



**HAL**  
open science

## **(Mis)uses of generative AI in research**

Ophélie Fraisier-Vannier

► **To cite this version:**

| Ophélie Fraisier-Vannier. (Mis)uses of generative AI in research. Doctoral. France. 2026. ⟨hal-04840973v8⟩

**HAL Id: hal-04840973**

**<https://hal.science/hal-04840973v8>**

Submitted on 3 Apr 2026

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-SA 4.0 - Attribution - Non-commercial use - ShareAlike - International License



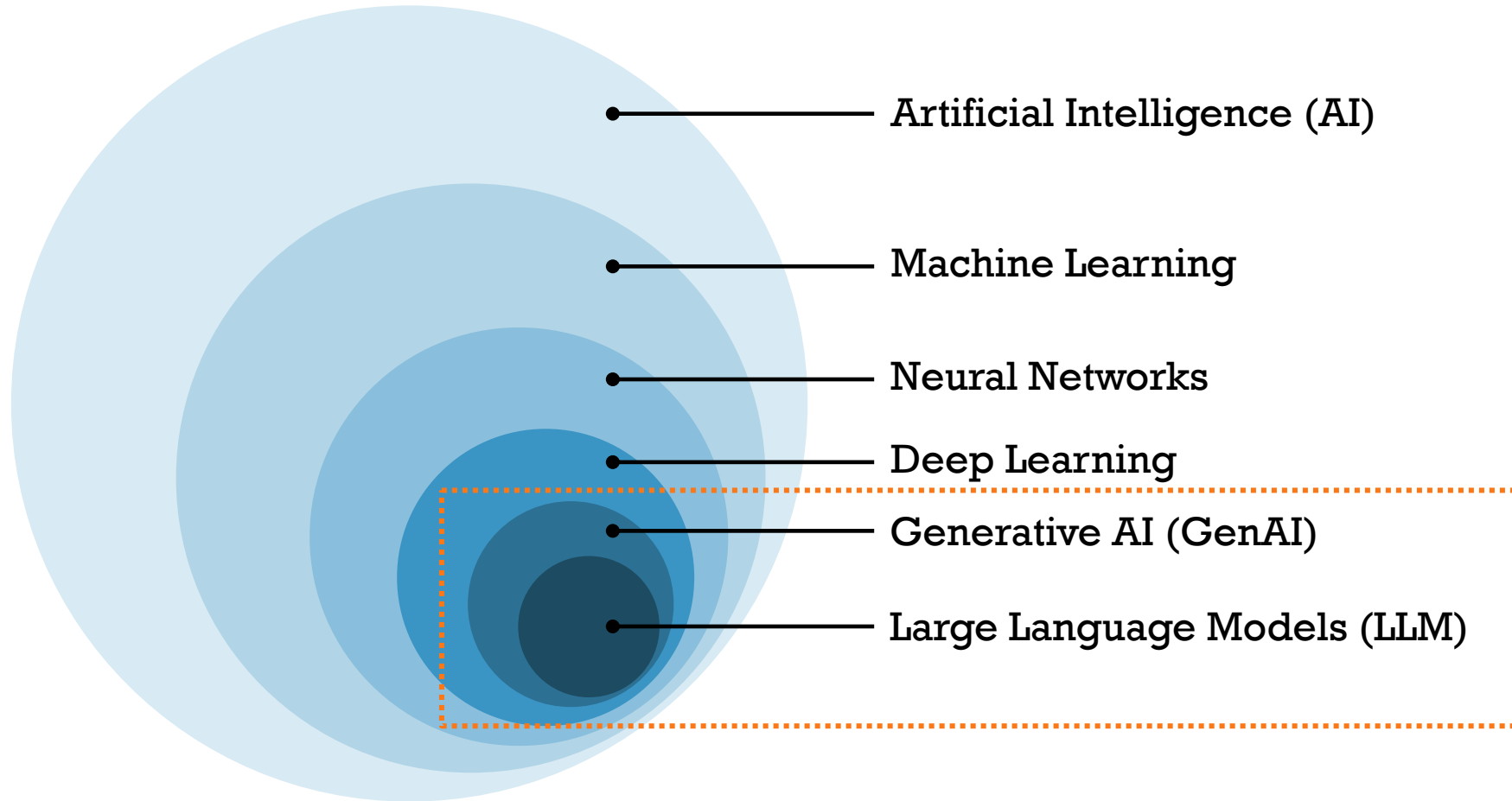
UNIVERSITÉ  
DE TOULOUSE

# (Mis)uses of generative AI in research

Ophélie FRAISIER-VANNIER, *Université de Toulouse*

Webinaire GTSO Données du Consortium Couperin, 03/04/2026

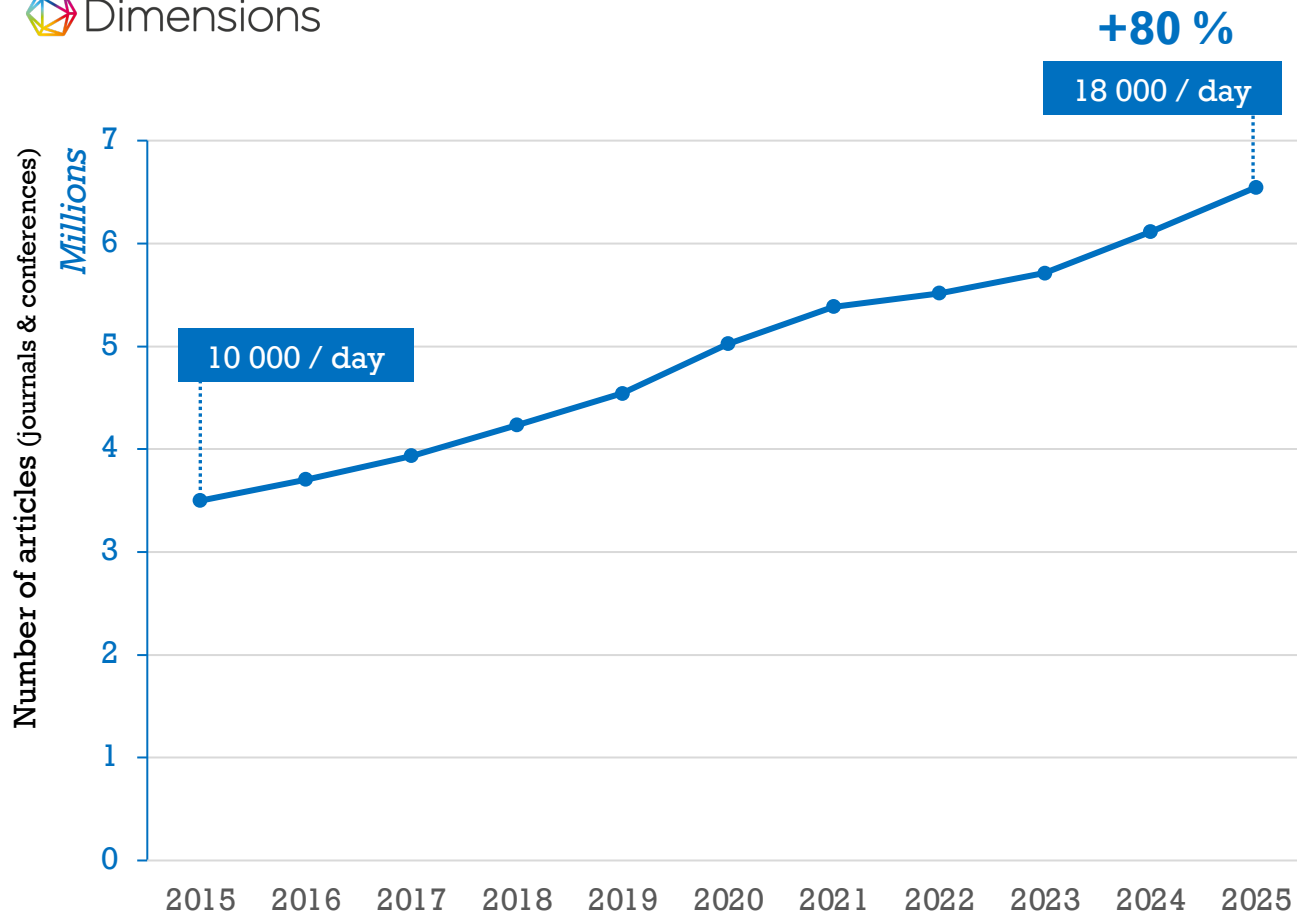
# Artificial Intelligence is a vast field



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

# A never-ending increase in scientific production

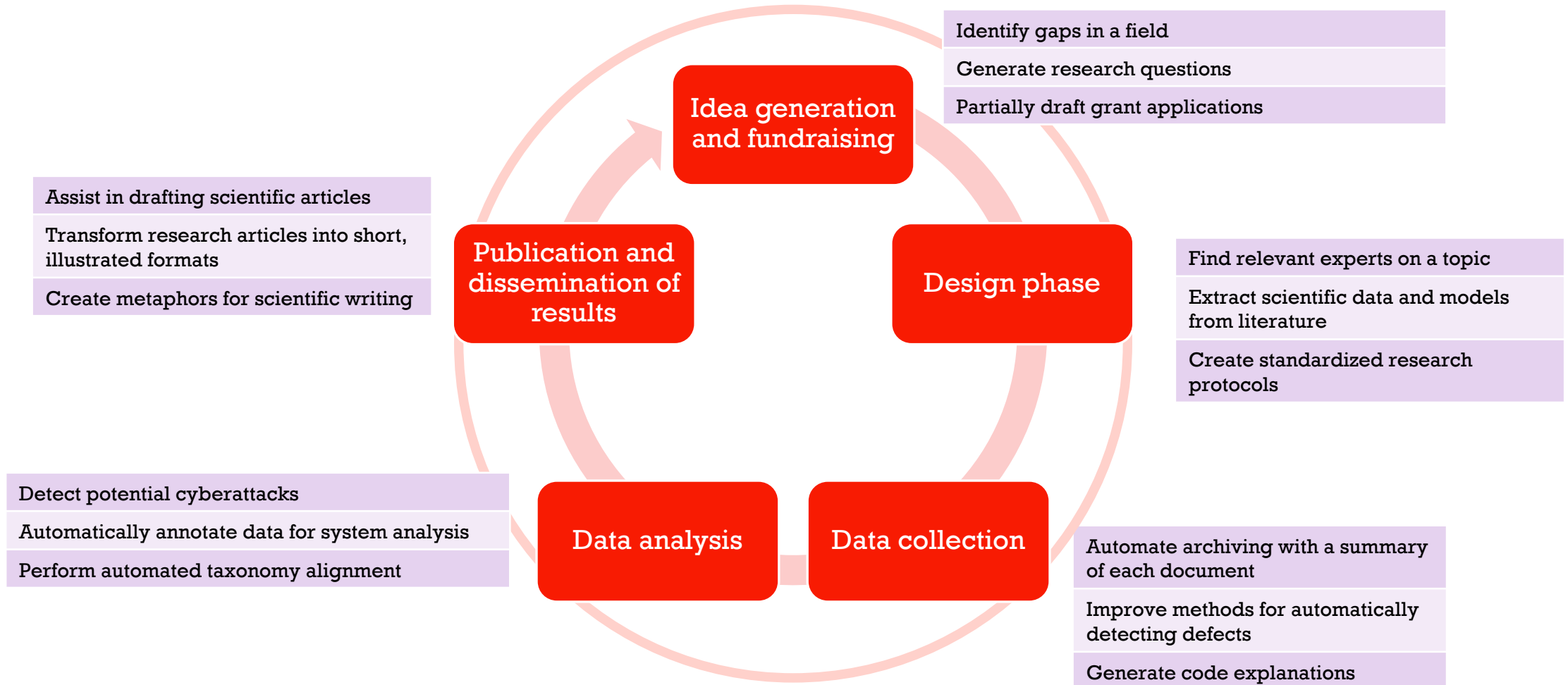
 Dimensions



‘The number of papers published each year is growing faster than the scientific workforce, with the **number of papers per researcher almost doubling between 1996 and 2022**. This reflects the fact that publishers’ commercial desire to publish (sell) more material has aligned well with the competitive prestige culture in which publications help secure jobs, grants, promotions, and awards.’

Beigel et al., 2025 (preprint)  
<https://doi.org/10.48550/ARXIV.2511.04820>

# GenAI tools can assist with scientific production



Using GenAI across different Research Phases – Cases, Potential and Risks  
Report to the Danish Council for Research and Innovation Policy by the Danish Centre for Studies in  
Research and Research Policy, Aarhus University, Sept. 2024  
<https://pure.au.dk/portal/en/publications/using-generative-artificial-intelligence-genai-across-different-r>

# GenAI and scientific (r)evolutions

## nature

NEWS | 09 October 2024

# Chemistry Nobel goes to developers of AlphaFold AI that predicts protein structures

**This year's prize celebrates computational tools that have transformed biology and have the potential to revolutionize drug discovery.**

By [Ewen Callaway](#)



AlphaFold model of phosphohistidine phosphatase overlaps closely with PDB structure.

<https://doi.org/10.1038/d41586-024-03214-7>

## IBM

# How generative AI models can fuel scientific discovery

Using generative models to come up with new ideas, we can dramatically accelerate the pace at which we can discover new molecules, materials, drugs, and more.

Similarly, another team at IBM Research used generative models, along with several other AI and high-performance computing advances, to come up with a new [photoacid generator \(PAG\)](#) — a material key to manufacturing semiconductors — a process that usually takes years and was completed in weeks.

<https://research.ibm.com/blog/generative-models-toolkit-for-scientific-discovery>

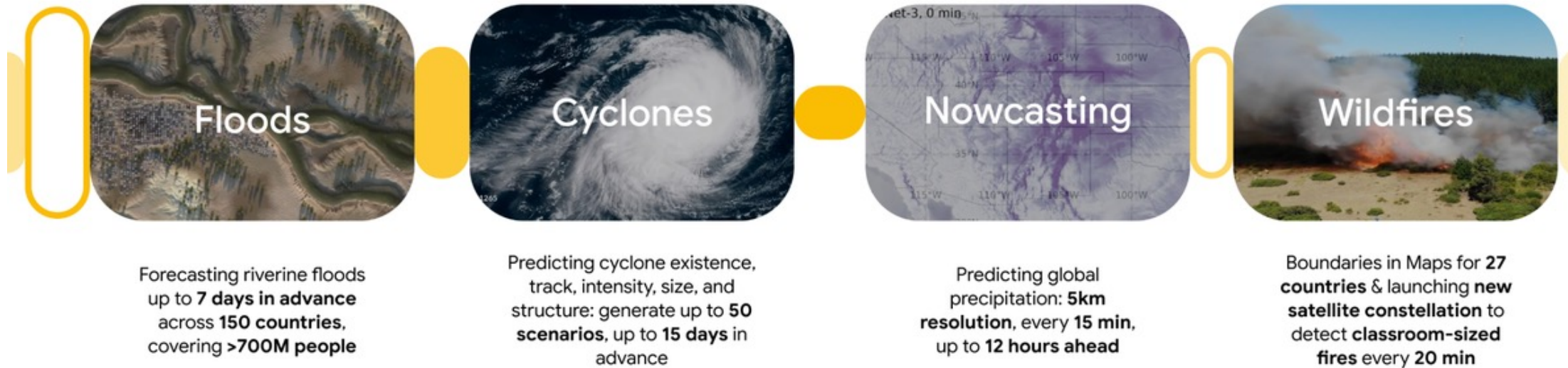
UC SAN DIEGO TODAY

## Accelerating Climate Modeling with Generative AI

A climate model combining generative AI and physics data is **25 times faster than the state of the art**

December 02, 2024

<https://today.ucsd.edu/story/accelerating-climate-modeling-with-generative-ai>

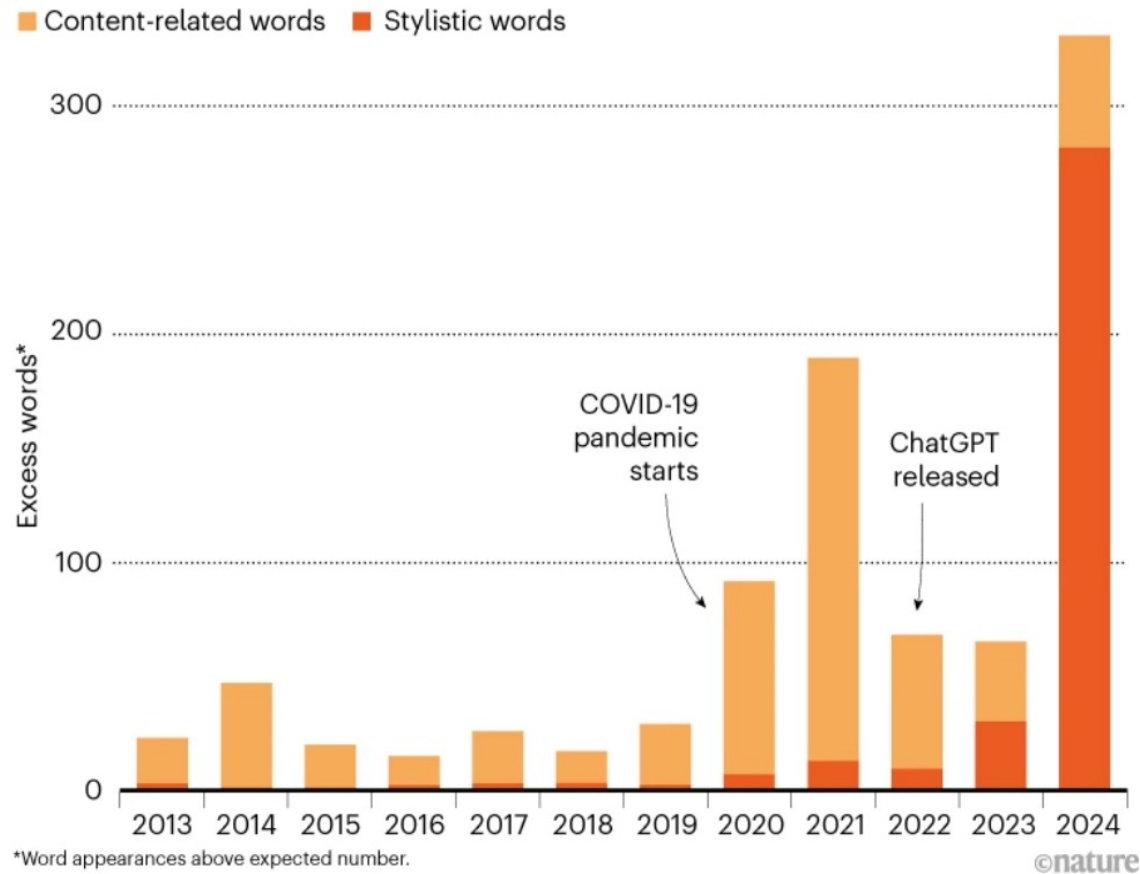


### Geospatial Reasoning

<https://research.google/blog/how-ai-is-helping-us-build-a-more-resilient-planet/>

# GenAI's broader influence on science

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’

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

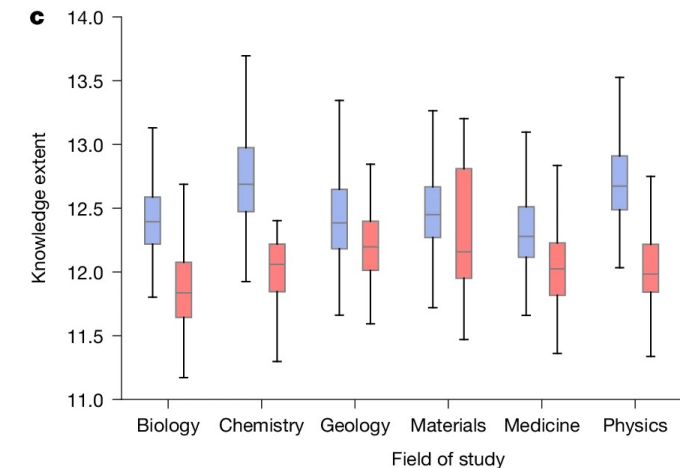
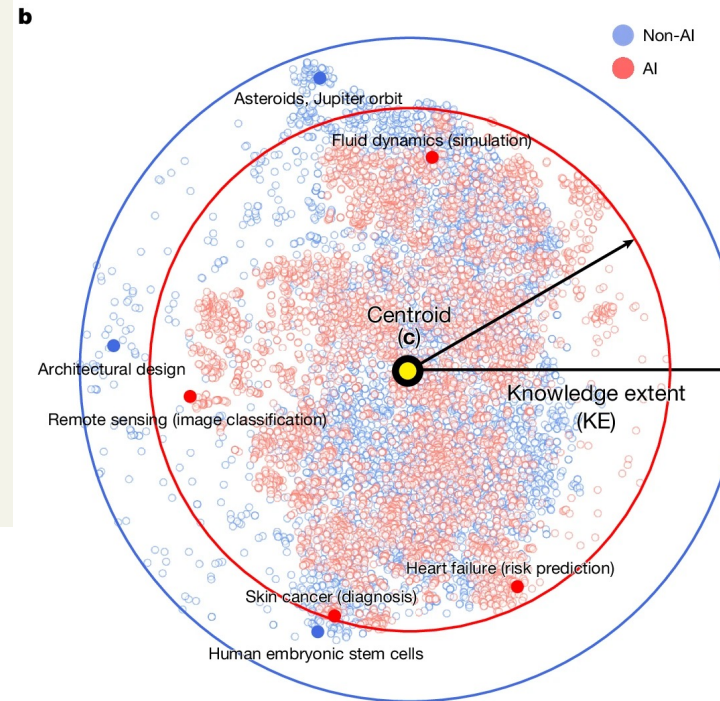
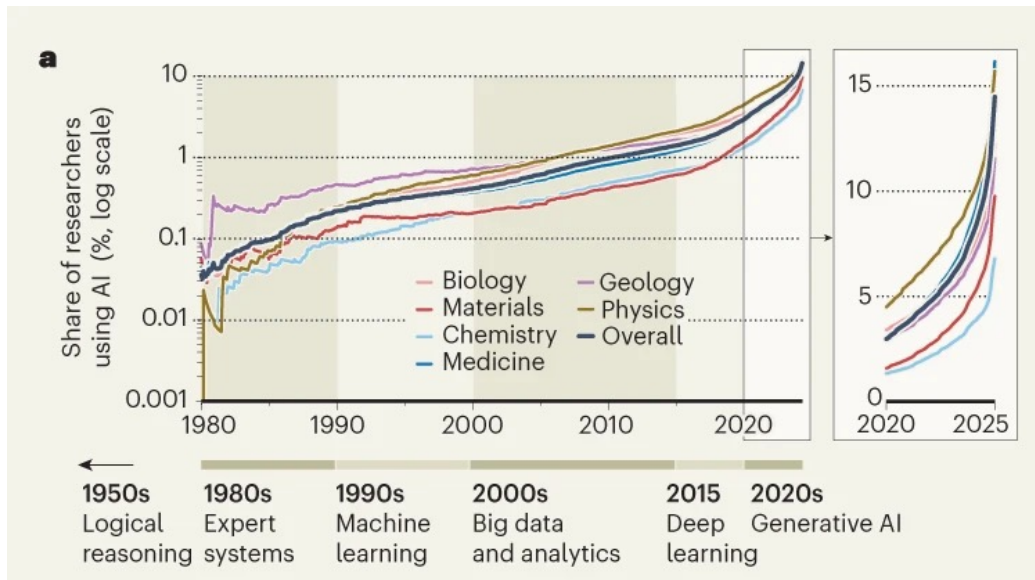
# GenAI's broader influence on science

Nature | Vol 649 | 29 January 2026

## Artificial intelligence tools expand scientists' impact but contract science's focus

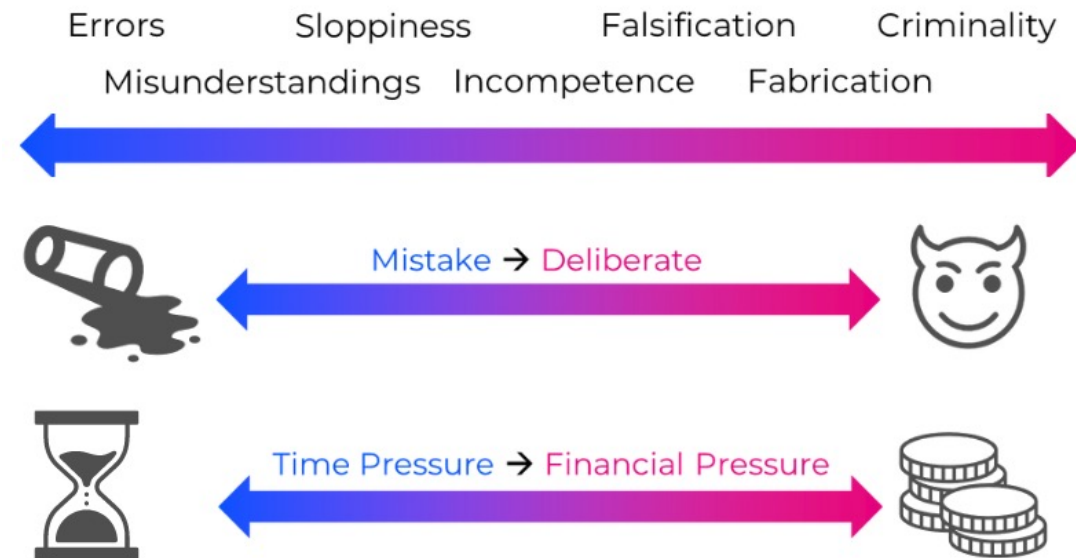
Qianyue Hao<sup>1</sup>, Fengli Xu<sup>1</sup>, Yong Li<sup>1,2</sup> & James Evans<sup>3,4</sup>

'Areas with an abundance of data are increasingly and disproportionately amenable to AI research, contributing to the observed concentration within knowledge space.'



# Use of GenAI in Questionable Research Practices

## Spectrum of Questionable Research Practices (QRPs)



Kolstoe, S.E. 2023

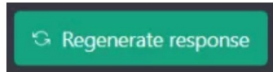
**UKRIO**  
RESEARCH INTEGRITY OFFICE

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

# LLM artefacts found in scientific articles

#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”:

Toxins 2023, 15, 199 5 of 12

proper guidance, leading to urine may deviate from the intended target. Although the Bo... through diffusion, the farther the injection site... stress urinary incontinence [24,25]. Therefore, it is essential to confirm to relax the urethral pressure and inc... likelihood of successful voiding in patie... targeted two areas of the urethra between and ten o'clock (Figure 3) to relax the vertic... the striated muscles on the lateral and ventral urethra [13].

**Regenerate Response**

According to most previous studies, the injection ranges from 50 units to 200 units to treat patient with detrusor sphincter dyssynergia, (FS), and poor relaxation of the external ureth... unclear what the appropriate dosage of Bo... be. In previous treatments, most studies u... that the effect is not significant in some... and Fowler's syndrome. In... that the effect... studies for treating detrusor sphincter dyssynergia (DSD) and Fowler's syndrome. In our study, the objective is to relax the external sphincter,

<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>



# LLM artefacts found in scientific articles

**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	11	122
Elsevier	75	0
Springer	36	20

 Abbas, 2023  
*Cancer Research Review*

Certainly! Here are 15 references related to "The role of artificial intelligence in revolutionizing cancer research".

 Abdel-Gabbar and Kordy, 2024  
*Genome Instability & Disease*

Different kinds of leukocytes are essential to this physiological process. There was no text provided by the user to alter. Blood leukocytes may be divided into two primary categories: myeloid and lymphoid.

# ChatGPT in peer reviews

#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.

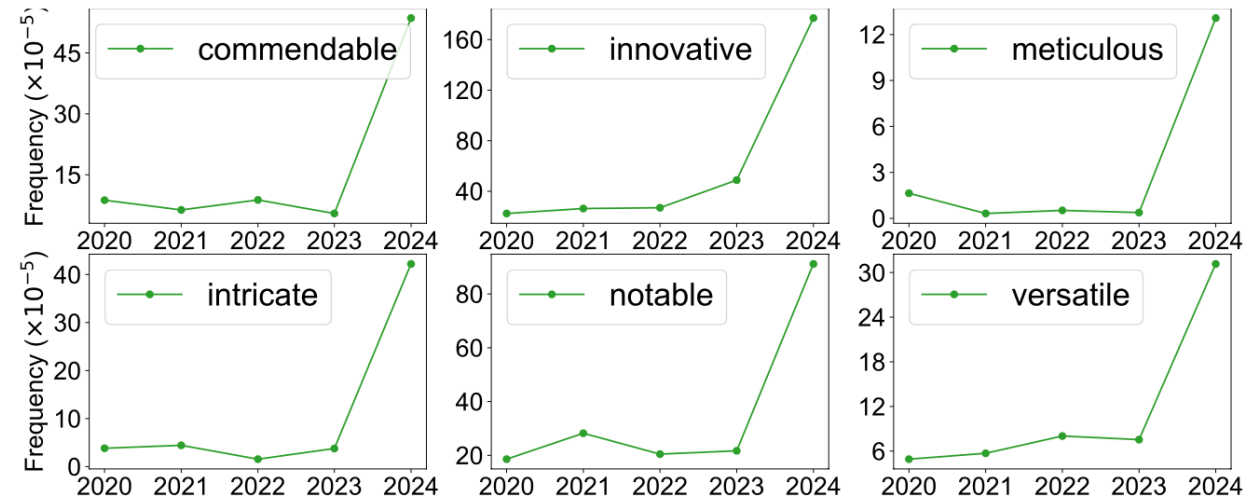
Regenerate response ←

### 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

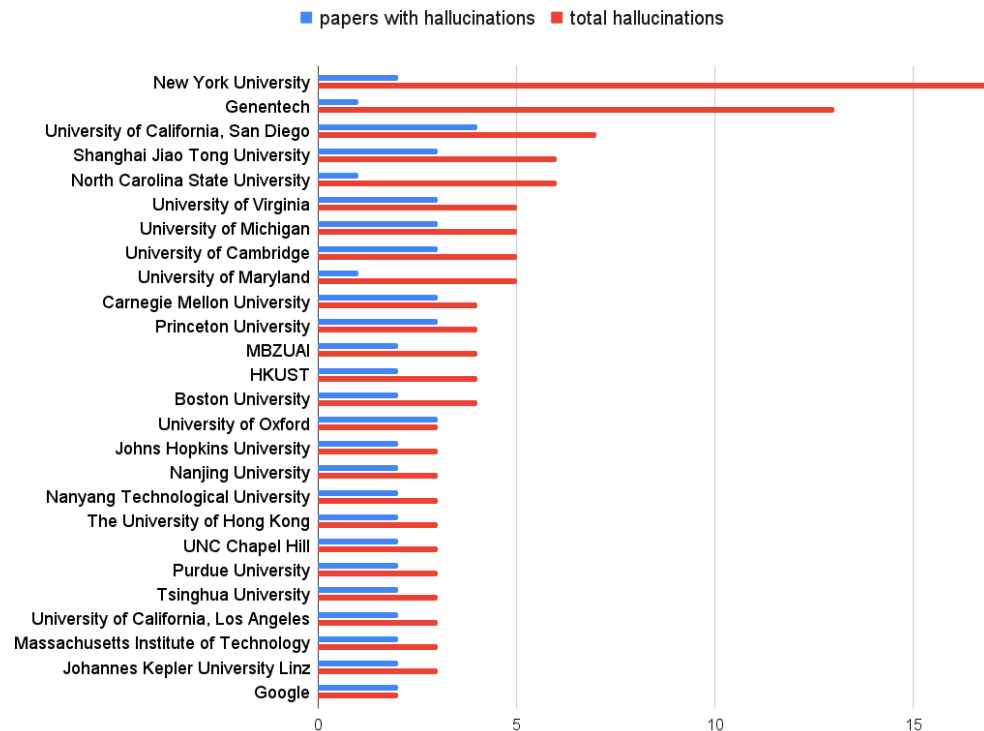
## GPTZero finds 100 new hallucinations in NeurIPS 2025 accepted papers



NAZAR SHMATKO, ALEX ADAM, PAUL ESAU

JAN 21, 2026 · 26 MIN READ

Number of hallucinations by author affiliation



## HalluCitation Matters: Revealing the Impact of Hallucinated References with 300 Hallucinated Papers in ACL Conferences

Yusuke Sakai, Hidetaka Kamigaito, Taro Watanabe  
Nara Institute of Science and Technology (NAIST), Japan  
{sakai.yusuke.sr9, kamigaito.h, taro}@is.naist.jp

Year	Venue	Main	Findings	S.I.D.	WS	Total
2024	NAACL	1 <sub>50.0</sub> <sup>0.18</sup>	0 <sub>0.00</sub> <sup>0.00</sup>	0 <sub>0.00</sub> <sup>0.00</sup>	1 <sub>50.0</sub> <sup>0.16</sup>	2 <sub>-</sub> <sup>0.12</sup>
2024	ACL	3 <sub>42.9</sub> <sup>0.31</sup>	1 <sub>14.3</sub> <sup>0.10</sup>	0 <sub>0.00</sub> <sup>0.00</sup>	3 <sub>42.9</sub> <sup>0.42</sup>	7 <sub>-</sub> <sup>0.26</sup>
2024	EMNLP	3 <sub>27.3</sub> <sup>0.24</sup>	4 <sub>36.4</sub> <sup>0.40</sup>	2 <sub>18.2</sub> <sup>1.16</sup>	2 <sub>18.2</sub> <sup>0.39</sup>	11 <sub>-</sub> <sup>0.37</sup>
2025	NAACL	6 <sub>17.1</sub> <sup>0.84</sup>	10 <sub>28.6</sub> <sup>2.11</sup>	2 <sub>5.71</sub> <sup>1.13</sup>	17 <sub>48.6</sub> <sup>3.04</sup>	35 <sub>-</sub> <sup>1.81</sup>
2025	ACL	14 <sub>16.3</sub> <sup>0.82</sup>	18 <sub>21.0</sub> <sup>1.30</sup>	10 <sub>11.6</sub> <sup>3.86</sup>	44 <sub>51.2</sub> <sup>3.78</sup>	86 <sub>-</sub> <sup>1.91</sup>
2025	EMNLP	60 <sub>39.0</sub> <sup>3.32</sup>	45 <sub>29.2</sub> <sup>3.20</sup>	13 <sub>8.44</sub> <sup>4.83</sup>	36 <sub>23.4</sub> <sup>5.26</sup>	154 <sub>-</sub> <sup>3.70</sup>
2024		7 <sub>35.0</sub> <sup>0.25</sup>	5 <sub>25.0</sub> <sup>0.22</sup>	2 <sub>10.0</sub> <sup>0.57</sup>	6 <sub>30.0</sub> <sup>0.33</sup>	20 <sub>-</sub> <sup>0.28</sup>
2025		80 <sub>29.1</sub> <sup>1.89</sup>	73 <sub>26.5</sub> <sup>2.24</sup>	25 <sub>9.09</sub> <sup>3.55</sup>	97 <sub>35.3</sub> <sup>4.03</sup>	275 <sub>-</sub> <sup>2.59</sup>
Overall		87 <sub>29.5</sub> <sup>1.24</sup>	78 <sub>26.4</sub> <sup>1.41</sup>	27 <sub>9.15</sub> <sup>2.56</sup>	103 <sub>34.9</sub> <sup>2.42</sup>	295 <sub>-</sub> <sup>1.65</sup>

Table 4: Number of papers including “HalluCitation”. We also report the proportion of papers with HalluCitation relative to the total number of papers, as well as track-specific proportions within each venue. We denote these values as  $\text{NUM}_{\text{Track}\%}^{\text{Venue}\%}$ .

<https://doi.org/10.48550/arXiv.2601.18724> (preprint)

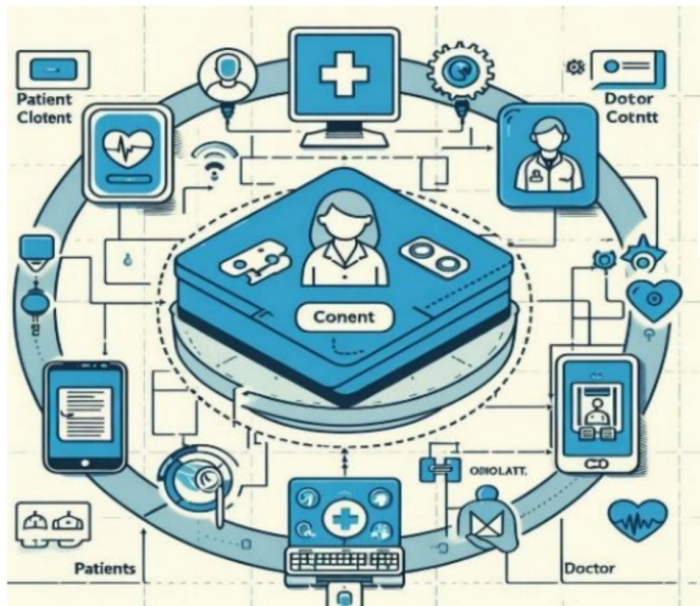
# More than text: GenAI images

## 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":



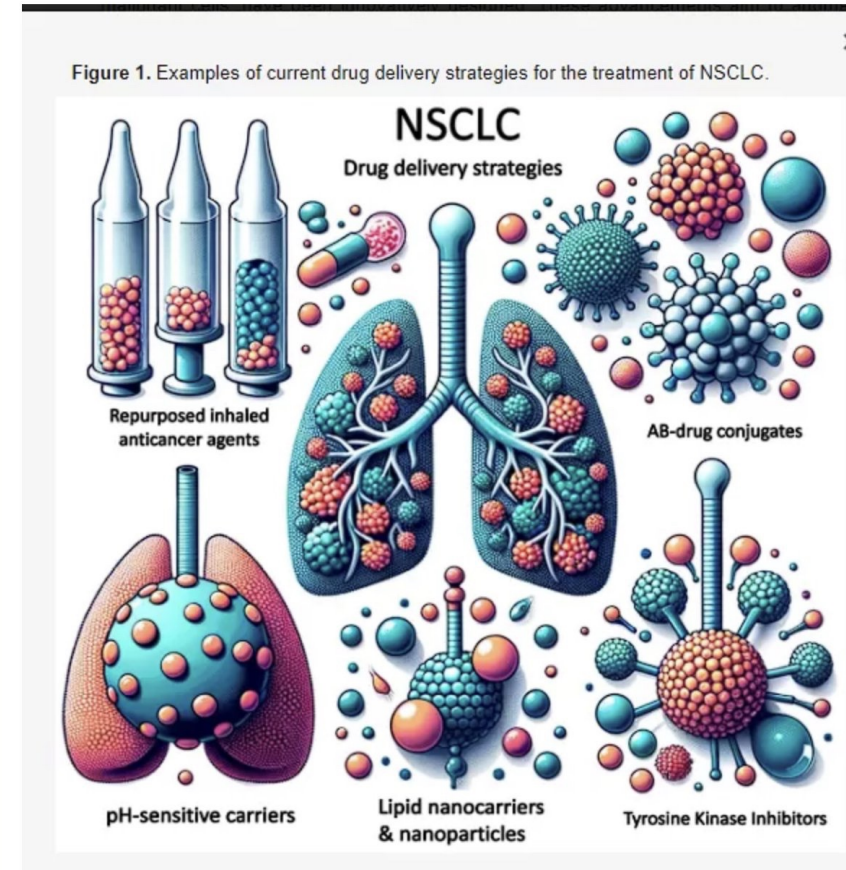
Dotor Cotntt  
Patient Clotent  
Conent

<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>

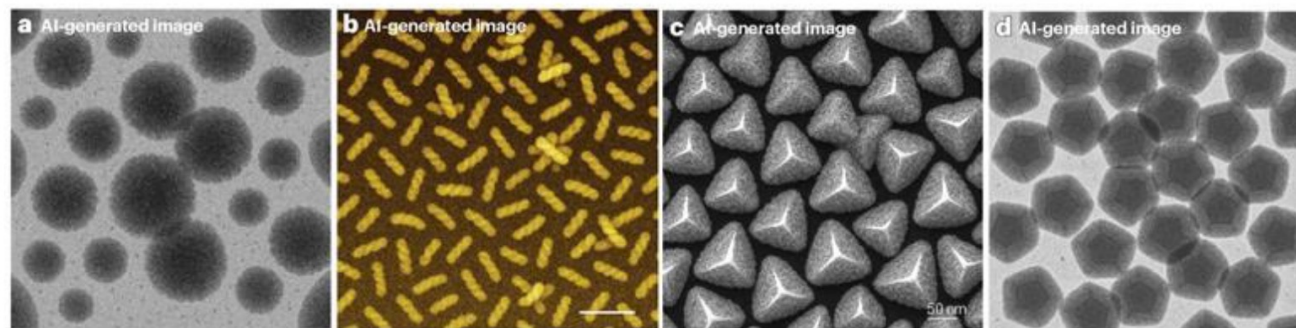


## More worrying: fabrication of experimental data presented as genuine

### CHEMISTRY\WORLD

## Materials scientists warn of threat posed by AI-generated experimental images

BY  KATRINA KRÄMER | 30 SEPTEMBER 2025



Source: © Nadiia Davydiuk et al, Nature Nanotechnology 2025

Fake microscopy images depicting four 'new' materials: nano-puffs, nano-cheetos, nano-dragon's teeth and nano-D12s. The images were generated by ChatGPT using simple text prompts

Besford, Faria and their teams surveyed 250 scientists, asking them to distinguish real from AI-generated microscopy images. The responses confirmed that this was essentially impossible, even for experts. The team is now calling for urgent action to avoid the literature becoming swamped with fake figures.

They suggest that publications should always include raw files, those created by the instrument. These are much harder to fake than images in a pdf. 'I think the solution needs to go much deeper, to institutional repositories of data', where raw data is kept safe from being modified or deleted, Rossner notes.

There should also be less pressure from journals and reviewers to create 'perfect' images. 'The community needs to get back to appreciating dirt,' says Besford. 'If you're making conclusions about chains of nanoparticles, and there's a couple of free nanoparticles, that doesn't mean your conclusions are invalid.'

<https://www.chemistryworld.com/news/fake-microscopy-images-generated-by-ai-are-indistinguishable-from-the-real-thing/4022215.article>

# The dangerous combination of GenAI and Paper Mills

# Generative AI is a new tool for Paper Mills

## 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>

nature

NEWS | 21 May 2025

## AI linked to explosion of low-quality biomedical research papers

Analysis flags hundreds of studies that seem to follow a template, reporting correlations between complex health conditions and single variables based on publicly available data sets.

By [Miryam Naddaf](#)

<https://doi.org/10.1038/d41586-025-01592-0>



National Health and Nutrition Examination Survey

# Collective action is essential to fight back

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.z6eukqo>

## 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/>



What can you do?

# Best practices for using LLMs in scientific production

Usage must be transparent and clearly indicated in the article

**Always** verify the information and references provided by the tool

Confidential data cannot be entered into a commercial tool

→ No use of commercial LLM (unless specific agreements exists)

(experimental data, manuscripts, ...)

**bibCheck**

 **GPTZero**

*References checkers*

 Université de Rennes

RAGaRenn

*Access via Renater*

 **Emmy**  
L'IA  
générative  
du CNRS

*Only CNRS agents*

NEWS | 25 March 2026

## Major conference catches illicit AI use – and rejects hundreds of papers

**The papers' watermarks allowed organizers to detect use of large language models in peer review.**

By [Elizabeth Gibney](#)

A major artificial-intelligence conference has rejected 497 papers – roughly 2% of submissions – whose authors violated AI-use policies in their peer reviews of other articles submitted to the meeting.

Conference organizers detected the illicit AI use by hiding watermarks in research papers distributed for review. If a researcher used an LLM to generate their peer review, instructions hidden in the watermark prompted the LLM to include telltale phrases in the review text. The presence of these phrases revealed that an AI model had been used to generate the review.

The issue was not AI tools' unreliability, but one of integrity, they added. "To be clear, we are not making a judgment call about the quality of flagged reviews or the reviewers' intentions," says the blog.

<https://www.doi.org/10.1038/d41586-026-00893-2>

# How to ensure the quality of your work & help clean the scientific literature

To check the quality of a journal / conference

## The Retraction Watch Hijacked Journal Checker

### Aletheia-Probe: A Tool for Automated Journal Assessment

Andreas Florath <sup>1</sup>

<https://doi.org/10.48550/arXiv.2601.10431>  
<https://github.com/sustainet-guardian/aletheia-probe>

To be up-to-date on science integrity topics

Retraction  Watch

To check the quality of an article



**Problematic Paper Screener**

Est. February 27<sup>th</sup>, 2021



**PUBPEER**

The online journal club

**bibCheck**

**Retraction Watch Database**

Plugin available

**zotero**

Plugin available



UNIVERSITÉ  
DE TOULOUSE

# (Mis)uses of generative AI in research

Ophélie FRAISIER-VANNIER, *Université de Toulouse*

Webinaire GTSO Données du Consortium Couperin, 03/04/2026