AI and the Erdős problems: AI-assisted proofs in research mathematics

IMPLEMENTED 2025–2026 · SCIENCE AND RESEARCH · GLOBAL · BAND C · PARTLY PROVED
VERSION 1.0 · PUBLISHED 20 NOVEMBER 2026 · UPDATED 20 NOVEMBER 2026 · FACTS CHECKED 20 NOVEMBER 2026

Thomas Bloom's erdosproblems.com, a database of problems posed by Paul Erdős, became the main proving ground for AI in research mathematics from late 2025. Early AI help was mostly literature search that found forgotten solutions; AI systems then produced new proofs, checked by mathematicians and in some cases in the Lean proof language, and in May 2026 an OpenAI model found a counterexample to Erdős's unit distance conjecture. Mathematicians report changed working habits and a growing burden of unchecked AI output; most Erdős problems are of narrow significance, and journal publication of AI-generated Erdős proofs is not reported in the sources opened.

ABOUT THIS SYNOPSIS

This is MarLuc's summary and grading of publicly available sources, for information only. It is not investment, legal, financial or professional advice, and should not be relied on for any decision. Readers must validate it against the original sources and, where appropriate, challenge its conclusions.

Facts and figures are as reported by the sources listed below, which may be incomplete, self-reported or out of date. Check the original source before relying on any figure.

The impact band, proof marker and SEAM view are MarLuc's opinion. The organisation is not a MarLuc client and has not reviewed or endorsed this synopsis. Names and trademarks belong to their owners.

NOTICE WORDING LAST UPDATED 16 NOVEMBER 2026

AT A GLANCE

Frontier case: AI used for a scientific, medical or engineering breakthrough. The transformation is potential and unproven; MarLuc grades what has changed in practice so far.

AI used: Large language models and AI agents, including OpenAI's GPT-5.2 Pro and internal models and Google DeepMind's Gemini, used for literature search and proof writing; some proofs formally checked in the Lean proof language with Harmonic's Aristotle.

Implemented: 2025–2026. Grade: band C, Process change, Partly proved.

Headline result, as reported: AI tools helped move about 100 Erdős problems to solved between October 2025 and February 2026, mostly through literature search; by May 2026 an OpenAI model had produced a counterexample to the 1946 unit distance conjecture that Tim Gowers said he would have recommended for acceptance in the Annals of Mathematics

1 · The case

The facts, as reported by the 5 sources listed at the end. Where a source is silent, the field says "Not reported".

Organisation

erdosproblems.com community (run by Thomas Bloom, University of Manchester), with AI systems from OpenAI, Google DeepMind and Harmonic. An online research community around erdosproblems.com, a public database of problems posed by Paul Erdős, built by Thomas Bloom of the University of Manchester. Participants range from leading mathematicians to amateurs, alongside teams at AI companies. Canvas block: 5. People and skills. Also 4. Operations and process.

Business challenge

The challenge. Per Quanta, Erdős posed thousands of problems scattered across more than 1,000 papers, and it was hard to track which were solved or forgotten. Per Scientific American, solving an obscure Erdős problem is rarely worth submitting to a journal and rarely cited, so many problems stayed open through obscurity.

How it was addressed. Bloom built a public database in 2023 and added comments in August 2025. From October 2025 mathematicians and amateurs used public AI models to search the literature and generate proofs, checking them by hand and in some cases formally. Google DeepMind ran Gemini over 700 open problems with human expert review, and OpenAI applied internal models to well-known problems. A community page started by Terence Tao classified AI contributions, with disclaimers. [1] [2] [3] [4] [5]

Who implemented it

erdosproblems.com (Thomas Bloom, University of Manchester) (University or research lab): Curates the database and forum where problems, proofs and claims are posted..
OpenAI (Model or cloud provider): Its public GPT-5.2 Pro model was used by independent solvers; an internal model produced the unit distance counterexample, and it released 372 proof claims in October 2026..
Google DeepMind (Model or cloud provider): Ran Gemini systematically over 700 open Erdős problems with human expert evaluation..
Harmonic (Software vendor): Its Aristotle tool was used to check the logic of the problem 728 proof..
Much of the work was done by independent mathematicians and amateurs using public models; Terence Tao started the community page recording AI contributions.

Technology (SEAM pillar E)

SEAM types: T1 Chat assistants, T7 AI agents. Deployment: Mixture. Vendors and products: OpenAI, Google DeepMind, Harmonic, GPT-5.2 Pro, Gemini, Aristotle, Lean (proof assistant).

Where it was built in: No single deployment: individuals prompt public chat models and post results to the erdosproblems.com forum; AI labs run internal models and agents over the problem list; results and their status are recorded on the community wiki.

What changed

Before: across 2024 and the first eight months of 2025, 111 problems on the list changed from open to solved, per Quanta. After: AI tools helped move about 100 problems to solved in the four months from October 2025, mostly by literature search, and later produced original proofs. Working habits changed: Wouter van Doorn says many of his recent papers should be mostly credited to AI, and Noga Alon says he has stopped trying to solve Erdős problems. [1] [2] [3] [4] [5]

What it cost

Not reported. Quanta quotes a DeepMind team saying its agent resolved problems at a per-problem cost of a few hundred dollars.

What it saved

Not reported. [1] [2] [3] [4] [5]

Governance and control

Controls are community norms: human checking, Lean formalisation, a public record of AI contributions with disclaimers, and curator notes treating AI releases as claims until checked. No journal or institutional standard is reported.

S2 · Accountable owner · Evidenced
Thomas Bloom curates the database and works out the most sensible version of ambiguously stated problems; AI developers are responsible for their own releases. [1]

A1 · Safe use · Evidenced
The community wiki publishes disclaimers, including that it is not a benchmark and that open status may reflect obscurity rather than difficulty; Bloom asks readers to treat AI proofs as claims. [4] [5]

A2 · Human in the loop · Evidenced
Google DeepMind's study used AI verification to narrow the search, followed by human expert evaluation of correctness and novelty; amateur users sought mathematicians to check their AI proofs. [1] [3]

A3 · Monitoring and assurance · Evidenced
A public wiki recorded each AI contribution with a status, including partial progress, incorrect work and unverified claims, until 30 June 2026. [4]

A4 · Vendor AI · Evidenced
The models are named in the sources; the unit distance result came from an internal OpenAI model not available to the public. [1] [3]

A5 · Agent controls · Not reported
Quanta reports a DeepMind agent that autonomously resolved 9 of 353 open problems; controls on such agents are not described. [1]

A6 · Standards and regulation · Not reported
No journal, society or institutional standard on AI-generated proofs is reported in the sources opened; formalisation in Lean is a voluntary check.

What worked, what did not, what is unproven

WORKED

– AI literature search found forgotten solutions, including an offhand remark in a 1981 paper that solved problem 1089, per Scientific American.

– Problem 728 was solved with GPT-5.2 Pro and checked with Aristotle; nobody could find prior work proving it, per Quanta.

– Tim Gowers said he would have recommended the unit distance paper for acceptance at the Annals of Mathematics without hesitation; related techniques were used to solve other problems

– Google DeepMind's study reports 5 seemingly novel autonomous solutions and 8 existing solutions found among 700 problems labelled open.

DID NOT

– Claimed AI solutions were repeatedly found to be already in the literature; problem 333 had been resolved by Erdős in a 1977 paper.

– In the First Proof challenge, Daniel Litt found many circulating AI proofs to be largely bogus and said the models were generating a huge amount of garbage, per Scientific American.

– Bloom describes long AI-generated papers that no human has read, and notes accusations of plagiarism around OpenAI's October 2026 release.

– DeepMind reclassified one of its autonomous solutions (problem 935) as an independent rediscovery, and its authors warn of subconscious plagiarism by AI.

UNPROVEN

– Whether AI-generated Erdős results pass peer review in major journals; Scientific American reported in February 2026 that no major mathematics journal had published a peer-reviewed proof citing the use of LLMs.

– Whether the results matter beyond the list; Andrew Sutherland says Erdős problems are mostly individual problems without broader implications.

– Whether the 25 Erdős-related proofs in OpenAI's October 2026 release are correct; Bloom had not yet checked them and expects most to be correct but in need of simplification.

– The longer-term effect on mathematicians' careers and training; Quanta reports leading mathematicians moving to AI companies.

Timeline

2023 Thomas Bloom launches erdosproblems.com as a personal list of Erdős problems [1]

Aug 2025 A commenting function is added and a community forms around the site [1]

Oct 2025 Mathematicians begin using ChatGPT to find forgotten solutions in the literature [2]

Dec 2025 A claimed first autonomous AI solution (problem 333) turns out to have been resolved by Erdős in 1977 [1]

Jan 2026 Problem 728 solved with GPT-5.2 Pro and the proof checked with Harmonic's Aristotle [1]

Jan 2026 Google DeepMind posts its Gemini case study on 700 open Erdős problems [3]

Feb 2026 About 100 problems moved to solved with AI help since October, per Scientific American [2]

May 2026 OpenAI announces an AI-found counterexample to Erdős's unit distance conjecture, with mathematicians commenting on its correctness [1]

Jun 2026 The community wiki of AI contributions stops updating, with data as of 30 June 2026 [4]

Oct 2026 OpenAI releases 372 proof claims; Bloom lists 25 relevant to Erdős problems and notes accusations of plagiarism [5]

2 · Grade and SEAM assessment

MarLuc's opinion, made on the evidence above. How cases are graded: marluc.co.uk/case-studies/how-we-rate

Level of change and impact

Claimed: A. Grade: C, Process change. Proof: Partly proved. Level of change: Process change.

Process change (Level 2, band C): problem solving in this community is now AI-assisted from literature search to proof, with people checking, formalising and simplifying. Claimed A: mathematicians quoted describe a dramatic change in how research is done; MarLuc treats that as potential and unproven.

Graded C. Within this research community the process of resolving problems has changed: AI systems search the literature, draft proofs and sometimes resolve problems outright, while people check, formalise and simplify the results. Independent science press and the mathematicians who keep the records document that change. The wider claim, voiced by Noga Alon, that AI is changing dramatically the way mathematical research is done, is treated as potential and unproven: Erdős problems are mostly narrow, results are largely unpublished in journals, and the sources describe a growing volume of unverified AI output.

The SEAM view and lessons learnt

Results came from capable public and internal models, a well-curated open problem list and a community willing to check proofs. The AI's part is visible because contributions are recorded case by case, though company resources and the many obscure, easier problems also shaped the counts. Checking, not generating, is now the bottleneck. Before relying on AI-generated analysis, a leadership team should ask who checks the output and how, and how prior work and credit are traced.

1. Keep a public record of what AI contributed and how each result was checked.
2. Search for prior work before announcing a breakthrough.
3. Budget expert time for checking, since generating output is now cheap.

Evidence

Score 2 of 4. Mixed: independent science journalism quoting named mathematicians; a developer preprint, not peer reviewed; community records kept by mathematicians with explicit disclaimers. Some proofs formally verified in Lean. Journal publication not reported.

Evaluation method. Community records and expert review: a public log of AI contributions classified by mathematicians, a developer case study with human expert evaluation, and formal verification of some proofs in Lean. No controlled comparison.

Sources disagree. Quanta reports mathematicians hailing recent results as a phase transition and quotes Tim Gowers praising the unit distance proof. In February 2026, before that result, Scientific American quoted Daniel Litt saying every individual result had been vastly overhyped

3 · Original sources

Read these in full. To challenge a fact or a grade, use the contact form at marluc.co.uk or the link on https://marluc.co.uk/case-studies/ai-erdos-problems-mathematics

[1] Quanta Magazine: How AI tore through a mathematical community (3 August 2026)
https://www.quantamagazine.org/how-ai-tore-through-a-mathematical-community-20260803/

[2] Scientific American: AI uncovers solutions to Erdős problems, moving closer to transforming math (12 February 2026)
https://www.scientificamerican.com/article/ai-uncovers-solutions-to-erdos-problems-moving-closer-to-transforming-math/

[3] arXiv (Google DeepMind and others, 2026): Semi-Autonomous Mathematics Discovery with Gemini: A Case Study on the Erdős Problems
https://arxiv.org/abs/2601.22401

[4] GitHub (community wiki started by Terence Tao): AI contributions to Erdős problems
https://github.com/teorth/erdosproblems/wiki/AI-contributions-to-Erd%C5%91s-problems

[5] Erdős Problems (Thomas Bloom): Erdős problems in the OpenAI proofs (8 October 2026)
https://www.erdosproblems.com/forum/thread/blog:10

erdosproblems.com community (run by Thomas Bloom, University of Manchester), with AI systems from OpenAI, Google DeepMind and Harmonic website: https://www.erdosproblems.com

Version 1.0 · Published 20 November 2026 · Updated 20 November 2026 · Facts checked against sources 20 November 2026 · Downloaded 26 November 2026. This PDF may be out of date: the latest version, and a way to report an error or a better source, is at https://marluc.co.uk/case-studies/ai-erdos-problems-mathematics

VERSION HISTORY

v1.0 20 November 2026 First published