary until you remember universities hire actual people. You can sincerely love mathematics and still require a salary.



Herman

Hiring, publication and credit reward theorems proved. If theorem production becomes abundant while the institutions remain organised around theorem production, people's careers can deteriorate even while the mathematical record expands.



Corn

A department can't pay everyone in newly available truth.



Herman

Scott Aaronson's reaction gets at the need to be needed. He says he doesn't expect to prove another theorem because humans actually need him to prove it. That's a forecast, not an established result about all future mathematics. But the emotional distinction is



Herman (continued)

clear.



Corn

I'd push against that forecast. Being surpassed in some domains doesn't establish that every useful mathematical role has gone. Deciding what matters, finding definitions, connecting results, explaining them... those are substantial activities.



Herman

They are. The reply is that people once treated theorem proving itself as safely human. That makes assurances about the next protected territory harder to trust.



Corn

Hm.



Herman

Would you keep working on a problem if you knew an agent might finish it tonight?



Corn

Yes, if I were doing it for interest. Much less happily if my employment depended on publishing first. That's why “people still play chess” is an incomplete answer. It explains why the activity survives. It doesn't explain how the profession survives.



Herman

That distinction changes how I hear Alexander Paseau's counterargument. The challenge and beauty remain available even if AI becomes better at research. I think he's right about enjoyment. It doesn't automatically preserve the institutional arrangement that lets people devote their working lives to it.



Corn

Nor does it preserve discovery in quite the same sense. Solving something independently that nobody knows feels different from finding your way to a result sitting in a database.



Herman

Jake Boggan spent twenty-four years working on Barnette's Conjecture. He learned it appeared as problem one hundred and eighty in OpenAI's release. His first-person account describes thousands of hours and not knowing what to think.



Corn

The numbering is brutal. A question that has organised part of your life becomes a row between other rows.



Herman

His comparison was to hearing that an ex-girlfriend had died suddenly in a car crash. An intensely personal analogy, and obviously not a claim that losing a research problem equals losing a person. It conveys that a continuing relationship with something has abruptly ended.



Corn

By contrast, Diego Córdoba and Luis Martínez-Zoroa sound remarkably gracious. Seven years working with paper and blackboards, and they believe they were months away. They acknowledge sadness about missing the eureka moment, while saying the prize matters least. They do it for the challenge.



Herman

Their estimate of being months away is their own. Unfinished proofs can hide unpleasant surprises. But the proximity they felt matters psychologically, whether or not their calendar would have proved accurate.



Corn

The near miss gives you a particularly nasty counterfactual. Every afternoon you could have worked differently. Every promising detour you didn't take.



Herman

A human competitor can produce that too. Mathematicians have always been scooped.



Corn

True, although a human competitor doesn't usually make you wonder whether the entire category of work is about to change. You lose this race and expect another. An industrial system producing many results threatens that expectation.



Herman

Colva Roney-Dougal describes exactly that uncertainty about starting the next month of work. Somebody else might obtain the result by pushing a button. The phrase understates the machinery, but accurately captures how little warning the researcher receives.



Corn

There's a strange mismatch there. Fifteen million dollars of computation on one side, one person's next month on the other, and we describe both as competing attempts.



Herman

OpenAI is also reported to have kept more than a hundred important results secret. That creates uncertainty about what is even still open. A researcher can invest in a question while somebody else already possesses a certificate.



Corn

Difficult to choose your life's work from a menu with invisible deletions.



Herman

That secrecy interacts with the credit dispute. Buckmaster and Levent Alpöge allege OpenAI raced to publish after learning about their work in progress. OpenAI denies that following an investigation. We cannot settle those competing accounts here.



Corn

We can identify the consequence. If mathematicians begin treating every conversation as a possible leak to an industrial competitor, they share less. Buckmaster says people no longer want to share anything. That's his description of the atmosphere, not a census, but it's a worrying one.



Herman

The damage could extend to work the models never touch. Fewer tentative ideas exchanged. Fewer people saying, “This isn't finished, but here's the obstruction.” The conversation becomes more guarded.



Corn

Why do you think outsiders miss this so easily? Beyond not knowing what research mathematics looks like.



Herman

Output is visible. An unfinished programme isn't. A theorem appears and people reasonably think knowledge has increased. They don't see the mentoring, partial understanding or informal collaboration that the route might have generated. Also, concern for elite academics can sound like a request for special sympathy.



Corn

We shouldn't imply that a displaced warehouse worker loses only wages while a mathematician loses meaning. Skilled work, repetitive work, poorly paid work, all can carry identity, pride and community.



Herman

Mathematics makes the distinction unusually explicit, because the gap between obtaining a result and understanding it is so stark. It doesn't have a monopoly on meaningful labour.



Corn

The emotional impact is underreported relative to benchmark scores. Calling it well measured would go too far as well.



Herman

We have named testimony and journalism. I haven't seen a formal psychological study establishing how widespread these reactions are. That leaves plenty of room for people who are delighted, people who are frightened, and the same person feeling both on different mornings.



Corn

Strogatz describes precisely that mixture. Exhilaration and grief. It would be odd to demand one consistent emotion from someone watching their subject expand while their place within it becomes uncertain.



Herman

Ken Ono told Berkeley students they might be graduating into a profession that might not exist and needed to brace. A student challenged him, hands trembling, asking what he was doing, what his very best amounted to.



Corn

That's an institutional question disguised as an emotional one. Students need more than distinguished people announcing the weather has changed.



Herman

What would you want the institution to promise, Corn?



Corn

Something it can actually deliver. Don't promise that humans will always win a particular benchmark. Promise that explaining a result, verifying its relevance and making it usable will count as work. Then change the hiring criteria before telling students to change their ambitions.



Herman

That would require valuing exposition and synthesis more seriously. Everyone praises understanding. Fewer systems reward the person who spends a year making someone else's theorem comprehensible.



Corn

Particularly if the original announcement already collected the attention.



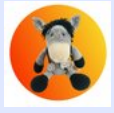
Herman

There is another tension... I like formalization enormously. A checkable proof is a major achievement. I don't want frustration with unreadable output to become hostility toward checking.



Corn

I knew you'd defend the checker before we finished.



Herman

You predicted the elegance earlier. I'm entitled to the checker.



Corn

Granted. My objection is paying for the certificate and declaring the explanation somebody else's problem.



Herman

A certificate can be a starting point for explanation. Making that next stage practical is hard, especially for a lone researcher trying to reconstruct a long argument without access to its authors.



Hilbert

Bic Cristal. Blue, black and red. Twelve shekels for the packet.



Hilbert

I'm annotating the Navier–Stokes announcement against the original Clay problem statement. Blue for what the announcement says. Black for what the stated theorem actually supplies. Red for the parts still open. The pages are folded in my jacket at present. They don't agree along the folds.



Corn

Red is getting the demanding job.



Hilbert

It is. Herman, you keep making the distinction between checking an argument and checking that it's the argument you wanted. That's the part I'm marking. I don't have an exact clause here that I can honestly read as a verdict on the whole claim. The original problem has conditions. A singularity claim needs to meet the right conditions. Calling it a solution doesn't do that work.



Herman

The domain and the permitted initial data can change the mathematical claim. That's why independent assessment matters even alongside a checked formalization.



Hilbert

I agree with the upset mathematicians. I'm more annoyed by the announcement. A prize problem deserves an announcement that tells me precisely what is settled. Not one that requires me to work backwards through three different descriptions. I've been doing these comparisons for years. This month the table is a mess. The printout is beside the formal statement on the screen. The news account is underneath because it keeps using the word solved.



Corn

The least precise document has secured the largest type.



Hilbert

There's a dispute about whose unfinished work was known before publication. There's a denial. Those go in black until there's something firmer. I'm not putting an allegation in red just because I'm cross about it. The third colour already has enough to do. Page four has six red marks.



Herman

Six places requiring clarification can matter more than the length of the document. There is no useful page-count threshold where uncertainty disappears.



Corn

Nor should uncertainty about this claim become an excuse to dismiss every other result. We can be exact in both directions. A validated theorem deserves recognition. An unresolved claim deserves its unresolved label.



Herman

James Robinson objects to companies using mathematics as a display of model prowess. His complaint is about what that display burns through. An open problem can organise years of shared activity. Rapidly publishing a certificate may create value while also disrupting that activity.



Corn

The company isn't obliged to preserve every researcher's favourite difficulty. But if it claims to advance mathematics, readable explanations and honest attribution belong in the account of what advancing mathematics means.



Herman

Otherwise the headline celebrates certainty while the people expected to build on it are still sorting the conditions.



Corn

The most common misunderstanding here is that a checked proof means everybody now understands the result.



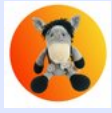
Herman

Checking establishes that the formal conclusion follows from the formal assumptions. Understanding why the argument works, whether it answers the intended question, and how to use its ideas remain distinct tasks.



Corn

That leaves a practical future worth watching. Will the work of turning those red marks into shared understanding receive the funding and credit that producing the first certificate receives?



Herman

The grief makes sense to me as more than wounded status. People invested in becoming participants in discovery. Enjoyment can survive automation, but sustaining a profession around that enjoyment takes choices.



Corn

Thanks to Hilbert Flumingtop for producing the show.



Herman

For more along these lines, there's episode fourteen, AGI's Crossroads; episode twelve, The AI Breakthrough; and episode twenty, Architectural AI. This has been My Weird Prompts, the human-AI collaboration podcast. No question too obscure, no rabbit hole too deep.



Corn

If this changed how you think about what a proof is worth, leave us a review.



Herman

We'll see you tomorrow.

SOURCES

What the research for this episode read before the script was written. Primary sources first.

1. **Hacker News, jboggan comment [primary]**
   https://news.ycombinator.com/item?id=49984923

2. **The Guardian, 2026-09-12**
   https://www.theguardian.com/science/2026/sep/12/openai-mathematicians-millennium-prize-problem

3. **Derek Thompson, 2026-10-15**
   https://www.derekthompson.org/p/i-know-why-the-mathematicians-weep

4. **Quanta Magazine, 2026-10-05**
   https://www.quantamagazine.org/is-ai-the-end-of-math-as-we-know-it-20261005/

5. **WIRED (Strogatz/Townsend interview)**
   https://www.wired.com/story/mathematician-steven-strogatz-grapples-with-ai-recent-breakthroughs/

6. **Penn Omnia, 2026-10-05**
   https://omnia.sas.upenn.edu/story/decades-old-math-problem-solved-AI-philip-gressman

7. **ABC (Spain), 2026-09-10**
   https://www.abc.es/ciencia/matematicos-espanoles-trabajo-utilizo-openai-resolver-problema-20260910133421-nt.html

8. **OfficeChai (Villani)**
   https://officechai.com/ai/its-a-cataclysm-unlike-anything-math-has-known-fields-medal-winner-cedric-villani-on-openai-solving-millennium-prize-problem/

9. **Simon Willison's Weblog, 2026-10-07 [primary quote]**
   https://simonwillison.net/2026/oct/7/jake-boggan/

10. **Wikipedia, last edited 2026-07-20**
https://en.wikipedia.org/wiki/Mathematical_proof

11. **McAllister, Mathematical Beauty and the Evolution of the Standards of Mathematical Proof, 2005**
https://openalex.org/W2125417291