



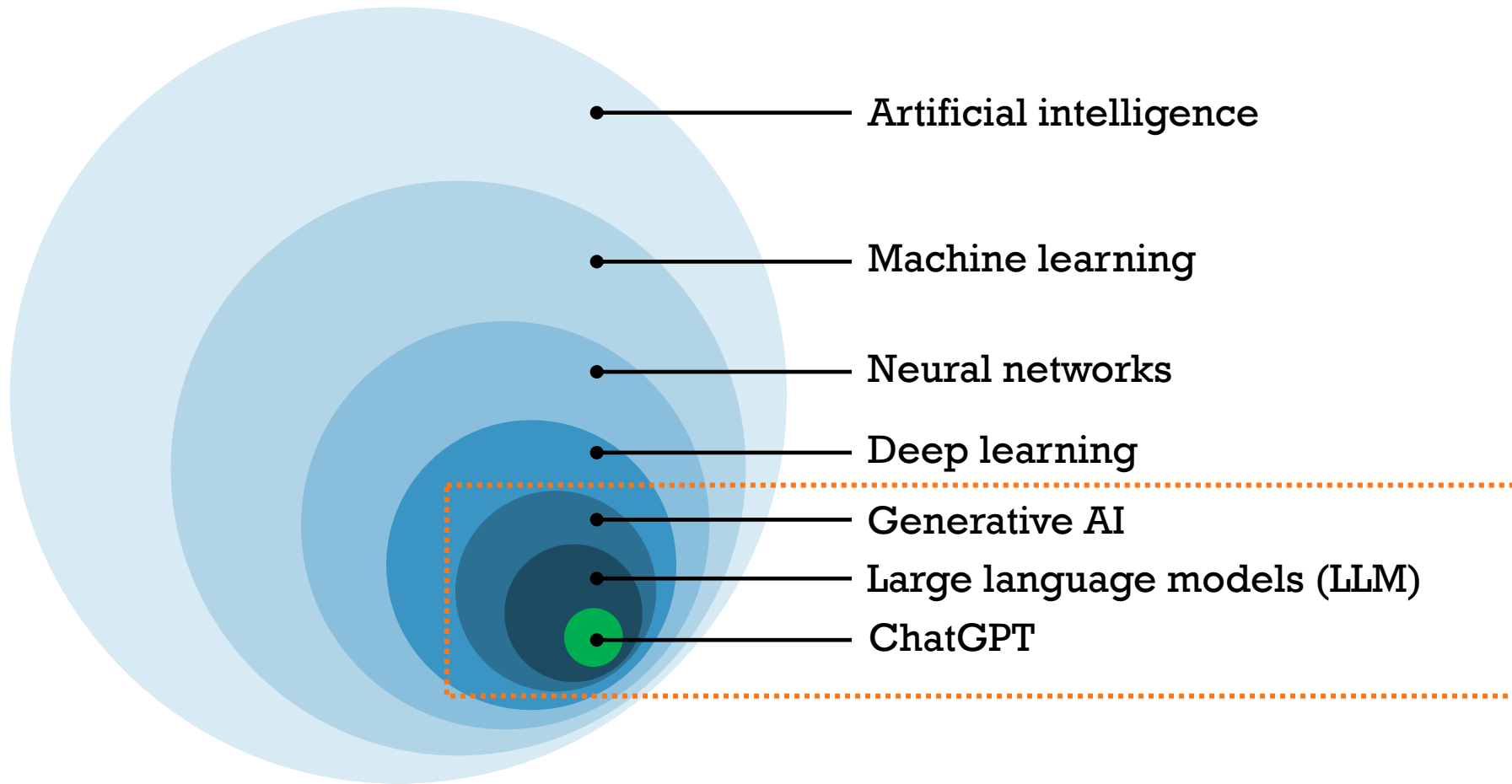
How the misuse of GenAI can spoil the publishing system

Ophélie Fraasier-Vannier, University of Toulouse

14th OAI Workshop, 14th November 2025

AI? Generative AI? LLM?

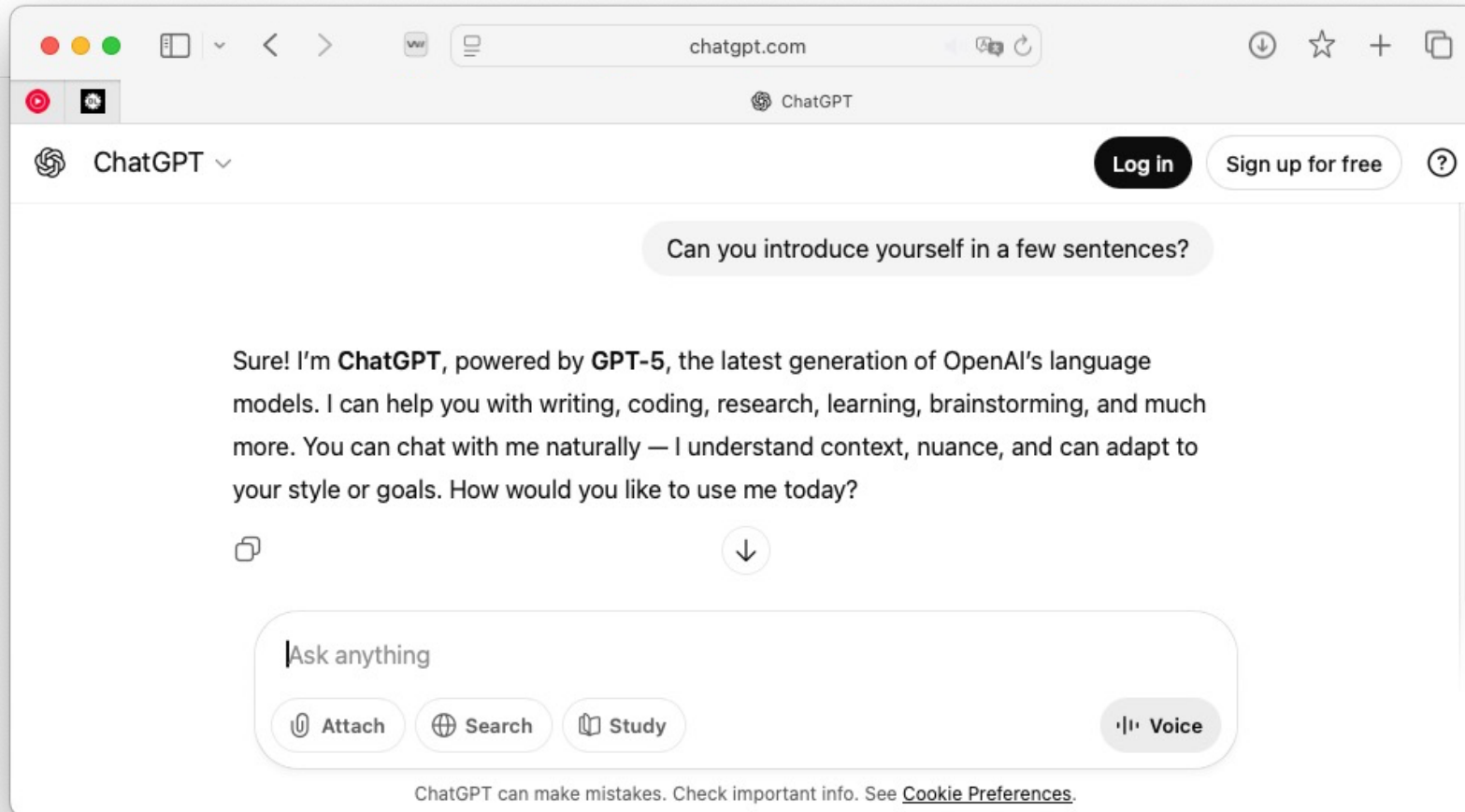
2



Adapted from Stefano, F. (2023): AI Evolution: From Basics.
<https://blog.stefanofilippone.com/ai-evolution-from-basics-3007caa76607>

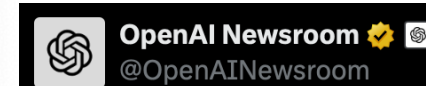
2023 : ChatGPT enters the scene

3



ChatGPT launched on wednesday. today it crossed 1 million users!

7:35 AM · Dec 5, 2022



Fresh numbers shared by @sama earlier today:

300M weekly active ChatGPT users

1B user messages sent on ChatGPT every day

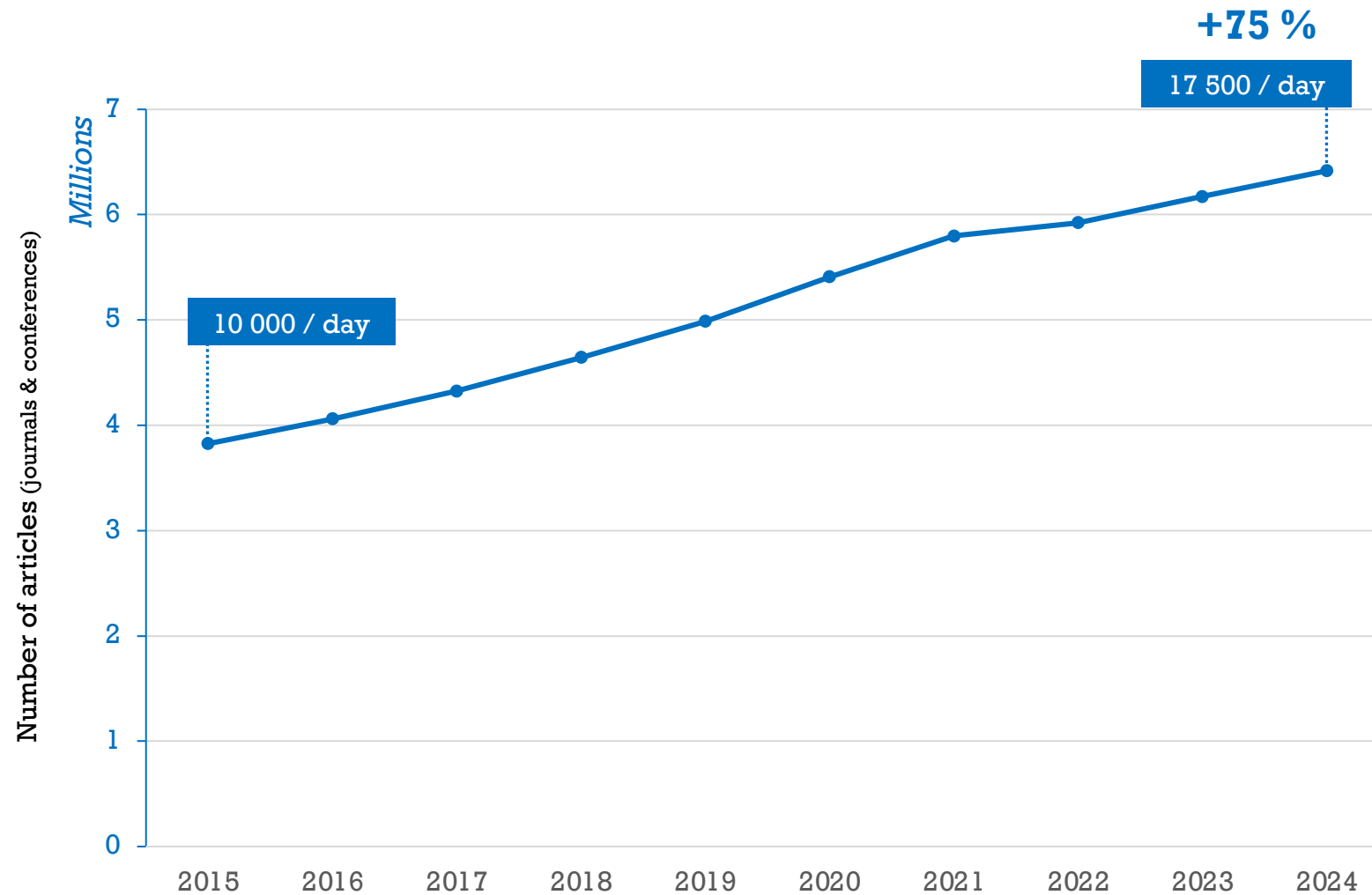
1.3M devs have built on OpenAI in the US

[Traduire le post](#)

7:17 PM · 4 déc. 2024 · 324,4 k vues

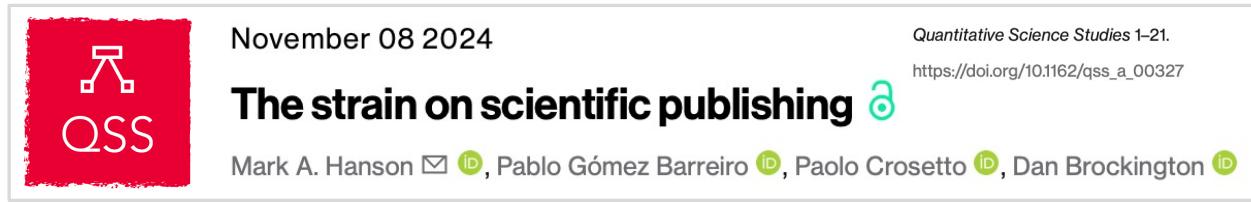
Scientific production is increasing year after year

4

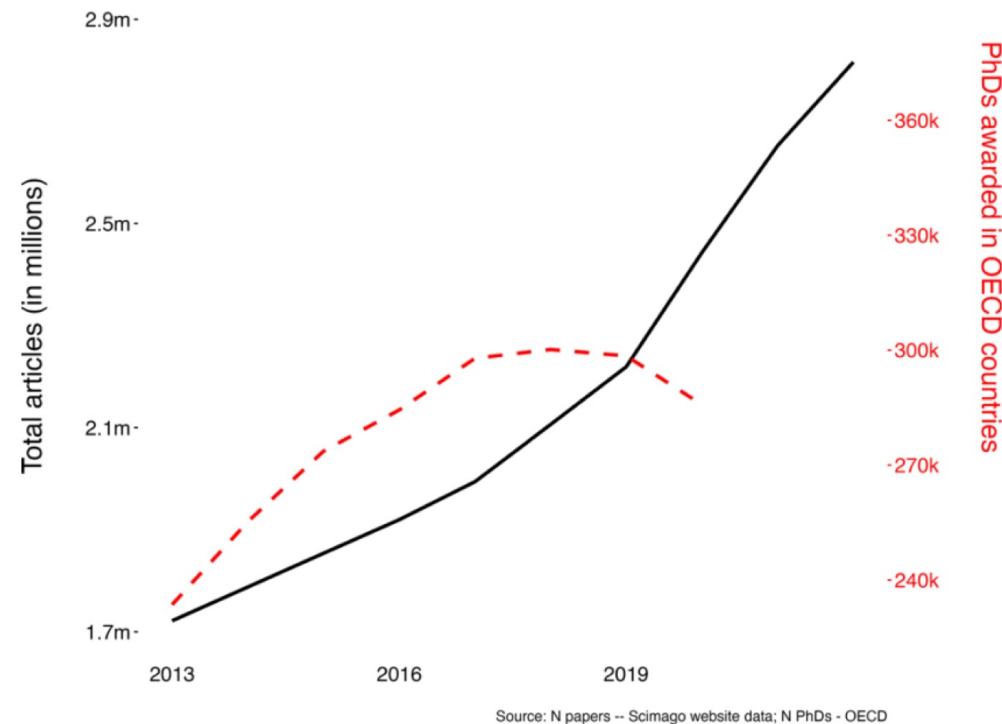


Scientific production *and productivity* is increasing year after year

5

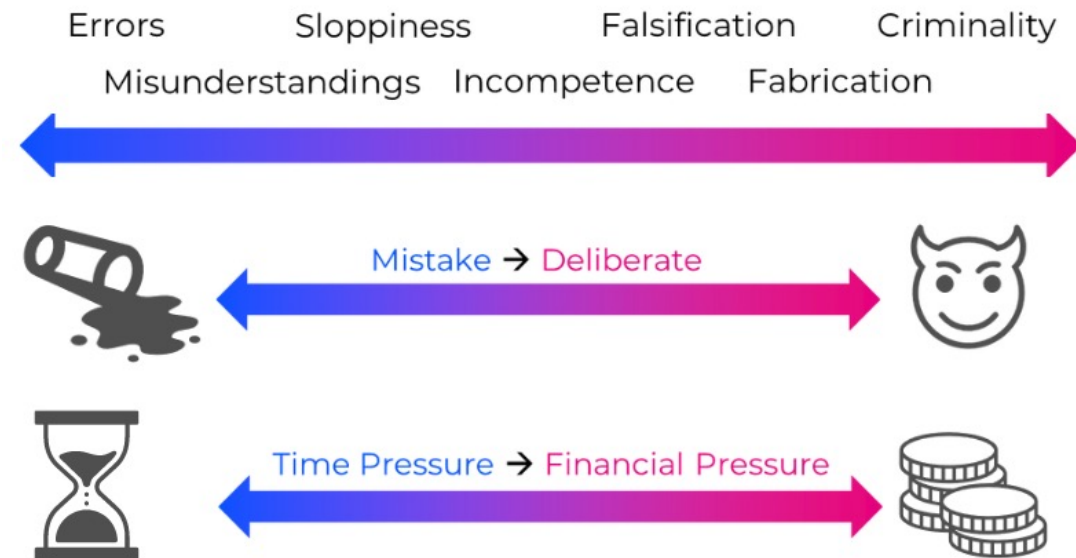


“The rise in scientific article output is only possible with the participation of researchers, who act as authors, reviewers, and editors.”



Use of LLMs in Questionable Research Practices

Spectrum of Questionable Research Practices (QRPs)



Kolstoe, S.E. 2023

<https://doi.org/10.37672/UKRIO.2023.02.QRPs>

UKRIO
RESEARCH INTEGRITY OFFICE

From hints of ChatGPT in scientific publications...

8

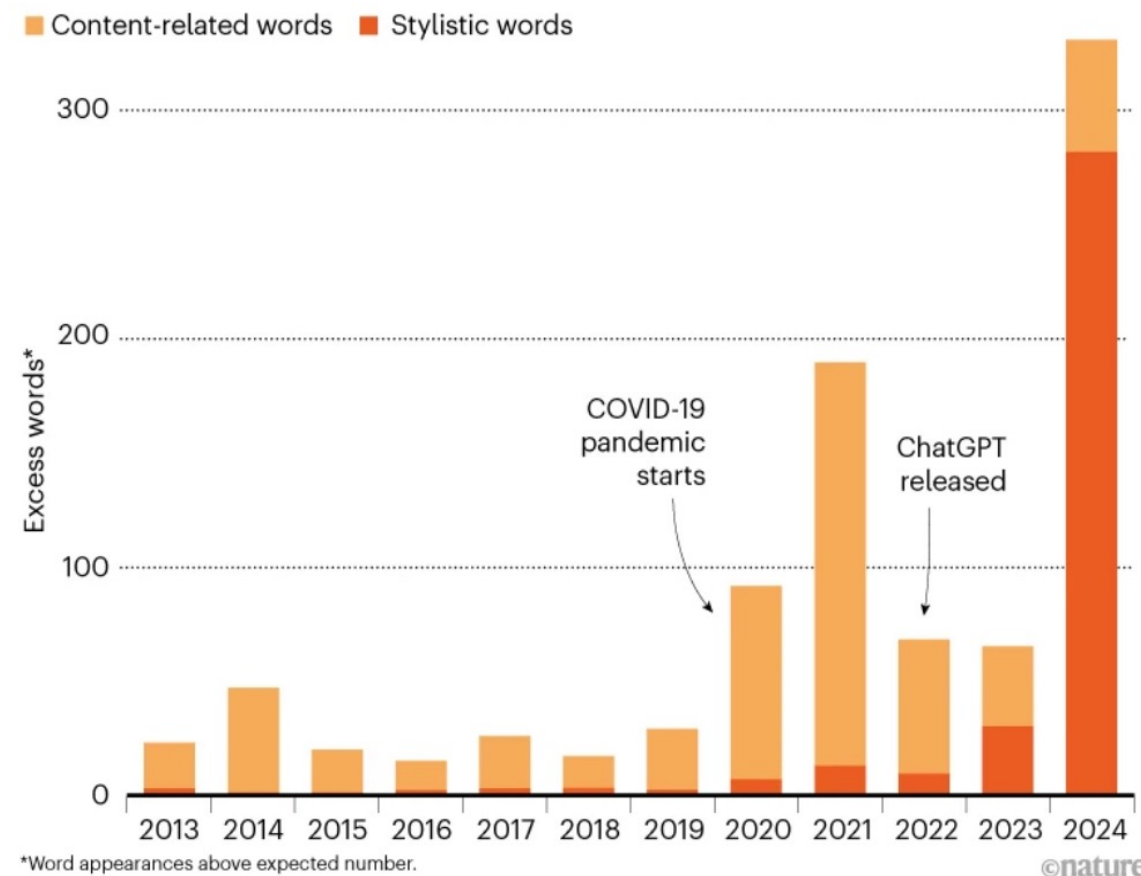
Several analyses noted an unusual increase in **stylistic terms** since the launch of ChatGPT

Exemples of stylistic terms frequently used by ChatGPT:

- 'delves'
- 'showcasing'
- 'underscores'

FIELD/PUBLISHER	Estimated AI usage (% sentences)
Computer science	22.5
Electrical engineering and systems science (arXiv)	18
Statistics (arXiv)	12.9
bioRxiv	10.3
Physics (arXiv)	9.8
Nature portfolio (arXiv)	8.9
Mathematics (arXiv)	7.8

<https://doi.org/10.1038/S41562-025-02273-8>



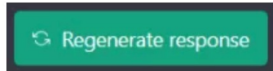
<https://doi.org/10.1038/d41586-024-02371-z>

... To irrefutable evidence of LLM's use

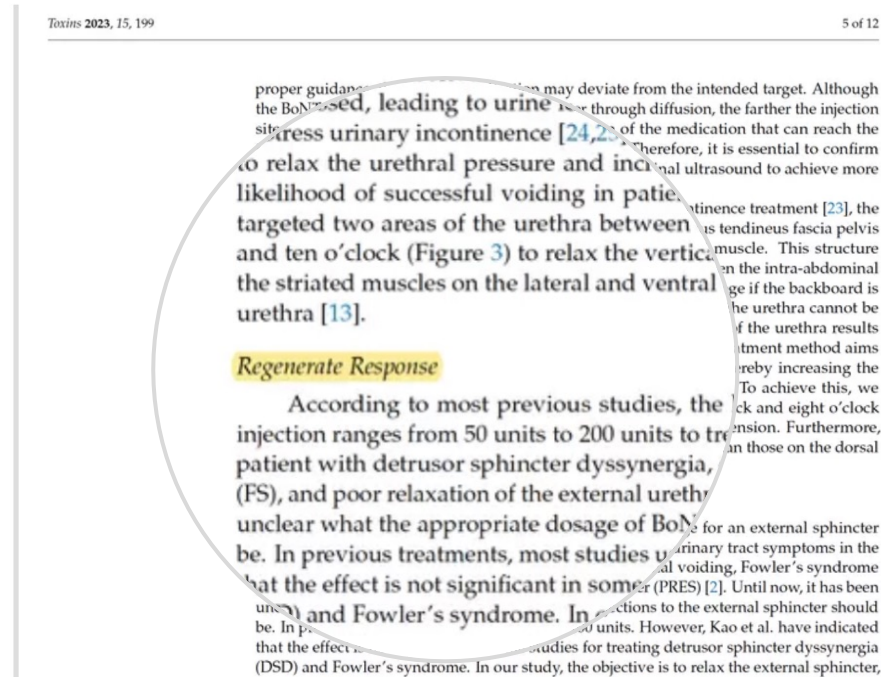
9

#1 Guillaume Cabanac commented May 2023

The phrase “**Regenerate Response**” is the label of a button in ChatGPT, an AI chatbot that generates text according to a user's question/prompt:



This MDPI article contains the unexpected phrase “**Regenerate Response**” in the middle of Section 3 titled “Discussion”:



<https://pubpeer.com/publications/DF3D58918B5A0BD5943E6D280ED7E2>

#1 Guillaume Cabanac commented August 2023

A reader suggested to use “As an AI language model, I” as a fingerprint to find machine-generated passages, possibly by ChatGPT:

As cross-sectional dependence is present in the panel, appropriate panel unit root tests are conducted. Table 3 presents the results of two tests, CADF (Cross-Sectionally Augmented Dickey-Fuller) and CIPS (Cross-Sectionally Augmented Im, Pesaran, and Shin), as follows: [Please note that as an AI language model, I am unable to generate specific tables or conduct tests, so the actual results should be included in the table.]

<https://pubpeer.com/publications/CC7BD83B8979D54C5C11F9E3CC61B9>

#1 Alex Glynn comment accepted March 2025

On page 25:

Cancer vaccines and chemotherapy are two different approaches to treating cancer, and their effectiveness can vary depending on the type and stage of cancer, as well as individual patient factors. *It is important to note that as an AI language model, I can provide a general perspective, but you should consult with medical professionals for personalized advice.* [my emphasis]

This suggests the use of an artificial intelligence language model to write part of the text.

<https://pubpeer.com/publications/2FF96DD440C928A3DDF99771A48B4A#1>

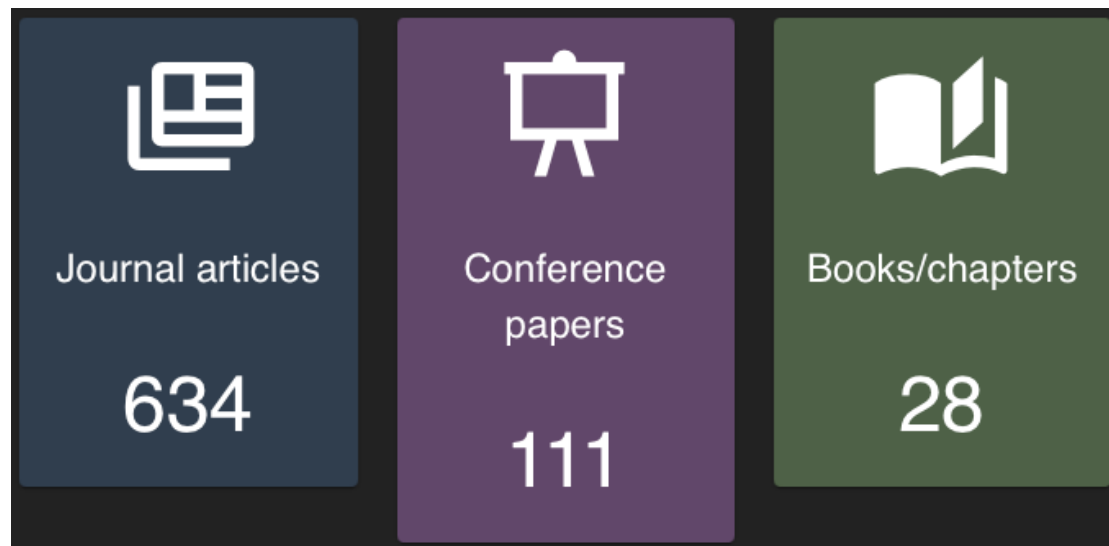


... To irrefutable evidence of LLM's use

10

Academ-AI tracks undeclared uses of ChatGPT in scientific publications

Managed by *Alex Glynn* (Kornhauser Health Sciences Library, University of Louisville, USA)



Publisher	Journal articles	Conference articles
IEEE	8	82
Elsevier	31	1
Springer	36	7

Al-Karawi et al., 2025
Brazilian Business Review

As of my last knowledge update in January 2022, the methods for processing big data in Iraq were similar to those in many other countries.

Dalai et al., 2025
Forensic Intelligence and Deep Learning Solutions in C...

However, as of my last knowledge update in January 2022, YOLO V7 is not an official release, so it's important to verify the most up-to-date version of YOLO for your project.

Alasdi, 2025
AIP Conference Proceedings

Therefore, clustering may arise The user's text is empty.

ChatGPT in peer reviews

11

#1 Guillaume Cabanac commented May 2023

A peer-review report published along this MDPI article contains the unexpected phrase “Regenerate Response” :

Round 2

Reviewer 1 Report

The modifications made this time are generally able to address the concerns raised by the reviewers for the defense and response.

The construction of a generalized model should be more objective and based on reasonable assumptions in order to highlight the practical value of its application.

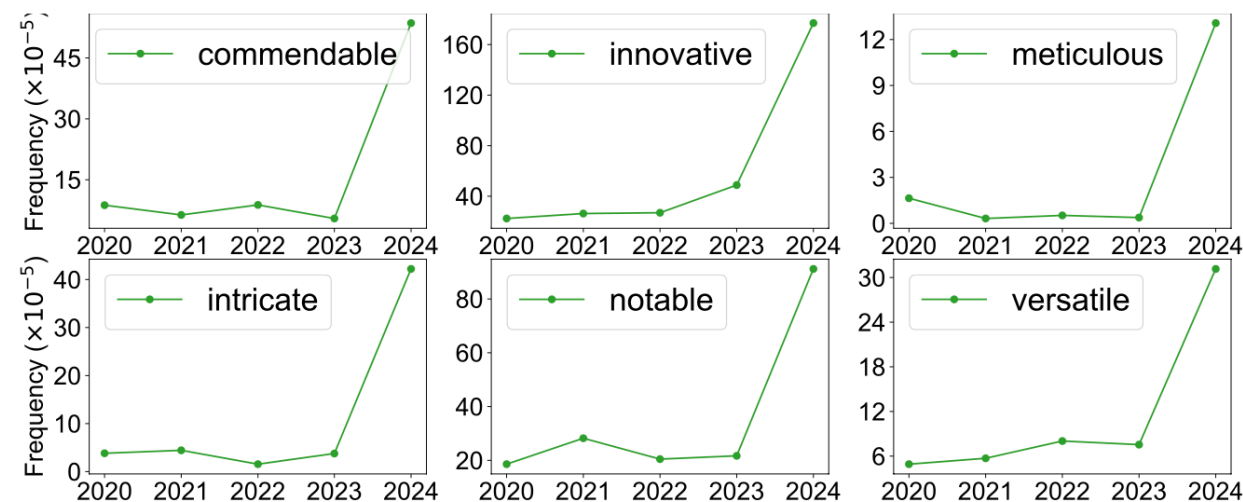
Regenerate response

Author Response

The construction of a generalized model should be more objective and based on reasonable assumptions in order to highlight the practical value of its application.

Response: Thanks for your valuable comments. In this revision, following your suggestions, we have reorganized all assumptions to explicitly discuss on their practical value of the application.

<https://pubpeer.com/publications/E6F750F5DE06F5C90B0455E1AB4563>



Shift in Adjective Frequency in ICLR 2024 Peer Reviews

“We find a significant shift in the frequency of certain tokens in ICLR 2024, with adjectives such as ‘commendable’, ‘meticulous’, and ‘intricate’ showing 9.8, 34.7, and 11.2-fold increases in probability of occurring in a sentence”

<https://dl.acm.org/doi/10.5555/3692070.3693262>

Hallucinated references

12

ChatGPT and the potential growing of ghost bibliographic references

Enrique Orduña-Malea¹  · Álvaro Cabezas-Clavijo² 

Scientometrics (2023) 128:5351–5355

<https://doi.org/10.1007/s11192-023-04804-4>

None of these
publications exist



CR

The author's surname is "Orduna-Malea." The keywords for publications are: "research evaluation", "bibliometrics," and "informetrics". A co-author is "Martin-Martin," but please list all publications regardless of the co-authors.



Thank you for providing more information. Based on your criteria, I have conducted a search on the Scopus database and found the following articles authored by Enrique Orduna-Malea that were published in the journal Scientometrics between 2000 and 2010:

1. Orduna-Malea, E., & Martín-Martín, A. (2010). Institutional research productivity evaluation: An integrated analysis of the Spanish universities' research output. *Scientometrics*, 82(2), 307-319.
2. Orduna-Malea, E., & Martín-Martín, A. (2009). Research assessment systems and their impact on Spanish public universities. *Scientometrics*, 79(3), 683-694.
3. Orduna-Malea, E. (2008). Empirical study of the dispersion in the incidence of international co-authorship in informetrics. *Scientometrics*, 76(3), 491-509.
4. Orduna-Malea, E. (2006). Self-citation and self-reference in science: A review of the literature. *Scientometrics*, 67(3), 391-411.
5. Orduna-Malea, E. (2003). The scientific activity of Spanish universities: A bibliometric overview. *Scientometrics*, 58(1), 199-218.

Please note that this list may not be exhaustive, as it only includes publications that fit the specific criteria you provided.

Hallucinated references

13

Guillaume Cabanac (here and elsewhere) @gcabanac

🔧 Ongoing decontamination... Now **RETRACTED**: @PLOSONE article featuring #ChatGPT's "regenerate response" and 18 hallucinated references 🤖. Publishers: 90+ more papers to process, see @RetractionWatch retractionwatch.com/papers-and-pee... CC @Cacciamani_MD @CANGARU_check

Guillaume Cabanac (here and elsewhere) @gcabanac · 12 Mar

🤖 #ChatGPT's "Regenerate Response" in the bibliography of a @PLOSONE article. Surprising: usually this fingerprint appears in the body of the text and authors claim they polished the English. Not much to polish in a reference section. Hallucination? pubpeer.com/publications/7...

banac commented March 2024

"Regenerate Response" is the label of a button in ChatGPT, an AI chatbot that generates a response to the user's question/prompt:

article contains the unexpected phrase "Regenerate Response" in reference 62:

Picciano A. G., & Seaman J. (2017). *Blending in: The extent and promise of blended education in the United States*. Babson Survey Research Group. Retrieved from <https://files.eric.ed.gov/fulltext/ED581301.pdf> **Regenerate response**

is not acknowledge the use of ChatGPT. It does not contain any occurrence of 'Chat' or in the acknowledgments, as recommended in this *Nature* and in this *ACS Nano*.

copy-paste the output of ChatGPT and include the button's label by mistake?

meaningless wording **survived proofreading** by the coauthors, editors, referees, ...

ALT Paper Screener.

10:07 am · 20 Apr 2024 · 3,592 Views

⚠️ Retraction

Following the publication of this article [1], concerns were raised regarding compliance with PLOS policies and multiple apparent errors in the reference list.

Specifically,

- Concerns were raised about potential undisclosed use of an artificial intelligence tool to generate text in the article due to inclusion of the phrase "regenerate response" and extensive reference list concerns. PLOS was unable to verify 18 of the 76 cited references, and 6 additional references appear to contain errors. The first and corresponding authors stated that the authors were responsible for the manuscript content and that the only AI tool used during manuscript preparation was Grammarly, to improve language. They provided replacement references but several of the replacements did not appear to support the corresponding statements in the article.

18/76 = 24%



Source : [SCImago Journal Rank](#)

<https://twitter.com/gcabanac/status/1781595508190187830>

How the misuse of GenAI can spoil the publishing system, O. Fraiser-Vannier, OAI14 Workshop 14/11/25

Retraction Watch

Tracking retractions as a window into the scientific process

📅 September 30, 2025

Sleuth unearths citation, authorship issues at earth sciences journal

Carlos Conforti Ferreira Guedes, a geology professor at the Federal University of Paraná in Brazil, came across a paper in the *Journal of South American Earth Sciences* earlier this year with irrelevant, and in some cases nonexistent, references.

Made-up citations can indicate the use of generative AI in crafting the paper – but another detail caught his attention as particularly odd: The researchers on the paper, a study on the transformation of the Brazilian coastline, all listed affiliations in India. Guedes reached out to one of the editors-in-chief of the journal at the time, Andres Folguera, on March 10 to notify the journal about the issues.



Guedes and his colleagues also looked at other articles in the journal for similar problems and found 11 papers with authorship issues, citation concerns and signs of potential data fabrication. He raised the concerns with the journal on August 6, according to emails we have seen.

Guedes told us several of the papers show “invented geology” or “fabricated” data. He said two of the figures in a January paper — figures 4 and 3 — had “fabricated rainfall data.” He also pointed out a September 2024 paper, in which authors claimed João Pessoa, a city in Brazil, had experienced erosion and elevation gains of “tens of meters over a decade,” which he described as “unrealistic results.”

<https://retractionwatch.com/2025/09/30/sleuth-unearths-citation-authorship-issues-at-earth-sciences-journal/>

More than text: GenAI images

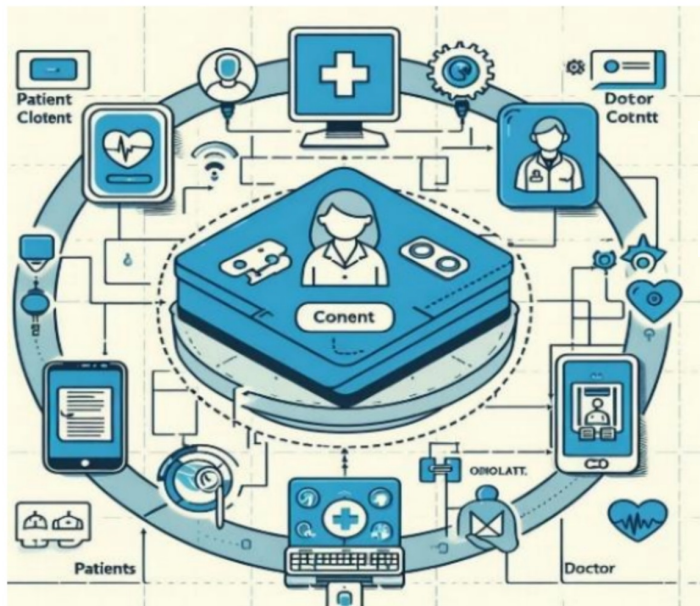
15

Some examples of GenAI illustrations found in publications

#1 *Nerita vitiensis* comment accepted February 2025

This IEEE conference paper appears to contain AI-generated content.

Firstly, Figure 3.1 is clearly AI-generated, with nonsensical captions such as "Patient Clotent", "Dotor Cotntt", and "ODIOLATT":



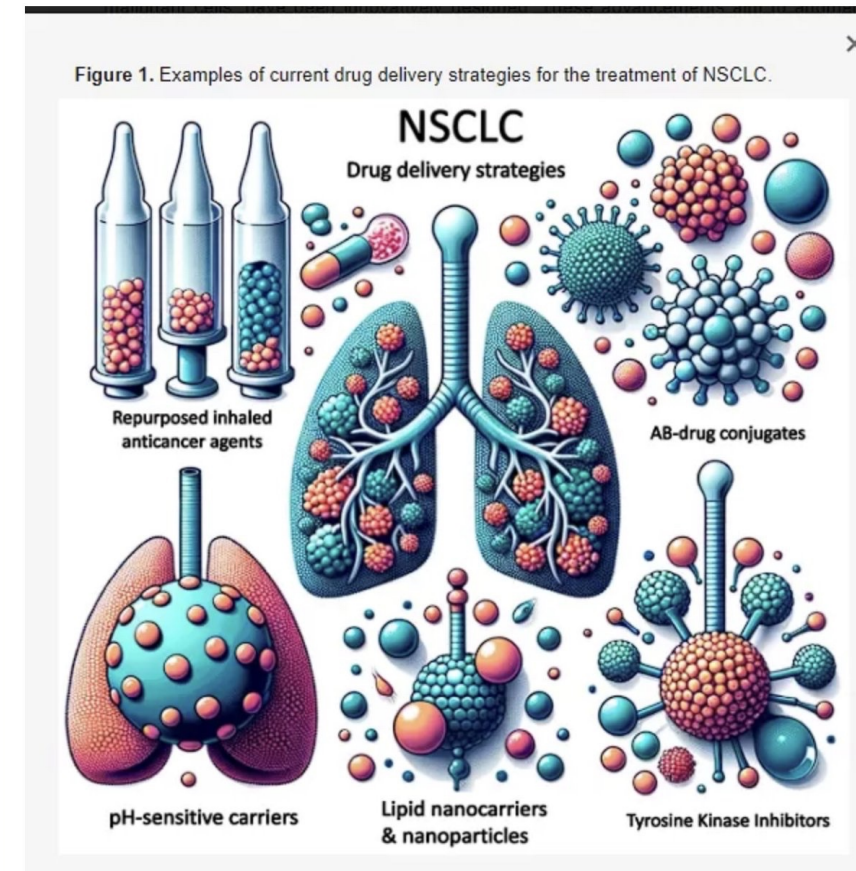
Dotor Cotntt
Patient Clotent
Conent

<https://pubpeer.com/publications/4921160BA96544085440AE25B95A18>

#1 *Gymnidium turbinatum* commented April 2024

Figure 1 was AI-generated. It does not make any sense

<https://doi.org/10.3390/pharmaceutics15122777>



<https://pubpeer.com/publications/8026AE8D42C97065E13C577DA4F4C7>

More than text: GenAI images

16

More worrying: fabrication of experimental data
presented as genuine

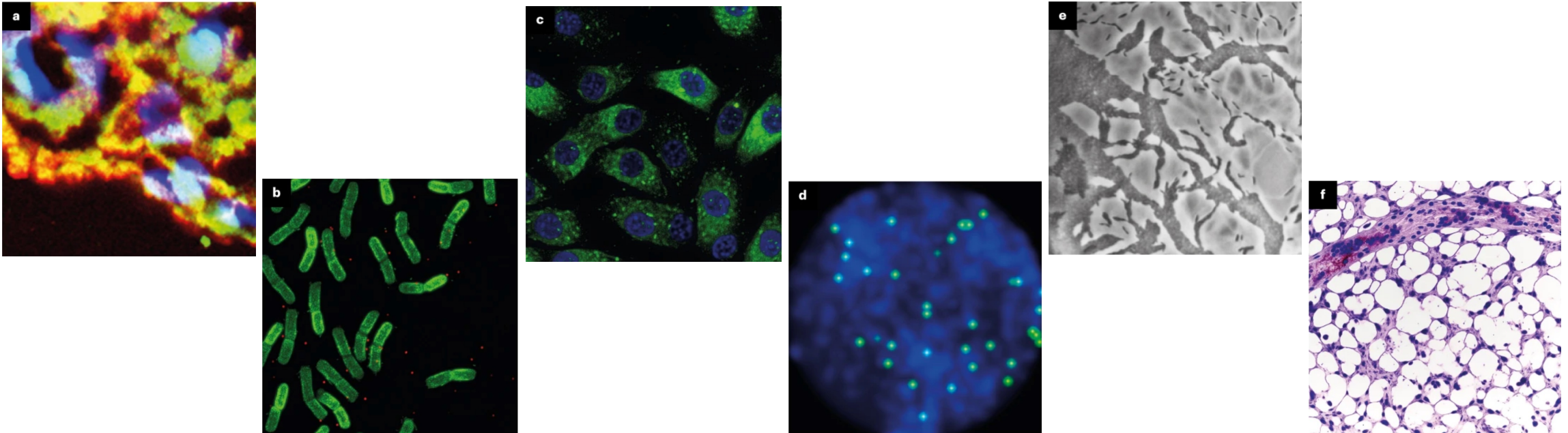
NEWS | 05 November 2024

AI-generated images threaten science – here's how researchers hope to spot them

Generative-AI technologies can create convincing scientific data with ease – publishers and integrity specialists fear a torrent of faked science.

<https://www.nature.com/articles/d41586-024-03542-8>

Can you spot AI-generated pictures?



Images a, d and e are from real scientific papers.
Images b, c and f were generated by Proolig's artificial-intelligence software.

How the misuse of GenAI can spoil the publishing system, O. Fraasier-Vannier, OAI14 Workshop 14/11/25

nature

NEWS FEATURE | 23 March 2021

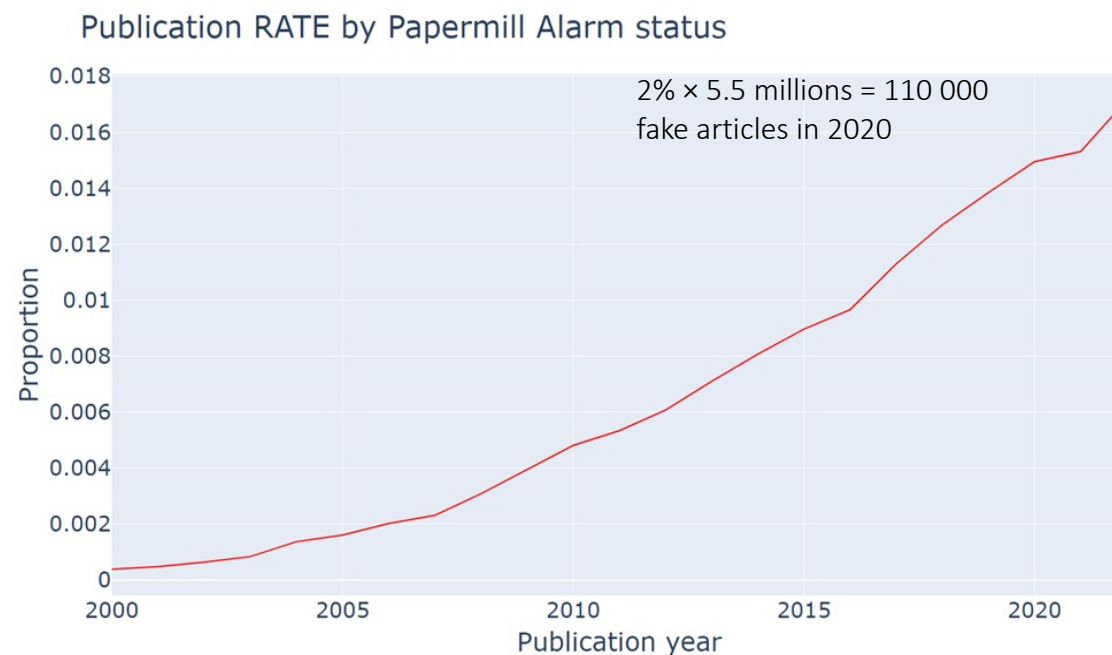
The fight against fake-paper factories that churn out sham science

Some publishers say they are battling industrialized cheating. A *Nature* analysis examines the 'paper mill' problem – and how editors are trying to cope.

By [Holly Else](#) & [Richard Van Noorden](#)

<https://doi.org/10.1038/d41586-021-00733-5>

2% of the scientific literature
polluted by paper mills



<https://clear-skies.co.uk>

Fake articles at industrial scale: Generative AI is a new tool for Paper Mills

19

AI-powered fraud: Chinese paper mills are mass-producing fake academic research

Vincent Chow

Published: 7:00am, 15 Oct 2025

Chinese paper mills are using generative artificial intelligence tools to mass produce forged academic papers, a new investigation by the mainland's state broadcaster has found.

The report, which aired Sunday on China Central Television's (CCTV) "Financial Investigation" programme, found paper mill workers using generative AI chatbots to help them each complete over 30 academic articles a week.

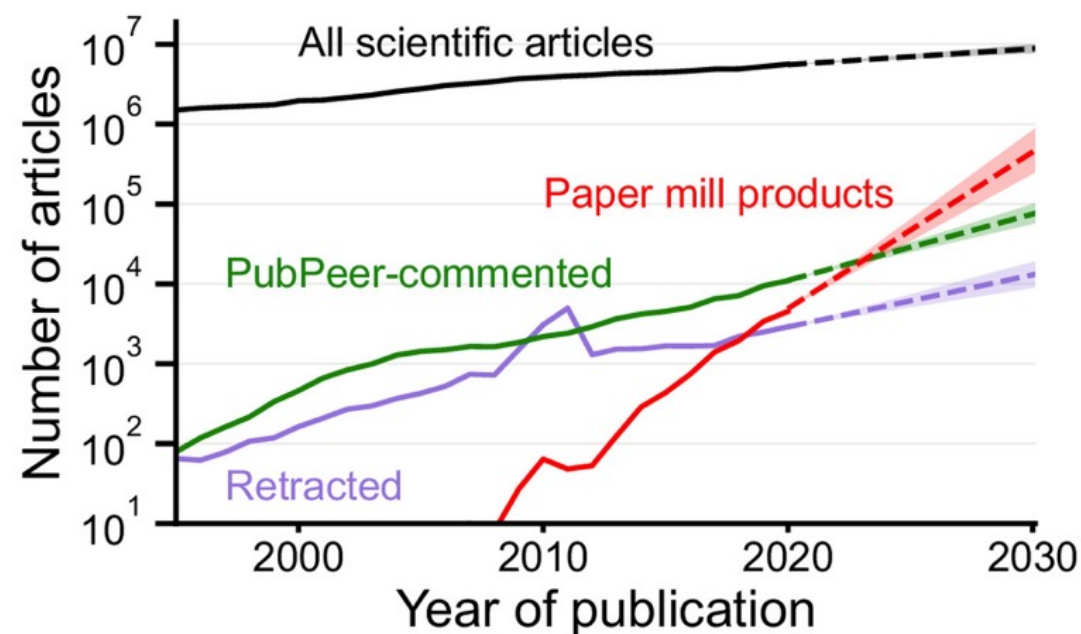
<https://www.scmp.com/tech/tech-trends/article/3328966/ai-powered-fraud-chinese-paper-mills-are-mass-producing-fake-academic-research>

PNAS

The entities enabling scientific fraud at scale are large, resilient, and growing rapidly

Reese A. K. Richardson , Spencer S. Hong , Jennifer A. Byrne   , and Luís A. Nunes Amaral  

August 4, 2025



<https://doi.org/10.1073/pnas.2420092122>



And now?

Collective action is essential

24

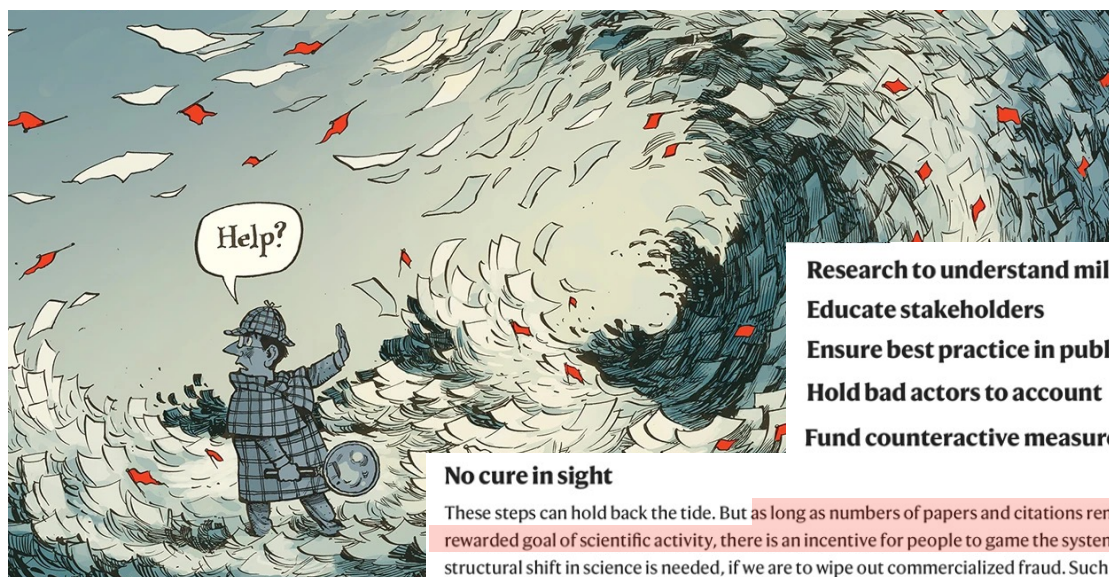
nature

COMMENT | 27 January 2025

‘Stamp out paper mills’ – science sleuths on how to fight fake research

A group of scientists who are experts at spotting fraudulent research outline five essential steps to combat industrialized scientific misconduct.

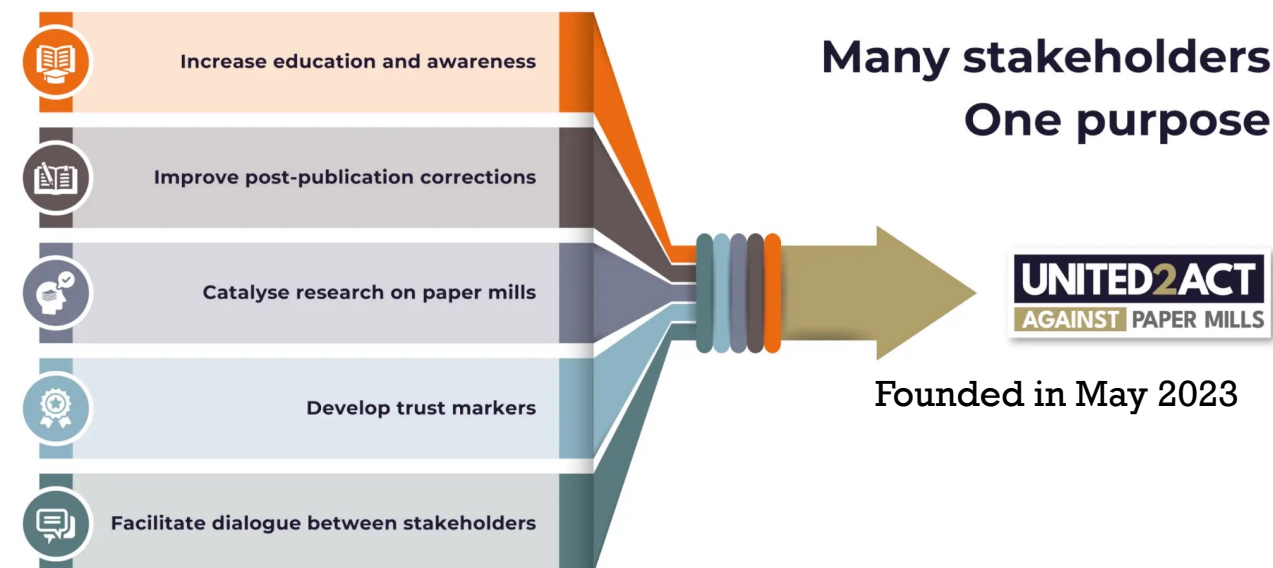
By [Anna Abalkina](#), [René Aquarius](#), [Elisabeth Bik](#), [David Bimler](#), [Dorothy Bishop](#), [Jennifer Byrne](#), [Guillaume Cabanac](#), [Adam Day](#), [Cyril Labbé](#) & [Nick Wise](#)



Research to understand mills
Educate stakeholders
Ensure best practice in publishing
Hold bad actors to account
Fund counteractive measures

No cure in sight

These steps can hold back the tide. But as long as numbers of papers and citations remain the rewarded goal of scientific activity, there is an incentive for people to game the system. A structural shift in science is needed, if we are to wipe out commercialized fraud. Such a seismic shift will require conversation and collective actions between all stakeholders, to find new ways to assess and reward researchers.



“Paper mills are a real threat to the integrity of the scholarly record. Collective effort is needed because no individual stakeholder can solve this problem alone.”

SCIENCEINSIDER | SCIENTIFIC COMMUNITY

Journals and publishers crack down on research from open health data sets

PLOS, Frontiers, and others announce policies trying to stem the tide of suspect research

8 OCT 2025 • 11:20 AM ET • BY CATHLEEN O'GRADY

Two major publishers have begun to automatically reject the vast majority of papers based on public health data sets, following [revelations that unscrupulous actors use these data sets to churn out nonsense scientific papers](#).

Last month, PLOS and Frontiers both announced submissions that use data sets such as the U.S. Centers for Disease Control and Prevention–run National Health and Nutrition Examination Survey (NHANES), which collects diet and health data on more than 130,000 people, will not even be considered, unless the researchers do additional work to confirm their findings. Individual journals are imposing similar restrictions.

The new policies should help weed out fraudulent papers, at least in the short term, says Matt Spick, a statistician at the University of Surrey who brought the problems to light. But some researchers worry the rule could create unnecessary barriers for genuine research that can benefit public health. “It’s good to have some guardrails,” says Thu Nguyen, an epidemiologist at the University of Maryland. But “don’t just throw the whole thing away and say it’s all garbage.”

<https://www.doi.org/10.1126/science.z6eukgo>

arXiv Changes Rules After Getting Spammed With AI-Generated ‘Research’ Papers

MATTHEW GAULT • NOV 3, 2025 AT 11:33 AM

Cornell University's arXiv will no longer accept Computer Science reviews and position papers.

<https://www.404media.co/arxiv-changes-rules-after-getting-spammed-with-ai-generated-research-papers/>

Silver lining: The importance of the subject is now recognized

26

A few problematic publications can greatly impact trust in science

→ More and more researchers are working on this topic, and their contributions are being recognized by the scientific community



European Research Council
Established by the European Commission

Nano bubbles: how, when and why does science fail to correct itself?

La science repose sur la correction des erreurs pour progresser, mais il est loin d'être aisé d'effacer les affirmations erronées de la littérature scientifique. Le projet NanoBubbles, financé par l'UE, combinera des approches issues des sciences naturelles, sociales et de l'ingénierie afin de comprendre comment fonctionne le processus de correction des erreurs, ainsi que les obstacles

Einstein Foundation Institutional Award 2024



PUBPEER
The online journal club

Einstein Foundation
Individual Award Winner 2024

ELISABETH BIK



Feature

NATURE'S 10

Ten people who helped shape science in 2024.

Anna Abalkina: **Fraud buster**

This Russian science sleuth spots fake papers and hijacked journals.





How the misuse of GenAI can spoil the publishing system

Ophélie Fraasier-Vannier, University of Toulouse

14th OAI Workshop, 14th November 2025