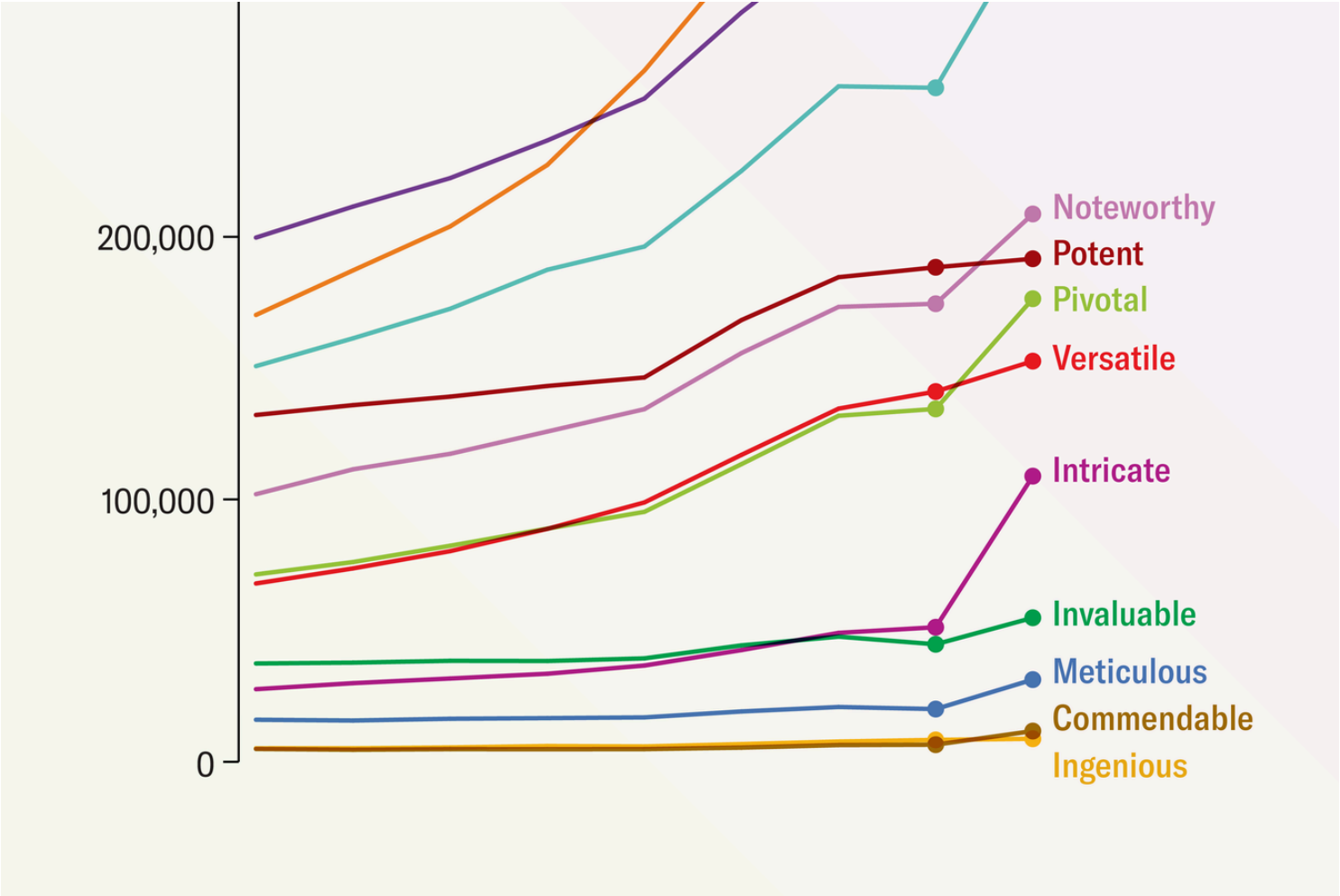# AI Chatbots Have Thoroughly Infiltrated Scientific Publishing

One percent of scientific articles published in 2023 showed signs of generative AI's potential involvement, according to a recent analysis

BY CHRIS STOKEL-WALKER   EDITED BY BEN GUARINO & SARAH LEWIN FRASIER

Amanda Montañez; Source: Andrew Gray

July 2024 Issue ⌄   Artificial Intelligence ⌄

Researchers are misusing ChatGPT and other artificial intelligence chatbots to produce scientific literature. At least, that's a new fear that some scientists have

raised, citing a stark rise in suspicious AI shibboleths showing up in published papers.

Some of these tells—such as the inadvertent inclusion of "certainly, here is a possible introduction for your topic" in a recent paper in *Surfaces and Interfaces*, a journal published by Elsevier—are reasonably obvious evidence that a scientist used an AI chatbot known as a large language model (LLM). But "that's probably only the tip of the iceberg," says scientific integrity consultant Elisabeth Bik. (A representative of Elsevier told *Scientific American* that the publisher regrets the situation and is investigating how it could have "slipped through" the manuscript evaluation process.) In most other cases AI involvement isn't as clear-cut, and automated AI text detectors are unreliable tools for analyzing a paper.

At least 60,000 papers may have used text generated by a large language model, according to librarian Andrew Gary's analysis.

Researchers from several fields have, however, identified a few key words and phrases (such as "complex and multifaceted") that tend to appear more often in AI-generated sentences than in typical human writing. "When you've looked at this stuff long enough, you get a feel for the style," says Andrew Gray, a librarian and researcher at University College London.

LLMs are designed to generate text—but what they produce may or may not be factually accurate. "The problem is that these tools are not good enough yet to trust," Bik says. They succumb to what computer scientists call hallucination: simply put, they make stuff up. "Specifically, for scientific papers," Bik notes, an AI "will generate citation references that don't exist." So if scientists place too much confidence in LLMs, study authors risk inserting

AI-fabricated flaws into their work, mixing more potential for error into the already messy reality of scientific publishing.
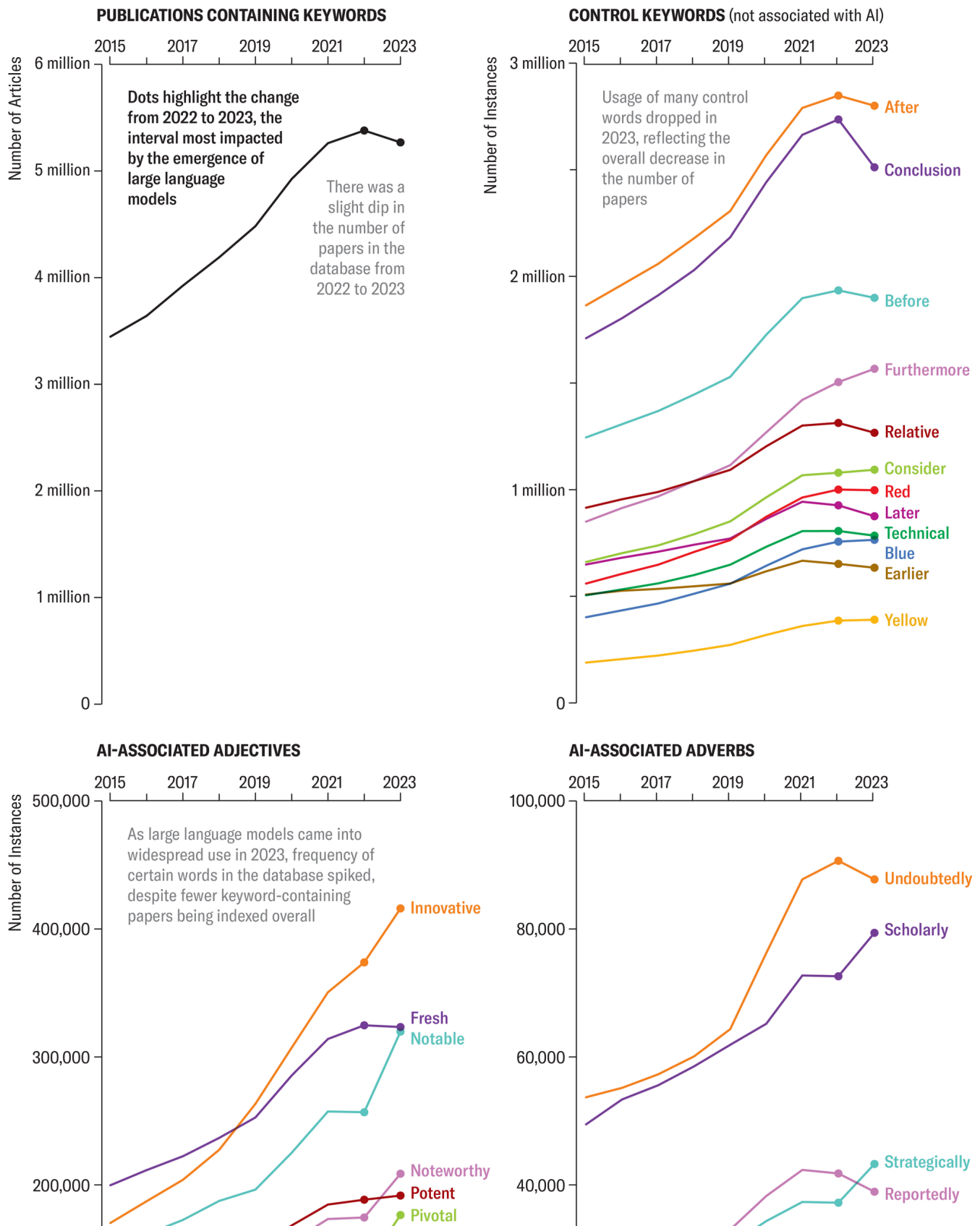
Gray recently hunted for AI buzzwords in scientific papers using Dimensions, a data analytics platform that its developers say tracks more than 140 million papers worldwide. He searched for words disproportionately used by chatbots, such as "intricate," "meticulous" and "commendable." These indicator words, he says, give a better sense of the problem's scale than any "giveaway" AI phrase a clumsy author might copy into a paper. At least 60,000 papers—slightly more than 1 percent of all scientific articles published globally last year—may have used an LLM, according to Gray's analysis, which was released on the preprint server arXiv.org and has yet to be peer-reviewed. Other studies that focused specifically on subsections of science suggest even more reliance on LLMs. One such investigation found that up to 17.5 percent of recent computer science papers exhibit signs of AI writing.
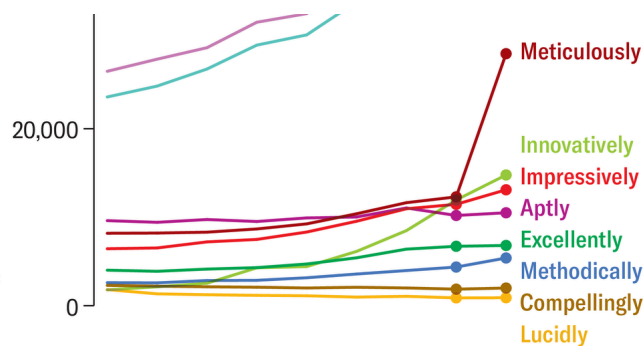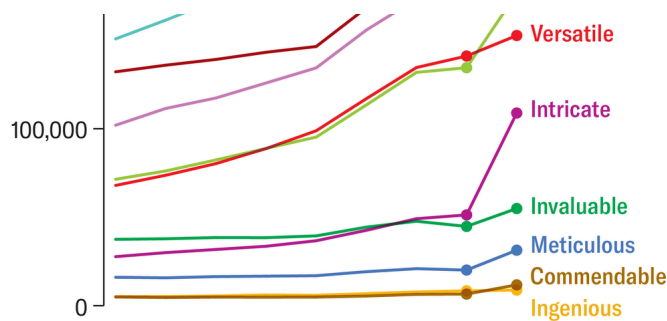
Suspicious Trends in Word Usage

For each year, starting with 2015, librarian and researcher Andrew Gray used the Dimensions database to count the number of published scientific papers that contained certain keywords and the total number of times each keyword was used. He used a set of 12 neutral "control" words, along with 12 adjectives and 12 adverbs that are associated with large language models such as ChatGPT.

Changes in Publishing Volume and Word Usage over Time

PUBLICATIONS CONTAINING KEYWORDS

Number of Articles

Dots highlight the change from 2022 to 2023, the interval most impacted by the emergence of large language models

There was a slight dip in the number of papers in the database from 2022 to 2023

2015 2017 2019 2021 2023

6 million
5 million
4 million
3 million
2 million
1 million
0

CONTROL KEYWORDS (not associated with AI)

Number of Instances

Usage of many control words dropped in 2023, reflecting the overall decrease in the number of papers

2015 2017 2019 2021 2023

3 million
2 million
1 million
0

After
Conclusion
Before
Furthermore
Relative
Consider
Red
Later
Technical
Blue
Earlier
Yellow

AI-ASSOCIATED ADJECTIVES

Number of Instances

As large language models came into widespread use in 2023, frequency of certain words in the database spiked, despite fewer keyword-containing papers being indexed overall

2015 2017 2019 2021 2023

500,000
400,000
300,000
200,000

Innovative
Fresh
Notable
Noteworthy
Potent
Pivotal

AI-ASSOCIATED ADVERBS

2015 2017 2019 2021 2023

100,000
80,000
60,000
40,000

Undoubtedly
Scholarly
Strategically
Reportedly

Versatile
Intricate
Invaluable
Meticulous
Commendable
Ingenious

100,000
0

Meticulously
Innovatively
Impressively
Aptly
Excellently
Methodically
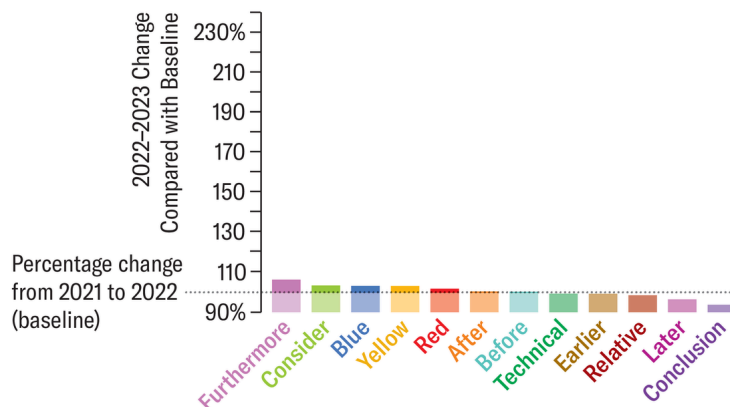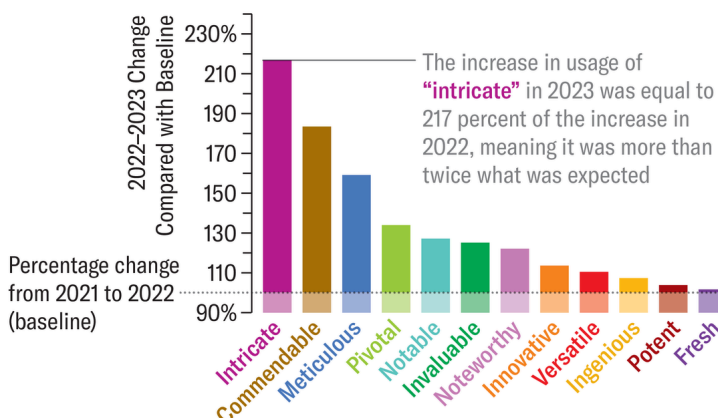**Compellingly**
Lucidly

20,000
0

Anomalies in Year-over-Year Changes

Gray adjusted the annual frequency of each word to account for the number of keyword-containing papers included in the database for that year. Then he compared each year-over-year increase or decrease with the previous one. By this metric, control words saw little varation in 2023, whereas some AI-associated ones far exceeded expectations.
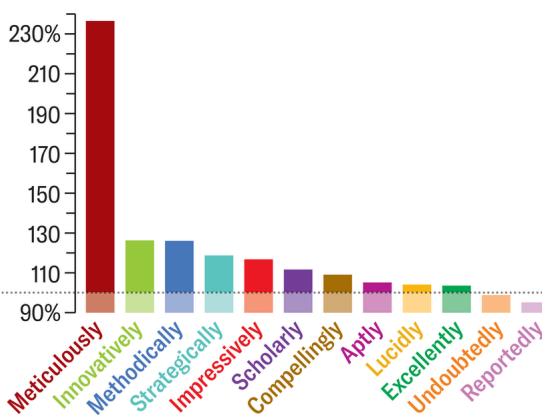
CONTROL WORDS

2022–2023 Change Compared with Baseline

230%
210
190
170
150
130
110
90%

Percentage change from 2021 to 2022 (baseline)

Furthermore · Consider · Blue · Yellow · Red · After · Before · Technical · Earlier · Relative · Later · Conclusion

AI-ASSOCIATED ADJECTIVES

2022–2023 Change Compared with Baseline

230%
210
190
170
150
130
110
90%

The increase in usage of **"intricate"** in 2023 was equal to 217 percent of the increase in 2022, meaning it was more than twice what was expected

Percentage change from 2021 to 2022 (baseline)

Intricate · Commendable · Meticulous · Pivotal · Notable · Invaluable · Noteworthy · Innovative · Versatile · Ingenious · Potent · Fresh

AI-ASSOCIATED ADVERBS

230%
210
190
170
150
130
110
90%

Meticulously · Innovatively · Methodically · Strategically · Impressively · Scholarly · Compellingly · Aptly · Lucidly · Excellently · Undoubtedly · Reportedly

Amanda Montañez; Source: Andrew Gray

Those findings are supported by *Scientific American*'s own search using Dimensions and several other scientific publication databases, including Google Scholar, Scopus, PubMed, OpenAlex and Internet Archive Scholar. This search looked for signs that can suggest an LLM was involved in the production of text for academic papers—measured by the prevalence of phrases that ChatGPT and other AI models typically append, such as "as of my last knowledge update." In 2020 that phrase appeared only once in results tracked by four of the major paper analytics platforms used in the investigation. But it appeared 136 times in 2022. There were some limitations

to this approach, though: It could not filter out papers that might have represented studies of AI models themselves rather than AI-generated content. And these databases include material beyond peer-reviewed articles in scientific journals.

Like Gray's approach, this search also turned up subtler traces that may have pointed toward an LLM: it looked at the number of times stock phrases or words preferred by ChatGPT were found in the scientific literature and tracked whether their prevalence was notably different in the years just before the November 2022 release of OpenAI's chatbot (going back to 2020). The findings suggest something has changed in the lexicon of scientific writing—a development that might be caused by the writing tics of increasingly present chatbots. "There's some evidence of some words changing steadily over time" as language normally evolves, Gray says. "But there's this question of how much of this is long-term natural change of language and how much is something different."

SYMPTOMS OF CHATGPT

For signs that AI may be involved in paper production or editing, *Scientific American*'s search delved into the word "delve"—which, as [some informal monitors] of AI-made text have pointed out, has seen an unusual spike in use across academia. An analysis of its use across the 37 million or so citations and paper abstracts in life sciences and biomedicine contained within the PubMed catalog highlighted how much the word is in vogue. Up from 349 uses in 2020, "delve" appeared 2,847 times in 2023 and has already cropped up 2,630 times so far in 2024—a 654 percent increase. Similar but less pronounced increases were seen in the Scopus database, which covers a wider range of sciences, and in Dimensions data.

Other terms flagged by these monitors as AI-generated catchwords have seen similar rises, according to the *Scientific American* analysis: "commendable" appeared 240 times in papers tracked by Scopus and 10,977 times in papers

tracked by Dimensions in 2020. Those numbers spiked to 829 (a 245 percent increase) and 20,536 (an 87 percent increase), respectively, in 2023. And in a perhaps ironic twist for would-be "meticulous" research, that word doubled on Scopus between 2020 and 2023.

MORE THAN MERE WORDS

In a world where academics live by the mantra "publish or perish," it's unsurprising that some are using chatbots to save time or to bolster their command of English in a sector where it is often required for publication. But employing AI technology as a grammar or syntax helper could be a slippery slope to misapplying it in other parts of the scientific process. Writing a paper with an LLM co-author, the worry goes, may lead to key figures generated whole cloth by AI or to peer reviews that are outsourced to automated evaluators.

These are not purely hypothetical scenarios. AI certainly has been used to produce scientific diagrams and illustrations that have often been included in academic papers—including, notably, one bizarrely endowed rodent—and even to replace human participants in experiments. And the use of AI chatbots may have permeated the peer-review process itself, based on a preprint study of the language in feedback given to scientists who presented research at conferences on AI in 2023 and 2024. If AI-generated judgments creep into academic papers alongside AI text, that concerns experts, including Matt Hodgkinson, a council member of the Committee on Publication Ethics, a U.K.-based nonprofit organization that promotes ethical academic research practices. Chatbots are "not good at doing analysis," he says, "and that's where the real danger lies."

A version of this article entitled "Chatbot Invasion" was adapted for inclusion in the July/August 2024 issue and the March 2025 special edition of Scientific American.

CHRIS STOKEL-WALKER is a freelance journalist in Newcastle, England.

More by Chris Stokel-Walker

View This Issue

It's Time to Stand Up for Science

If you enjoyed this article, I'd like to ask for your support. *Scientific American* has served as an advocate for science and industry for 180 years, and right now may be the most critical moment in that two-century history.

I've been a *Scientific American* subscriber since I was 12 years old, and it helped shape the way I look at the world. *SciAm* always educates and delights me, and inspires a sense of awe for our vast, beautiful universe. I hope it does that for you, too.

If you **subscribe to *Scientific American***, you help ensure that our coverage is centered on meaningful research and discovery; that we have the resources to report on the decisions that threaten labs across the U.S.; and that we support both budding and working scientists at a time when the value of science itself too often goes unrecognized.

In return, you get essential news, **captivating podcasts**, brilliant infographics, **can't-miss newsletters**, must-watch videos, **challenging games**, and the science world's best writing and reporting. You can even **gift someone a subscription**.

There has never been a more important time for us to stand up and show why science matters. I hope you'll support us in that mission.

Thank you,
David M. Ewalt, Editor in Chief, *Scientific American*

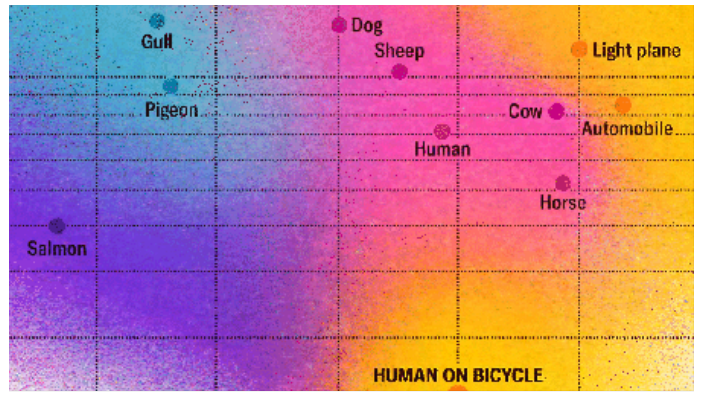**Subscribe**

Popular Stories

ASTRONOMY | OCTOBER 17, 2025

How to Catch a Rare Comet and a Meteor Shower

A comet visible to the naked eye will make its closest approach to Earth on October 21

DAN FALK

ENGINEERING | OCTOBER 14, 2025

A Classic Graphic Reveals Nature's Most Efficient Traveler

A famous graphic, now updated, compares locomotion in the animal kingdom
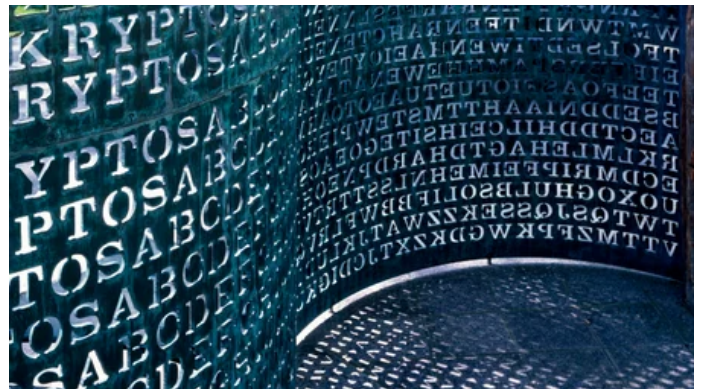
ALLISON PARSHALL, DTAN STUDIO

CLIMATE CHANGE | OCTOBER 16, 2025

Can We Bury Enough Wood to Slow Climate Change?

Wood vaulting, a simple, low-tech approach to storing carbon, has the potential to remove 12 billion tons of carbon dioxide from the atmosphere every year—and some companies are already trying it

SYRIS VALENTINE

GAMING | OCTOBER 16, 2025

A Solution to the CIA's Kryptos Code Is Found after 35 Years

After decades of speculation, two writers uncovered the answer to the Kryptos code's final cipher

HUMBERTO BASILIO

Subscribe to *Scientific American* to learn and share the most exciting discoveries, innovations and ideas shaping our world today.

Subscription Plans

Give a Gift Subscription