Eight Became Late

A missed mathematics checkpoint, a limited token budget, and a narrower research plan for Hedegreen Research.

2026.10.28 15:10 DENNIS HEDEGREEN CORRECTIONS V1.0
HTTPS://HEDEGREENRESEARCH.COM/ARTICLES/EIGHT-BECAME-LATE/

A missed deadline, a changing research frontier, and a different promise.

On October 20, I was supposed to return to a mathematical investigation I started in July. I had made the date public, preserved the original data, and promised to examine what had happened since. The plan was straightforward: take another snapshot, compare the results, and see whether the growing involvement of artificial intelligence in mathematics was changing the frontier of unsolved problems.

Today is October 28. I haven't done it.

I have been thinking about that more than I expected. There is something uncomfortable about making a public commitment and then watching the date pass without delivering. Nobody was standing over my shoulder demanding the next report. I chose the deadline myself. Perhaps that is precisely why missing it bothers me. If I want Hedegreen Research to take public commitments seriously, that has to include the commitments I make when nobody else is enforcing them.

The strange part is that I am not particularly disappointed with how I spent my time. Quite the opposite. October has been full of things that mattered, both personally and professionally. I spent time with my son, dealt with the usual complications of life, and continued rebuilding much of the infrastructure behind Hedegreen Research. I would not trade those days with my son for a more punctual publication calendar. But that doesn't make the unfinished work disappear from my mind.

And meanwhile, the world has been moving at a frankly ridiculous speed. I am falling behind on something I genuinely want to investigate, partly because the very technology I am investigating keeps producing new things worth investigating. It is frustrating, fascinating, and, if I am being fair, rather wonderful.

When Nine Became Eight

In July, I published *Nine Became Eight*, an investigation into nine mathematical problems that had reportedly been resolved through AI-assisted research. The original announcement was exciting, but I wanted to know exactly what the number nine represented. Were these nine previously unsolved problems, nine valid mathematical results, or nine cases that had been classified as new discoveries?

That distinction turned out to matter. One of the problems, Erdős problem 846, had a computer-checked proof, but an earlier published theorem already implied the result. The proof could be valid without the underlying mathematical conclusion being new. Under the classification I was using, the count of newly resolved problems therefore fell from nine to eight.

It was a small correction to a number, but an important reminder of why scientific claims need more than impressive headlines. When an AI system produces a mathematical proof, checking the logic is only part of the work. Someone must also establish what was already known, whether the result addresses the original question, and what kind of contribution has actually been made.

I wanted to follow those nine cases over time. Not because nine problems could tell us everything about AI and mathematics, but because a small, well-documented sample might reveal something about how mathematical discoveries are verified, classified and eventually accepted. I marked October 20 for the next examination.

I still think that was a worthwhile idea. I just didn't deliver the second measurement.

The World Didn't Wait

While I was working on other things, developments in AI-assisted mathematics accelerated dramatically.

On October 6, OpenAI published a collection of 722 mathematical manuscripts grouped into 372 result families. The scale alone was extraordinary compared with the nine cases I had been investigating. And even those numbers were not static. Three manuscripts were subsequently withdrawn, others were revised, and two more were added later in the month. By October 28, the collection stood at 721 manuscripts in 373 families.

The withdrawals were not merely administrative changes. OpenAI reports that a sign error invalidated an argument and the construction used by two dependent papers, leading to three withdrawals. Other changes included proof repairs, clearer assumptions and updated references. Together, they show why the collection needs a history: a manuscript can be added, revised or withdrawn, and a count taken on one date may no longer describe it three weeks later. None of these numbers should be interpreted as a count of newly solved open mathematical problems without further examination.

That is precisely the kind of distinction I had been trying to investigate in July. Except now, rather than examining nine cases, I was looking at hundreds of manuscripts and an increasingly complicated landscape of research claims, formal verification, independent assessment and prior mathematical literature.

Other people have started building systems to track this landscape. New registries are attempting to distinguish open problems from submitted solutions, contested claims and accepted results. Some of this work overlaps with what I had imagined for my own Problem Frontier Tracker, and some goes considerably further.

My first reaction was a mixture of admiration and the uncomfortable realization that I could not realistically keep up. But after thinking about it, I began questioning why I should try.

If other researchers are building better and more comprehensive registries, that is good news. I don't need to duplicate their work simply because I had an idea for something similar a few months earlier. I can use their findings, examine their methods, challenge their classifications when there is reason to, and contribute where I have something distinct to offer.

There is something genuinely exciting about reaching a point where the problem is no longer finding enough interesting scientific developments to investigate. The problem is choosing among more developments than I can reasonably follow.

I don't think I have ever been so pleased to be falling behind.

Twenty-Four Hours, Apparently Still the Limit

There is another reason the October deadline passed. Actually, several.

Much of my recent work has gone into V3, a substantial rebuild of Hedegreen Research itself. I have been developing the systems that support articles, sources, editorial histories, tools and the relationships between them. It is not especially glamorous work from the outside. A reader visiting the site may see another article, while behind that article sits an increasingly complicated attempt to make the research more transparent, reproducible and maintainable.

The irony is difficult to miss. I missed a deadline for an investigation about scientific verification while building infrastructure intended to improve how my own work can be verified.

I could present that as a clever justification, but it would be too convenient. The infrastructure may be valuable, and I believe it is worth building, but the decision still had consequences. I allocated time to one thing instead of another. The promised follow-up didn't happen. Both facts can be true without one cancelling the other.

There is also a practical limit behind the research: my available token budget. Using AI to investigate mathematical claims means spending resources on reading papers, checking interpretations, exploring approaches and revisiting answers that may be wrong. I do not currently have the token budget to sustain that process across the wider field at the level of care I would want.

If Hedegreen Research eventually becomes the research environment I hope to build, with the resources to support sustained investigation and independent checking, that calculation might change. More tokens would not make a mathematical claim correct. They would give me more room to investigate it properly. For now, I need to choose questions I can afford to examine carefully.

And then there is the rest of my life, which stubbornly refuses to organize itself around my research calendar.

I am also working toward finishing my HF education, part of the Danish upper-secondary system. I have four assignments left to complete, followed by preparation for my exams. So somewhere between mathematical conjectures, AI agents, software development and parenting, I also need to sit down and finish my homework.

There is something wonderfully absurd about that situation. I am studying within an education system while researching technology that may fundamentally change how knowledge is produced, evaluated and taught. AI is beginning to solve difficult mathematical problems, generate scientific hypotheses and perform tasks that once required substantial specialist work. Meanwhile, I still have four assignments waiting for me.

I sincerely hope AI can refrain from changing anything too fundamental about education until after my exams. It would be nice to finish the qualification before the world starts seriously reconsidering what parts of it should look like.

Of course, I am joking. Mostly.

Learning to understand something is not the same as having a machine produce an answer, and that distinction may become more important rather than less. But I do sometimes wonder how quickly education can adapt when the tools available outside the classroom are changing faster than the curriculum inside it.

Perhaps that is another research question for another day. Preferably after I have handed in those four assignments.

A Different Kind of Follow-Up

I am putting the planned tracker follow-up on hold. I am also setting aside the broader ambition of developing a comprehensive, continuously updated register of AI-generated mathematical discoveries.

That is a change in research direction, not a change to the historical findings of *Nine Became Eight*. The original July investigation remains a record of what I examined and how I classified it at the time. I have not performed the promised October re-audit, so I cannot claim to know whether every classification would remain unchanged today.

I still want to return to those nine cases. There is value in examining a defined group of claims over time, especially when the original evidence has been preserved. But I no longer think that following the entire mathematical frontier is the best use of my limited research capacity.

Instead, I want to explore smaller questions in greater depth. We might take a mathematical problem that has attracted little attention, examine the existing literature, document an AI-assisted attempt, and independently verify whatever the system produces. Perhaps it discovers something interesting. Perhaps it reproduces a result that mathematicians already knew. Perhaps it fails entirely.

I recently tried a modest experiment of this kind with Erich Friedman's *Polyform Floes*, a mathematical puzzle involving random movement across a grid of connected squares. The immediate goal was to find a connected shape of squares in which the expected survival time is a whole number at every square, with at least one square having a value of 11. I could verify a smaller construction with a maximum value of 9, but I did not find the requested 11. That does not establish that 11 is impossible, and it certainly isn't a mathematical breakthrough.

Yet I enjoyed the experiment. There was a clearly stated problem, a method, a computational attempt, and a result whose limitations could be described without exaggeration. It was something I could investigate rather than another enormous collection of claims I was trying to keep up with.

And perhaps that is where Hedegreen Research can make a more useful contribution. Not by trying to count everything AI discovers, but by investigating particular claims carefully enough to understand what they actually mean.

The Uncomfortable Part

I am still bothered by the missed deadline. I could have quietly removed the date from my plans, published something else, and assumed nobody would notice. Perhaps nobody would have. But the problem with that approach is not whether the audience notices. It is that a public research record becomes less useful when inconvenient parts of its history quietly disappear.

I would rather acknowledge that I made a commitment, explain why I didn't meet it, and say what I intend to do differently. That is not quite the same as correcting a mathematical error. It is correcting the public account of what I said I would do.

At the same time, I don't want this article to become an apology for having a life outside research or for failing to compete with the pace of increasingly powerful AI systems. I am one person building an independent research project. I have a son, an education to finish, other responsibilities, limited resources, and far more ideas than I can investigate. Those constraints are real, and ignoring them would not make the research more rigorous.

What matters is being honest about the choices they force.

Looking back over October, I find myself with two apparently contradictory feelings. I feel guilty about the work I promised but didn't complete. And I feel almost absurdly optimistic about the scientific developments that helped make the original plan less sensible.

AI systems are beginning to produce mathematical arguments and scientific hypotheses at a scale that would have sounded extraordinary not long ago. Whether the most ambitious claims survive independent examination remains to be seen, but the pace of experimentation is remarkable. Researchers are being challenged not only to solve problems, but also to develop better ways of evaluating an expanding body of proposed solutions.

I want to be part of investigating that change. I just don't think the best way to do it is to run after every new result with a clipboard, desperately trying to keep the total current.

In July, I questioned whether nine really meant nine and ended up with eight. In October, I discovered that even following eight can be more complicated than expected. I missed my deadline, changed my plan, and found myself more excited about the future of scientific research than I was when I first began the investigation.

I would still like to catch up eventually. Just not at the expense of everything else worth doing.

And if the world could avoid another fundamental scientific revolution until I have finished my HF exams, I would appreciate it.

Although, knowing myself, I would probably be disappointed if it did.

— Dennis Hedegreen, updated publicly

Sources and research notes

- *Nine Became Eight* (https://hedegreenresearch.com/articles/nine-became-eight/) — Original investigation, July 21, 2026.
- *Corrections Need Memory* (https://hedegreenresearch.com/articles/corrections-need-memory/) — Editorial approach to corrections and historical records, October 8, 2026.
- OpenAI: Sharing AI Progress in Mathematics (https://openai.com/index/sharing-ai-progress-in-mathematics/) — October 2026 mathematics release.
- OpenAI mathematics initial catalogue (https://github.com/openai/math/blob/adc7f1241b42e322a6451854ab7e4b4c146bf78a/README.md) — October 6 release: 722 manuscripts in 372 families.
- OpenAI mathematics repository (https://github.com/openai/math) — Manuscripts and classifications; catalogue checked on October 28, 2026.
- OpenAI mathematics release history (https://github.com/openai/math/blob/main/history.md) — October 7 withdrawals and revisions; October 26 additions.
- Plasma AI: An Open Record of the Erdős Problems (https://www.plasma.ai/research/erdos) — Independent problem-tracking initiative.
- Erich Friedman: Polyform Floes (https://erich-friedman.github.io/mathmagic/0416.html) — Original mathematical puzzle. The experimental findings described above are preliminary and do not resolve the open case.

RELATION MEMORY

RETURNS TO Nine Became Eight · Returns to the missed October 20 follow-up commitment and revises the research plan; does not correct or re-audit the July mathematics.

NEARBY Corrections Need Memory · Provides the editorial distinction between historical records, factual corrections and visible public self-revision.

CANONICAL URL: HTTPS://HEDEGREENRESEARCH.COM/ARTICLES/EIGHT-BECAME-LATE/