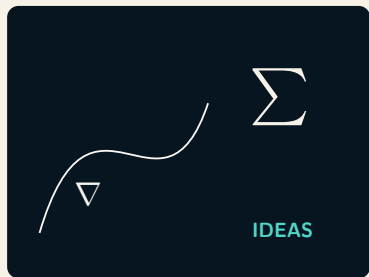


Digital Tools and AI for Contemporary Mathematical Productivity

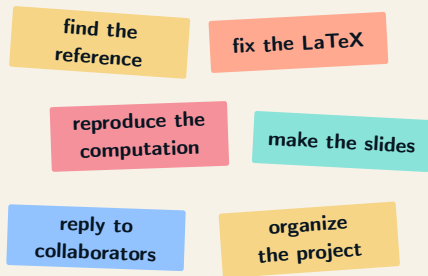
From reducing friction to changing the research frontier

Tomás S. R. Silva IMECC — University of Campinas (UNICAMP) August 17th, 2026

The blackboard is only the visible part



THE PICTURE OF MATHEMATICS



WHERE THE WEEK ACTUALLY GOES

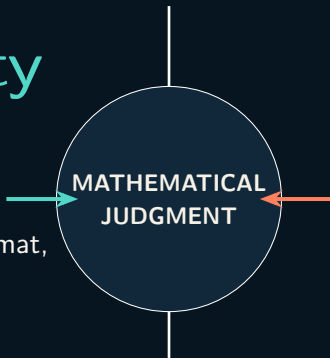
The first promise of AI was to remove **friction**.

The new question is whether it can also change the **mathematics**.

There are now two AI stories in mathematics

Productivity story

Search, summarize, code, format, organize, communicate.

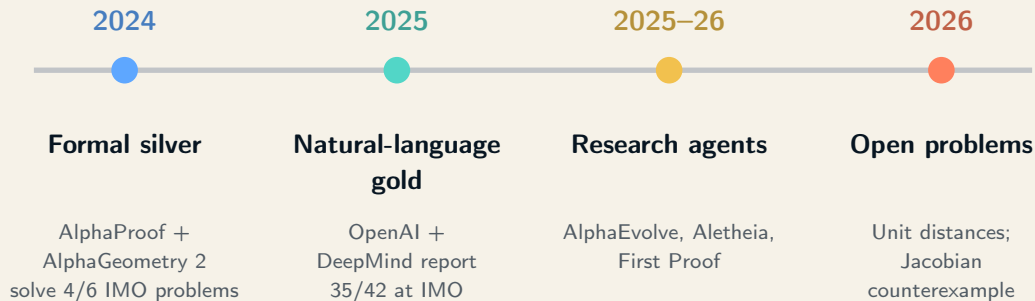


Discovery story

Conjecture, search, compute, prove, verify, generalize.

A serious workflow needs both stories—without confusing them.

Why rethink the talk now? The frontier moved



The relevant question is no longer “*Can a language model do mathematics?*”
It is “**Which mathematical workflows make it capable—and trustworthy?**”

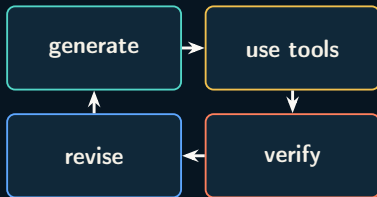
Sources: Google DeepMind, 25 Jul 2024; Google DeepMind, 21 Jul 2025; OpenAI, 20 Feb 2026; OpenAI, 20 May 2026; Gao, arXiv:2608.00222.

The question changed from capability to conditions

THE OLD MENTAL MODEL



THE EMERGING SYSTEM



The model matters. The **pipeline** matters just as much.

AI crossed from exercises to notable open problems

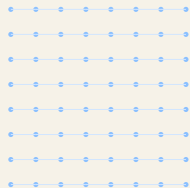


A threshold—not the end of the story

A few verified examples establish **possibility**. They do not yet establish reliability, breadth, taste, or autonomous theory-building.

Context: Epoch AI, FrontierMath: Open Problems overview, Jul 2026.

Case study: AI crossed a disciplinary silo



square-grid intuition



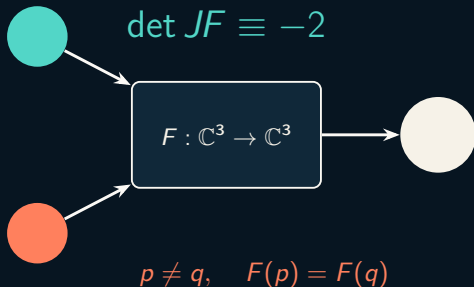
number-field construction

OPENAI — PLANAR UNIT DISTANCES

- An internal general-purpose model produced an **infinite family** with at least $n^{1+\epsilon}$ unit distances.
- This refuted the prevailing $n^{1+o(1)}$ grid-optimality conjecture.
- The key move connected elementary geometry to **algebraic number theory**.
- External mathematicians checked and digested the proof.

Sources: OpenAI, 20 May 2026; companion remarks by Gowers, Litt, Sawin et al., May 2026.

Case study: the Jacobian conjecture fails for $n = 3$



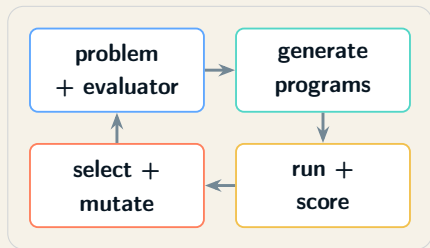
ALPÖGE + CLAUDE FABLE 5 — JULY 2026

- An explicit polynomial map has constant nonzero Jacobian but is not injective.
- The conjecture is therefore false for $n \geq 3$.
- The plane case $n = 2$ remains open.
- Exact-arithmetic checks and independent expositions followed quickly.

Sources: Alpöge's announcement and attribution are documented in *A Counterexample to the Jacobian Conjecture*, 20 Jul 2026; independent account and extensions: Gao, arXiv:2608.00222. The original $n = 2$ case remains open.

Search plus a verifier is becoming a discovery engine

ALPHAEVOLVE



- explored **67 problems** across analysis, combinatorics, geometry, and number theory;
- recovered best-known constructions in most cases;
- improved several known bounds and constructions.

The sweet spot is a large search space with a **cheap, objective test**.

Source: Georgiev, Gómez-Serrano, Tao, Wagner, Mathematical exploration and discovery at scale, arXiv:2511.02864 (2025).

Benchmarks are evidence, not research

WHAT BENCHMARKS CAN SHOW

closed answer

new problem

long reasoning

specialized proof

WHAT RESEARCH ADDITIONALLY REQUIRES

choose a meaningful question

recognize novelty

survive expert scrutiny

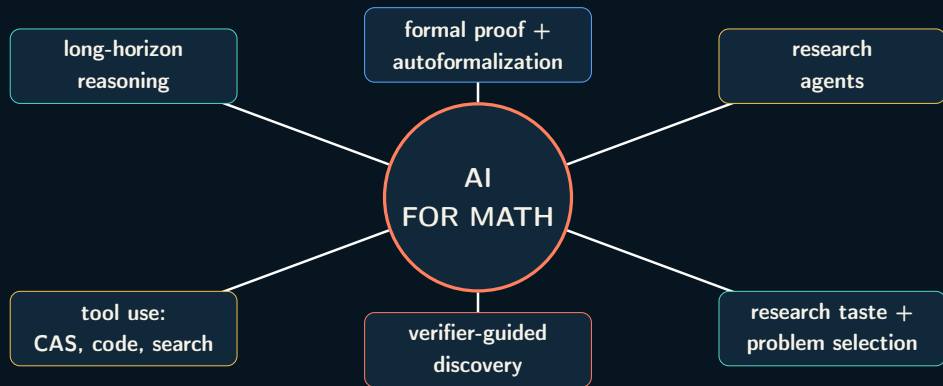
open a useful direction

The First Proof warning

OpenAI reported at least five promising proof attempts; one initially believed likely correct was later judged incorrect. Expert checking remained decisive.

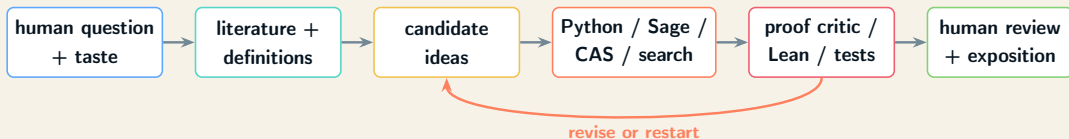
Sources: OpenAI, First Proof submissions, 20 Feb 2026; Epoch AI, FrontierMath: Open Problems overview, Jul 2026.

What is “hot” in AI for mathematics in 2026?



Synthesis based on Artificial Intelligence for Mathematical Reasoning: An Integrated Survey (2026); Google DeepMind, Aletheia (2026); and Lean AI evaluation.

A mathematical agent is a system, not a chat box



Capability comes from the loop

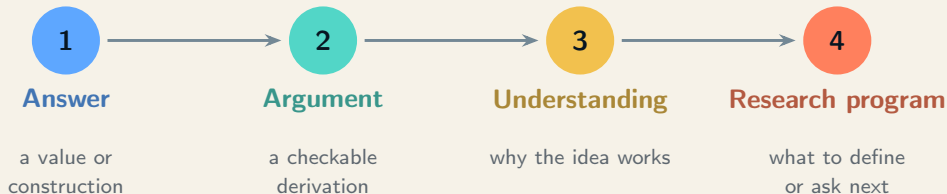
Iteration lets the system test ideas, expose errors, and allocate more effort to promising branches.

Correctness still has an owner

Self-critique is useful evidence. It is not an independent certificate.

Agent pattern illustrated by Aletheia, Google DeepMind, Feb 2026 and AlphaEvolve, 2025.

Four different claims hide behind “the AI solved it”

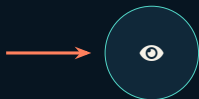


Recent systems have crossed levels 1–2 in striking cases.
Levels 3–4 remain the deeper test of mathematical agency.

The durable insight: attention is the bottleneck



many plausible directions



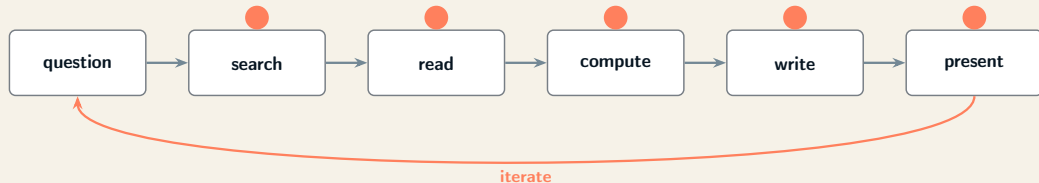
focused attention

AI can expand the number of
ideas tried,
connections noticed, and
objects checked.

That can matter even before it
becomes a fully autonomous
mathematician.

Related reflections: companion remarks on the unit-distance result, especially Litt and Gowers, May 2026.

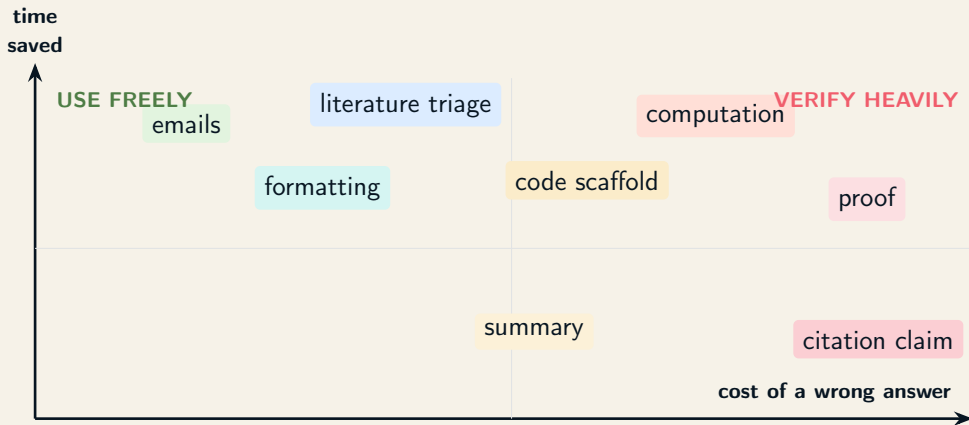
Everyday workflows remain the most reliable gain



Let AI absorb the **mechanical layer**.

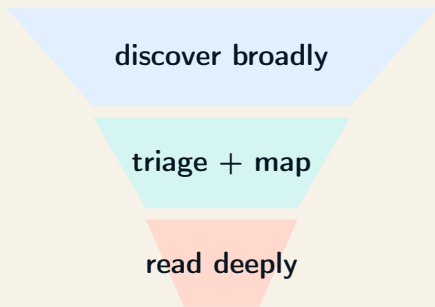
Keep mathematical intent, standards, and final judgment visible at every node.

Choose tasks by leverage and cost of error



The model may be the same. The required protocol should not be.

Reading: use AI as a funnel, not a substitute



A ROBUST LITERATURE LOOP

1. Search real indexes: arXiv, MathSciNet, zbMATH.
2. Store metadata and PDFs in Zotero.
3. Ask for a map: definitions, dependencies, points of disagreement.
4. Open every cited source; annotate the load-bearing passages yourself.

Rule

Never cite a paper you have not opened. Never inherit an AI summary as your understanding.

Writing: accelerate revision, not authorship



Excellent uses

Abstract variants, referee responses, email, translation, shortening, structure, accessibility.

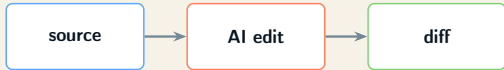
Inspect closely

Changed quantifiers, softened hypotheses, inflated novelty, altered notation, invented citations.

The text you sign should be text you can defend.

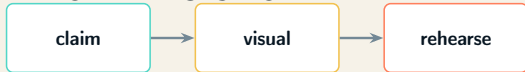
Delegate form; retain semantic control

LATEX



Generate TikZ, macros, tables, and minimal fixes—but review every notation change.

TALKS AND LECTURES



Use AI to test narrative, create variants, and expose density—not to choose what is mathematically essential.

A visual can make a correct idea memorable.
A wrong visual can make an error **persuasive**.

Build a durable stack before adding an AI layer

one AI workspace connected to the material

LaTeX · Sage/Python/GAP/Macaulay2 · Lean/mathlib

Zotero · notes · project conventions · reproducible environments

FILES + BACKUPS + VERSION CONTROL

AI amplifies the quality of the system beneath it—including its disorder.

Prompting is specification

CONTEXT

TASK

MATERIAL

FORMAT

CONSTRAINTS

TESTS

“Here is the theorem, notation, failed attempt, and exact claim.
Find a counterexample first. If none appears, audit each implication.
Return a proof sketch, unresolved gaps, and executable checks.”

Good prompting does not guarantee truth. It makes failure easier to diagnose.

Ask for adversarial work, not agreement

Weak

“Is this proof correct?”

↓ invites completion, deference, and repair

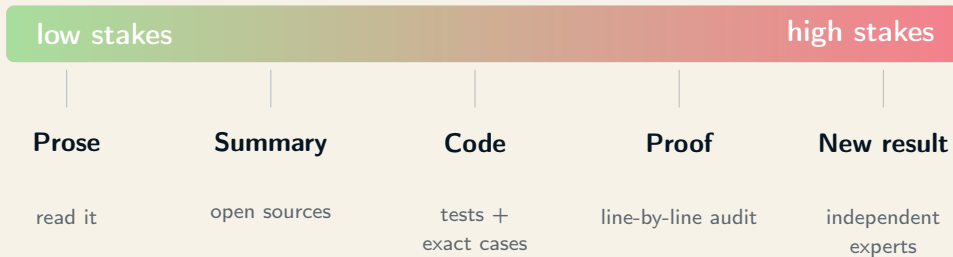
Stronger

“Try to falsify the claim. Check edge cases. List every hidden assumption. For each step, state the exact lemma used. Separate verified steps from conjectural ones.”

↓ creates inspectable failure modes

Do not ask the model to be confident.
Ask it to make **uncertainty legible**.

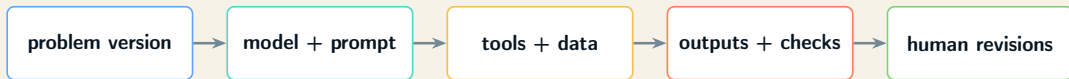
Verification should scale with mathematical stakes



For a claimed theorem

Preserve the transcript; rerun computations; search for prior art; isolate lemmas; seek independent review; formalize critical steps when practical.

Provenance is part of the mathematics



Record

Dates, model/version, prompts, files supplied, code, seeds, outputs, failed attempts, verification steps.

Privacy changes which tools are appropriate



Local or approved environment

Unpublished manuscripts · referee reports · student data · confidential collaboration material

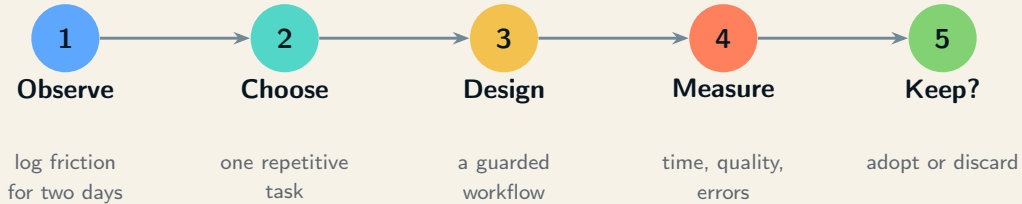


Cloud assistant

Public papers · generic examples · redacted text · material covered by an explicit data policy

Convenience is not consent—from you, your coauthors, or your students.

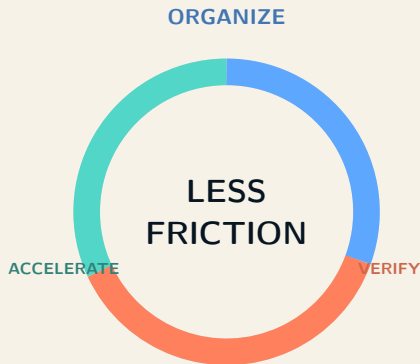
A practical experiment for the next seven days



Do not adopt a tool because it is impressive.
Adopt it because it improves a **measured workflow**.

A minimum viable system for a mathematician

1. **A reference manager** connected to real indexes.
2. **Version control + backups** for text, code, and figures.
3. **A reproducible computation environment** with tests.
4. **One AI workspace** used with explicit constraints.
5. **A verification layer** matched to the stakes.



What should remain unmistakably human?



AI may generate a proof. A mathematical community decides what it establishes, why it matters, and where to go next.

The synthesis



Exoskeleton

Reduce friction around mathematical thinking.

Discovery engine

Search more ideas and connect distant mathematics.

Both require a **verification culture**.

The goal is not to automate judgment—it is to spend more of it where it matters.

Questions worth discussing

PRACTICE

Which part of your week should be delegated first?

EPISTEMOLOGY

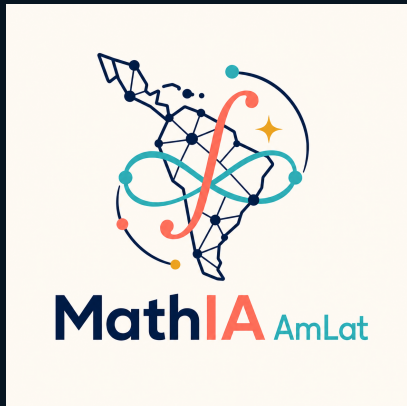
When does a verified proof become understanding?

COMMUNITY

How should credit, training, and standards change?

Join the MathIA AmLat community

MATHEMATICS AND AI IN LATIN AMERICA



MathIA AmLat brings together researchers, students, and educators interested in the uses, limits, and implications of artificial intelligence in mathematical research, education, and collaboration.

Interested in joining?

Send me an email:



tomas@ime.unicamp.br

Thank You!