

Fields Medalists Are Wrong: Why AI Is an Accelerant, Not a Bane, to Progress in Mathematics

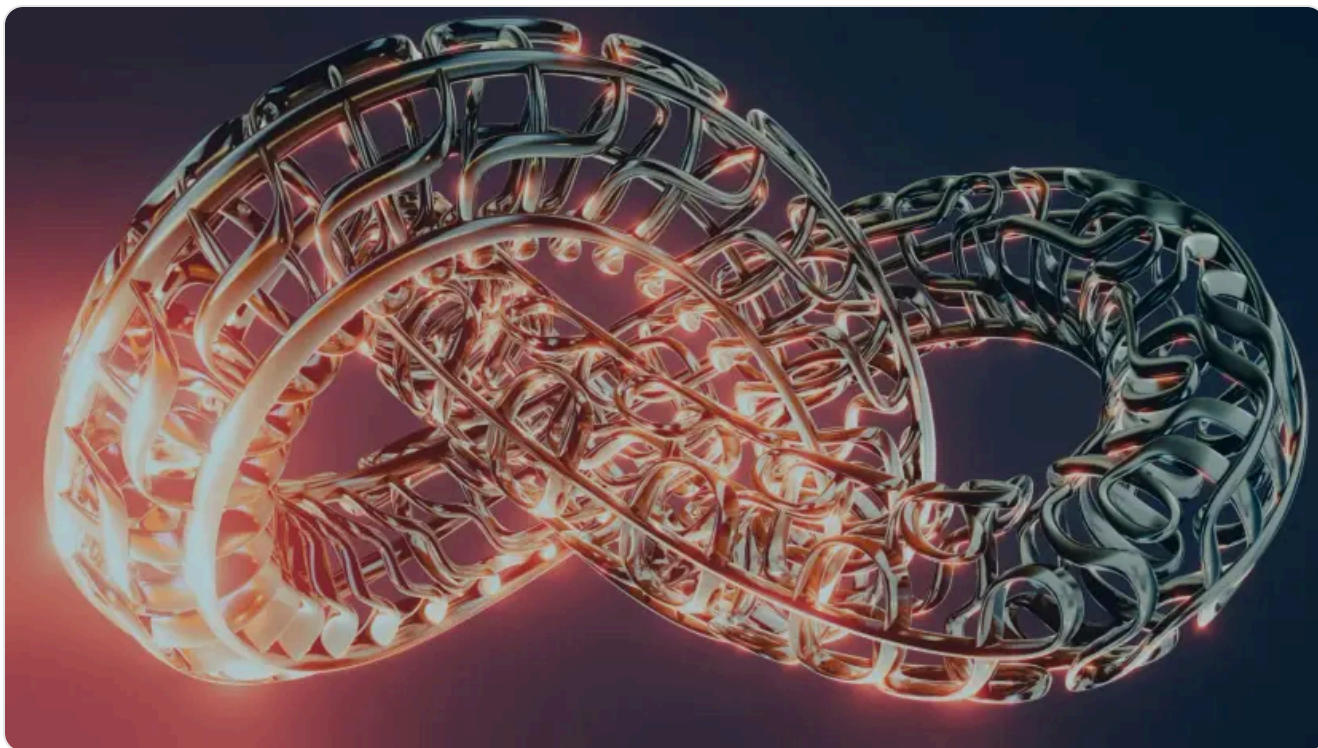


Photo by Steve A Johnson on Unsplash

On the heels of OpenAI having solved the Navier–Stokes Millennium Prize Problem, about half of the living Fields Medalists (a Fields Medal is one of the most prestigious prizes that one can earn in mathematics) signed an open letter titled “A Severe Misalignment of AI in Mathematics”. As of October 4th, 2026, the letter was ultimately signed by over 8,000 (and counting) mathematicians and academics in related fields.

Conjecture Institute Fellow and mathematician Paul Raymond-Robichaud and President Logan Chipkin disagree with the spirit of the letter, which is essentially that AI is a *problem* for mathematics. On the contrary, it is a *catalyst* that opens entirely new vistas to mathematicians that were heretofore unreachable.

Below is Paul and Logan's response to the letter. Passages from the letter appear in shaded boxes; the responses follow in regular text.

FROM THE LETTER

Over the last few months, the mathematical capabilities of LLMs have improved dramatically, to the point that they can solve major outstanding problems in many fields of mathematics. However, the push by AI companies to solve mathematical problems as a benchmark is detrimental to the science of mathematics, and to the mathematical community. The goals of the AI companies and the goals of the mathematical community are severely misaligned. We see these as part of broader alignment issues impacting other scientific and creative professions, as well as the whole of society.

An AI company is entitled to advertise its product's remarkable ability to solve a mathematical problem, and such a company's aspiration to solve mathematical problems does not inhibit mathematicians from doing math—as we will discuss.

Theoretically, the effectiveness of AI could be detrimental to the mathematical community *as the community is currently constructed*. No subculture is entitled to a static environment, including those that make up academia. As the culture of mathematicians currently stands, AI is indeed bound to be disruptive, and the culture will have to adapt to reap its benefits in full. But this is an opportunity for *improvement*, not a reason for *despair*.

There is currently a misalignment between the values mathematicians currently hold and the actual purpose of doing mathematics. The fact that machines are now able to prove or disprove conjectures by trying a wide array of known methods (within known theories) shows us that we have an obsolete criterion for what constitutes the role of a mathematician and the source of genuine progress in mathematics. The culture of

mathematics can no longer rationally reward solving mathematical problems that can be solved by AI mindlessly applying known methods. At one time, rewarding such activity made sense, but now that this sort of work can be automated, the culture of mathematics should instead reward discovering, understanding, and explaining those portions of the abstract world to which we have access.

Brute-Force Theorem Proving

In standard mathematics, proof verification is purely mechanical. By definition, a statement S is a theorem if there exists a finite proof P that derives S from the axioms under fixed rules of inference. Crucially, verifying whether an arbitrary text qualifies as a valid proof is an algorithmically decidable task, meaning there is a computer program that is guaranteed to finish in a finite number of steps and give a definitive “yes” or “no” answer by checking each step against the formal deduction rules. Because a proof is just a finite sequence of symbols, a machine can automate theorem proving through brute-force enumeration, which entails generating every candidate sequence of symbols in lexicographical order (from shortest to longest) and running this verification program on each one. The machine is guaranteed to eventually find and confirm a valid proof for S , provided S is indeed a theorem.

There exists a physically realizable algorithm that can find the proof of every mathematical theorem.

Mathematicians have long known that proving theorems could, in principle, be automated. This has not been cause for despair, since mathematicians have made progress on all of the *non*automatable aspects of their field, such as creating new and fruitful methods, theories, and criteria for what constitutes an interesting mathematical object or problem and explaining why. Implicitly, mathematicians have appreciated the dichotomy between proving theorems, which is automatable, and making mathematical progress, which is not.

Meanwhile, the range of theorems that AI can reach is narrower than that of brute-force search, as AI relies on pattern recognition and other heuristics rather than exploring every possible sequence of symbols. So if mathematicians think that the advent of AI spells the death of mathematics, then they should certainly feel that way about brute-force search! And yet they, rightly, have not given up on their pursuits in the face of what brute-force search can do (in principle). AI indeed accelerates the proving of a subset of all theorems, but it remains true that mathematics is about deepening *our understanding* of the mathematical world, and proving mathematical theorems is merely one aid that mathematicians exploit on their quest for progress.

FROM THE LETTER

Research mathematics deals with understanding basic structures of shapes, numbers, and natural phenomena. Over the course of generations, it has built a large corpus of sophisticated ideas, methods, abstractions, and other tools to comprehend the mathematical landscape. In turn, modern technologies and sciences are based on mathematical tools.

To date, AI has solved problems by executing man-made methods on man-made axioms, running over as many possibilities as its resource constraints will allow. But it cannot create *new* methods or identify interesting new axioms—that is still the job of mathematicians. AI can only perform the automatable task of running every already known method against every set of already known axioms and telling the user of the results. Moreover, AI will never pause upon discovering an interesting theorem or axiom but will instead blindly go on to generate the next one. A person, on the other hand, may (with some luck) realize that he has discovered something fruitful and investigate it rather than unthinkingly prove the next theorem in line.

What Is a Method?

A method is neither a theorem nor an exact algorithm, but an idea for converting an intractable problem into a solvable one. To give an extremely simple example: in combinatorics, one might prove that an object of a certain type exists by showing there is an odd number of them. The underlying theorem states that if a natural number n is odd, then $n > 0$. By itself, this theorem tells you almost nothing. Its true value lies entirely in building the intuition to recognize when this trick applies and how to deploy it. An AI could learn the pattern of which problems the method has previously solved, but it would still not know exactly when to apply and not apply the method.

One of the most important tasks of the mathematician is to create new theories, or sets of axioms and their implications, through which they can solve old problems that could not be solved in previous frameworks. For example, the problem of ‘squaring a circle’—given a circle, how can one construct a square with the same area as the circle?—was first asked by the thinkers of ancient Greece, well-known for their study of geometry. Mathematicians would continue to work on the problem for the next two thousand years *within the framework of geometric theory* without success. The problem was finally solved in 1882, when German mathematician Ferdinand von Lindemann stepped *outside* of geometry and approached the problem from *abstract algebra*. The puzzle of squaring a circle had its origins in geometry, but its solution had its origins in an altogether different mathematical field—and one that the Greek geometers had never even imagined.

AI could not have helped the ancient Greeks in solving this problem, trapped as they were in the prevailing mathematical paradigm. No amount of automation could have discovered the solution within geometry, since the solution was not *in* that theory but in an altogether different theory whose discovery would require genuine creativity.

In short: if a problem requires *either* a new method *or* a new theory, then AI as it is currently constituted is hopeless to solve it—and no amount of increasing its efficiency or computational power can overcome this limitation.

Before Navier–Stokes was solved by AI, we did not know whether it *could* be solved with known methods or whether it would need new ones. It turns out that it could be

solved with known methods, and that's why AI was able to solve it. But whether or not we need new methods to solve a given problem is in general unknown before we solve it.

Even if some future computer program is able to discover new axiomatic systems and theorems by brute force alone in short order, mathematicians would *still* have their most important work ahead of them—no program can explain the significance of any axioms it discovers, nor could it distinguish interesting axioms from boring ones in all cases (and the criteria by which mathematicians make such a distinction are themselves subject to improvement). As for theorems, the overwhelming majority of all mathematical theorems are trivial or meaningless, much as the overwhelming majority of all possible books are sheer babble. The creative mind of the mathematician alone can sift through the computer's database of theorems and, in light of his mathematical ideas, identify the interesting ones and explain their significance.

As Hans Hermes and Werner Markwald wrote in *Fundamentals of Mathematics*:

“Mathematicians are interested not only in theoretical insight and profound theorems but also in general methods for solving problems, methods whereby certain classes of problems can be handled in such a systematic way that the actual process of solution becomes, so to speak, automatic. Every newly discovered method represents an advance in mathematics, although the problems that are solvable by this method thereby become trivial and cease to form an interesting part of creative mathematics.”

FROM THE LETTER

Famous problems have often served as landmarks and lighthouses against which one can measure an improved understanding of this landscape. Solving one of these problems has been a certain sign of new insights and interesting methods, which would then be studied by a community of mathematicians, through a long and arduous process of talks, discussions, simplifications. At the end of this process, one will ideally find a

textbook presentation of the results suitable for any graduate or even undergraduate student to study. Some of the mathematical ideas pursue their journey even further to become, decades or centuries after, tools that are understood and used by the whole population.

Any problem, including those we think are ‘landmarks and lighthouses’, can only be efficiently solved by AI if no new methods are required (they could also be *inefficiently* solved by brute force search). Such problems may not have much to teach us, despite our expectations to the contrary. Our judgment of which problems will be fruitful and which will be boring cannot be known before we solve them and is itself liable to error. Thanks to AI, our ability to identify the non-lighthouses and non-landmarks in math has dramatically accelerated.

FROM THE LETTER

The mathematical community functions, in many ways, as a miniature version of humanity. It consists of individuals using a wide variety of different approaches, joined by core values. The most precious resources of our profession are students and ideas, and these we nurture with great care. We feel responsible to let them grow to their full potential, until they can live a life of their own in the mathematical world. For students we often suggest problems with the core intention of developing skills making them well-positioned for advances in research and elsewhere. Our ideas we disseminate in talks, private discussions and careful writeups, connecting them to the previous ideas of others.

Knowledge and people are indeed the most precious resources in all of reality. The rest of the universe consists of raw materials that people can transform once they acquire the knowledge of how to do so. For instance, we have learned how to transform silicon and other earthly substances into subservient automatons that explore the space of all known mathematical methods and axioms and their logical implications.

It is wonderful that some mathematicians let their students work on whatever they want, although this is not yet a universal norm. The ideal culture in mathematics (and,

indeed, every other academic discipline) would grant students the freedom to work on whatever they want, and professors would simply make suggestions and offer problems that the students may find interesting.

FROM THE LETTER

These processes invariably take time and are based on human interaction.

Everything takes time. The invention of pencil and paper accelerated progress in mathematics, as did the calculator and the computer. None of these inventions spelled disaster for mathematicians or their objects of study, even though things ranging from arithmetic operations to calculating certain integrals have all been automated, their execution made faster. It is still up to the creative mind of the mathematician to wield these tools in the pursuit of discovering interesting mathematical theories, definitions, axioms, frameworks, conjectures, and even new conceptions of mathematics itself.

In addition to *physical* inventions, mathematicians have been inventing automating algorithms (*software* inventions) since at least the time of the Babylonians. The creation of formulae for solving quadratic equations, adding numbers together, and countless other mathematical tasks has not prevented mathematicians from finding interesting problems—on the contrary, they are able to solve *more* and *deeper* problems with these shortcuts in hand.

The situation is the same with AI, except that now, for the first time, technology has exposed the gap between *proving* a mathematical theorem—which, as we have written, can be automated—and *conjecturing* a new and fruitful mathematical idea, including explaining the significance of a machine-generated proof—which can only be done by a creative mind.

The letter implies that the mathematics community may soon consist of many isolated individuals using AI. But AI can no more replace the fruitful collision of ideas between individuals than it can replace the creation of new methods, frameworks, and explanations. For example, an isolated mathematician may want feedback for his new

idea. An AI could only offer criticisms that could be deduced from the body of already known mathematical arguments. An actual mathematician colleague, on the other hand, could offer genuinely novel criticisms that could not be deduced from humanity's body of knowledge.

FROM THE LETTER

In recent months, the success of AI in solving major mathematical problems has made headlines even outside mathematical circles. But solving problems is only a tool and proxy for achieving the primary goal of conceptual understanding and insight. Forgetting this in the world of AI may turn the tool against the primary goal.

“Forgetting this” would mean that mathematicians choose to never try to conjecture new methods, explain why some axioms are more interesting than others, or create new theories. It would also mean that their use of AI would be forever constrained by their knowledge of axioms, methods, and theories at the time that the mathematicians throw up their hands. Within these confines, AI could still help them make interesting discoveries for a while, but eventually AI would only have boring work to do (such as discovering pointless theorems). At that point, mathematicians would face a choice: end the human project of mathematics as we know it, or else choose to wonder about unexplained mathematical regularities, new candidate axioms and methods, and the rest of the questions that AI could not answer. Given their interest in deepening their understanding of mathematics, it seems implausible that mathematicians would not pursue the latter.

Already, it is not the case that mathematicians are sitting idly by while AI solves certain problems. The solution to the Navier–Stokes problem may have been discovered by AI, but mathematicians are already poring over the 166-page solution so that they may come to *understand* it, so that they may glean *original insights* as to how and why the proof works.

The recently solved Navier–Stokes problem may have been solvable by AI, but only mathematicians could *choose* to direct AI towards it and *explain* why this was worth

solving in the first place—and that explanation is itself hard-won and valuable knowledge.

Consider a future in which all problems that can be solved by known methods are solved by AI. Rewarding mathematicians who blindly insist on solving such problems by hand may seem pointless, antiquated, or even embarrassing (doing so *may* lead to new ideas if done creatively, but to dogmatically reject technological assistance is guaranteed to inhibit progress). Instead, mathematicians may come to appreciate their colleagues' work in developing new mathematical theories and structures—regardless of how much AI explores these things *after* the humans have created them! By that point, it may seem obvious to the community that they should be awarding the creation of new mathematical ideas beyond what current machines can do, and not what has become grunt work in the face of technology.

As Paul recently wrote:

“AI can solve the Navier–Stokes existence and smoothness problem. But it is nowhere near capable of discovering ideas like this:

$$i^2 = j^2 = k^2 = ijk = -1.$$

Proving conjectures with known methods is a small part of what mathematicians do.

Creative vision is everything.”

The above equation describes the so-called quaternion numbers and is of great significance in mathematics. To date, AI has not discovered a single idea of such magnitude.

Indeed, the mass production at faster and faster pace of “true/false” statements could destroy fertile ground instead of breathing life into new ideas.

No, it is plowing ground that’s already been made fertile by a mathematician. This then frees up mathematicians to create yet more fertile ground in unpredictable regions of land.

AI’s extensive solution to the Navier–Stokes problem has *provided* mathematicians material from which they may deepen their understanding of mathematics. This material may have taken mathematicians untold hours of manpower to generate themselves. Thanks to AI, they can skip the ‘gruntwork’ of *generating* the proof and work immediately on the fun, creative, and edifying work of *understanding* it.

Consider chess: has the rise of machines that have computed a host of solutions to chess games been detrimental or beneficial to the chess playing community? The skills of chess players and the quality of games have been *rising* in response to computers’ newfound solutions. Chess players study the solutions, develop novel explanations for why they work, and earn greater understanding of how to play chess more effectively.

FROM THE LETTER

Often these solutions are announced in a rush, leaving no time for a proper writeup, the isolation of new methods and ideas, and citing relevant previous work of others. As in all creative professions, this raises severe attribution and plagiarism questions.

Citing relevant previous work is a reasonable concern, but it’s no reason to ban AI. Rather, academic culture will have to adjust—as it did with the advent of the Internet. Perhaps expectations will evolve such that citations are not expected to coincide with publication but rather are expected to be added in due time. Alternatively, maybe AI itself will take over more and more of the citation process.

When is a solution ‘rushed out’, and when is it hidden from the community for ‘too long’? Every creative project is idiosyncratic, so there can be no universal optimized

timeline across all work in mathematics.

Was AI's solution to the Navier–Stokes problem 'rushed out'? If mathematicians discover fatal errors in it, then perhaps so. On the other hand, such errors were discovered in published works well *before* the age of AI! Mistakes are inevitable, and so the crucial feature of academic work is not to ensure that errors are never made but to have institutions in place that have a chance at detecting and eliminating them.

If a mathematician publishes a sequence of papers with nothing but worthless AI content, then his reputation in the community will suffer, which will damage his career prospects, reduce his access to journals and conferences, and otherwise make it more difficult for him to take advantage of the opportunities that the institutions of the academy have to offer. So mathematicians have an incentive to include something original in their papers, something of interest to their peers.

Consider the Internet, which has resulted in an explosion of content of all stripes. This exponential increase in people's choices has not rendered the typical Internet user a mindless consumer of every random webpage he stumbles across. On the contrary, Internet users *curate* (and rely on third parties to curate for them) so that their browsing habits serve their interests.

The mathematics community already does this, choosing which manuscripts are worth their time and which colleagues are worth following. AI may change the details of how the community curates, but it will not be a change in kind.

The mathematicians could create two databases: one that contains all AI-generated proofs, and one containing the creative explanations of mathematicians that expound on those proofs and other insights. Mathematicians could pull from the former database as they develop their original ideas, much as chemists look up the specific heat of the substance they are investigating.

FROM THE LETTER

Moreover, without the willing mathematicians who must take care of their development and integration into the mathematical canon, AI-conceived ideas would never become fully alive and the crucial human transmission chain between mathematicians would be lost.

Just because AI is producing theorems faster than we can absorb them, that does not take anything away from the mathematician's ability to make progress. On the contrary, it means that he only has *more* theorems at his disposal, should he ever want to search for a particular theorem (or one with specified attributes) in the database. That theorems may lie in our databases, inert and ignored, is neither a travesty nor a roadblock to progress.

Once again, consider chess. We now have a massive database of endgame positions. Most such positions are not currently 'alive' inside any human. Still, any chess player is free to draw upon any of these inert positions, study them, and improve their chess skills accordingly. Far from rendering chess a 'finished' game, this database empowers chess players to consider *more* chess ideas than they otherwise would have.

FROM THE LETTER

We are witnessing a general threat to intellectual work, with misalignment between the outcome of the use of AI and its initial purpose.

In the age of AI, one of the tasks of the mathematician is to solve problems for which known methods turn out to be inadequate. Most generally, a mathematician's job is to do things that have not yet been reduced to algorithms, i.e., to do what computers cannot. So long as we continue to improve our technology, the set of mathematical tasks that machines can do will expand. But such an expansion empowers mathematicians in turn to pursue problems that had previously lain outside their reach. In other words, the more *machines* can do, the more that *mathematicians plus machines* can do. The set of all non-algorithmizable mathematical discoveries is boundless, and, therefore, the work of the mathematician is without end.

The misalignment is not between the capabilities of AI and the task of the mathematician, nor is it between the outcomes of AI and progress in mathematics. Rather, the misalignment is between one of the purported aspirations of mathematics—solving problems that can be solved by known methods or even brute force—and its actual purpose—deepening our understanding of the world of abstractions.

FROM THE LETTER

In many fields and activities, years of training have traditionally served not only to produce a final answer or product, but also to develop understanding and the ability to formulate new questions and ideas. However, building on a vast body of previous human work, AI systems are becoming increasingly capable of producing the results of such work directly, and these goals cease to align. The issues the mathematical community faces now are similar to issues that other scientific and creative professions are facing, and indicate issues that all of humanity might face: how to make sure that, as AI changes the way work is done, we do not lose sight of what that work was meant to achieve in the first place.

Professors often give students problems that have already been solved as part of the students' training. This does not have to change in the era of AI. If anything, professors will have *more* pedagogical content to work with.

AI does not prevent goal alignment, unless your goal is to earn credit for automatable work. If your goal is to discover and earn credit for original and fruitful mathematics, then AI can only assist you in that endeavor.

AI can show us some of the implications of our knowledge, but only mathematicians can come up with said knowledge (in the form of fruitful methods, explanations, axioms, or any other original mathematical idea that lives beyond the previous boundary of our knowledge) in the first place.

FROM THE LETTER

AI offers the potential of enhancing and accelerating genuine mathematical study and understanding. Mathematics as a profession will need to adapt to these changes in several ways. However, whether these changes ultimately benefit the field or have a destructive effect will in large part be determined by the decisions of the humans in control of this new technology.

But there is no centralized control of AI at the moment. AI companies have their own incentives, and right now they are very good at solving problems with known methods. In a free market, you benefit from offering a good or service that the culture wants. If your good or service is rejected by the culture, then you have to adjust to meet their wants or risk going out of business in the face of competitors. AI companies make their money through voluntary subscriptions, so people have to want to use their products for these companies to stay in business. That can only happen on a regular basis if the product is useful to the customers—including those mathematicians (and enthusiasts of mathematics) who *want* to accelerate their progress in mathematical research. So much the worse for mathematicians who wish it were otherwise.

Furthermore, the letter's framing casts mathematicians as having no agency, which is false. Their culture is not a prison, and their standards and practices can evolve.

FROM THE LETTER

These issues must be addressed urgently, in the mathematical community, by the companies developing these technologies and, more broadly, by a society that will confront similar problems in many other forms of intellectual work.

As we have explained, AI is not an issue in mathematics. Rather, the issue is in what the community rewards.

The AI companies do not have a responsibility to cater to the demands of those mathematicians who do not want their colleagues to use AI. AI companies have a right to sell their products, and mathematicians have a right to employ said products

towards the generation of mathematical theorems—or to stick to older, inferior technologies.

The signatories of the letter want something to be done. Ironically, something *is* being done—namely, problems are being solved by AI, and culture is acclimating. So something is happening with AI, it just doesn't align with their vision for the field.

AI is currently proving *some* theorems with known methods (and able to try them in far more combinations than a human can). As of this writing, it has solved only one out of the seven Millennium Prize problems. Nor has AI solved all (or any) of the problems in other famous catalogues, such as Hilbert's problems (about thirteen remain unsolved) or Erdős's problems (about six hundred remain unsolved). It is fair to suppose that, as AI improves, it may solve some of these famous open problems if their solution does not require new theories or methods. But, as we have explained, even if AI ends up solving all of them, there will always be more creative work to do for the mathematician.

To date, AI has not achieved anything conceptually new across some of the most important activities in mathematics. AI has not:

- Discovered a new mathematical theory (such as category theory or non-Euclidean geometry),
- Discovered a new methodological framework (such as the probabilistic method or generating functions),
- Identified a profound open problem (such as P vs. NP),
- Found a new interpretation of an existing theory (such as viewing algebraic equations through symmetry groups), or
- Made a discovery in metamathematics (such as Gödel's incompleteness theorems).

The optimistic implications of AI on the other hand, are manifold: we will be able to criticize our ideas more rapidly, explore the consequences of our axioms more thoroughly, and put our methods to use more widely. But perhaps most fundamentally, we will finally learn to respect and reward the very thing that mathematics was about

all along—not mechanically deriving solutions from known methods, but creating new theories, methods, and interpretations, making discoveries in metamathematics, and identifying new and interesting problems in the abstract but very real world of mathematics.