Material on AI for students

Daniele Coslovich

May 24, 2026

Contents

1 Attitudes

A personal attempt to classify attitudes towards AI (and technology in general):

- enthusiast
- solutionist
- pragmatist
- critic
- luddist

About the last one, check out In praise of Luddism, David Edgerton, Nature (2011):

> Two centuries on from the Luddite insurrection, David Edgerton celebrates today's most important opponents to new ideas, inventions and innovations: scientists.

Try to identify which attitude have the authors of the papers below!

2 AI in teaching

2.1 Help or Hype? Students' Engagement and Perception of Using AI to Solve Physics Problems

{2509.25642} Help or Hype? Students' Engagement and Perception of Using AI to Solve Physics Problems, Qurat-ul-Ann Mirza, N. Sanjay Rebello (2025)

With the rise of large language models such as ChatGPT, interest has grown in understanding how these tools influence learning in STEM education, including physics. This study explores how students use ChatGPT during a physics problem-solving task embedded in a formal assessment. We analyzed patterns of AI usage and their relationship to student performance. Findings indicate that students who engaged with ChatGPT generally performed better than those who did not. Particularly, students who provided more complete and contextual prompts experienced greater benefits. Further, students who demonstrated overall positive gains collectively asked more conceptual questions than those who exhibited overall negative gains. However, the presence of incorrect AI-generated responses also underscores the importance of critically evaluating AI output. These results suggest that while AI can be a valuable aid in problem solving, its effectiveness depends significantly on how students use it, reinforcing the need to incorporate structured AI-literacy into STEM education.

2.2 How undergraduate physics students use generative AI for computational modeling

{2603.06342} How undergraduate physics students use generative AI for computational modeling, Karl Henrik Fredly, Tor Ole B. Odden, Benjamin M. Zwickl (2026)

Generative artificial intelligence (genAI) is becoming increasingly prevalent and capable in physics, particularly for programming-related tasks. How, then, does genAI affect students' computational modeling? We interviewed 19 undergraduate students who had recently completed an open-ended computational assignment that encouraged the use of genAI, asking them how they used it. We then conducted a thematic analysis of these interviews using a framework for computational modeling in physics. We found that genAI significantly impacts several aspects of students' computational modeling, such as the planning, implementing, and debugging of computational models. GenAI can also help students find resources and introduce them to new computational tools. Productive use of genAI was associated with students limiting its use to small steps in the modeling process and consistently double-checking the formulas, explanations, and code it provided. We also identified challenges students faced due to an over-reliance on genAI, such as working from false model assumptions and not spending time learning the fundamentals of computational modeling, especially debugging. Finally, we discuss implications for teaching, such as the need to teach students how to use genAI productively and to urge them to plan before they code. We also highlight the continued value of low-stakes assessment and teaching assistants for teaching computational modeling, as the task remains difficult even with the introduction of genAI.

2.3 How Physics Professors Use and Frame Generative AI Tools

{2511.11317} How Physics Professors Use and Frame Generative AI Tools, Vidar Skogvoll, Tor Ole Odden (2025)

Generative AI is rapidly reshaping how physicists teach, learn, and conduct research, yet little is known about how physics faculty are responding to these changes. We interviewed 12 physics professors at a major Scandinavian research university to explore their uses and perceptions of Generative AI (GenAI) in both teaching and research. Using the theoretical framework of epistemic framing, we conducted a thematic analysis that identified 19 overlapping practices, ranging from coding and literature review to assessment and feedback. From these practices, we derived six overlapping epistemic frames through which professors make sense of GenAI: as a threat to genuine learning and assessment, a source of knowledge, a discussion partner, a text-processing tool, a coding tool, and a labor-saving device. While the latter five position GenAI as a useful tool in the physicists' toolbox, the threat frame represented an overarching concern that colored all other frames. These findings reveal how GenAI is beginning to transform what it means to be a physicist, highlighting both opportunities for innovation and challenges for academic integrity and learning.

2.4 The Boiling-Frog Problem of Physics Education

{2508.08842} The Boiling-Frog Problem of Physics Education, Gerd Kortemeyer, Phys. Teach. (2026)

It is astonishing how rapidly general-purpose AI has crossed familiar thresholds in introductory physics. Comparing outputs from successive models, GPT-5 Thinking moves far beyond the plug-and-chug tendencies seen earlier: on a classic elevator problem it works symbolically, notes when variables cancel, and verifies results; attempts to prompt novice-like behavior mainly affect tone, not method. On representation translation, the model scores 24/26 (92.3%) on TUG-Kv4.0. In a card-sorting proxy using two of my comprehensive finals (60 items), its categories reflect solution method rather than surface features. Solving those same exams, it attains 27/30 and 25/30, with most misses in ruler-based ray tracing and circuit interpretation. On epistemology, five independent CLASS runs yield 100 percent favorable, indicating a simulated expert-like stance. Framed as a ”boiling frog” problem, the paper argues for a decisive jump: retire credit-bearing unsupervised closed-response online assessments; grade process evidence; use paper, whiteboarding; shift weight to modeling, data, and authentic labs; require transparent, citable AI use; rebuild problem types; and lean on research-based instruction and peer learning. The opportunity is to foreground what AI cannot substitute for: modeling the world, arguing from evidence, and making principled approximations.

Advisory notes

Written with an LLM.

Quotes

> Do not pretend AI is absent; require disclosure. Ask students to include representative prompts and outputs in an appendix, then critique them: What is correct? What is spurious? How was the output verified or improved? And, before you ask: yes, GPT-5 Thinking assisted in polishing this opinion piece (grammar, flow, structure).

3 AI in academia

3.1 Meet the academics refusing to use generative AI

Meet the academics refusing to use generative AI, Nature (2026)

Researchers say they have their reasons for avoiding AI tools — and they're sick of arguing about it.

Quotes

> Danielle Crowley is getting tired of people telling her to use generative artificial intelligence (genAI). As a marine zoologist at Bangor University, UK, she says that she is pretty much the only PhD student in her cohort who does not use it. She has seen colleagues use genAI tools for coding and for getting the tone of e-mails right. On one occasion, she was even encouraged by a lecturer to use it to generate a conference poster.
>
> She says her colleagues are often surprised to hear she hasn't tried it and have suggested she uses it for applications such as coding. "I've had a lot of people go like 'oh but you have to use it'," she recalls. But Crowley has her reasons. She has concerns about the ethics of copyright, what she calls a lack of transparency from companies about how they're using the data, the environmental effects of AI tools and the accuracy of what genAI models spit out.
>
> She also thinks that using the tools would be counterproductive to her studies. "Coding is a skill I want to learn and develop, because it's not the thing I'm the most confident in," she says. She would rather try and do it herself, learning from her mistakes.

3.2 The indiscriminate adoption of AI threatens the foundations of academia

{2602.10165} The indiscriminate adoption of AI threatens the foundations of academia, R. Trotta (2026)

Artificial intelligence offers much promise, but its use in scientific research should be restrained so that the primary aim of academia – advancing knowledge for humans – is safeguarded.

Notes

Roberto Trotta is an physicist at SISSA that advocates for ”scientific alignment” in the use of AI in research.

Quotes

> The repercussions on the next generation of scientists might be even more severe: if the difficult, uncomfortable process of learning to reason like a scientist is replaced by a prompt to a chatbot, the critical thinking and depth of expertise that enables human scientists to supervise the output of LLM agents will be diminished. Research students might become little more than prompt engineers. Systemic changes to the role and funding of higher education and research institutions might ensue, as expensive PhD fellowships might swiftly be replaced by far cheaper API credits to run an army of AI agents on commercial platforms. Scientific literacy might wane within years, not decades.

3.3 Against the Uncritical Adoption of ’AI’ Technologies in Academia

Against the Uncritical Adoption of ‘AI’ Technologies in Academia, Guest et al. (2025)

Under the banner of progress, products have been uncritically adopted or even imposed on users — in past centuries with tobacco and combustion engines, and in the 21st with social media. For these collective blunders, we now regret our involvement or apathy as scientists, and society struggles to put the genie back in the bottle. Currently, we are similarly entangled with artificial intelligence (AI) technology. For example, soft- ware updates are rolled out seamlessly and non-consensually, Microsoft Office is bun- dled with chatbots, and we, our students, and our employers have had no say, as it is not considered a valid position to reject AI technologies in our teaching and research. This is why in June 2025, we co-authored an Open Letter calling on our employers to reverse and rethink their stance on uncritically adopting AI technologies. In this position piece, we expound on why universities must take their role seriously to a) counter the technology industry’s marketing, hype, and harm; and to b) safeguard higher education, critical thinking, expertise, academic freedom, and scientific integrity. We include pointers to relevant work to further inform our colleagues.

Quotes

> Students have always cheated. Bending and breaking the rules is human nature. And by the same token, educators are not police. We are not here to obsessively surveil our students — education is based on mutual trust. Therefore, our duty is to build mutually shared values with our students and colleagues. Especially when education is not valued, we as educators are obliged to show our students 11O. Guest et al. that they are not just here to receive a degree: education is more than qualification (Biesta 2021). It is about preparing students to become a capable and active members of society.

Cited from Compulsive Technology: Computer as Culture, Tony Solomonides, Les Levidow (1985):

> ”The culture of AI is imperialist and seeks to expand the kingdom of the machine. The AI community is well organized and well funded, and its culture fits its dreams: it has high priests, its greedy businessmen, its canny politicians. The U.S. Department of Defense is behind it all the way. And like the communists of old, AI scientists believe in their revolution; the old myths of tragic hubris don’t trouble them at all.”

3.4 The Agentification of Scientific Research: A Physicist's Perspective

{2604.14718} The Agentification of Scientific Research: A Physicist's Perspective, Xiao-Liang Qi [2026]

This article argues that the most important significance of the AI revolution, especially the rise of large language models, lies not simply in automation, but in a fundamental change in how complex information and human know-how are carried, replicated, and shared. From this perspective, AI for Science is especially important because it may transform not only the efficiency of research, but also the structure of scientific collaboration, discovery, publishing, and evaluation. The article outlines a gradual path from AI as a research tool to AI as a scientific collaborator, and discusses how AI is likely to fundamentally reshape scientific publication. It also argues that continuous learning and diversity of ideas are essential if AI is to play a meaningful role in original scientific discovery.

Advisory note

WTF

Quotes

> To understand what AI brings to research, we must review common problems currently facing scientific enquiry. While challenges vary by field, several are universal:
>
> 1. Time Costs: Understanding industry progress and learning from others' work requires immense time.
> 2. Loss of Tacit Knowledge: A vast amount of "intermediate" experience and data accumulated during research is not reflected in papers, forcing other scholars to explore from scratch.
> 3. Collaboration Limits: The scale of research collaboration is constrained by human communication costs, making large-scale and cross-disciplinary cooperation difficult.
> 4. Administrative Burden: Significant time is consumed by non-creative tasks, such as writing papers, peer review, writing grant proposals, and explaining work to others after the research is completed.

4 AI in writing

4.1 Delving into LLM-assisted writing in biomedical publications through excess vocabulary

{2406.07016} Delving into LLM-assisted writing in biomedical publications through excess vocabulary, Dmitry Kobak, Rita González-Márquez, Emőke-Ágnes Horvát, Jan Lause, Science Advances (2025)

Large language models (LLMs) like ChatGPT can generate and revise text with human-level performance. These models come with clear limitations: they can produce inaccurate information, reinforce existing biases, and be easily misused. Yet, many scientists use them for their scholarly writing. But how widespread is such LLM usage in the academic literature? To answer this question for the field of biomedical research, we present an unbiased, large-scale approach: we study vocabulary changes in over 15 million biomedical abstracts from 2010–2024 indexed by PubMed, and show how the appearance of LLMs led to an abrupt increase in the frequency of certain style words. This excess word analysis suggests that at least 13.5% of 2024 abstracts were processed with LLMs. This lower bound differed across disciplines, countries, and journals, reaching 40% for some subcorpora. We show that LLMs have had an unprecedented impact on scientific writing in biomedical research, surpassing the effect of major world events such as the Covid pandemic.

Quotes

> Our analysis is performed on the corpus level and cannot identify individual abstracts that may have been processed by an LLM. Still, the following examples from three real 2023 abstracts illustrate the LLM-style flowery language:
>
> • By **meticulously delving** into the **intricate** web connecting [. . .] and [. . .], this **comprehensive** chapter takes a deep dive into their involvement as **significant** risk factors for [. . .]. • A **comprehensive** grasp of the **intricate interplay** between [. . .] and [. . .] is **pivotal** for effective therapeutic strategies. • Initially, we **delve** into the **intricacies** of [. . .], **accentuating** its indispensability in cellular physiology, the enzymatic labyrinth governing its flux, and the **pivotal** [. . .] mechanisms.

Notes

Here is my black list of excessive jargon (mostly ML-related) prior to the advent of LLM

> opens up, paves the way, paves a new avenue, unveil, unravel, reveal, outperform, surpass, leverage, harness, showcase, unearth (!), meticulously, state-of-the-art, power, potent, intricate, unprecedented, innovative, hidden (depends on context), unique (depends on context), novel, insight, drive, control

4.2 Impact of ChatGPT on the writing style of condensed matter physicists

{2408.17325} Impact of ChatGPT on the writing style of condensed matter physicists, Shaojun Xu, Xiaohui Ye, Mengqi Zhang, Pei Wang (2024)

We apply a state-of-the-art difference-in-differences approach to estimate the impact of ChatGPT's release on the writing style of condensed matter papers on arXiv. Our analysis reveals a statistically significant improvement in the English quality of abstracts written by non-native English speakers. Importantly, this improvement remains robust even after accounting for other potential factors, confirming that it can be attributed to the release of ChatGPT. This indicates widespread adoption of the tool. Following the release of ChatGPT, there is a significant increase in the use of unique words, while the frequency of rare words decreases. Across language families, the changes in writing style are significant for authors from the Latin and Ural-Altaic groups, but not for those from the Germanic or other Indo-European groups.

> The results indicates that the release of ChatGPT does not lead to a statistically difference in writing quality between English and Germanic groups. [...] This suggests that Germanic speakers do not experience a significant improvement in their writing skills compared to English native speakers. A more probable reason could be that, since English is a sub-branch of the Germanic language and shares similarities with it, Germanic speakers may not feel a strong need to rely on ChatGPT. Similarly, in column (3), we observe that the speakers from other Indo-Euro language families do not show a significant improvement in their writing skills [...]. In contrast, for the Latin group (see column (2) of Tab. VI), we observe a significant improvement following the release of ChatGPT, compared to the English group. [...] A similar significant improvement is also noted for the Ural-Altaic group (see column (4) of Tab. VI).

4.3 Is ChatGPT Transforming Academics' Writing Style?

{2404.08627} Is ChatGPT Transforming Academics' Writing Style?, M. Geng, R. Trotta (2024)

Based on one million arXiv papers submitted from May 2018 to January 2024, we assess the textual density of ChatGPT's writing style in their abstracts through a statistical analysis of word frequency changes. Our model is calibrated and validated on a mixture of real abstracts and ChatGPT-modified abstracts (simulated data) after a careful noise analysis. The words used for estimation are not fixed but adaptive, including those with decreasing frequency. We find that large language models (LLMs), represented by ChatGPT, are having an increasing impact on arXiv abstracts, especially in the field of computer science, where the fraction of LLM-style abstracts is estimated to be approximately 35%, if we take the responses of GPT-3.5 to one simple prompt, "revise the following sentences", as a baseline. We conclude with an analysis of both positive and negative aspects of the penetration of LLMs into academics' writing style.

Quotes

> How could the frequencies of words like "significant" grow significantly together? Another striking example is the frequency change of the words "are" and "is". The counts in 10,000 abstracts of these two words were quite stable before 2023. However, the frequency of these two terms has dropped by more than 10% in 2023.
>
> These examples, anecdotal as they are, may represent the tip of the iceberg of a wider and growing phenomenon: the rapid increase in the usage of ChatGPT or other LLMs. The rise and fall in frequency of specific technical nouns may well be related to the changing popularity of certain research topics, but that a research trend is responsible for the change in usage of adjectives appears implausible – even less so for words like "is" and "are".

5 AI in publishing

5.1 Scientific production in the era of Large Language Models

{2601.13187} Scientific production in the era of Large Language Models, Keigo Kusumegi, Xinyu Yang, Paul Ginsparg, Mathijs de Vaan, Toby Stuart, Yian Yin, Science (2025)

Large Language Models (LLMs) are rapidly reshaping scientific research. We analyze these changes in multiple, large-scale datasets with 2.1M preprints, 28K peer review reports, and 246M online accesses to scientific documents. We find: 1) scientists adopting LLMs to draft manuscripts demonstrate a large increase in paper production, ranging from 23.7-89.3% depending on scientific field and author background, 2) LLM use has reversed the relationship between writing complexity and paper quality, leading to an influx of manuscripts that are linguistically complex but substantively underwhelming, and 3) LLM adopters access and cite more diverse prior work, including books and younger, less-cited documents. These findings highlight a stunning shift in scientific production that will likely require a change in how journals, funding agencies, and tenure committees evaluate scientific works.

Notes

Thank you to all those who publish x1.5 or x2 more papers that I will *not* have time to read.

Quotes

> For peer reviewers and journal editors, this represents a significant issue. As a shortcut to (imperfectly) screen scientific research, writing characteristics are fast becoming uninformative signals just as the quantity of scientific communication surges. As traditional heuristics break down, editors and reviewers may increasingly rely on status markers such as author pedigree and institutional affiliation as signals of quality, ironically counteracting LLMs' democratizing effects on scientific production. One potential response is to leverage the same technology to assist in evaluating manuscripts. Specialized "reviewer agents" could flag methodological inconsistencies, verify claims, and even assess novelty. Whether this scalable approach will help editors and reviewers focus on substance over surface-level signals or introduce new and unforeseen challenges to the scientific process is a critical uncertainty.

> Second, in non-LLM-assisted papers across all three repositories, writing complexity is positively associated with manuscript quality as approximated by the probability of publication in a peer-reviewed venue (logistic regressions, Fig. 3E-G). These results confirm prior research showing a positive association between writing complexity and scientific merit (32). Third, and critically, we find a reversal in the relationship between writing complexity and peer-review outcomes for LLM-assisted manuscripts. For these documents, increases in writing complexity are associated with lower peer assessments of scientific merit (Fig. 3E-G).

5.2 Publish and Perish: How AI-Accelerated Writing Without Proportional Verification Investment Degrades Scientific Knowledge

{2604.05714} Publish and Perish: How AI-Accelerated Writing Without Proportional Verification Investment Degrades Scientific Knowledge, Seok Joon Kwon (2026)

Artificial intelligence tools are accelerating manuscript production far faster than peer review capacity can expand. Applying the theory of constraints from manufacturing science, we formalize this asymmetry through a minimal two-variable ordinary differential equation model coupling review queue evolution and verification quality degradation via an endogenous, queue-pressure-driven review AI adoption mechanism. The causal chain is: writing AI adoption increases submissions, growing the review queue, which drives reviewer AI adoption under pressure, degrading verification quality and reducing net knowledge output. Under empirically informed parameters (writing acceleration $\{\gamma\} = 2.0$, review acceleration $\{\delta\} = 0.5$), the model predicts a deceptive honeymoon where knowledge output peaks at 1.10K0 (circa 2026), followed by paradox onset at t = 6 years (2028) and long-term degradation to 0.68K0 (32% loss), approaching a steady state of 0.60K0 (40% loss). The critical condition for net benefit is $\{\delta\}$ ¿ $\{\gamma\}$; the current operating point lies deep in the paradox regime. Empirical validation against NeurIPS, ICLR, arXiv, and bioRxiv submission data shows qualitative consistency with observed post-ChatGPT acceleration patterns. Policy analysis reveals that only combined interventions such as review infrastructure investment paired with institutional quality standards can restore positive knowledge production.

Quotes

> Goldratt's theory of constraints demonstrates that optimizing a non-bottleneck is not merely wasteful but destructive. Our minimal model shows that AI-accelerated writing, unmatched by proportional investment in review infrastructure, produces a paradoxical decrease in verified knowledge output through a clear causal mechanism: queue pressure drives reviewer AI adoption, which degrades verification quality faster than throughput increases.

Notes

From the author of the paper:

> Hi Daniele, Thanks for reading my paper. Hope all of us can go through this unprecedent hard time for academia. cheers, sjk

5.3 LLM hallucinations in the wild: Large-scale evidence from non-existent citations

{2605.07723} LLM hallucinations in the wild: Large-scale evidence from non-existent citations, Zhenyue Zhao, Yihe Wang, Toby Stuart, Mathijs De Vaan, Paul Ginsparg, Yian Yin (2026)

Large language models (LLMs) are known to generate plausible but false information across a wide range of contexts, yet the real-world magnitude and consequences of this hallucination problem remain poorly understood. Here we leverage a uniquely verifiable object - scientific citations - to audit 111 million references across 2.5 million papers in arXiv, bioRxiv, SSRN, and PubMed Central. We find a sharp rise in non-existent references following widespread LLM adoption, with a conservative estimate of 146,932 hallucinated citations in 2025 alone. These errors are diffusely embedded across many papers but especially pronounced in fields with rapid AI uptake, in manuscripts with linguistic signatures of AI-assisted writing, and among small and early-career author teams. At the same time, hallucinated references disproportionately assign credit to already prominent and male scholars, suggesting that LLM-generated errors may reinforce existing inequities in scientific recognition. Preprint moderation and journal publication processes capture only a fraction of these errors, suggesting that the spread of hallucinated content has outpaced existing safeguards. Together, these findings demonstrate that LLM hallucinations are infiltrating knowledge production at scale, threatening both the reliability and equity of future scientific discovery as human and AI systems draw on the existing literature.

Quotes

> Once published, hallucinated citations diffuse through the bibliographic sources that researchers and AI systems treat as factual. When authors or automated tools attempt to verify a reference, they consult scholarly search engines and citation databases. As non-existent references surface in these systems as standalone bibliographic entries, in addition to appearing in the full text of citing papers, they become part of the permanent record. Cross-validating unmatched references against Google Scholar (SI S2.11), we do indeed find a growing number of entries that cannot be linked to any real publication yet already appear as references in other papers (Fig. 3c). The pattern replicates across all four corpora (Fig. S10). The accumulation of hallucinated citations within bibliometric databases may begin to erode the mechanisms we have to detect them.

6 AI in research funding

6.1 Scientific Text Analysis with Robots applied to observatory proposals

{2407.02992} Scientific Text Analysis with Robots applied to observatory proposals, T. Jerabkova, H.M.J. Boffin, F. Patat, D. Dorigo, F. Sogni, F. Primas (2024)

To test the potential disruptive effect of Artificial Intelligence (AI) transformers (e.g., ChatGPT) and their associated Large Language Models on the time allocation process, both in proposal reviewing and grading, an experiment has been set-up at ESO for the P112 Call for Proposals. The experiment aims at raising awareness in the ESO community and build valuable knowledge by identifying what future steps ESO and other observatories might need to take to stay up to date with current technologies. We present here the results of the experiment, which may further be used to inform decision-makers regarding the use of AI in the proposal review process. We find that the ChatGPT-adjusted proposals tend to receive lower grades compared to the original proposals. Moreover, ChatGPT 3.5 can generally not be trusted in providing correct scientific references, while the most recent version makes a better, but far from perfect, job. We also studied how ChatGPT deals with assessing proposals. It does an apparent remarkable job at providing a summary of ESO proposals, although it doesn't do so good to identify weaknesses. When looking at how it evaluates proposals, however, it appears that ChatGPT systematically gives a higher mark than humans, and tends to prefer proposals written by itself.

Quotes

> Most importantly, one should not forget that these models will generally provide non-substantiated statements, and when asked to provide scientific references they may sometimes just make them up. In one case, we asked Gemini to provide a reference about the importance of binary stars in the formation and evolution of planetary nebulae. It correctly mentioned Boffin & Jones (2019)[9], but with a completely wrong title. When prompted to provide the full reference, it gave a completely different – and still wrong – title, and a totally wrong reference. When asked why it did that, its excuse was that the original reference (a book) was not readily accessible, so it decided to provide a link to a paper to which it had access. It is doubtful that such an argument would convince a proposal reviewer.

> Figure 3 clearly indicates that ChatGPT reviewed the original and ChatGPT-adjusted proposals almost the same (with no statistically significant difference as indicated in the top panel of Fig. 4), although there is a slight indication that it would favor ChatGPT-adjusted proposals. Thus, although humans may not seem to see any improvements in the proposals, the tool seems to think it did a good job!

6.2 The Rise of Large Language Models and the Direction and Impact of US Federal Research Funding

{2601.15485} The Rise of Large Language Models and the Direction and Impact of US Federal Research Funding, Yifan Qian, Zhe Wen, Alexander C. Furnas, Yue Bai, Erzhuo Shao, Dashun Wang (2026)

Federal research funding shapes the direction, diversity, and impact of the US scientific enterprise. Large language models (LLMs) are rapidly diffusing into scientific practice, holding substantial promise while raising widespread concerns. Despite growing attention to AI use in scientific writing and evaluation, little is known about how the rise of LLMs is reshaping the public funding landscape. Here, we examine LLM involvement at key stages of the federal funding pipeline by combining two complementary data sources: confidential National Science Foundation (NSF) and National Institutes of Health (NIH) proposal submissions from two large US R1 universities, including funded, unfunded, and pending proposals, and the full population of publicly released NSF and NIH awards. We find that LLM use rises sharply beginning in 2023 and exhibits a bimodal distribution, indicating a clear split between minimal and substantive use. Across both private submissions and public awards, higher LLM involvement is consistently associated with lower semantic distinctiveness, positioning projects closer to recently funded work within the same agency. The consequences of this shift are agency-dependent. LLM use is positively associated with proposal success and higher subsequent publication output at NIH, whereas no comparable associations are observed at NSF. Notably, the productivity gains at NIH are concentrated in non-hit papers rather than the most highly cited work. Together, these findings provide large-scale evidence that the rise of LLMs is reshaping how scientific ideas are positioned, selected, and translated into publicly funded research, with implications for portfolio governance, research diversity, and the long-run impact of science.

Quotes

> Overall, these results indicate that the consequences of LLM adoption for proposal selection are agency-dependent: while LLM use does not systematically advantage proposals at NSF, it is associated with a higher likelihood of being funded at NIH, even for the same investigators.

7 AI in physics and maths

7.1 What Has a Foundation Model Found? Using Inductive Bias to Probe for World Model

{2507.06952} What Has a Foundation Model Found? Using Inductive Bias to Probe for World Models, Keyon Vafa, Peter G. Chang, Ashesh Rambachan, Sendhil Mullainathan (2025)

Foundation models are premised on the idea that sequence prediction can uncover deeper domain understanding, much like how Kepler's predictions of planetary motion later led to the discovery of Newtonian mechanics. However, evaluating whether these models truly capture deeper structure remains a challenge. We develop a technique for evaluating foundation models that examines how they adapt to synthetic datasets generated from some postulated world model. Our technique measures whether the foundation model's inductive bias aligns with the world model, and so we refer to it as an inductive bias probe. Across multiple domains, we find that foundation models can excel at their training tasks yet fail to develop inductive biases towards the underlying world model when adapted to new tasks. We particularly find that foundation models trained on orbital trajectories consistently fail to apply Newtonian mechanics when adapted to new physics tasks. Further analysis reveals that these models behave as if they develop task-specific heuristics that fail to generalize.

Quotes

> Taken together, our results provide a direction for understanding the deficiencies of foundation models: if a model's inductive bias isn't toward a known model of reality, what is it toward? We explore this question by examining whether these foundation models have alternative inductive biases. Our analysis reveals that these models instead behave as if they develop task-specific heuristics that fail to generalize. For physics, rather than learning one universal physical law, the foundation model applies different, seemingly nonsensical laws depending on the task it's being applied to.

Notes

Check out this amazing precursor: ”Se Simplicio avesse avuto un Cray” by Preparata (in Italian).

7.2 Mathematics: the Rise of the Machines

{2511.17203} Mathematics: the Rise of the Machines, Yang-Hui He (2025)

We argue how AI can assist mathematics in three ways: theorem-proving, conjecture formulation, and language processing. Inspired by initial experiments in geometry and theoretical physics in 2017, we summarize how this emerging field has grown over the past years, and show how various machine-learning algorithms can help with pattern detection across disciplines in the mathematical sciences. At the heart is the question how does AI help with theoretical discovery, and the implications for the future of mathematics.

Advisory note

WTF

Quotes

> Even in a hypothetical distant future where complete mechanization is achieved, where AI will (dis-)prove every major conjecture and propose new problems and continuously map out new areas of research, human mathematicians are still of great value. We will simply become priests to oracles, and interpret the results to the rest of humanity. Think of the philosophy departments in the world, centuries are spent in analyzing and critiquing Plato, or the literature departments, over Shakespeare. Perhaps one day in the far future, mathematics departments will consist of experts digesting the (Mathlib-verified) proofs that AI produces. Still, I think a number of my colleagues agree with the exclamation [HLF] that "I don't care whether it is God, AI, or Terry Tao who finds the proof of the Riemann Hypothesis, I just want to know."

7.3 Perfect score on IPhO 2025 theory by Gemini agent

{2603.03352} Perfect score on IPhO 2025 theory by Gemini agent, Yichen Huang (2026)

The International Physics Olympiad (IPhO) is the world's most prestigious and renowned physics competition for pre-university students. IPhO problems require complex reasoning based on deep understanding of physical principles in a standard general physics curriculum. On IPhO 2025 theory problems, while gold medal performance by AI models was reported previously, it falls behind the best human contestant. Here we build a simple agent with Gemini 3.1 Pro Preview. We run it five times and it achieved a perfect score every time. However, data contamination could occur because Gemini 3.1 Pro Preview was released after the competition.

Quotes

> Gemini 3.1 Pro Preview was released after IPhO 2025. Thus, IPhO 2025 problems could occur in the training dataset of the model, and the perfect performance by our Gemini agent should be interpreted with caution. However, it should be clear that our result is still meaningful. The second highest reported result 87.7% was achieved by Gemini 3 Deep Think [undefm]. In the release note [undefu], the Gemini team indicates that Gemini 3.1 Pro is the "upgraded core intelligence that makes [the Gemini 3 Deep Think] breakthroughs possible." Thus, the risk of data contamination for Gemini 3 Deep Think is at the same level as that for our agent built on Gemini 3.1 Pro Preview.

8 AI in society

8.1 Generative artificial intelligence reduces social welfare through model collapse

{2604.21853} Generative artificial intelligence reduces social welfare through model collapse, Fabian Baumann, Erol Akçay, Joshua B. Plotkin (2026)

Generative artificial intelligence (genAI) is rapidly reshaping how knowledge and culture are produced and consumed. Yet generative models are vulnerable to model collapse: when trained on data generated by earlier versions of themselves, their outputs can lose diversity and accuracy. This creates a social dilemma, because delegating tasks to genAI can be individually beneficial in the short term even as widespread adoption degrades future model performance. Here we develop a parsimonious model of behavior in collaborative interactions in which individuals can either exert human effort, rely on genAI, or refrain from work altogether. The welfare consequences of genAI are organized by a simple two-dimensional taxonomy: the strength of the incentive to perform the task without AI, and the severity of model collapse. Within this framework, the introduction of genAI – while initially beneficial at the individual level – will reduce social welfare for the most important types of tasks. In addition, habit formation around genAI use can couple otherwise separate domains, so that adoption in low-stakes tasks spills over into high-value tasks and amplifies welfare losses. Together, these results identify a general pathway by which, in the absence of intervention, individually rational adoption of genAI will assuredly and profoundly reduce collective welfare.

Notes

Check out the scenario of strong model collapse Ib in fig.2b!

Quotes

> Although our model is intentionally stylized, its qualitative implication is robust: when present-day AI use degrades future performance, individually rational adoption can undermine collective welfare precisely in the domains where human contribution is most valuable.

8.2 Is Artificial Intelligence the great filter that makes advanced technical civilisations rare in the universe?

{2405.00042} Is Artificial Intelligence the great filter that makes advanced technical civilisations rare in the universe?, Michael Garrett

This study examines the hypothesis that the rapid development of Artificial Intelligence (AI), culminating in the emergence of Artificial Superintelligence (ASI), could act as a ”Great Filter” that is responsible for the scarcity of advanced technological civilisations in the universe. It is proposed that such a filter emerges before these civilisations can develop a stable, multiplanetary existence, suggesting the typical longevity (L) of a technical civilization is less than 200 years. Such estimates for L, when applied to optimistic versions of the Drake equation, are consistent with the null results obtained by recent SETI surveys, and other efforts to detect various technosignatures across the electromagnetic spectrum. Through the lens of SETI, we reflect on humanity's current technological trajectory - the modest projections for L suggested here, underscore the critical need to quickly establish regulatory frameworks for AI development on Earth and the advancement of a multiplanetary society to mitigate against such existential threats. The persistence of intelligent and conscious life in the universe could hinge on the timely and effective implementation of such international regulatory measures and technological endeavours.

Notes

Check out Stanislaw Lem's book ”Fiasco” written in 1986 (in Italian: Il pianeta del silenzio), in particular the (fictional) Ortega-Nilssen hypothesis about the evolution of civilizations.