

Report

Preprint counts and AI policies in the AI-driven research wave

A monthly report on preprint numbers, AI milestones, and policies

Klas Markström

October 6, 2026



UMEÅ UNIVERSITY

Abstract

This report provides a regularly updated view of the number of preprints on arXiv and other preprint servers, as well as publication policies related to the use of AI tools. The aim is to make visible the impact that general-purpose AI tools, primarily LLMs, are having on research output, preprints and publication.

Suggested citation. Klas Markström, *Preprint counts and AI policies in the AI-driven research wave: A monthly report on preprint numbers, AI milestones, and policies*, technical report, Umeå University, version dated 6 October 2026, http://abel.math.umu.se/~klasm/Uppsatser/LLM_monthly.pdf. Please include the version date when citing figures or results.

Contents

1	Purpose and data	2
1.1	Reactions	3
1.2	Consequences and challenges	4
1.3	The LLMs	5
2	Arxiv preprints	5
2.1	The arXiv main categories	5
2.2	Selected sub-categories	8
3	Preprint counts in other archives	10
3.1	bioRxiv	10
3.2	ChemRxiv	10
3.3	medRxiv	10
3.4	NBER	11
3.5	SocArXiv	11
4	Publisher's policies on use of AI	13
5	Some of the influential breakthroughs	14
5.1	LLMs as programming tools in research	15
5.2	Capability demonstrations	15
5.3	Research breakthroughs	15
5.4	Mathematical formalization and verification	17

1 Purpose and data

General purpose research tools in the form of LLMs and variations on AI agents have in 2026 become strong enough to materially change the way much research is being done. AI tools have been good programming tools for some time and special-purpose machine learning tools like AlphaFold [42] have several years ago transformed narrow research areas. However, now general-purpose LLM-based tools have reached research level reasoning, as seen in a string of highly visible mathematical breakthroughs during the spring and summer of 2026. The combination of mathematical ability and automated programming also make it possible to perform tasks in research fields where data analysis, statistics, or simulation are valuable. Web enabled tools can also perform information search and retrieval, often with a speed which is hard for a human to match.

These capabilities, both already present and still improving, have received attention via the highly publicized mathematical breakthroughs made by the research groups at major AI companies, but the wider effects of publicly available tools have not been visible in the same way. However, following the number of preprints uploaded to the arXiv preprint server makes the impact clearly visible.

The aim of this report is to provide a regularly updated collection of plots of monthly or weekly preprint counts from arXiv and other preprint servers, together with some of the influential research breakthroughs achieved by use of AI. Rather than including all forms of AI and machine learning when listing concrete research results, we focus on those achieved using general-purpose LLMs.

Here we should also point out that currently there is a large gap in the capabilities of free and paid versions of many AI tools. Hence much, but not all, of the research activity visible here comes from those with access to such paid tools.

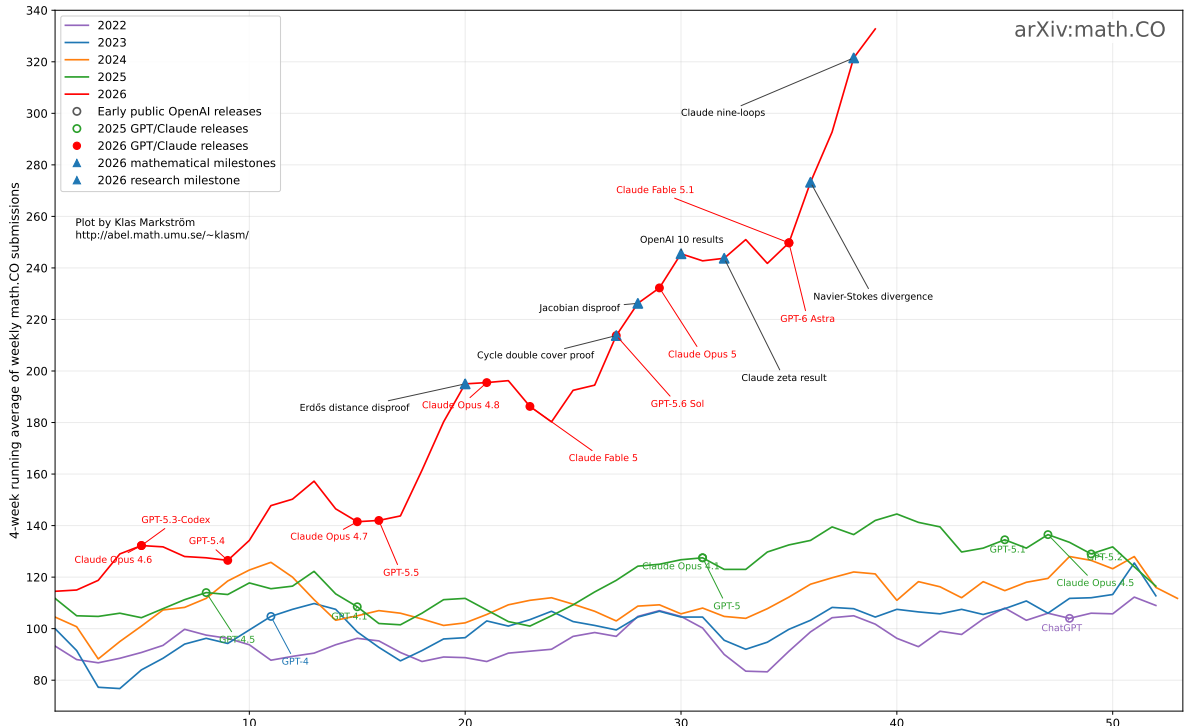


Figure 1: Weekly arXiv submissions in math.CO; the curve is a four-week moving average.

Figure 1 will serve as our main example, demonstrating both a broad trend, and the way

our later plots are built. Here, the x -axis corresponds to the weeks during a calendar year, starting in January. There are separate curves for each year, showing a running four-week average of the number of preprints in the combinatorics subsection of the mathematics section at arXiv. Superimposed on these curves we have added the official release dates for GPT and Claude, two of the major publicly available LLMs. These two have been chosen as representatives for the high-end LLMs since they have so far had the strongest mathematical ability. Additionally, we show, as blue triangles, the approximate announcement dates of influential research breakthroughs, made either autonomously by AI agents, or with strong assistance by such.

Here it is visible that while the years 2022 to 2024 have seen variability, they are roughly similar, with an upward fluctuation in the second half of 2025. The curve for 2026 on the other hand, looks nothing like the earlier ones. The first ten weeks of 2026 could initially be seen as a fluctuation, but the later part of the year is rapidly outpacing the corresponding month in earlier years. At the start of September, the weekly number of preprints was roughly twice that of the same week in 2025.

Pure preprint counts are of course only indirect evidence of rapid AI-driven research. In Figure 2 we show the fraction of the paper in the `math.CO` which acknowledge use of AI tools in their work. This was estimated by sampling 40 papers each week. This type of disclosure is entirely voluntary

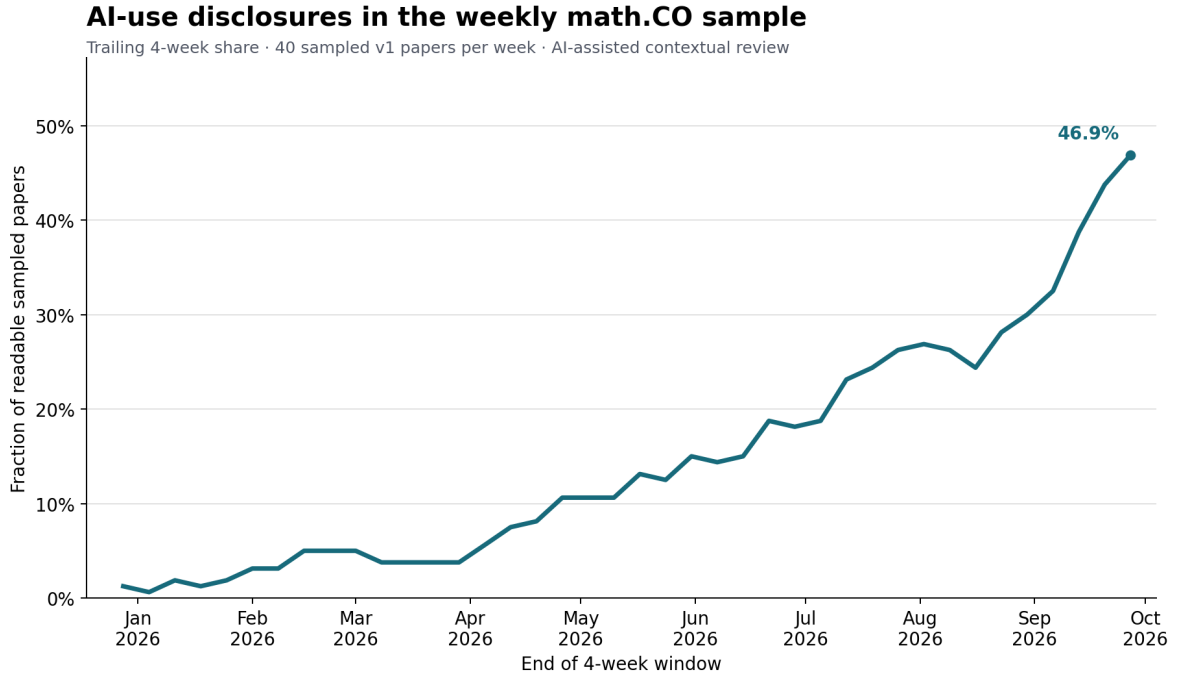


Figure 2: Fraction of the papers in `math.CO` which acknowledge AI use.; the curve is a four-week moving average.

1.1 Reactions

So far the most visible reactions to the impact of LLMs on research has come from the mathematics community. Already last year workshop on the potential impact of AI in mathematics was held, leading to the Leiden Declaration [7]. In anticipation of coming improvements in this technology, the authors listed several high-level risks, emphasised the importance of disclosure, and encouraged a wider discussion around AI driven research.

Following a string of new mathematical results, by different AI companies, during the summer of 2026, the OpenAI theorem on the Navier-Stokes existence problem made the debate more publicly visible. A group of Fields medalists wrote "A Severe Misalignment of AI in Mathematics" [47] highlighting how company publication practices could have a negative impact on mathematics and the mathematics community; despite providing formally correct solutions to long-standing problems. In response, OpenAI formed an independent advisory group of high-level mathematicians [54]. One of the initial tasks for this group was to advise on how to release a large number of mathematical results developed by OpenAI's research side. On the 29th of September 2026 they published their guidelines on "Responsible Release of AI-Generated Mathematics".

At the end of September 2026 the Simons Institute hosted a working group on how the research community in Theoretical Computer Science(TCS) can adapt to AI. They released the report "AI and TCS: The Next Six Months" [2], with recommendations regarding e.g AI use disclosure, practices for conferences, and job interviews.

Fields medalist Terrence Tao has on his blog [74] hosted a long list of well-known experts in different subfields of mathematics, each one writing a blog-post essay on their views on AI and its impact on both Mathematics and research more widely.

The rapid development has also forced major publishers and conferences to update their codes of conduct or create specific policies on AI use and disclosure. A clear example is the arXiv's new limit [21] of two papers per calendar month for each author. In Section 4 we briefly describe several examples of such policies.

1.2 Consequences and challenges

The underlying technology for these AI tools is currently improving rapidly. The early research breakthroughs in 2026 have been dominated by Anthropic and OpenAI, but as e.g. benchmark suits [29] there are many other companies whose tools are at level similar to what the top two had 6 to 12 months ago.

Already at the current level some challenges are visible.

- The balance between paper production and human refereeing capacity has been lost. A sudden increase in the number of written papers does not correspond to an equal increase in the number of reviewers. When some of the material is written by AI it becomes even more important to check that the results are presented in a human readable form. This is already a challenge for journals and conferences.
- Human understanding lags behind the production of new results. This is visible in several ways. In some cases, completely AI written papers convey little understanding, despite being formally correct. Additionally, the sheer number of papers, even when well-written, makes it hard for humans to keep up to date.
- Research education faces both the task of teaching students about a collection of rapidly developing tools and the issues around assessment and AI assisted writing as undergraduate teaching does. The recent "Summit on PhD Math Education in the Age of AI" [40] produced a report discussing this. Currently, mathematics also faces the issue of choosing good problems for a new student to work on. Here,, one must now balance learning for the student, with and without AI assistance, and the impact of a research field changing rapidly.

- Grant applications can now be written with AI assistance, or entirely by AI. This will lead to more applications for funding agencies to evaluate, much like for journals. It will also become harder to tell whether a written application corresponds to the skill and understanding of the applicant, or instead reflects the strength of their available AI tools.

1.3 The LLMs

The general-purpose tools which are behind most results in this report are built around Large Language Models. At the time of writing, the models behind the most visible breakthroughs are Claude, by Anthropic, and GPT, by OpenAI. Other models such as Gemini, by Google DeepMind, and DeepSeek, by its eponymous owner, have produced impressive results on benchmarks and have many users. There are also special-purpose research LLMs, mainly aimed at mathematics, which at this moment have a narrower user base, but this may change.

Statement on AI use

Writing has been done by the author. The GPT app was used to run the data collection scripts and manage the database of preprint records and plots. The list of research milestones was imported from a subset of a local database to LaTeX using GPT.

2 Arxiv preprints

In this section we show the weekly preprint counts for each of the main categories at arXiv. Physics is not an official main category but we have created a joint count for all the main categories in the original physics set up, in order to compare with the large mathematics and computer science categories.

As we will see, there are different visible patterns depending on research field, where e.g. some are strongly influenced by conference deadlines. There are also visible annual effects due to holidays.

2.1 The arXiv main categories

In Figure 3 we show the total counts for the three volume-wise largest categories on Arxiv: physics, mathematics, and computer science. All three show clearly increasing trends, but the pattern for computer science is different. The peaks in the computer science graphs roughly correspond to deadlines associated with groups of major conferences, either initial paper submission or later process deadlines for accepted contributions.

In Figure 4 the corresponding graphs for Economics, Electrical Engineering and Systems Science, Quantitative Biology, Quantitative Finance, and Statistics are shown. Here we see increasing trends for Economics, Quantitative Finance, and Statistics. The latter superimposes this on a periodic pattern. Here Electrical Engineering and Systems Science, and Quantitative Biology stand out in having less visible upward movement the last two years. We speculate that the natural dependence of these fields on data collection and hardware provides a natural inertia, when a smaller fraction of the research process can make use of AI tools.

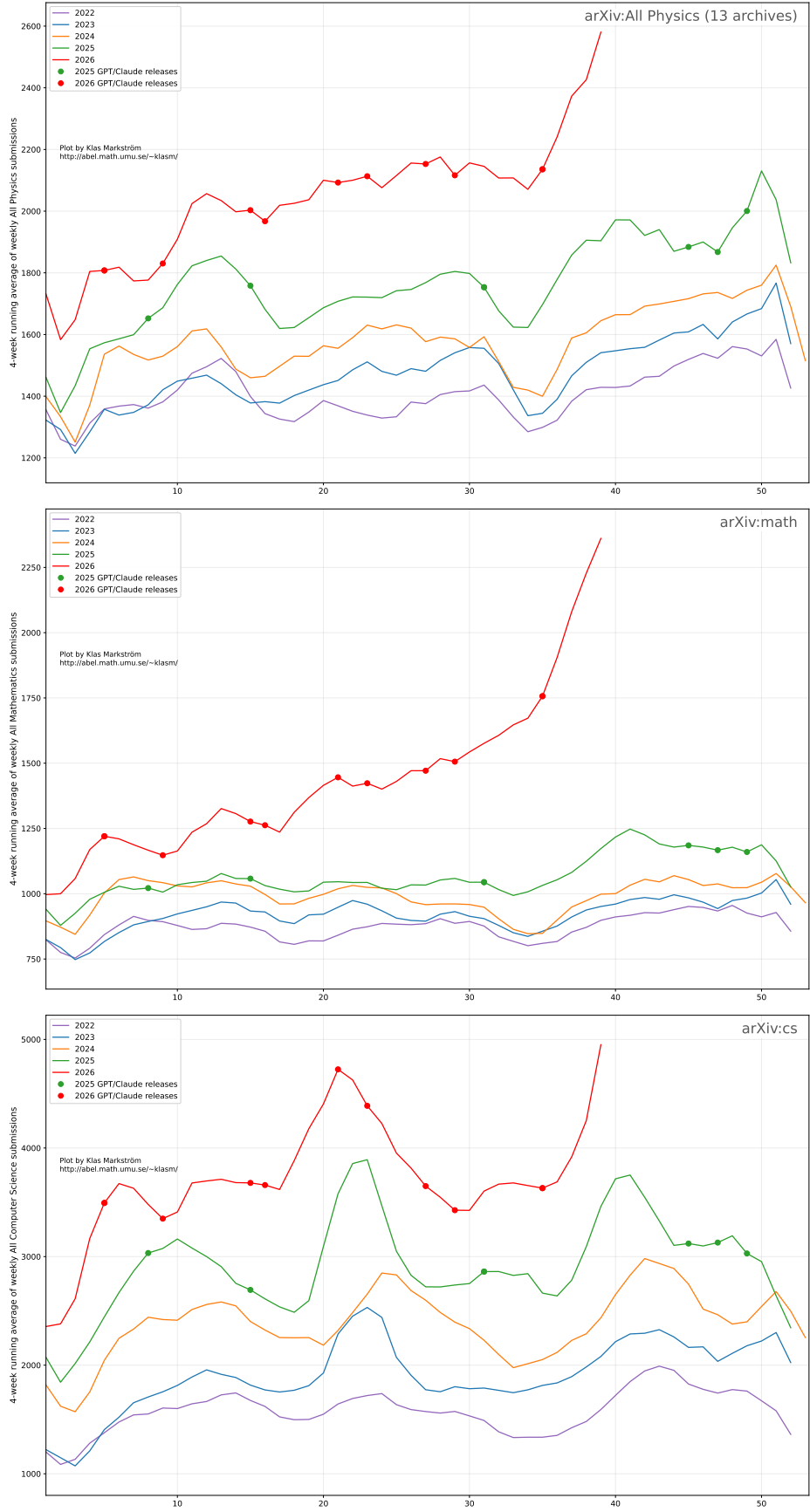


Figure 3: Weekly arXiv preprint counts for All Physics (top), Mathematics (middle), and Computer Science (bottom).

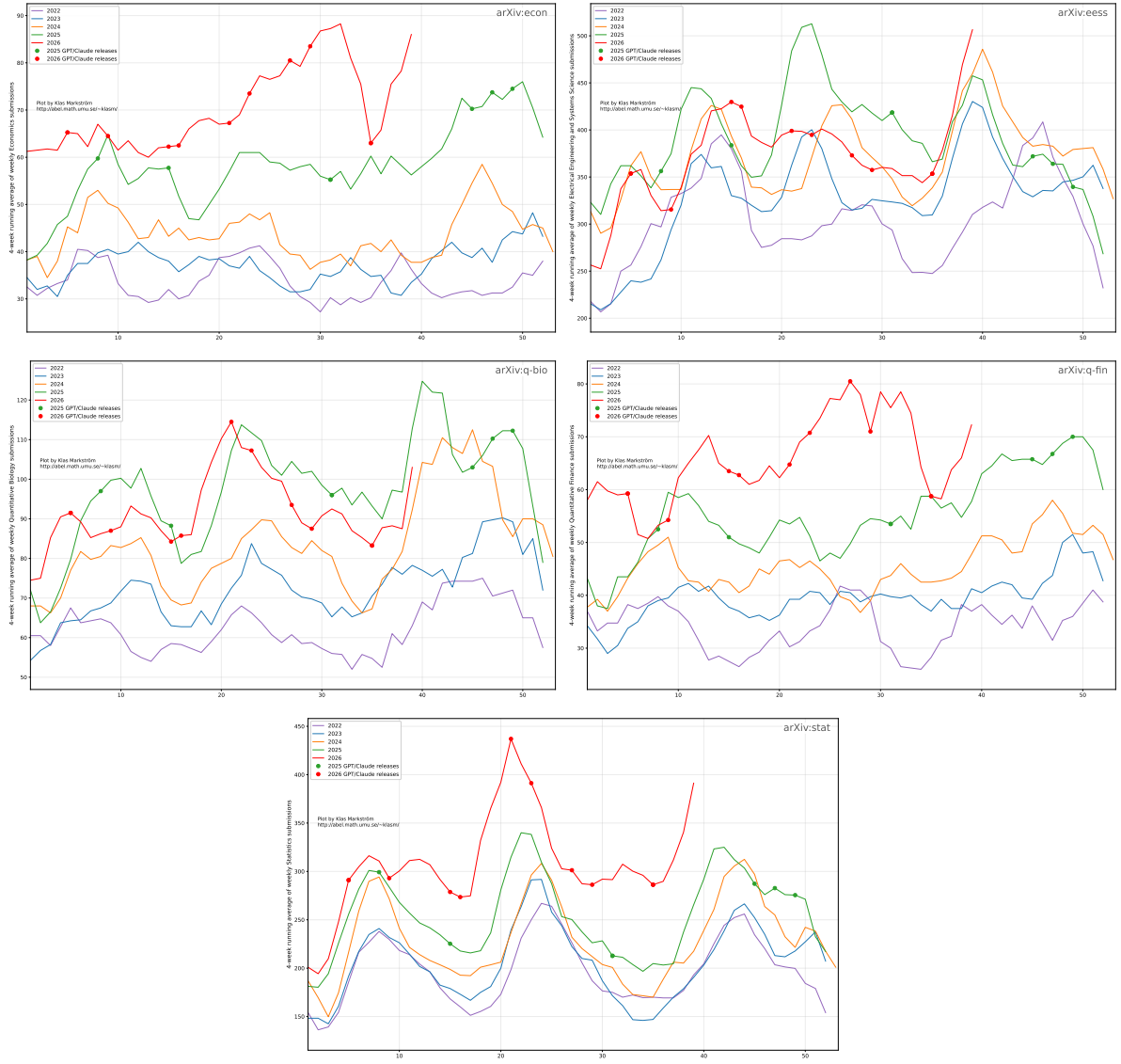


Figure 4: Weekly arXiv preprint counts for Economics, Electrical Engineering and Systems Science, Quantitative Biology, Quantitative Finance, and Statistics.

2.2 Selected sub-categories

Since subfields within the major categories, especially in physics and computer science, depend on experiments and hardware to different degrees, and have different publication patterns due to conferences, here we show the graph for four different subcategories of each. As already mentioned, the plots for all subcategories are shown in Appendix A.

As one may expect, the two physics categories with a larger proportion of theoretical papers show a clearer upward trend here. But for Astrophysics there has also been a clear increase in 2025 and 2026. Some may relate to new observations becoming available, but another tentative explanation is faster analysis of new data, and re-analysis of existing data sets.

The different subcategories of mathematics have roughly the same increasing trends, with some difference with respect to when the field became visibly affected.

For computer science the annual conference peaks are visible, though for Computational Complexity the qualitative shape of the 2026 curve clearly deviates from previous years. This might be because the more mathematical nature of this subfield makes the recent improvements in LLM mathematical ability more impactful here. The Robotics category also show a clear increase in 2026. This might be due to a mix of AI tools and rapid hardware development, so some caution regarding the cause of the increase is warranted.

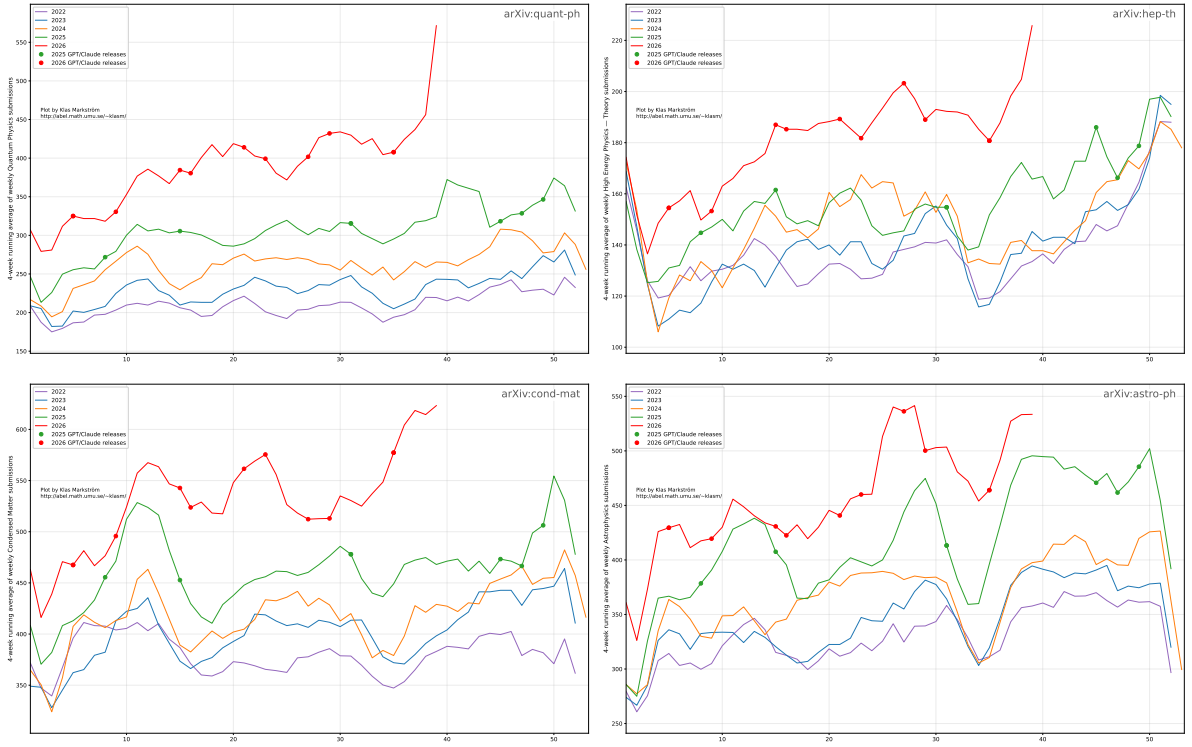


Figure 5: Weekly arXiv preprint counts for Quantum Physics, High Energy Physics–Theory, Condensed Matter, and Astrophysics.

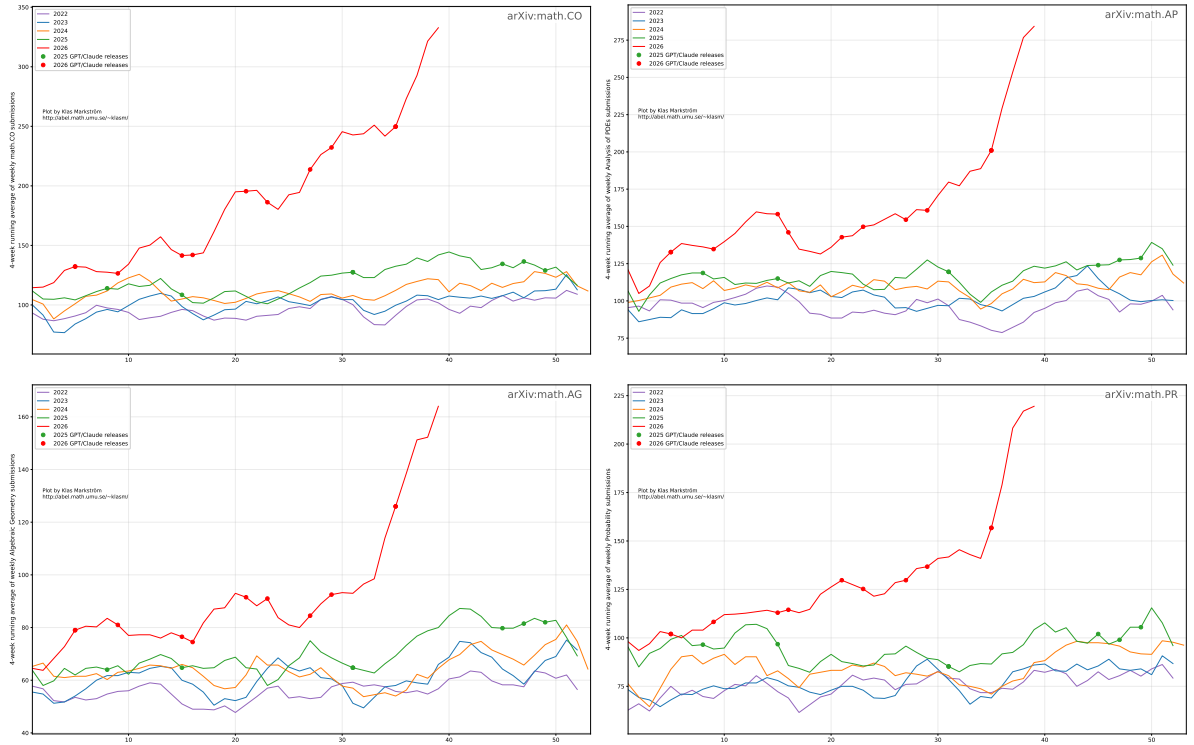


Figure 6: Weekly arXiv preprint counts for Combinatorics, Analysis of PDEs, Algebraic Geometry, and Probability.

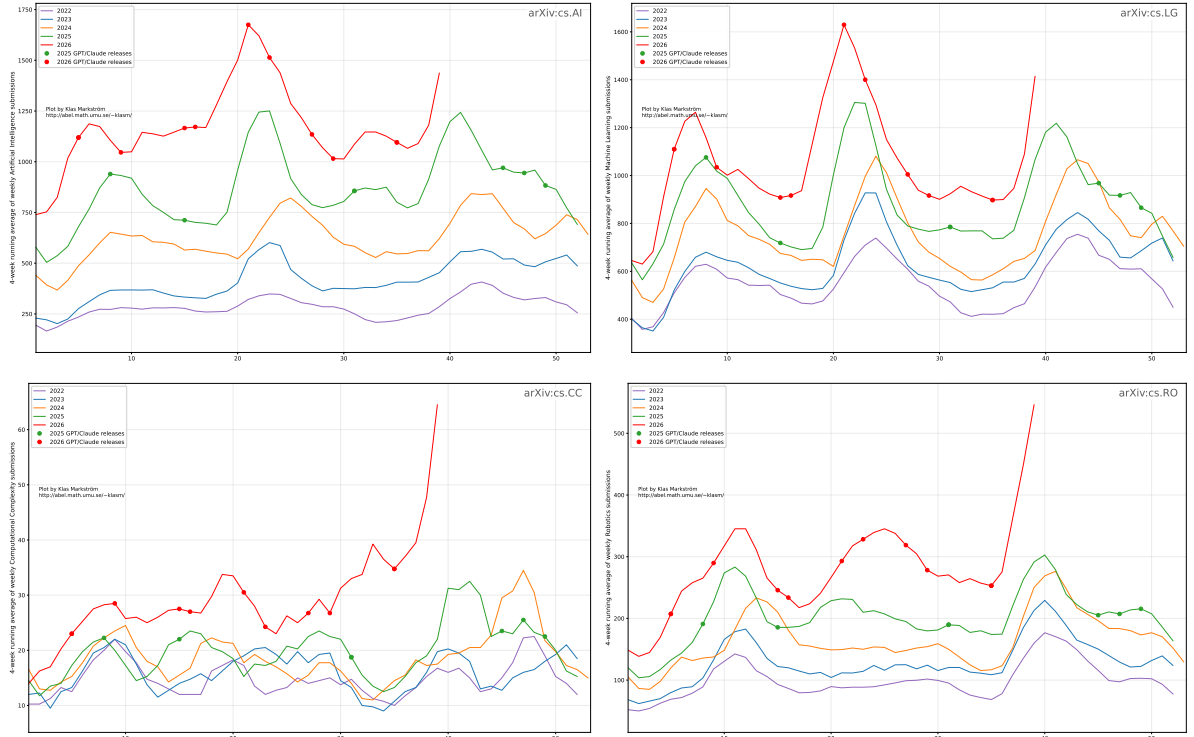


Figure 7: Weekly arXiv preprint counts for Artificial Intelligence, Machine Learning, Computational Complexity, and Robotics.

3 Preprint counts in other archives

The arXiv is our main source of data but it is far from the only preprint archive. In this section we display similar figures for several other archives, focused on research fields not covered by arXiv. Our data consists of weekly or monthly counts of new preprints, depending on the source archive. For each of these archive we first show their own official description of the area they cover.

3.1 bioRxiv

"bioRxiv is a free online archive and distribution service for unpublished preprints in the life sciences. "

3.2 ChemRxiv

"ChemRxiv is a free submission, distribution, and archive service for unpublished preprints in chemistry and related areas. "

3.3 medRxiv

"medRxiv is a free online archive and distribution server for complete but unpublished manuscripts (preprints) in the medical, clinical, and related health sciences. "

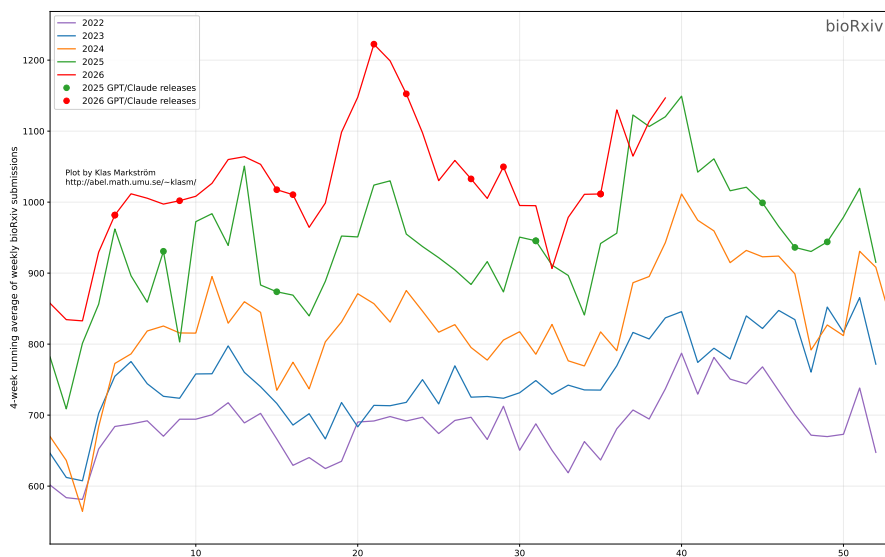


Figure 8: Weekly bioRxiv preprint counts.

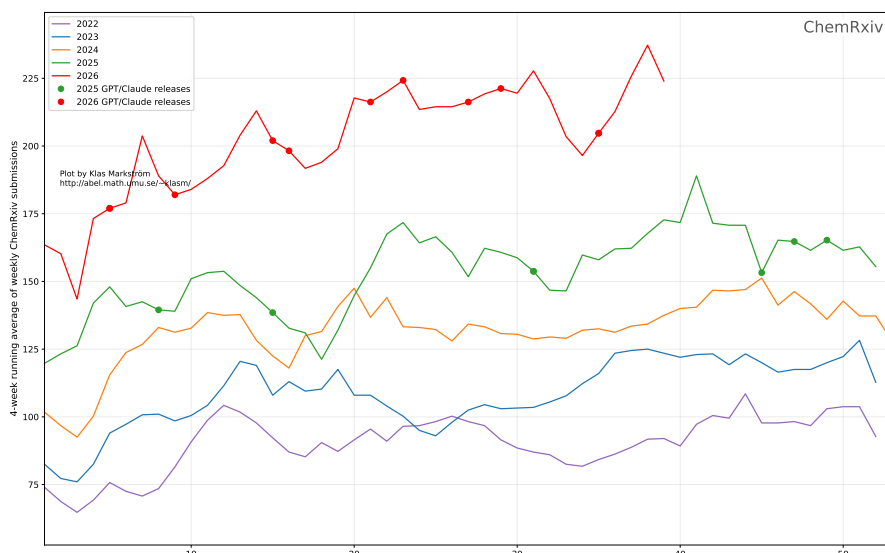


Figure 9: Weekly ChemRxiv preprint counts.

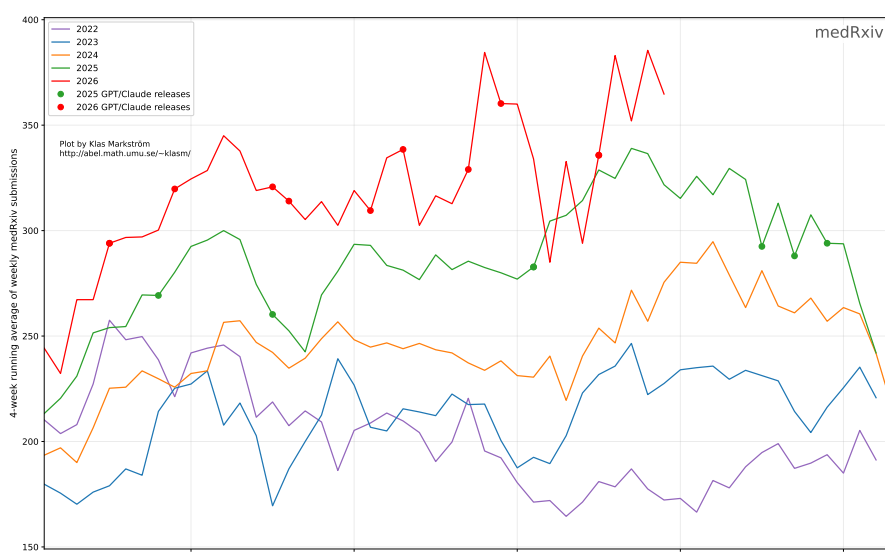


Figure 10: Weekly medRxiv preprint counts.

3.4 NBER

"The National Bureau of Economic Research (NBER) is a private, nonprofit organization that facilitates cutting-edge research on and analysis of major economic issues. It is nonpartisan and refrains from making policy recommendations, focusing instead on providing background studies and data that underlie decision-making in both the public and private sectors."

3.5 SocArXiv

"SocArXiv, open archive of the social sciences, provides a free, non-profit, open access platform for social scientists to upload working papers, preprints, and published papers, with the option to link data and code. "

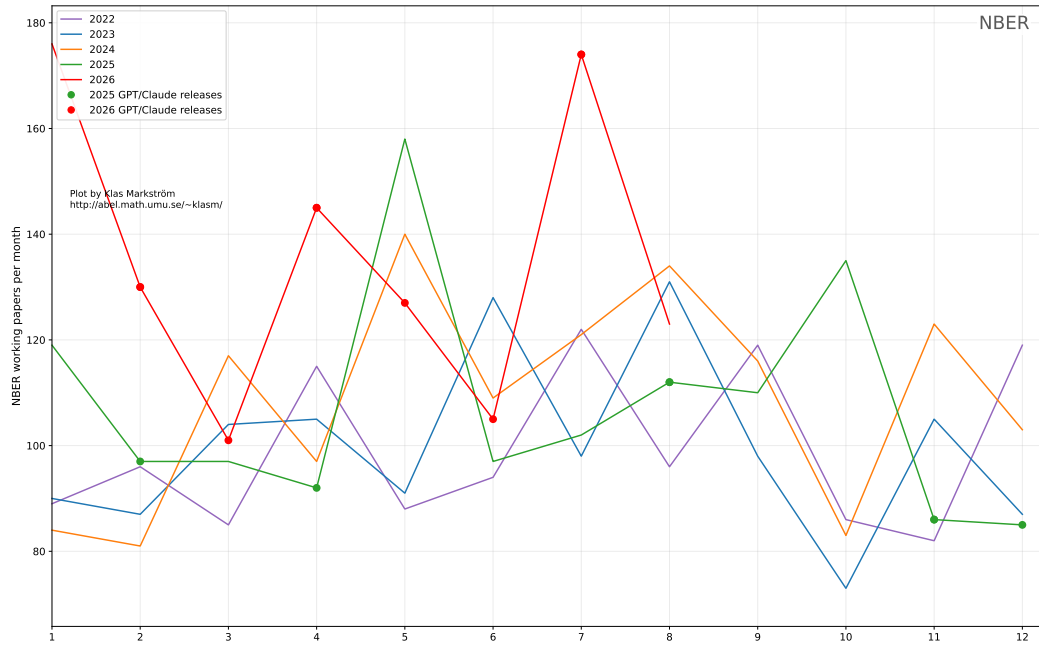


Figure 11: Monthly NBER working-paper counts.

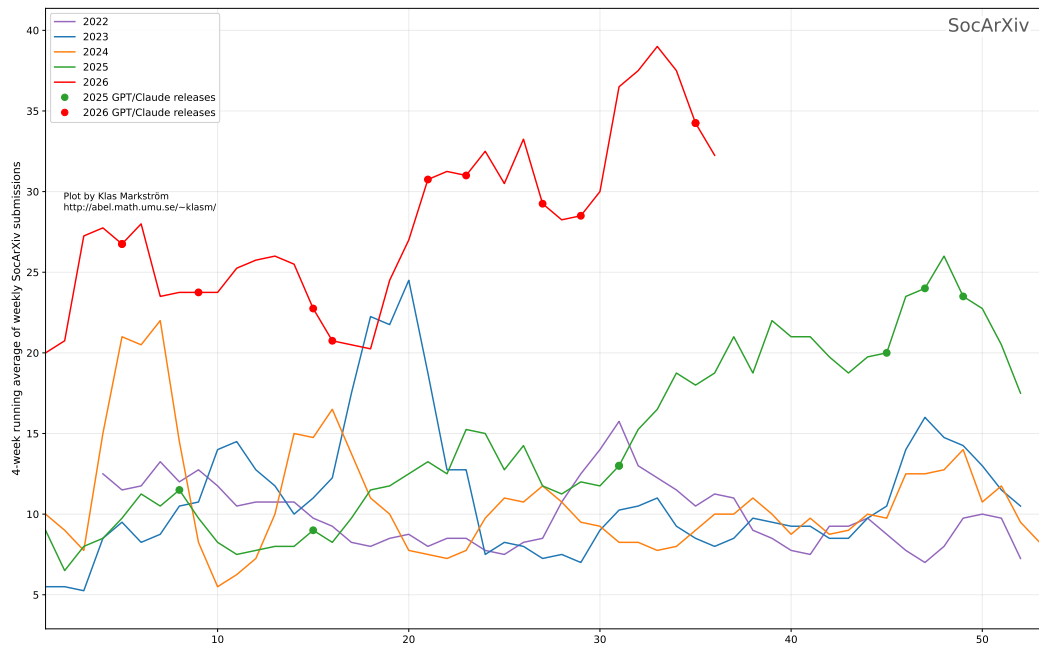


Figure 12: Weekly SocArXiv preprint counts.

4 Publisher’s policies on use of AI

The pure increase in the number of papers, and the shift of some of the creative process and understanding from the human author to the AI assistant, have pushed many publishers and conferences to respond. Some have updated codes of conduct, demanded different form of disclosure of AI use, or added steps aimed at judging the author’s understanding of a paper. Here we list some examples of such responses.

- **arXiv.** The Code of Conduct requires honesty and accountability in submissions, but does not specify a strict AI-disclosure rule [16]. However, the moderation policy [15] now states that AI use should "...be reported consistent with subject standards for methodology". In online commentary [27] one of the arXiv moderators stated that papers with obvious LLM hallucinations will lead to a one year ban for the author. On the 1st of October 2026 [21] the arXiv also implemented a cap of at most two submitted papers from each author every calender month.
- **ACM.** The Association for Computing Machinery permits AI assistance but requires disclosure when AI tools make a substantive contribution to the research, such as in coding, analysis, or experimental work. Routine use of AI for improving the writing does not in general require disclosure [17].
- **ACS Publications.** ACS permits AI tools in research and manuscript preparation under author supervision. AI-generated text or images must be disclosed, and substantial uses affecting the research should be described in the Methods section. [1].
- **AIP Publishing** Permits AI assistance in manuscript preparation without disclosure. Uses that may affect findings or conclusions must be described in the Methods section. Editors and reviewers may not upload submitted material to unauthorized AI tools [3].
- **AMS.** The American Mathematical Society journal policy specifies that AI tools cannot be authors. Authors must disclose AI use in writing, images, or data analysis and remain responsible for the work; editors and referees must not upload manuscripts under review to an LLM [9].
- **APS** Permits AI assistance in research, writing, and peer review, while requiring human responsibility. Substantive uses must be disclosed while light text editing is exempt. Reviewers may not upload confidential material to unrestricted AI tools [10].
- **BMJ.** BMJ permits AI assistance provided that authors remain responsible for the work and disclose relevant uses. AI used in carrying out the research should be described in the Methods section, and editors and reviewers must not compromise manuscript confidentiality by uploading unpublished material to inappropriate AI systems [20].
- **De Gruyter Brill.** De Gruyter Brill permits AI-assisted research and writing but requires authors to remain fully responsible for the work. Significant AI contributions must be disclosed, while routine proofreading and language editing are exempt [25].
- **Elsevier.** The Elsevier policy states that AI may assist authors under human oversight. AI use in manuscript preparation must be declared, while research uses must be described in the Methods section. [28].
- **IOP Publishing** Permits AI-assisted research and writing, but the author is responsible for the content. Material uses require an AI statement, with research uses also described in Methods. Language editing is exempt [41].

- **Oxford University Press.** OUP permits limited AI assistance without disclosure, including spelling and grammar correction, while more substantial uses such as rewriting, translation, code generation, data analysis, or image generation must be declared. Authors remain responsible for all submitted material [64].
- **PLOS.** PLOS permits responsible use of AI tools but requires transparency about substantive use in research or manuscript preparation. Authors remain responsible for the accuracy and integrity of all AI-assisted content, and AI tools cannot qualify for authorship [65].
- **Sage.** Sage distinguishes between assistive, generative, and prohibited uses of AI. Routine assistance with language, grammar, or structure does not require disclosure, while generative AI that contributes substantively to the research or manuscript must be declared. Editors and reviewers are not allowed to upload manuscripts to LLMs [67].
- **SIAM.** The Society for Industrial and Applied Mathematics state that authors must disclose AI use and be able to defend their submissions. Ordinary grammar and spelling checking is exempt. Referees and editors are not allowed to upload manuscripts under review to AI tools [69].
- **Springer Nature.** Its approach depends on how AI is used: language assistance is acceptable with author oversight, while AI cannot replace authorship or independently supply scholarly claims. Authors remain responsible for checking and disclosing AI-assisted work [72].
- **STOC.** The Symposium on Theory of Computing, one of the major conferences in theoretical computer science, have added several steps to their 2027 call for papers, explicitly in response to generative AI. The papers must now first be uploaded to arXiv, and a 20 minute video talk on the paper by the author must be uploaded when the paper is submitted to the conference. They also demand that "substantive" use of AI tools must be declared. [73]
- **Wiley.** Wiley permits the use of AI tools under human oversight. Authors must disclose relevant generative AI use, with research-related uses such as analysis or methodology described in the appropriate part of the paper. Editors and reviewers must preserve manuscript confidentiality when using AI tools [76].

5 Some of the influential breakthroughs

This section lists the research breakthroughs indicated in some of our plots, plus a few more which were left out in order to keep the figures readable. The main criterion is that a general-purpose language model or agent made a material contribution to a research result. The section is not restricted to mathematics but the first group of examples is mainly in mathematics and computer science.

We have also included a subsection 5.4 on formalization of proofs, mostly using Lean [51]. Right now this primarily concerns results in mathematics and computer science, but we see it as a precursor of tools for fact-checking and reviewing in other areas.

5.1 LLMs as programming tools in research

This is a small sample showing how LLMs have been used for programming in research before they reached more independent problem solving.

December 2023–January 2024: FunSearch. FunSearch used a large language model to generate programs inside an evolutionary search loop, producing new cap-set constructions and online bin-packing heuristics [66]. Humans supplied the problem representation and the program evaluator.

May–June 2025: AlphaEvolve. AlphaEvolve extended the same broad pattern to larger programs: Gemini models generated and revised code, while external evaluators selected successful variants. Among its reported results was a rank-48 algorithm for multiplying 4×4 complex matrices [37, 52].

5.2 Capability demonstrations

Here we list a few examples of capability demonstrations which were not new research results.

July 2025: Gemini Deep Think at the IMO. Google DeepMind announced that Gemini Deep Think had reached gold-medal-level performance at the 2025 International Mathematical Olympiad [36].

17 September 2025: Gemini Deep Think at the ICPC World Finals. Google DeepMind announced that an advanced version of Gemini 2.5 Deep Think had achieved gold-medal-level performance at the 2025 International Collegiate Programming Contest (ICPC) World Finals, held on 4 September. The system solved ten of the twelve problems within the five-hour competition window [46].

1 December 2025: DeepSeek-V3.2-Speciale on competition problems. DeepSeek reported that DeepSeek-V3.2-Speciale had achieved gold-medal-level performance on problems from the 2025 International Mathematical Olympiad, Chinese Mathematical Olympiad, International Olympiad in Informatics, and ICPC World Finals. The accompanying release made the model weights publicly available [26].

5.3 Research breakthroughs

The entries in this subsection vary in type. Some are new research results achieved with little human input, some are examples of human–AI collaboration

January 2026: Erdős Problem #728. GPT-5.2 Pro and Aristotle produced a proof resolving Erdős Problem #728. The accompanying write-up cautiously describes it as the first Erdős problem regarded as fully resolved autonomously by an AI system [70]. The proof was produced and checked in Lean, as discussed in the formalization subsection.

February 2026: Aletheia and autonomous research mathematics. The Aletheia/Gemini Deep Think work reported both autonomous and human–AI results on research-level mathematical problems [32, 39].

1 May 2026: primitive sets and von Mangoldt chains. A Markov-chain method suggested by GPT-5.4 Pro was developed by a human team into proofs of the Erdős–Sárközy–Szemerédi conjectures #1196 and #1217, together with related results [5]. This is an example of an AI-suggested idea subsequently developed and checked by human mathematicians.

20 May 2026: the planar unit-distance conjecture. OpenAI released a model-generated construction showing that, for some fixed $\delta > 0$ and infinitely many n , the maximum number $u(n)$ of unit-distance pairs satisfies $u(n) \geq n^{1+\delta}$. This contradicted the proposed upper bound $u(n) = n^{1+O(1/\log \log n)}$. Human mathematicians checked, reorganized, and extended the proof [60, 55, 6, 68].

10 July 2026: the Cycle Double Cover Conjecture. OpenAI released a proof that every finite bridgeless undirected graph has a cycle double cover. Independent expositions followed, but no journal publication has been published [53, 34, 63].

19 July 2026: the Jacobian conjecture in dimensions at least three. Levent Alpöge and Claude Fable 5 found an explicit counterexample in dimension three, which also gives counterexamples in every higher dimension; the two-dimensional case remains open. Exact-arithmetic checks, human expositions and generalizations, and an independent elementary re-derivation followed the original announcement [71, 35, 33, 31].

1 August 2026: ten claimed advances. OpenAI released a portfolio of ten claimed new results across mathematics and theoretical computer science [61].

10–13 August 2026: zeros of the Riemann zeta function. Anthropic reported that Claude had found an unconditional proof that more than two thirds of the nontrivial zeros, counted with multiplicity, are simple and lie on the critical line; the optimized constant is 0.6725. The result appears in a human-verified preprint and was followed by a human proof [14, 8, 43].

8 September 2026: forced Navier–Stokes blow-up. OpenAI released a finite-time blow-up construction for the three-dimensional Navier–Stokes equations with a smooth external force [59, 57]. The claim concerns the forced alternatives C and D in the Clay formulation, not unforced blow-up. Broad independent validation had not been established by the cutoff date.

25 September 2026: Nine-loop super Yang–Mills Anthropic announced ?? that they have succeed in performing a nine-loop calculation in super-symmetric Yang–Mills theory. This was a challenge problem posted by Matt von Hippel as an example of a problem where traditional methods would be too time-consuming to use. Claude Fable 5.1 gave two separate solutions.

5.4 Mathematical formalization and verification

Formalization has a narrower and more mathematics-specific role than general-purpose LLMs. A proof assistant such as Lean checks whether a formal proof follows from its encoded assumptions. This verifies the mathematical correctness of a result, but it does not check presentation and readability as a human reviewer would normally do. Hence a result may be verified as formally correct, but presented in a way which is more or less unreadable for a human. Human-led formalization projects such as that for Fermat’s last theorem, led by Kevin Buzzard, typically have aims beyond just checking correctness. Auto-formalized proofs produced by LLMs have so far been longer than they might need to be, and narrowly aimed at checking correctness.

2024–2025: human-led formalization for Fermat’s Last Theorem. The Imperial College project established a human-led program for formalizing the proof of Fermat’s Last Theorem, while a separate project completed the classical regular-prime case in Lean [23, 22, 19]. These projects provide important context, but have not yet reached a complete formalization of the full theorem.

January 2026: Erdős Problem #728. The GPT-5.2 Pro and Aristotle result was expressed as a Lean proof, making the kernel-checked artifact part of the original research result rather than a later verification step [70].

21 May 2026: AI-driven formal proof search. A study reported autonomous Lean resolutions of 9 of 353 selected open Erdős problems and 44 of 492 selected OEIS conjectures [75]. These figures describe the study’s selected benchmark sets, not open mathematics as a whole.

May–June 2026: formalizing the unit-distance result. The formalization proceeded in stages with different logical strength. A formal specification first encoded the main result and related challenges, but left central proofs open [38]. A 26 May Lean development proved the fixed-power result conditional on two deep number-theoretic inputs [30]; an 11 June development unconditionally proved the weaker literal negation of the proposed uniform-constant bound [50, 49]; and a 26 June submission unconditionally formalized the stronger fixed-power result [4, 44]. These stages should not be conflated.

July 2026: verification alongside new results. The Cycle Double Cover proof was released with a Lean artifact and was followed by a separate formalization [56, 18]. The Grothendieck group-scheme construction and proof were reviewed and merged as a Mathlib contribution [48]. The explicit Jacobian counterexample was also checked in Lean against a pre-existing formal statement [45].

August 2026: portfolios and theorem verification. OpenAI accompanied its ten claimed advances with Lean certificates [62]. The new zeta-zero result was likewise accompanied by a Lean formalization [12]. These artifacts provide machine-checked evidence for the encoded statements, while that encoding still requires human scrutiny.

4 September 2026: Fermat’s Last Theorem in Lean. Anthropic released a full AI-led, machine-checked formalization of the established Frey–Serre–Ribet–Wiles/Taylor–Wiles route to Fermat’s Last Theorem [13, 11, 24]. This is an autoformalization breakthrough, not the discovery of a new proof of Fermat’s Last Theorem. The released artifact was self-assessed by Anthropic, and later also checked by Kevin Buzzard.

8 September 2026: the Navier–Stokes artifact. OpenAI accompanied its forced blow-up claim with a Lean repository [58]. This supplies a formal verification layer for the construction, as encoded in Lean.

References

- [1] ACS Publications. *Artificial Intelligence (AI) Best Practices and Policies at ACS Publications*. <https://researcher-resources.acs.org/publish/aipolicy>. Updated 13 December 2024; accessed 1 October 2026. 2024.
- [2] AI + TCS Working Group. *AI and TCS: The Next Six Months*. <https://simons.berkeley.edu/news-publications-videos/ai-tcs-working-group>. Simons Institute for the Theory of Computing; released 28 September 2026. Sept. 2026.
- [3] AIP Publishing. *AI Policy*. <https://publishing.aip.org/resources/researchers/policies-and-ethics/ai-policy/>. Accessed 1 October 2026.
- [4] Boris Alexeev. *Formal Lean Proof of OpenAI’s 2026 Counterexample to the Erdős Unit-Distance Conjecture*. Lean 4 formalization. Lean-Eval accepted commit; Codex with a human in the loop. June 2026. URL: <https://github.com/plby/Erdos90/tree/2c7ace40d681d9696fbc15c54c9edca9a2ad9f07> (visited on 09/20/2026).
- [5] Boris Alexeev et al. *Primitive Sets and von Mangoldt Chains: Erdős Problem #1196 and Beyond*. 2026. DOI: 10.48550/arXiv.2605.00301. arXiv: 2605.00301 [math.NT]. URL: <https://arxiv.org/abs/2605.00301>.
- [6] Noga Alon et al. *Remarks on the Disproof of the Unit Distance Conjecture*. 2026. DOI: 10.48550/arXiv.2605.20695. arXiv: 2605.20695 [math.CO]. URL: <https://arxiv.org/abs/2605.20695>.
- [7] Jarod Alper, Michael Barany, Alain Chavarri Villarello, et al. *Leiden Declaration on Artificial Intelligence and Mathematics*. Published 2 June 2026. June 2026. DOI: 10.5281/zenodo.20302944. URL: <https://leidendeclaration.ai/> (visited on 09/22/2026).
- [8] Levent Alpöge and Ralph Furman. *More than Two Thirds of the Zeta Zeros Are Simple and on the Critical Line*. Proof discovered autonomously by Claude; verified and communicated by the listed authors. 2026. DOI: 10.48550/arXiv.2608.13637. arXiv: 2608.13637 [math.NT]. URL: <https://arxiv.org/abs/2608.13637>.
- [9] American Mathematical Society. *Use of Artificial Intelligence*. <https://www.ams.org/publications/journals/policies/UseofArtificialIntelligence>. Accessed 22 September 2026.
- [10] American Physical Society. *Appropriate Use of AI Tools*. <https://journals.aps.org/authors/appropriate-use-ai-tools>. Accessed 1 October 2026.
- [11] Anthropic. *Fermat’s Last Theorem in Lean 4*. Lean 4 formalization. Pinned research artifact, commit aa2d8b3; Lean 4.33.1 and Mathlib v4.33.0. Sept. 2026. URL: <https://github.com/anthropics/fermats-last-theorem/tree/aa2d8b34692b16c70f699536de0d8e75b9a3e9ef> (visited on 09/20/2026).
- [12] Anthropic. *Formal Verification of the Two-Thirds Zeta-Zero Result*. Lean 4 formalization. Aug. 2026. URL: <https://github.com/anthropics/formal-math/tree/main/zeta23> (visited on 09/20/2026).
- [13] Anthropic. *Formalizing Fermat’s Last Theorem*. Published 4 September 2026. Sept. 2026. URL: <https://www.anthropic.com/research/formalizing-fermats-last-theorem> (visited on 09/20/2026).
- [14] Anthropic. *Learning More about Claude’s Mathematical Capabilities*. Published 10 August 2026. Aug. 2026. URL: <https://www.anthropic.com/research/riemann-zeta> (visited on 09/20/2026).

- [15] arXiv. *Policy for Authors' Use of Generative AI Language Tools*. <https://info.arxiv.org/help/moderation/index.html#policy-for-authors-use-of-generative-ai-language-tools>. Accessed 2 October 2026. 2026.
- [16] arXiv. *arXiv Code of Conduct*. https://info.arxiv.org/help/policies/code_of_conduct.html. Accessed 22 September 2026.
- [17] Association for Computing Machinery. *ACM Policy on Authorship*. <https://www.acm.org/publications/policies/new-acm-policy-on-authorship>. Updated 14 May 2026; accessed 1 October 2026. 2026.
- [18] Vaibhav Bajpai, Johannes Carmesin, and Utku Okur. *CDC_lean: A Lean 4 Formalisation of the Cycle Double Cover Conjecture*. Zenodo research artifact. July 2026. DOI: 10.5281/zenodo.21410822. URL: <https://doi.org/10.5281/zenodo.21410822>.
- [19] Alex J. Best et al. "A Complete Formalization of Fermat's Last Theorem for Regular Primes in Lean". In: *Annals of Formalized Mathematics* 1 (2025), pp. 103–132. DOI: 10.46298/afm.14586. arXiv: 2410.01466 [cs.FL]. URL: <https://doi.org/10.46298/afm.14586>.
- [20] BMJ. *AI Use*. <https://www.bmj.com/content/ai-use>. Current policy; accessed 1 October 2026. 2026.
- [21] Kat Boboris. *Fair Moderation, Equitable Access, and AI: arXiv's Updated Rate Limit Policy*. <https://blog.arxiv.org/2026/10/01/updated-rate-limit-policy/>. Published 1 October 2026; accessed 2 October 2026. 2026.
- [22] Kevin Buzzard et al. *FLT: Ongoing Lean Formalisation of the Proof of Fermat's Last Theorem*. GitHub repository. Imperial College London; EPSRC grant EP/Y022904/1. 2024. URL: <https://github.com/ImperialCollegeLondon/FLT> (visited on 09/20/2026).
- [23] Kevin Buzzard. *The Fermat's Last Theorem Project*. Published 30 April 2024. Apr. 2024. URL: <https://leanprover-community.github.io/blog/posts/FLT-announcement/> (visited on 09/20/2026).
- [24] Shuze Chen et al. *Prove2Me: An Open Collaborative Platform for Scaling Math Formalization*. 2026. DOI: 10.48550/arXiv.2608.28433. arXiv: 2608.28433 [cs.AI]. URL: <https://arxiv.org/abs/2608.28433>.
- [25] De Gruyter Brill. *AI Guidelines for Authors*. <https://www.degruyterbrill.com/publishing/for-authors/author-policies/artificial-intelligence>. Updated September 2025; accessed 1 October 2026. 2025.
- [26] DeepSeek. *DeepSeek-V3.2 Release*. Dec. 1, 2025. URL: <https://api-docs.deepseek.com/news/news251201/> (visited on 09/22/2026).
- [27] Thomas G. Dietterich. *Bluesky Post on Unchecked LLM Output in arXiv Submissions*. <https://bsky.app/profile/tdietterich.bsky.social/post/3mlqt6k5ix22p>. Posted 13 May 2026; accessed 29 September 2026. May 2026.
- [28] Elsevier. *Generative AI Policies for Journals*. <https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals>. Updated June 2026; accessed 22 September 2026. 2026.
- [29] Epoch AI. *FrontierMath Tiers 1–4*. [https://epoch.ai/frontiermath/tiers-1-4?tab=release-date&tier=Tier+4+\(v2\)&view=graph](https://epoch.ai/frontiermath/tiers-1-4?tab=release-date&tier=Tier+4+(v2)&view=graph). Tier 4 (v2), release-date graph; accessed 2 October 2026.

- [30] Erdős Unit-Distance Formalization Contributors. *Lean 4 Formalization of the Disproof of Erdős’s Planar Unit-Distance Conjecture*. Lean 4 formalization. Conditional on the Golod–Shafarevich inequality and Shafarevich relation-rank bound. May 2026. URL: <https://github.com/logical-intelligence/erdos-unit-distance> (visited on 09/20/2026).
- [31] Arno van den Essen. *An Elementary Way to Find a Counterexample to the Jacobian Conjecture*. 2026. DOI: 10.48550/arXiv.2609.17795. arXiv: 2609.17795 [math.AG]. URL: <https://arxiv.org/abs/2609.17795>.
- [32] Tony Feng et al. *Towards Autonomous Mathematics Research*. 2026. DOI: 10.48550/arXiv.2602.10177. arXiv: 2602.10177 [cs.AI]. URL: <https://arxiv.org/abs/2602.10177>.
- [33] Shuhong Gao. *Counterexamples to the Jacobian Conjecture in Dimensions Greater than Two*. 2026. DOI: 10.48550/arXiv.2608.00222. arXiv: 2608.00222 [math.AG]. URL: <https://arxiv.org/abs/2608.00222>.
- [34] Jim Geelen. *OpenAI’s Proof of the Cycle Double Cover Theorem*. 2026. DOI: 10.48550/arXiv.2607.15399. arXiv: 2607.15399 [math.CO]. URL: <https://arxiv.org/abs/2607.15399>.
- [35] Liam Giannini. *The Alpöge–Fable Counterexample to the Jacobian Conjecture: Verification, Geometry, Dynamics, and an Equivariant Program for the Plane Case*. Zenodo publication. Version 1.1; independent exact-arithmetic verification; not peer reviewed. July 2026. DOI: 10.5281/zenodo.21461572. URL: <https://doi.org/10.5281/zenodo.21461572>.
- [36] Google DeepMind. *Advanced Version of Gemini with Deep Think Officially Achieves Gold-Medal Standard at the International Mathematical Olympiad*. Published 21 July 2025. July 2025. URL: <https://deepmind.google/blog/advanced-version-of-gemini-with-deep-think-officially-achieves-gold-medal-standard-at-the-international-mathematical-olympiad/> (visited on 09/20/2026).
- [37] Google DeepMind. *AlphaEvolve: A Gemini-Powered Coding Agent for Designing Advanced Algorithms*. Published 14 May 2025. May 2025. URL: <https://deepmind.google/blog/alphaevolve-a-gemini-powered-coding-agent-for-designing-advanced-algorithms/> (visited on 09/20/2026).
- [38] Google DeepMind. *Erdős Problem 90: Formal Statement and Challenge Variants*. Lean 4 source in the Formal Conjectures repository. Statement and benchmark scaffold; principal declarations contain proof holes. 2026. URL: <https://github.com/google-deepmind/formal-conjectures/blob/main/FormalConjectures/ErdosProblems/90.lean> (visited on 09/20/2026).
- [39] Google DeepMind. *Gemini Deep Think: Redefining the Future of Scientific Research*. Published 11 February 2026. Feb. 2026. URL: <https://deepmind.google/blog/accelerating-mathematical-and-scientific-discovery-with-gemini-deep-think/> (visited on 09/20/2026).
- [40] Harvard Center of Mathematical Sciences and Applications. *Summit on PhD Math Education in the Age of AI*. https://cmsa.fas.harvard.edu/aimathphd_summit/. Summit held 17–18 September 2026; accessed 2 October 2026. 2026.
- [41] IOP Publishing. *Generative AI Tools*. <https://publishingsupport.iopscience.iop.org/questions/generative-ai-tools/>. Accessed 1 October 2026.

- [42] John Jumper et al. “Highly Accurate Protein Structure Prediction with AlphaFold”. In: *Nature* 596 (2021). Published online 15 July 2021, pp. 583–589. DOI: 10.1038/s41586-021-03819-2. URL: <https://doi.org/10.1038/s41586-021-03819-2>.
- [43] Youness Lamzouri. *A New Proof That More Than $2/3$ of the Zeros of the Riemann Zeta Function Are Simple and on the Critical Line*. 2026. DOI: 10.48550/arXiv.2609.02882. arXiv: 2609.02882 [math.NT]. URL: <https://arxiv.org/abs/2609.02882>.
- [44] Lean Eval. *Erdős’s Unit-Distance Conjecture Is False*. Benchmark statement and accepted-submission record. 2026. URL: https://lean-lang.org/eval/problems/erdos_unit_distance_conjecture_false/ (visited on 09/20/2026).
- [45] Paul Lezeau. *Lean Verification of the Alpöge–Fable Counterexample to the Jacobian Conjecture*. Lean 4 contribution to the Formal Conjectures repository. Merged pull request. July 2026. URL: <https://github.com/google-deepmind/formal-conjectures/pull/4474> (visited on 09/20/2026).
- [46] Hanzhao (Maggie) Lin and Heng-Tze Cheng. *Gemini achieves gold-medal level at the International Collegiate Programming Contest World Finals*. Google DeepMind. Sept. 17, 2025. URL: <https://deepmind.google/blog/gemini-achieves-gold-medal-level-at-the-international-collegiate-programming-contest-world-finals/> (visited on 09/22/2026).
- [47] Math and AI. *A Severe Misalignment of AI in Mathematics*. Declaration published 11 September 2026. Sept. 2026. DOI: 10.5281/zenodo.22737750. URL: <https://mathandai.org/> (visited on 09/22/2026).
- [48] Akhil Mathew. *A Finite Free Group Scheme of Order Four Not Killed by Four*. Lean 4 contribution to Mathlib. AI-assisted construction and formalization using Codex and Claude under the author’s direction; merged pull request. July 2026. URL: <https://github.com/leanprover-community/mathlib4/pull/41748> (visited on 09/20/2026).
- [49] Kim Morrison. *Comparator Verification for the Uniform-Constant Erdős Unit-Distance Theorem*. Independent Lean comparator artifact. June 2026. URL: <https://github.com/kim-em/erdos-unit-distance-comparator> (visited on 09/20/2026).
- [50] Kim Morrison. *The Uniform-Constant Erdős Unit-Distance Conjecture Is False—Formalized*. Lean 4 formalization. Formalization of an argument by Levent Alpöge; completed 11 June 2026. June 2026. URL: <https://github.com/kim-em/erdos-unit-distance> (visited on 09/20/2026).
- [51] Leonardo de Moura and Sebastian Ullrich. “The Lean 4 Theorem Prover and Programming Language”. In: *Automated Deduction – CADE 28*. Ed. by André Platzer and Geoff Sutcliffe. Vol. 12699. Lecture Notes in Computer Science. Cham: Springer, 2021, pp. 625–635. DOI: 10.1007/978-3-030-79876-5_37. URL: https://doi.org/10.1007/978-3-030-79876-5_37.
- [52] Alexander Novikov et al. *AlphaEvolve: A Coding Agent for Scientific and Algorithmic Discovery*. 2025. DOI: 10.48550/arXiv.2506.13131. arXiv: 2506.13131 [cs.AI]. URL: <https://arxiv.org/abs/2506.13131>.
- [53] OpenAI. *A Proof of the Cycle Double Cover Conjecture*. Technical manuscript. Released July 2026. July 2026. URL: https://cdn.openai.com/pdf/04d1d1e4-bc75-476a-97cf-49055cd98d31/cdc_proof.pdf (visited on 09/20/2026).
- [54] OpenAI. *Advisory Group on Mathematics and Artificial Intelligence*. Published 21 September 2026. Sept. 2026. URL: <https://openai.com/index/advisory-group-on-mathematics-and-ai/> (visited on 09/22/2026).

- [55] OpenAI. *An OpenAI Model Has Disproved a Central Conjecture in Discrete Geometry*. Published 20 May 2026. May 2026. URL: <https://openai.com/index/model-disproves-discrete-geometry-conjecture/> (visited on 09/20/2026).
- [56] OpenAI. *CDC Lean: Formalization of the Cycle Double Cover Theorem*. Lean 4 formalization. July 2026. URL: <https://github.com/openai/cdc-lean> (visited on 09/20/2026).
- [57] OpenAI. *Finite Time Blowup for Navier–Stokes*. Technical manuscript. Sept. 2026. URL: <https://cdn.openai.com/pdf/32d9f210-8b73-45e0-91bc-82a30aef8a9a/navier-stokes.pdf> (visited on 09/20/2026).
- [58] OpenAI. *Navier–Stokes and Euler: Formal Proof Artifacts*. Lean 4 formalization repository. Sept. 2026. URL: <https://github.com/openai/NavierStokesAndEuler> (visited on 09/20/2026).
- [59] OpenAI. *On the Navier–Stokes Millennium Prize Problem*. Published 8 September 2026. Sept. 2026. URL: <https://openai.com/index/navier-stokes-solution/> (visited on 09/20/2026).
- [60] OpenAI. *Planar Point Sets with Many Unit Distances*. Technical manuscript. Announced 20 May 2026. May 2026. URL: <https://cdn.openai.com/pdf/74c24085-19b0-4534-9c90-465b8e29ad73/unit-distance-proof.pdf> (visited on 09/20/2026).
- [61] OpenAI. *Ten Advances in Mathematics and Theoretical Computer Science*. Technical manuscript. Updated 6 August 2026. Aug. 2026. URL: <https://cdn.openai.com/pdf/ten-proofs-oai.pdf> (visited on 09/20/2026).
- [62] OpenAI. *Ten Proofs: Lean Certificates for Ten Advances in Mathematics and Theoretical Computer Science*. Lean 4 formalization repository. Aug. 2026. URL: <https://github.com/openai/ten-proofs> (visited on 09/20/2026).
- [63] Sang-il Oum. *A Proof of the Cycle Double Cover Conjecture by OpenAI: An Exposition*. 2026. DOI: 10.48550/arXiv.2607.16356. arXiv: 2607.16356 [math.CO]. URL: <https://arxiv.org/abs/2607.16356>.
- [64] Oxford University Press. *Policy on AI Use & Disclosure for OUP Journals Authors*. <https://academic.oup.com/pages/for-authors/journals/preparing-and-submitting-your-manuscript/ai-principles-for-oxford-journals/policy-on-ai-use-and-disclosure-for-oxford-journals-authors>. Updated 13 July 2026; accessed 1 October 2026. 2026.
- [65] PLOS. *Artificial Intelligence Tools and Technologies*. <https://journals.plos.org/plosone/s/ethical-publishing-practice>. Policy updated and centralized in 2026; accessed 1 October 2026. 2026.
- [66] Bernardino Romera-Paredes et al. “Mathematical Discoveries from Program Search with Large Language Models”. In: *Nature* 625 (2024), pp. 468–475. DOI: 10.1038/s41586-023-06924-6. URL: <https://doi.org/10.1038/s41586-023-06924-6>.
- [67] Sage. *Artificial Intelligence Policy*. <https://www.sagepub.com/journals/publication-ethics-policies/artificial-intelligence-policy>. Updated policy in force in 2026; accessed 1 October 2026. 2026.
- [68] Will Sawin. *An Explicit Lower Bound for the Unit Distance Problem*. 2026. DOI: 10.48550/arXiv.2605.20579. arXiv: 2605.20579 [math.CO]. URL: <https://arxiv.org/abs/2605.20579>.
- [69] Society for Industrial and Applied Mathematics. *SIAM Publications – Editorial Policy on Artificial Intelligence*. <https://epubs.siam.org/artificial-intelligence>. Version 2.0, effective May 2026; accessed 22 September 2026. 2026.

- [70] Nat Sothanaphan. *Resolution of Erdős Problem #728: A Writeup of Aristotle’s Lean Proof*. 2026. DOI: 10.48550/arXiv.2601.07421. arXiv: 2601.07421 [math.NT]. URL: <https://arxiv.org/abs/2601.07421>.
- [71] David E. Speyer. *The New Counterexample to the Jacobian Conjecture*. Secret Blogging Seminar. July 2026. URL: <https://sbseminar.wordpress.com/2026/07/20/the-new-counterexample-to-the-jacobian-conjecture/> (visited on 09/20/2026).
- [72] Springer Nature. *AI Use in Manuscript Preparation*. <https://www.springernature.com/gp/policies/editorial-policies/ai-manuscript-preparation>. Accessed 22 September 2026.
- [73] STOC 2027. *STOC 2027 Call for Papers*. <https://acm-stoc.org/stoc2027/stoc2027-cfp.html>. Includes the “Use of generative AI and LLMs” policy; accessed 29 September 2026.
- [74] Terence Tao. *What’s new*. <https://terrytao.wordpress.com/>. Research and expository blog; accessed 1 October 2026.
- [75] George Tsoukalas et al. *Advancing Mathematics Research with AI-Driven Formal Proof Search*. 2026. DOI: 10.48550/arXiv.2605.22763. arXiv: 2605.22763 [cs.AI]. URL: <https://arxiv.org/abs/2605.22763>.
- [76] Wiley. *Using AI Tools in Your Research: A Guide for Research Authors, Editors, and Reviewers*. <https://www.wiley.com/en-ae/publishing/article/ai-guidelines>. Research-specific guidelines released 28 October 2025; accessed 1 October 2026. 2025.