AI and the Evolving Shape of Mathematics

Proof, interpretation, and the work that comes before and after

torsorlab · Danny Krashen × OpenAI Codex

Last revised: November 4, 2026

CREDITS

Produced using the torsorlab infrastructure by
Danny Krashen, in collaboration with an
OpenAI Codex agent running GPT-6.1 Sol.

torsor.org

Contents

Mathematics begins before the proof 1

What proof contributes 3

AI enters the whole cycle 4

A result still has to be placed 5

Interpretation must become uptake 7

The transition has two clocks 8

AI is changing mathematics. This is true, vague, and therefore almost useless. A calculator changes mathematics. A notation changes mathematics. A funding crisis changes mathematics too, mostly by ensuring that less of it happens. The real question is which parts of mathematical life are moving, which parts are becoming more valuable, and which parts have been hiding in plain sight because the finished theorem was easier to point at.

Four questions usually arrive together. How is AI changing the way mathematics is done? Where is its role in research heading? How will teaching and learning change? What should early-career researchers do about any of this?

They are not four separate questions. They are four views of the same shift.

Mathematics begins before the proof

A useful starting perspective is deliberately broad:

> Mathematics is the development of conceptual frameworks which are internally consistent, communicable, and independently verifiable.

This is not a border-control definition. Mathematics has survived several thousand years without agreeing on what mathematics is, and it would be uncharacteristic to settle the matter now. The definition is a declaration of what this account values.

The important word is *development*.

Finished mathematics has a polished stillness. There is a statement. There is a proof. The definitions have stopped wriggling. Notation appears to have been chosen by a calm intelligence rather than assem-

bled during a week in which three plausible conventions fought to the death. This is how finished work is supposed to look. It is also misleading if treated as a picture of how the work was done.

Before the proof came curiosity, examples, failed definitions, diagnostic computations, half-remembered analogies, literature searches, objections, and the occasional realization that the original question had been phrased in the wrong language. The statement itself had to be found. So did the world in which the statement became worth making.

A conceptual framework includes propositions, but it also includes the objects being discussed, the distinctions that matter, the examples that keep everyone honest, the notation that gathers recurring structure, and the questions that make the framework move. A theorem secures a relation inside this framework. It does not constitute the whole framework any more than a bridge constitutes the terrain it crosses.

The three requirements then do different work.

Internal consistency asks whether the framework hangs together. This includes formal consistency where formal consistency is available, but it also includes the daily mechanics of mathematics: hypotheses actually hold, conventions do not drift halfway through an argument, and a correction propagates to whatever depended on the corrected claim.

Communicability asks whether another mind can enter. A private intuition can begin mathematics. It cannot finish the job. A framework must eventually be teachable, criticizable, revisable, and available for extension. Exposition is not gift wrapping applied after discovery. It is one of the places where the discovery is tested.

Independent verification asks whether the claims can be checked

without trusting their source. Proof is central here. So are exact computation, formalization, fresh mathematical reading, and other forms of scrutiny. These forms are not interchangeable, and their independence is never absolute, but they make the framework answerable to more than confidence and reputation.

Mathematics, on this account, is not merely a warehouse of true statements. Warehouses are useful. Nobody asks a warehouse to understand its inventory.

What proof contributes

Proof contributes something extraordinary: if the assumptions hold, the conclusion follows. The implication becomes public. Another participant can inspect the route, locate a gap, repair a local error, or decide that the whole thing has driven into a lake.

This is indispensable.

It is not interpretation.

A proof does not decide whether the statement captured the intended phenomenon. It does not explain why the theorem matters, why this formulation is better than a nearby one, or whether the argument exposes the governing idea. It does not teach the surrounding theory. It does not even guarantee that a formalization says what the informal statement was meant to say.

Proof is a public receipt for an implication. The receipt matters. It does not explain why anyone went shopping.

After proof come interpretation, communication, learning, use, and new questions. Sometimes these activities begin before the proof and help produce it. Often they force a return to the statement. A result

can be correct and still reveal, through the misery of trying to explain it, that the framework has chosen the wrong basic objects.

This produces a striking asymmetry. Mathematics has public standards for what happens after a proof is found: the argument can be checked. It has no comparably standard, self-declared account of how proofs are found, how the right statements are selected, or how competing proof routes are triangulated.

The claim is not that mathematicians have no methods. They have an unruly abundance of them. Compute examples. Change coordinates. Seek an invariant. Strengthen the hypothesis. Weaken the conclusion. Draw the picture everyone pretends not to need. Ask what the theorem would say in the smallest case. These practices are taught locally, often by imitation and apprenticeship.

What mathematics lacks is something analogous in cultural role to the scientific method: not one literal sequence, but a declared vocabulary in which exploratory methods can be compared, criticized, and improved. Finished mathematics is recognizable. Mathematical discovery mostly recognizes itself in retrospect, after it has put on a jacket.

AI enters the whole cycle

AI is commonly introduced into this picture as a theorem prover. That is an understandable entrance. Proofs are visible, outcomes can sometimes be checked, and benchmark designers also have families and would like to go home.

But AI is already entering the whole research cycle. It searches literature, generates examples, compares definitions, suggests statements, proposes proof routes, writes code, formalizes claims, explains argu-

ments, and remembers more of a project than fits in one person's working attention. It can do several of these badly in one afternoon, which is still an impressive kind of scale.

This does not make the research cycle automatic. It makes many moves within the cycle cheaper.

Cheap production changes the economics of attention. Candidate arguments, computations, explanations, and connections can arrive faster than a person or community can understand them. Some will be good. Some will be subtly wrong. Some will be correct answers to questions that nobody meant to ask. Volume does not discriminate among these cases. It enlarges the discrimination problem.

The analogy with gene sequencing is imperfect but useful. Sequencing technology transformed the amount and availability of genetic evidence. It did not eliminate genetics. Sequences did not choose the biological questions, connect themselves to phenotype, or decide which variation was explanatory. The technology changed what evidence could be produced and at what scale. It therefore changed what interpretation had to do.

Proof automation casts a similar mechanizable shadow. It makes certain parts of mathematical implication available at a new scale. The danger is not mechanization. The danger is looking at the shadow and deciding that the shadow was the organism.

A result still has to be placed

A result can be correct without yet having a place in mathematical understanding. It may formalize the wrong informal statement, answer a nearby question, depend on a mechanism nobody has recognized, or remain unusable because its surrounding framework is unavailable.

Abundant production multiplies this work. AI can help with it too, but generating a candidate and placing it within a conceptual framework are different achievements. Placement requires orientation, judgment, continuity, and interpretation. Which question is alive? Which example separates two conceptual pictures? Is a failed proof a local mishap or evidence that the framework is wrong? Which explanation opens the subject, and which merely replaces unfamiliar notation with familiar fog?

These questions are not outside mathematics. They organize the mathematics that gets made.

The distance between a human mathematical statement and its Lean formalization makes the point sharply. The informal statement carries ambient conventions, intended scope, motivating examples, and a sense—sometimes an unreliable sense—of what phenomenon is being isolated. The formal proposition records something precise in a particular environment. A formal proof can certify that proposition. It cannot, by itself, certify fidelity between the two.

Nor is the human statement transparent merely because a human wrote it. A statement can compress an entire mathematical culture. The Langlands conjectures are not difficult only because their proofs are missing. An entire course can be devoted to understanding what a conjecture says, why the objects on opposite sides belong in the same sentence, and why anyone expects a correspondence at all.

Suppose a proof appeared tomorrow. The proof would be an event of enormous mathematical importance. It would not instantly teach the statement to people who could not understand it the day before. Mathematical implication does not arrive with interpretation included in the box.

An answer to a question nobody knows how to ask is not necessar-

ily worthless. It may provoke the later invention of the question. But until a framework exists in which the answer can be understood, criticized, and used, it remains socially inert. Correct, perhaps. Transformative, not yet.

Interpretation must become uptake

Interpretation is not the final step either. A result can be interpreted by a small group and remain unavailable to the wider mathematical culture. It can be available in principle and still be nearly impossible to learn. It can be learnable and still fail to reorganize anyone's thought.

There are as many layers here as the situation requires. Four to six usually provide a workable map—the rule of five—not because knowledge has five floors, but because a map with one layer hides the terrain and a map with 900 layers becomes the terrain, now with worse typography.

A useful progression is: result, interpretation, conceptual framework, explanation, learning path, cultural uptake. Each item can be opened again. Each can also be compressed by better understanding.

Cardano's cubic equation makes the compression visible. Reading the original account requires entry into a rhetorical and notational world that modern algebra has reorganized. Modern notation does not merely abbreviate the old prose. It gathers structure so that relationships can be seen and manipulated together. Centuries of understanding changed the cost of communication.

The same may happen to theories that currently require long initiation. A course may be needed to state Langlands properly using the concepts available now. That does not prove the course-length interface is permanent. Deeper concepts may eventually make the

same world easier to enter without making it shallower.

This is why learnability becomes a frontier. AI can help construct adaptive routes through a subject, generate examples calibrated to a learner's current picture, compare explanations, and maintain several levels of exposition at once. The aim is not faster answer delivery. The aim is a framework that can enter human thought and change what questions become possible there.

The transition has two clocks

Technical capability moves quickly. Search, formalization, proof, synthesis, and explanation improve on a timescale measured in model releases and startling weekends.

Cultural adaptation moves through curricula, careers, funding, institutions, prestige, and identity. These are slower systems. They contain committees. Some contain committees that report to other committees, a defensive adaptation whose evolutionary purpose remains disputed.

The mismatch creates opportunity. A larger space of questions becomes reachable. Small groups can sustain investigations that once exceeded their memory, time, or range of technique. Mathematical methods that remained tacit may become explicit because agents need to be told what an example is meant to test and why one direction deserves another hour.

The mismatch also creates risk. If institutions believe theorem production was the mathematician's whole vocation, then automated theorem production looks like automated mathematicians. Funding can contract before the broader work of conceptual development becomes legible. Career exits can remove mentorship, tacit knowl-

edge, and the very communities needed to interpret the new supply of mathematics.

Early-career researchers therefore face an unpleasantly concrete version of the philosophical problem. The useful preparation is not a retreat to tasks machines cannot yet do. That territory will keep moving, and careers built on negative space tend to discover that the space has opinions.

The durable work is broader: learning how frameworks are built, how questions are formed, how examples diagnose, how proof strategies express different ideas, how formalization preserves or loses intent, how exposition changes understanding, and how knowledge can be organized so another mind can enter. Technical fluency matters. So do judgment, communication, pedagogy, and the ability to steer a mixed human–AI inquiry toward actual curiosity.

AI is changing how mathematics is done by entering the whole developmental cycle. Its research role is heading beyond proof production toward sustained participation in exploration, memory, formalization, criticism, and explanation. On the working assumption developed here, teaching should move away from answer delivery toward adaptive paths into conceptual frameworks. This account locates the mathematician's continuing value in the intellectual development of culture: enlarging what can be asked, understood, checked, communicated, and carried forward.

The theorem still matters.

Now the rest of mathematics is becoming harder to ignore.