Comment

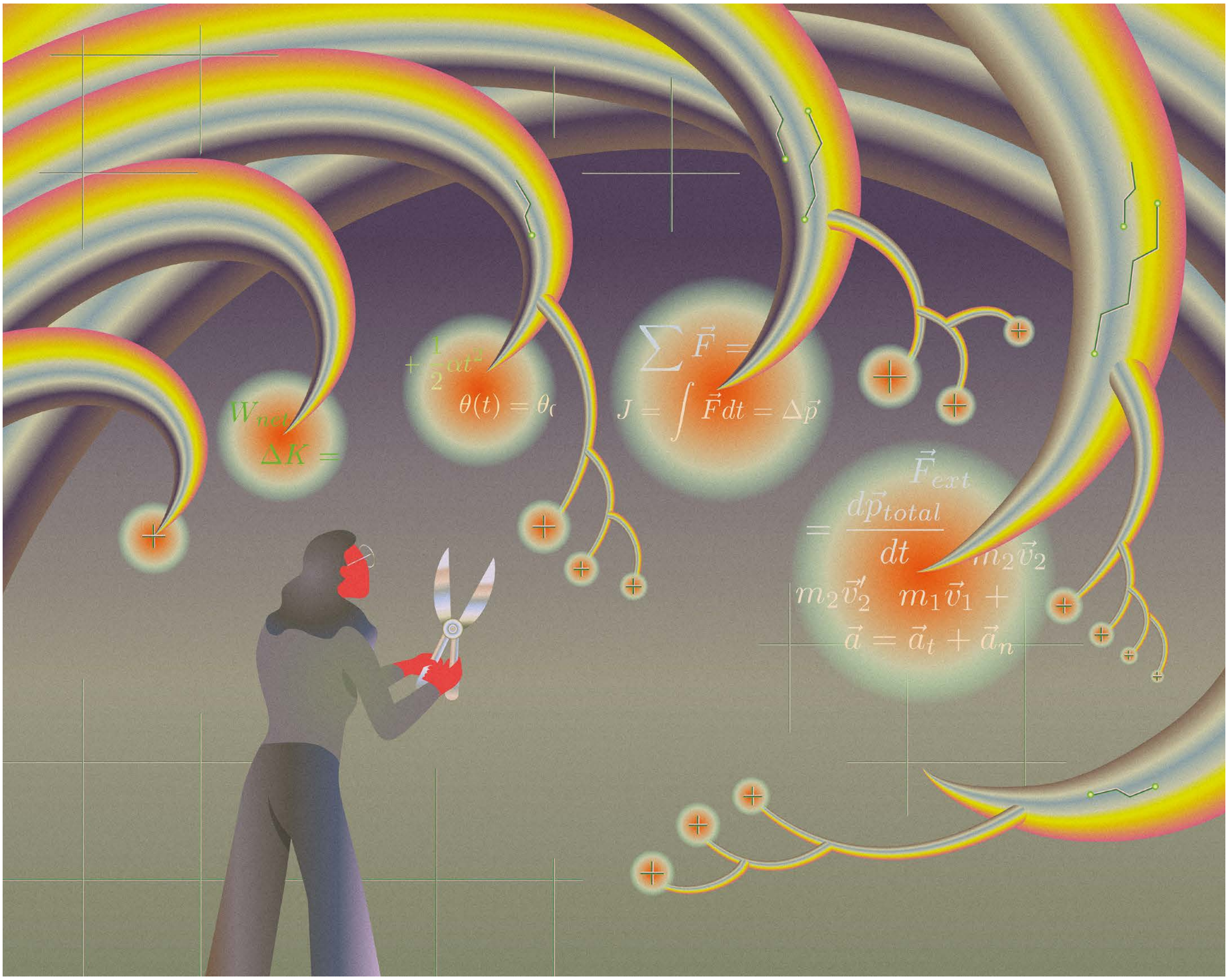


ILLUSTRATION: ANA KOVA

How AI is reshaping discovery in maths and physics

Mikhail Burtsev, Yang-Hui He, Evgeny Sobko, Thore Graepel & Ananyo Bhattacharya

Artificial intelligence is not replacing human intuition in these fields, but reimagining how questions are asked, explored and understood.

Among mathematicians and theoretical physicists, artificial intelligence provokes a range of reactions. Some see it as irrelevant to their work; others fear it could encroach on the most creative, intellectually rewarding aspects of their fields. Yet, the truth that's emerging, from the work our team is doing at the London Institute for Mathematical Sciences and elsewhere, is subtler.

Rather than displacing human creativity in mathematical sciences, AI is augmenting it. Software can now check proofs line by line and catch errors that would once have taken months of human scrutiny to find. It can search systematically for counterexamples – testing whether a conjecture truly holds or fails in an unexpected way. And it can propose intermediate steps in an argument, suggesting useful auxiliary results that help to bridge the gap between what is known and what still needs to be shown.

In experimental fields, prototype 'AI scientists' are beginning to automate parts of the discovery cycle, but they remain constrained by the demands of the physical world: mixing reagents, culturing cells, waiting for reactions and contending with noise in the data. Mathematics and theoretical physics face many fewer bottlenecks. 'Experiments' are cheap, fast and digital, and mathematical data – from prime numbers to the properties of abstract structures, such as manifolds – are clean and abundant[1].

Companies developing AI systems tailored to mathematical reasoning have reported steady progress in the past year. Aristotle, a system from software company Harmonic in Palo Alto, California, has helped to solve several problems posed by the prolific mathematician Paul Erdős – questions that are easy to state but notoriously hard to crack. Axiom Math, a start-up company in Palo Alto, has announced that its AI tool found solutions to many research-level problems that professional mathematicians had not yet solved. Meanwhile, models from technology firms OpenAI in San Francisco, California, and Google DeepMind in London have solved several challenges from the First Proof Project, a set of difficult mathematical problems that test whether AI systems can generate new and verifiable results.

Here, we give examples of progress in the past few years in this rapidly evolving area, outline the opportunities that AI presents to scientists and mathematicians in theoretical domains – and invite researchers to lean in to using AI in their work.

The research pipeline

In theoretical physics and maths, researchers weave together creative insight and rigorous logical reasoning to make discoveries – but this process is only partly understood, and there is no single explanation for how breakthroughs happen. For clarity – without putting forth a definitive model – we break the process into several overlapping phases: setting the agenda, formalizing ideas, proposing conjectures and solving and verifying results. This framework is imperfect, but it provides a useful way to assess where AI is already contributing, where challenges lie and how they might be addressed.

Setting the agenda. One of the most distinctly human acts in research is deciding which questions are worth asking in the first place. These might arise from outside the field – through real-world problems or contact with neighbouring disciplines – or from within it, in that theories evolve according to their own internal logic and aesthetic standards[2,3]. These sources are intertwined: concrete problems can generate new concepts, and abstract theory can reshape and deepen the original question.

Today's AI systems have only limited access to this broader context. As a result, they lack intuition and 'taste': a sense of where questions come from, what makes them timely and how they fit into a field's evolving structure. For instance, physicist Albert Einstein developed his special theory of relativity after noticing a contradiction in how light waves were treated in classical mechanics and in Maxwell's equations, which describe the interplay of electricity and magnetism.

"Mathematicians can benefit from a system that insists every inference be made explicit."

One promising but under-explored direction is to build AI systems that help to sort and prioritize potential problems using criteria selected by researchers. For example, AI could follow those criteria when scanning large mathematical databases, such as the On-Line Encyclopedia of Integer Sequences, or preprint repositories, including arXiv, to identify overlooked connections and structural parallels between fields. Used in this way, AI might sharpen our understanding of how scientists identify fertile directions for discovery.

Formalizing ideas. Many important ideas take shape before they can be precisely defined. A classic example is the path integral, introduced by theoretical physicist Richard Feynman, which describes quantum systems by imagining all the ways something could happen and combining them. Although this idea has never been fully pinned down in a strict mathematical sense, it has shaped modern physics and inspired new tools in maths[4] – for example, ways to distinguish between different types of knots and methods for counting shapes in complex geometries.

Turning an informal, prose-style argument into a form that a computer can process often demands substantial effort: reconstructing omitted steps, filling in seemingly obvious gaps and making tacit assumptions explicit. But this process can deepen understanding and expose errors. For example, when mathematician Terence Tao at the University of California, Los Angeles, ran an argument in one of his own papers through a proof assistant (Lean4) to check it, he spotted a subtle gap in the logic. A step that had seemed clear had not been rigorously justified.

Even the most accomplished mathematicians can benefit from a system that insists every inference be made explicit. Reducing the human labour involved in formalization would lead to larger, higher-quality bodies of verified maths, which in turn could be used to train better AI models. Fully automating formalization is the long-term goal.

Progress has been substantial[5], but human input is still required. For example, the Xena project, led by mathematician Kevin Buzzard at Imperial College London, has mobilized university students to systematically digitize all the proofs in the undergraduate maths curriculum.

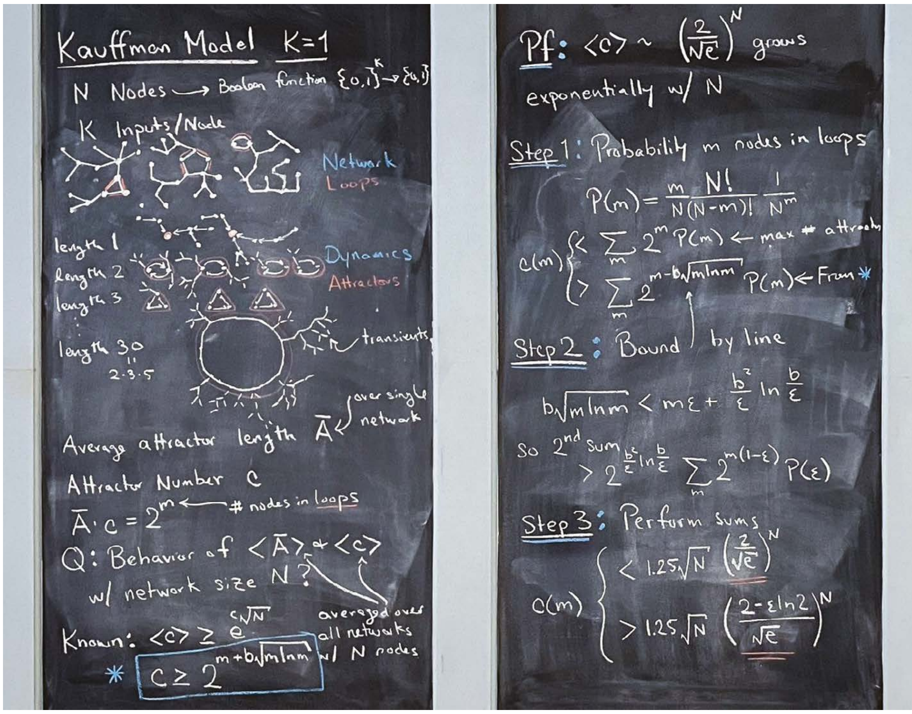
AI is beginning to help scale up such tasks. Computer scientist and mathematician Josef Urban at Chalmers University of Technology in Gothenburg, Sweden, used a large language model to formalize theorems in topology – the study of the properties of shapes when they are stretched or twisted.

Proposing conjectures. A conjecture is a plausible answer to a well-posed problem–that is, an educated guess that seems likely to be true but has not yet been proven. AI can now generate conjectures, but its role remains tentative and tightly coupled to human oversight.

This is not a new area for computational approaches. Early specialized computer programs – such as Graffiti[6] and the Ramanujan Machine[7] – showed that algorithms can indeed suggest new mathematical ideas, not just check existing ones.

Graffiti, for instance, found unexpected patterns in networks – simple diagrams of connected points – that later proved useful in chemistry, in which molecules can be understood in terms of how their atoms are linked. The Ramanujan Machine proposed surprisingly simple formulae for fundamental mathematical constants. Similar approaches are now being applied in theoretical physics, helping researchers to uncover hidden patterns and exact formulae[8–10].

In practice, however, AI generates many conjectures, most of which are trivial, previously known results or false. Human experts still decide which ones are worth pursuing. For example, in 2021, AI helped to narrow down a broad hypothesis concerning the algebraic and



The deepest advances in maths and physics still require human creativity and judgement.

ROMAN RYBIANSKY/LONDON INSTITUTE FOR MATHEMATICAL SCIENCES

geometric structure of mathematical 'knots' to a single rigorously defined conjecture, which was then proved by humans[11].

In 2022, researchers who used AI to analyse large data sets of elliptic curves — important mathematical objects in number theory, or the study of integers — noticed an unexpected pattern in how some key properties vary. When they plotted the data, they saw that it was not randomly scattered but formed wave-like bands that resembled the flocking behaviour of starlings, known as murmurations[12]. Uncovering such patterns might prove transformational in many maths fields[9].

The next step could be linking AI-enhanced generation of conjectures with agenda-setting. Rather than working blindly in a fixed domain, AI systems could first map the existing body of maths knowledge to identify bottlenecks, gaps and unexpected parallels, and then generate conjectures to bridge them.

Solving and verifying results. In 2025, DeepMind released AlphaEvolve[13], a coding agent that can propose, test and refine algorithmic solutions to open problems. Soon after, experts tested it on 67 challenges; in most cases, it rediscovered the best-known solutions and, in several cases, improved them[14].

AlphaEvolve integrates the generative reasoning of Google's AI model Gemini with automated systems that assess candidate solutions, using an 'evolutionary search' strategy to iteratively develop the most promising ones. It has demonstrated the ability to advance mathematical knowledge by, for example, discovering improved algorithms for multiplying matrices (used in a variety of areas in physics, data science and computer science). Meanwhile, in May, OpenAI announced that it had used a large language model to disprove the unit distance problem, a geometry conjecture first posed by Erdős in 1946 — perhaps the first major mathematical result produced by a machine. These successes are notable, and although the overall state of the art remains limited, they suggest that the pace of progress is accelerating.

Using AI to check proofs — or verify them — is a more-developed application. Proof assistants can already check complex arguments line by

"For now, the decisive creative leaps are still made by humans. The real promise lies in partnership."

line, and their growing libraries provide a structured foundation for AI-assisted reasoning. Formal verifications of complex theorems show that these tools are approaching routine use at the research frontier[15].

Rather than a single all-purpose 'AI mathematician', progress is likely to come from ecosystems of specialized agents — generators, refuters, surveyors, educators — whose interplay produces reliable knowledge. Future AI tools might go further, experimenting with how to approach a problem and judging which strategies lead to quicker, cleaner proofs.

Looking ahead

AI systems that suggest proof steps, uncover hidden patterns and solve competition-level problems now assist mathematicians in ways that were unimaginable just five years ago. Yet the deepest advances in maths and physics often demand radically new concepts or paradigms, and no AI system has yet been able to invent them. For now, the decisive creative leaps are still made by humans. The real promise lies in partnership.

AI can search vast spaces and surface unexpected regularities; humans bring judgement, taste and the ability to invent fresh ways of thinking. This collaboration is already producing new results. Theory is not an assembly line of solved problems; it is an expanding map of human understanding. Earlier tools, such as calculators and computer algebra systems, did not diminish the field of mathematics — they elevated it. AI can do the same, extending our cognitive reach much as the telescope once extended our sight.

Future systems must explain their insights, coach researchers entering new areas and help to organize growing bodies of knowledge. The task now is to build these systems with care and ambition. If they can make the frontier more navigable — and more deeply interconnected — they will accelerate discovery, not replace discoverers.

The authors

Mikhail Burtsev is an AI fellow at the London Institute for Mathematical Sciences in London, UK. **Yang-Hui He** is a fellow at the London Institute for Mathematical Sciences in London, UK. **Evgeny Sobko** is a fellow at the London Institute for Mathematical Sciences in London, UK. **Ananyo Bhattacharya** is chief science writer at the London Institute for Mathematical Sciences in London, UK. **Thore Graepel** is a research scientist at Google DeepMind, London, UK.
e-mail: ab@lims.ac.uk

1. Fink, T. *Nature* **629**, 505 (2024).
2. Hilbert, D. *Bull. Am. Math. Soc.* **8**, 437–439 (1900).
3. von Neumann, J. in *Works of the Mind* Vol. I 180–196 (Univ. Chicago Press, 1947).
4. Atiyah, M., Dijkgraaf, R. & Hitchin, N. *Phil. Trans. A Math. Phys. Eng. Sci.* **368**, 913–926 (2010).
5. Weng, K. *et al.* Preprint at arXiv https://doi.org/10.48550/arXiv.2505.23486 (2025).
6. Fajtlowicz, S. in *Annals of Discrete Mathematics* Vol. 38 (eds Akiyama, J., Egawa, Y. & Enomoto, H.) 113–118 (Elsevier, 1988).
7. Raayoni, G. *et al. Nature* **590**, 67–73 (2021).
8. He, Y.-H. *The Calabi–Yau Landscape* (Springer, 2021).
9. Douglas, M. R. *Nature Rev. Phys.* **4**, 145–46 (2022).
10. He, Y.-H. *Nature Rev. Phys.* **6**, 546–553 (2024).
11. Davies, A. *et al. Nature* **600**, 70–74 (2021).
12. He, Y.-H., Lee, K.-H., Oliver, T. & Pozdnyakov, A. *Exp. Math.* **34**, 528–540 (2025).
13. Novikov, A. *et al.* Preprint at arXiv https://doi.org/10.48550/arXiv.2506.13131 (2025).
14. Georgiev, B., Gómez-Serrano, J., Tao, T. & Wagner, A. Z. Preprint at arXiv https://doi.org/10.48550/arXiv.2511.02864 (2025).
15. Gowers, W. T., Green, B., Manners, F. & Tao, T. *Ann. Math.* **201**, 515–549 (2025).

The authors declare competing interests; see go.nature.com/3s8jut for details.