AI as a Scientific Collaborator

From biology to black holes, ChatGPT is accelerating research

January 2026

OpenAI

Introduction: Why AI for Science Matters

8.4 million average weekly messages
on advanced topics in the hard sciences and mathematics

Roughly 1.3 million weekly ChatGPT users
focus on advanced math & science topics worldwide

The number of monthly advanced science messages
grew nearly 50%
last year

OpenAI is building tools to help researchers generate insights, accelerate scientific discovery and translate those insights into real-world impact. Across ChatGPT, researchers, students, STEM faculty and engineers already use AI to read and synthesize technical literatures, debug and write code, analyze data, and plan experiments. Each week, ChatGPT sees almost 8.4 million messages on advanced topics in the sciences and mathematics. These now come from roughly 1.3 million weekly users worldwide.

Only about 0.1 percent of the global population identifies as scientists, according to UNESCO, and yet they have an outsized impact. Scientific research drives the engine of progress toward a healthier, more prosperous, and more resilient future. New medicines, new technologies, and new industries come from new knowledge put to practical use. A small group of early twentieth-century physicists laid the foundations of quantum mechanics through abstract research that, decades later, would underpin much of the modern digital economy, now measured in the tens of trillions of dollars. Basic research – work done before the payoff is clear – was the source of that knowledge. In 1947, scientists at Bell Labs created the first working transistor, based on insights from quantum physics. The transistor became a building block of computers, phones, and today's digital technology. The Global Positioning System (GPS) relies on Einstein's insights into relativity to guide our cars and keep atomic clocks aligned.

Yet in many domains, it is getting harder to keep making progress. Economists and research analysts point to falling "research productivity," meaning more people, time, and money are required to produce the same number of insights. Semiconductors offer a well-known example: sustaining Moore's Law to double the number of transistors in an integrated circuit every two years has required a dramatic increase in effort, with the number of researchers needed today estimated at more than 18 times what was needed in the early 1970s. As knowledge grows more complex, each new generation of researchers faces a heavier burden just to reach the frontier, which lengthens their training time and narrows their specializations. Institutionally, research has shifted toward larger teams, with growing overhead for grant proposals, compliance, reporting, and coordination costs.

In medicine, scientific advances have saved countless lives. Worldwide life expectancy rose from roughly 32 years in 1900 to about 73 years in 2023 (and to more than 78 years in the United States). But the remaining burden of disease is heavy. But the World Health Organization reports that noncommunicable diseases such as stroke, heart disease, cancer, and diabetes still account for about 74% of global deaths. Even when progress is rapid, turning new ideas into available treatments takes time. On average, it takes 10-15 years from target discovery to regulatory approval of a new drug in the United States, a lag imposed on patients who need new and better treatments.

Making progress faster will save lives and improve them. AI is already helping to address the bottlenecks that slow science down. Modern research is fragmented across disciplines and constrained by limits that are both cognitive and logistical: reading and digesting enormous literatures to determine what is known, translating ideas into mathematics and code, setting up analyses and simulations, checking calculations, searching huge design spaces, and deciding which future experiments are the most promising. Used well, AI can serve as a high-throughput partner for thought, computation, and structured reasoning, shortening the cycle from hypothesis to test and increasing the capacity of researchers working alone and in teams, even across disciplinary barriers.

Kevin Weil, VP of OpenAI for Science, describes the opportunity this way: *"AI is increasingly being used as a scientific collaborator, and we're seeing its impact grow in real research settings. More researchers are using advanced reasoning systems to make progress on open problems, interpret complex data, and iterate faster in experimental work. That usage has been growing quickly over the past year, and the results are starting to show up across fields. We're still early, but the pace of adoption and the quality of the work suggest science is entering a new acceleration phase."*

OpenAI is proud to work with research partners across government agencies, national laboratories, academia, and medicine, including the U.S. Department of Energy, Lawrence Livermore National Laboratory, the U.S. Centers for Disease Control and Prevention, Harvard University, Massachusetts Institute of Technology, the University of Oxford, Texas A&M University, and Boston Children's Hospital.

This report details:

1. How AI tools are already being used in day-to-day research workflows, including literature synthesis, code generation and debugging, data analysis, simulation support, and experiment planning
2. What early results suggest about AI's potential to support new breakthroughs
3. How individual scientists across multiple disciplines have used ChatGPT to make progress in their field
4. Policy suggestions to support continued AI progress in science and math

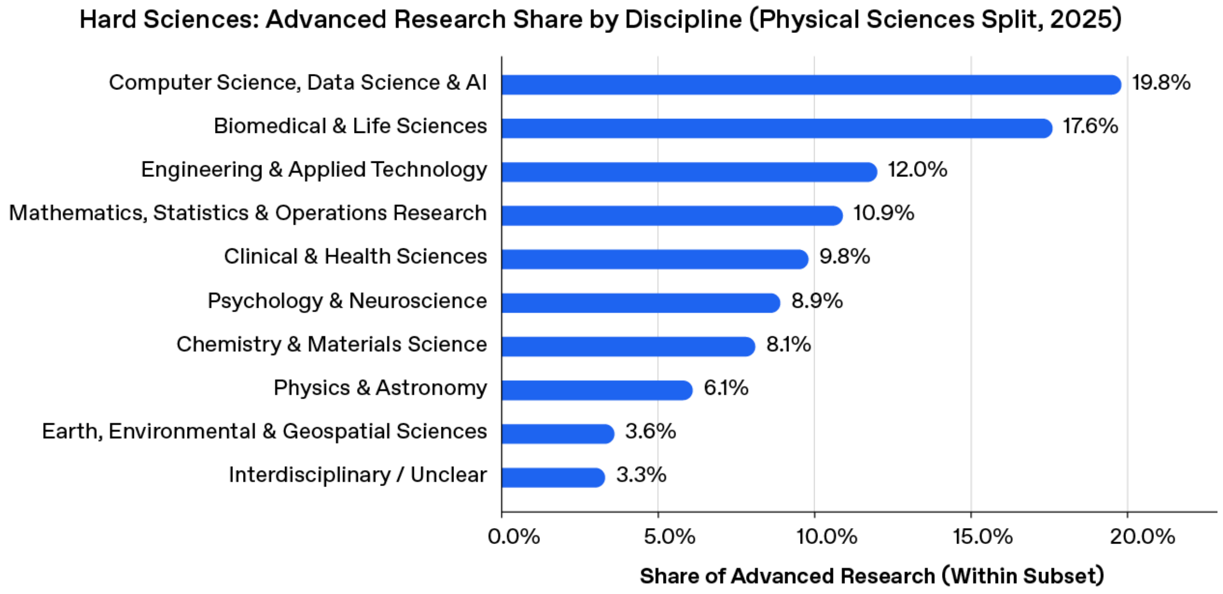
The scale of scientific and mathematical work on ChatGPT

Across ChatGPT, a small but consequential cohort of researchers uses OpenAI's AI models for sophisticated tasks ranging from technical derivations, advanced mathematics, engineering simulation and modeling, and other advanced problem-solving. This includes scientists and mathematicians spanning PhD candidates and post-docs to working researchers and STEM faculty.

Based on an internal analysis of a full random sample of anonymized ChatGPT conversations from January through December 2025, average weekly message counts on advanced science and math topics grew about 47%, from 5.7 million messages to nearly 8.4 million messages over the course of 2025. As of January 2026, there are nearly 1.3 million weekly users discussing advanced topics in science and math.

Together, these signals show how ChatGPT is accelerating advanced research: with tens of millions of advanced hard-science and math prompts each month, generated by a large and growing cohort using the system for serious scientific and engineering work to benefit society and support economic growth.

Here's how usage breaks down across disciplines among our cohort of research-focused users:



What scientists and mathematicians actually do with ChatGPT

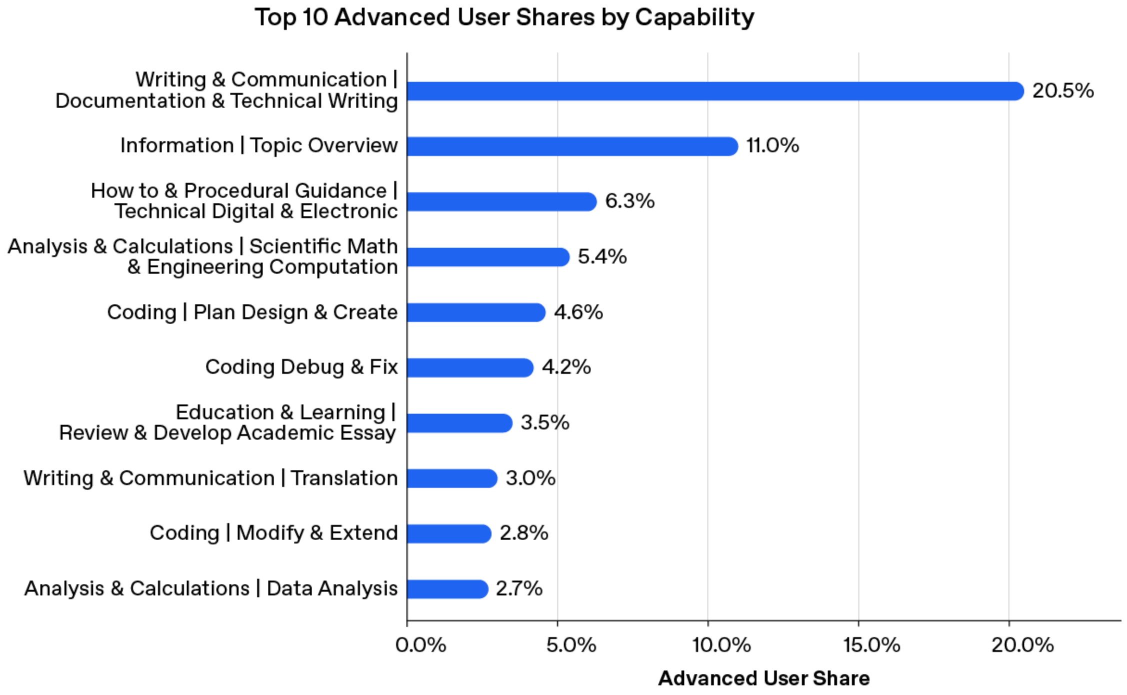
Scientists, mathematicians, and engineers use ChatGPT as a highly available technical collaborator: a tool with which they can iterate on calculations, translate ideas into code, interrogate assumptions, and compress complex materials into workable mental models. In OpenAI's analysis, "advanced" hard science prompts are defined as those being oriented toward research, and requiring graduate-level or research-level expertise to answer competently. Within that cohort, behavior differs from typical users in ways that map directly onto modern research workflows.

Research tasks cluster in domains such as coding (drafting, reworking, and debugging code), data analysis (cleaning and merging datasets, running statistics, interpreting results), mathematical reasoning (derivations, proof strategies, algebraic checking, long calculations, translating between formalisms), and literature review and synthesis (finding references, understanding recent work).

Contrasted with typical users of ChatGPT, advanced science and math users:

- Send roughly 3.5× more messages than the baseline
- Send coding-related messages nearly 12× more often
- Average 9 informational-overview prompts per week vs. 1.5 prompts

Categorized by most common tasks, here is how this cohort tends to use ChatGPT,:



AI at the frontier of math and science

Frontier AI capabilities in mathematics

Recent progress

Over the last two years, large language models have progressed from early, uneven performance on basic arithmetic to handling multi-step mathematical reasoning that can be useful in real mathematical work. Much of that improvement came from methods that encourage step-by-step reasoning, and from tighter integration with tools like calculators and code execution for exact computation. As models improved, benchmarking also shifted toward harder tests designed to measure deeper reasoning while reducing "pattern matching" wins.

In 2025 and early 2026, the greatest impact has come from test-time compute scaling, or "slow thinking." Instead of committing quickly to one path, a model will spend more computation exploring alternatives and self-checking. At the same time, approaches to training that reward verifiable outcomes, such as producing a correct final answer or executable code, have pushed math and coding to become more reliable, and correct often enough to be useful with human guidance.

One sign of this shift came with International Mathematical Olympiad coverage in July 2025, when an OpenAI model achieved gold-level performance on the 2025 problem set alongside DeepMind.

Current capabilities

GPT-5.2's steady advance in mathematical capabilities stems from stronger long-horizon reasoning, more systematic verification habits, and better use of checkable tools. AIME, the American Invitational Mathematics Examination, is designed to test multi-step problem solving: GPT-5.2 Thinking achieved a perfect score on AIME 2025 without external tools. The GPT-5.2 series (e.g. Thinking and Pro) has progressed past competition-level performance toward mathematical discovery, including work on established open problems.

On research-style benchmarking, the "Google-proof" FrontierMath problem set has been constructed to be accessible only to true experts; i.e. even a smart PhD student in math cannot solve them in a few hours of work. On that benchmark, GPT-5.2 Thinking has solved 40.3% of problems in Tiers 1–3.

Performance still drops on the hardest tier, where GPT-5.2 Pro has scored 31% on FrontierMath Tier 4 on a set of problems that can be described as "mini research projects."

A second major capability leap has occurred as GPT-5.2 is increasingly paired with formal verification workflows. In one prominent integration, GPT-5.2 generates natural language proofs and uses Aristotle, a third-party LLM, to formalize those proofs in Lean, which is a proof assistant where proofs are written in a form a computer checks step by step, with the system detecting and correcting gaps during formalization.

These integrations matter because one longstanding failure mode for language models in mathematics happens when a solution "looks right", but isn't: i.e. arguments that appear plausible, but contain subtle gaps. Lean-checked proofs substantially raise the standard of confidence in a proof by forcing explicit, mechanically checked steps under a stated formalization.

Erdős problems, AI solutions

Paul Erdős (1913–1996) was a globe-trotting Hungarian mathematician who lived out of a suitcase, moving from campus to campus as he sought out fellow math researchers. The "Erdős problems" are the vast set of questions and conjectures he posed, ranging from deceptively simple puzzles to problems that still resist the best techniques we have. The problems have acted like trail markers for modern mathematics: clarifying what we don't yet understand, seeding whole research programs, and drawing generations of mathematicians toward the edges of the known.

In early 2026, GPT-5.2 has contributed to solutions to several open Erdős problems with the help of tools like Aristotle and Lean, and with the solutions validated by Terence Tao. Problems #281, #728, #729 are now listed as proved, and #397 as disproved. While mathematicians caution that Erdős problems vary enormously in difficulty, these solutions point to the increasing capability of OpenAI models to do real mathematical work and make novel contributions with minimal guidance.

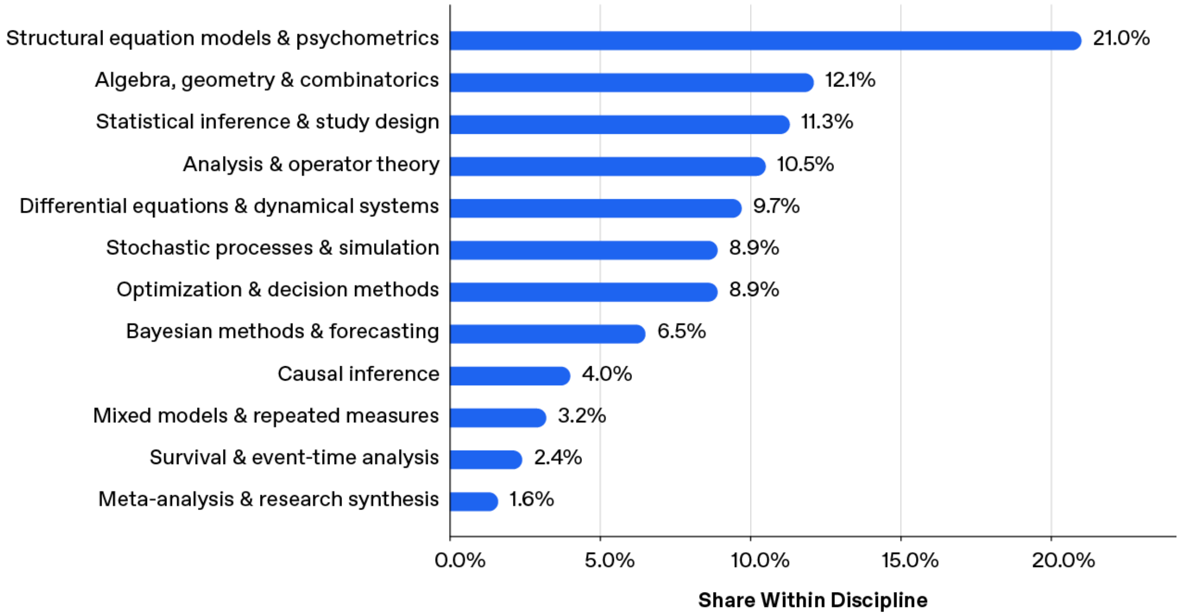
Near-term potential

Some significant mathematical discoveries can take the form of stitching known methods together to find the correct argument. GPT 5.2 can do this now in many cases. Other discoveries involve inventing entirely new kinds of math, as Newton invented calculus to understand dynamics (the forces, mass and energy that explain changes in motion). That is beyond current AI models. But a third type of significant discovery involves establishing connections between two fields, and bringing the known machinery, results and tools of one field to the other (e.g. algebraic geometry, which arose from abstract algebra and classical geometry). Modest examples of this have already occurred with AI, and we believe the significance of those connections, and the subfields created from them, will increase in the near future.

At the same time, much of AI's near-term value will be in transforming workflows. GPT-5.2 can propose and iterate on solution paths, while external tools enforce correctness through exact computation or formal checking. This aligns with broader trends toward hybrid approaches and auto-formalization, where informal math is translated into formal languages like Lean so that correctness can be verified mechanically. GPT-5.2 is already useful for literature review (surveying what is known) to locate the edge of knowledge and surface unexpected or obscure references. It is useful in coming up with proofs, critiquing them, simplifying them and suggesting proof strategies.

If this trajectory continues, GPT-5.2's near-term impact is likely to show up as a broad productivity upgrade for mathematical researchers, as well as science and engineering teams (since math is a core component of much scientific and engineering work): this will manifest as faster translation from a messy problem description to a clean mathematical statement, fewer dropped constraints in multi-step derivations, more reliable debugging of calculations and proofs, and a growing share of results that can be backed by formal verification.

Hard Sciences – Mathematics, Statistics & Operations Research: Topic Share Within Discipline (Physical Sciences Split, 2025)



Frontier AI capabilities in science

Across disciplines such as physics, chemistry, and biology, ChatGPT-class LLMs increasingly support technical reasoning and tool-mediated research workflows, as well as scientific writing. Benchmarks like GPQA, a graduate-level set of “Google-proof” questions authored by domain experts, initially showed a significant human advantage, with experts reaching 65% accuracy while the GPT-4 baseline reached 39%. OpenAI now reports GPQA Diamond accuracy of 93.2% for GPT-5.2 Pro and 92.4% for GPT-5.2 Thinking (with no tools enabled and at maximum reasoning effort), suggesting a higher baseline for graduate-level scientific question answering across many scientific disciplines. In parallel, automating the unspoken drudgery of research – reference hunting, bibliography assembly, and routine administrative reporting – frees scientists’ scarce attention for higher-value work. Paired with information retrieval and executable tools that enable checkable calculations and stepwise validation, these models are becoming reliable workflow orchestrators for scientific planning, analysis, and documentation, teeing up acceleration across many fields.

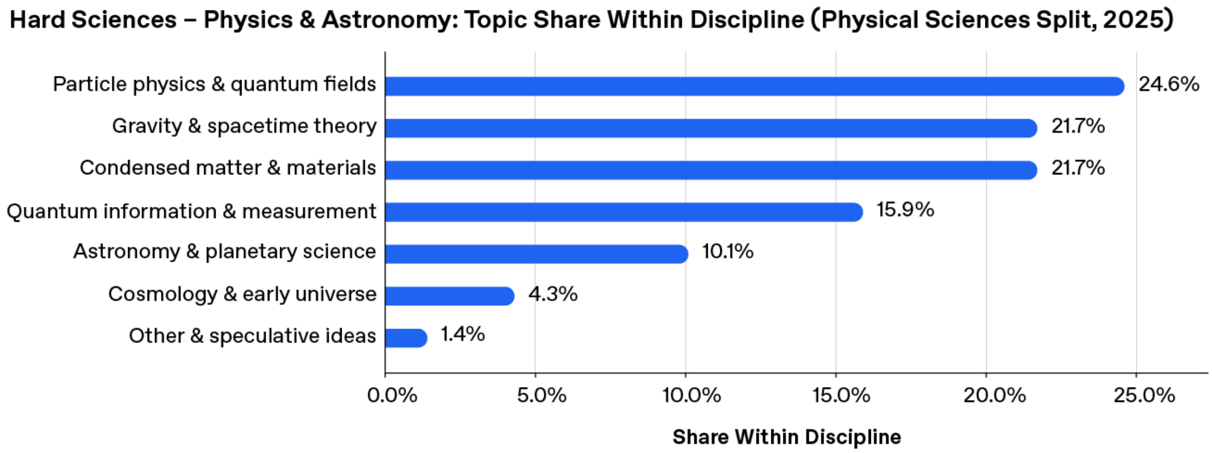
Physics

Last month, OpenAI announced a memorandum of understanding with the US Department of Energy to support collaboration on AI and advanced computing in order to advance DOE initiatives including the Genesis Mission, with applications in energy such as fusion research.

In physics, LLMs are being used across major facilities, including many US national labs, as a unifying layer over complex operations stacks and internal knowledge bases, accelerating analysis and decision-making under strict constraints alongside existing machine-learning tools for simulation, real-time data reduction, and experimental control. LLMs can digest shift logs and alerts, answer questions from internal documentation, and help route work to the right analysis, simulation, or control tool, all under strict safety, timing, and resource constraints. This augments a longer track record of specialized machine learning in physics: neural “surrogate” models that approximate equation-governed simulations when full computation is too slow, real-time filtering and reconstruction in particle detectors that see tens of millions of collisions per second, and machine-learning supported controllers that coordinate the many electromagnets in tokamak fusion experiments while staying within hardware and safety limits.

Near-term gains in physics are likely to concentrate in high-throughput, decision-heavy settings where expert attention and turnaround time are the bottlenecks. AI assistants that can reference a lab’s internal documentation and run automated checks can turn live experiment alerts, notes, and log files into prioritized next-step research plans and repeatable analytic outputs, such as notebooks, scripts, and reports.

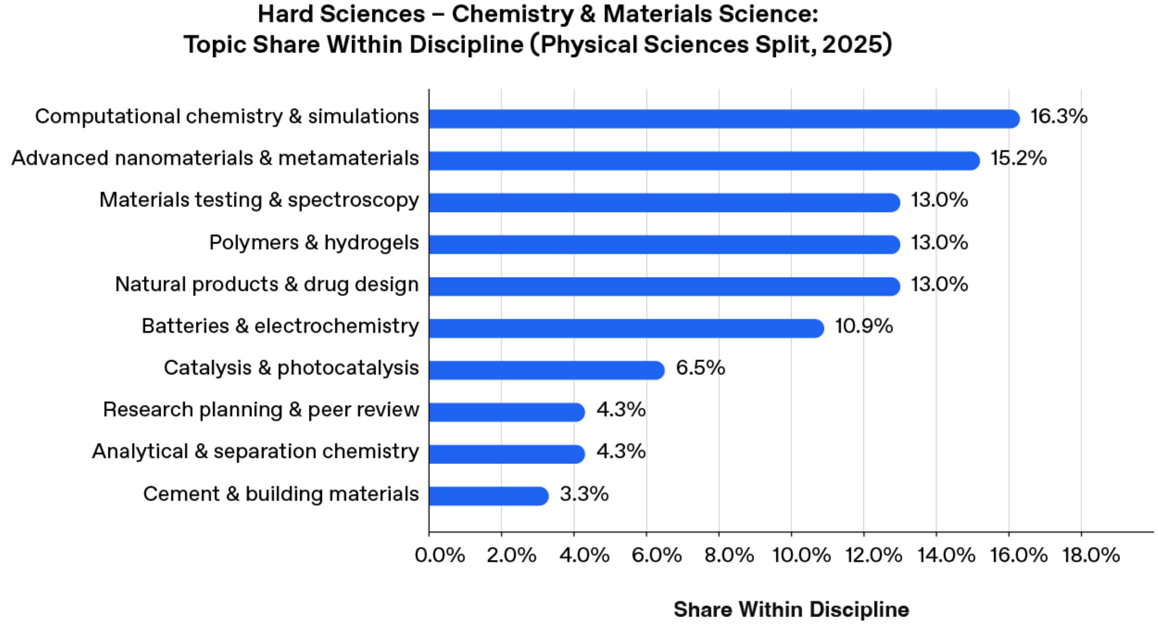
In theoretical physics, LLMs will continue to deliver value as thought partners compressing researchers' cognitive overhead while expanding the space of exploration. When a complex calculation hits a roadblock, LLMs can surface ways to reframe the problem, suggest intermediate steps, and provide quick consistency checks that reduce time that physicists spend being "stuck." Sometimes this produces insights like a missing condition, a cleaner formulation, or a useful relationship between expressions that can become a meaningful ingredient of a paper. The larger multiplier may come from synthesizing research at scale, where models scan the literature across papers and subfields to surface new connections automatically.



Chemistry

ChatGPT's applications in chemistry have moved past one-shot question answering toward multi-step workflows that translate between natural language and chemical representations, and rely on external tools for verification and retrieval. ChemBench, published in Nature Chemistry in 2025, curated more than 2,700 expert-written questions and found that leading models outperformed human chemists on average, while still struggling on some basic tasks and producing overconfident errors.

Leading AI systems in chemistry increasingly use a hybrid workflow: a general-purpose LLM helps to plan multi-step work and coordinate tools, while specialized models that understand molecular structure handle prediction and simulation. A key example is state-of-the-art graph neural networks (GNNs): these models treat a molecule like a network, with atoms as nodes and bonds as connections, so the system can learn how local changes affect the whole structure. Newer GNNs are designed so their predictions remain consistent when a molecule is rotated or shifted in 3D space, which makes them well suited for learning the energy rules needed to run fast, accurate molecular simulations. As molecular graph models scale with more data and pretraining, results continue to improve, but tougher benchmarks for chemical reasoning, including organic reaction mechanism tasks such as oMeBench, underscore the need for human oversight.

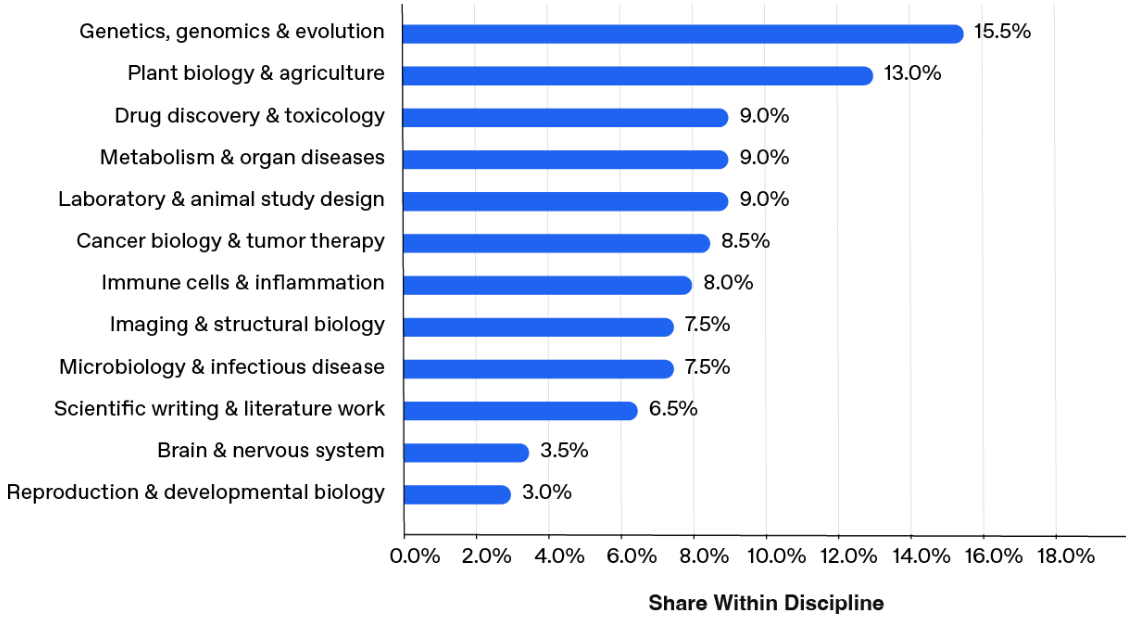


Biology

ChatGPT's applications in biology increasingly extend into multi-step workflows that combine natural-language questions with structured scientific sources such as genomics databases, protein repositories, and the biomedical literature, often with code and retrieval tools used for traceability and verification. GeneTuring, a 2025 genomics benchmark in Briefings in Bioinformatics, curated 1,600 questions across 16 task types and manually evaluated 48,000 answers from multiple model configurations. The strongest results came from a tool-augmented setup that paired a general-purpose model with direct access to National Center for Biotechnology Information (NCBI) APIs, reinforcing that reliability improves when language models are connected to authoritative reference data and can show their work.

As with chemistry, state-of-the-art AI-enabled research in biology relies on hybrid stacks: general-purpose language models help plan and coordinate analysis, while specialized foundation models trained on biological sequences and structures power prediction and design. In protein science, AlphaFold 3 represents a step toward unified biomolecular modeling by predicting the joint 3D structure of complexes that can include proteins, DNA and RNA, and small molecules within a diffusion-based architecture.

Hard Sciences – Biomedical & Life Sciences:
Topic Share Within Discipline (Physical Sciences Split, 2025)



Use case profiles

Ernest Ryu - Mathematician

Ernest Ryu picked up ChatGPT out of curiosity in 2023, and saw it advance until it could generate a publishable result last year. Ryu's academic work has focused on optimization: the math behind efficient, reliable algorithms that support modern economies, from planning logistics to keeping aircraft wings stable.

When large language models were surging in popularity in 2023, Ryu began his first experiments: could a model translate real-world "word problems" into precise optimization models, including all the hidden constraints, and then hand them to a solver? Scheduling a baseball season, for example, requires hard constraints (e.g. no team plays two games at the same time) and softer ones (e.g. travel rest days that can be violated if necessary). That early model struggled with the careful constraint handling this work demands, sometimes omitting constraints and failing on larger, realistic schedules.

The inflection point came last year, after the arrival of reasoning models and OpenAI's winning gold at the International Mathematical Olympiad. The same class of scheduling problems Ryu had tested before were reliably solved. That success led Ryu to apply LLMs to everyday mathematical work: while writing lectures, Ryu began asking ChatGPT for proofs of results he knew were true but didn't have top of mind.

Finally, he tried it on research. Ryu chose a problem related to Nesterov acceleration, a well-known technique for speeding up optimization, and picked a version of the problem that was open long enough that others had attempted it, yet simple enough that a short proof might exist. For three consecutive evenings, after his son went to bed, he worked from 8pm to midnight, and by the third night he had guided AI to the point where it cracked the problem.

Their collaboration looked like real research. The model produced an initial proof with a calculation mistake, so Ryu began to iterate: he corrected the error, preserved the correct intermediate steps in a growing prompt, abandoned dead ends, and pushed the model into other approaches. Ryu describes this as maze running, where you turn down corridors and open doors, sometimes only to find them empty, while keeping a mental map of what fails and what seems promising. ChatGPT helped Ryu accelerate how fast he ran the maze by 3x to 10x.

On the third night, the model made a small but meaningful leap that "looked different" enough to unlock the proof. Ryu said he "more than triple-checked" the argument, then had a student verify it, before he shared it publicly to an optimization community that reacted with surprise and excitement. From there, the continuous-time result was translated into the discrete-time algorithm statement with a single prompt, leaving a short, one-page novel core that met the standard for a publishable advance.

Since then, Ryu has joined OpenAI's synthetic data team, where his core focus is improving the model's mathematical capability.

Alex Lupsasca - Physicist

Alex Lupsasca came to AI the way many physicists come to bold claims: with polite skepticism and a bunch of tests. In early 2025, he tried ChatGPT, and he found it useful for the routine administrative tasks that regularly pop up in academia. But he did not see it as a tool for the hard part of the job: turning the laws of physics into concrete and verifiable predictions.

A common reality of academic publishing, Lupsasca says, is that a research project may often result in a draft equations and rough connective tissue, and a professor's work becomes that of a copy editor, especially when collaborators are writing outside their first language. ChatGPT can take a rough draft and turn it into cleaner scientific prose, saving substantial time. Despite his experience, Lupsasca says AI tools are almost always better at writing than he is.

Yet as the models improved, he kept testing for greater capabilities. What changed his mind was watching the system finally do physics at the level of a graduate problem set at speed. Lupsasca described a simple-to-state question from general relativity: starting from a well-known textbook model for electromagnetic fields near a black hole (a so-called "Wald solution"), what is the magnetic field strength right at the black hole's edge, the event horizon? A beginning graduate student might need hours to work it through, but the model produced the full derivation in seconds.

Then he noticed something stranger: the model often arrived with correct answers, but expressed them in unfamiliar mathematical language, as if it had learned multiple dialects of physics and math and could choose whichever compressed the result best.

The bigger test came from Lupsasca's own research. He had recently derived new "hidden symmetries" of an equation that governs a black hole's tidal response, roughly the black hole analogue of ocean tides raised by the moon. Those symmetries explain why a famous tidal effect vanishes. When he handed that equation to GPT-5 Pro with minimal guidance, it thought for about 18 minutes and returned the same symmetry generators he had spent years building the skills to find, and months working on directly. "I think this is just incredible, and it's clearly going to change everything that we do," he said.

He has since repeated the pattern with skeptical colleagues: at CERN, a colleague offered a homework-style problem he would normally give PhD students a week to solve, and the model produced a detailed solution in minutes; in Aspen, astrophysicist Elliot Quataert tried to trick it with a puzzle-like transient signal, and it correctly identified a magnetar (a neutron star with an extreme magnetic field), then laid out follow-up observations and a draft abstract.

That arc, moving from skepticism to contagious enthusiasm, led Lupsasca to join OpenAI. He is now pushing beyond one-off wins toward repeatable scientific acceleration: better tools for reading and explaining papers, stronger workflows than a single chat window, and training setups that embed frontier physics in the model's capabilities. His goal is to unfold the consequences of physics insights faster, so researchers spend less time stuck in algebra and more time identifying the next question worth asking, to eventually crack the biggest mysteries haunting his discipline.

RetroBioSciences - Biology

RetroBioSciences is pursuing a core longevity challenge: making cell rejuvenation practical and scalable. A key lever is cellular reprogramming using the four "OSKM" factors, which can reset aspects of cellular age, but in practice the process is slow and inefficient, especially in older cells. Traditional protein engineering can improve these factors, yet progress has historically taken years of trial-and-error, exemplified by prior "super" variants that required long, labor-intensive lab work to discover.

OpenAI partnered with RetroBio to compress that timeline by building a protein-focused foundation model, GPT-4B Micro, designed to support broad protein engineering tasks, including targeted protein editing via in-filling and 3D structure generation and protein folding. Instead of training on sequences alone, the model was mid-trained on multiple biological "modalities," including protein sequences, tokenized three-dimensional structures, co-evolutionary information from multiple sequence alignments, protein-protein interaction data, and relevant scientific text. This combination aimed to create a single model that could flex across common protein design workflows: generating sequences, performing fill-in-the-middle edits to preserve critical regions while redesigning others, and producing structure tokens that correspond to plausible 3D conformations.

In the RetroBio collaboration, the model was used in a lab-in-the-loop workflow to engineer improved OSKM factors. A prompting strategy based on evolutionary series made the system steerable: it learned patterns across related proteins and could propose new sequences conditioned on examples. In practice, the model generated thousands of candidate sequences, which were filtered to preserve essential domains and to maintain diversity, then narrowed to a few hundred variants that a typical lab could feasibly test. RetroBio synthesized DNA for these variants, delivered them into human fibroblasts via lentiviral constructs, reprogrammed cells for roughly 10–14 days, and used cell-surface markers plus sequencing to rank which variants enriched in more successfully reprogrammed populations. Early readouts showed accelerated morphology changes and improved pluripotency-marker expression, with notably high hit rates: for some targets, about half of the model-generated variants outperformed wild-type proteins, even when sequences were dramatically different from their natural counterparts.

Follow-on experiments on top individual variants showed performance comparable to, and in some cases exceeding, prior best-in-class engineered factors, with larger and more numerous stem-cell-like colonies. RetroBio also reported downstream validation steps relevant to translation, including checks that reprogrammed cells retained expected differentiation capacity and showed no obvious chromosomal abnormalities in basic safety screens. Separate assays suggested some variants improved resilience to induced DNA damage, a proxy linked to rejuvenation-related mechanisms, despite that property not being explicitly screened for during design.

All in all, the collaboration illustrates a concrete pattern for AI-enabled biology: OpenAI models help RetroBio search vast protein design spaces more efficiently, raise experimental success rates, and shorten iteration cycles from years to months, supporting the kind of real-world scientific acceleration that show how AI acts as a collaborator in research workflows.

Policy: Accelerating AI-enabled science

Below we focus on strengthening American scientific innovation with AI to ensure that 2026 is the Year of AI and Science. OpenAI recommends four pillars for sustained leadership: AI skilling, data access, modern AI infrastructure, and giving the scientific community scaled access to frontier AI systems. We described these four pillars in greater detail in a recent submission to the White House Office of Science and Technology Policy (OSTP).

Scale AI-skilling to prepare America's workforce and future US scientists. Launch a National AI Workforce Program that supports K-12 AI curricula, expands AI degree and certificate programs at community colleges and universities, and creates short-term training for mid-career workers. Congress and agencies could authorize grants for curriculum development and instructor training in AI and data science for schools in every state, alongside an AI Skills Corps or grant fund that brings free AI workshops and training to communities through libraries, job centers, and schools.

Open up data and expand open research partnerships to accelerate discovery. AI-assisted scientific research can compress decades of discovery into years, especially when paired with broad exploratory access to frontier AI tools and newly opened data. OpenAI supports establishing the National AI Research Resource as a shared platform providing academic and non-profit researchers access to large-scale compute and high-quality datasets, and urges swift passage and implementation with robust funding and governance. Agencies should identify high-value datasets and make them available in machine-readable form for AI R&D, including a centralized AI Training Data Catalog, while maintaining privacy and security with de-identification for sensitive data and a default of open access for bona fide research use.

Modernize AI infrastructure as a strategic national asset, including energy, compute, and chips. Compute is a critical instrument for AI-enabled discovery, and ensuring abundant compute on US soil depends on modern infrastructure and sufficient additions to the US power grid so American scientific progress is not constrained by energy scarcity. Federal policy should establish AI Infrastructure Hubs using authorities in the CHIPS and Science Act, designating AI innovation zones with priority support for high-capacity data centers and related energy investments, paired with streamlined permitting. Public-private partnerships can help procure capacity on cutting-edge AI systems for federally funded researchers, alongside efforts to strengthen the semiconductor supply chain needed to site, build, and manufacture AI chips in the United States.

Provide broad, scaled access to frontier AI systems and strengthen the innovation ecosystem. The federal government should establish a National Frontier AI Access Allocation that gives researchers across universities, the government's own national laboratories, and nonprofit institutions access to advanced AI systems at a scale sufficient for sustained experimentation, method development, and validation; access should be broad-based, lightweight to obtain, and designed to support open-ended exploration alongside defined research projects, treating AI usage as a shared national research resource analogous to telescope time or supercomputing hours. To reinforce diffusion into the broader economy, federal policy should expand Small Business Innovation Research (SBIR) programs to cover AI adoption and establish regional innovation incubators that pair AI technologists with local industry needs.

About OpenAI

Artificial intelligence is an innovation like electricity—it will change how we live, how we work, and how we engage with one another. OpenAI's mission is to ensure that artificial general intelligence benefits all of humanity. We're building AI to help people solve hard problems because by helping with the hard problems, AI can benefit the most people possible—through more scientific discoveries, better healthcare and education, and improved productivity. We're off to a strong start, creating freely available intelligence being used by more than 800 million people around the world, including 4 million developers. We believe AI will scale human ingenuity and drive unprecedented productivity, economic growth, and new freedoms that help people accomplish what we can't even imagine today.

Cover image created with ChatGPT Images