# the rise of digital fossils in scientific research poses challenges for ai and academic publishing



In recent years, the scientific community has encountered a curious and concerning phenomenon linked to the rise of artificial intelligence (AI) in research and academic publishing. An investigation highlighted the emergence of what researchers are calling “digital fossils”: erroneous, nonsensical terms that have inadvertently become embedded in scientific literature and AI knowledge bases. One prominent example is the phrase “vegetative electron microscopy,” a term that, despite being scientifically meaningless, has appeared multiple times across different research papers and contexts.

The origin of this term dates back to digitisation errors in the 1950s. During the scanning and optical character recognition (OCR) of papers published in Bacteriological Reviews, the word “vegetative” and the phrase “electron microscopy” were positioned in adjacent columns. The OCR software, malfunctioning, merged the two into the nonsensical hybrid. This error was compounded by later mistranslations in contemporary research papers published in 2017 and 2019, where subtle differences in Farsi between the words for “vegetative” and “scanning” led to the incorrect substitution of “vegetative” for “scanning.”

As reported by Retraction Watch and analysed further in a detailed examination by Aaron J. Snoswell and colleagues for The Conversation, the term “vegetative electron microscopy” proliferated after 2020, amplified by AI systems trained on a vast array of existing published content. The repeated presence of this erroneous term in various papers led AI text generators, including prominent large language models such as GPT-4o and Anthropic’s Claude 3.5, to treat it as a legitimate scientific concept. Consequently, when researchers used these AI models for drafting or assisting the writing of academic papers, these systems occasionally produced this nonsensical term, thereby propagating the error further.

This phenomenon highlights a critical challenge facing scientific research in the AI era. Despite some efforts to identify and correct such mistakes, including automated tools like the Problematic Paper Screener—which scans millions of articles weekly and has identified numerous AI-related anomalies—addressing these errors remains difficult. There are an estimated three million papers published annually, many now using AI-assisted writing, making detection and correction a daunting task. Journals also face dilemmas, as illustrated by Elsevier’s initial attempt to justify the use of “vegetative electron microscopy” before ultimately issuing a correction, indicating the complexities involved in maintaining publication integrity.

The persistence of “vegetative electron microscopy” in AI knowledge bases suggests it could become a permanent fixture in the scientific corpus, repeatedly resurfacing in future AI-generated content. Researchers warn that such entrenched errors compromise the integrity of scientific knowledge, which fundamentally relies on accurate, incremental advancements. If flawed concepts become embedded in foundational knowledge pools, the long-term implications for research and its applications could be significant.

This episode illuminates the interplay between digitisation, linguistic nuances, and AI training processes, underscoring inherent vulnerabilities in the increasingly automated landscape of academic research and publishing. The immortalisation of this digital fossil illustrates not only the challenges in detecting and correcting misinformation but also the evolving nature of knowledge creation and dissemination in an AI-influenced world.

Source: [Noah Wire Services](https://www.noahwire.com)

## References

* <https://www.zmescience.com/science/news-science/ai-wrong-science-term/> - This article discusses the phenomenon of AI-generated erroneous scientific terms like 'vegetative electron microscopy' appearing in multiple papers and the implications for the integrity of scientific knowledge, corroborating the core claim about digital fossils in AI-assisted research.
* <https://world.edu/a-weird-phrase-is-plaguing-scientific-papers-and-we-traced-it-back-to-a-glitch-in-ai-training-data/> - This source traces the origin of the phrase 'vegetative electron microscopy' to OCR errors from digitised Bacteriological Reviews and details subsequent propagation issues including mistranslations and AI amplification, confirming the article's narrative on the term's emergence and spread.
* <https://world.edu/a-weird-phrase-is-plaguing-scientific-papers-and-we-traced-it-back-to-a-glitch-in-ai-training-data/> - This article also covers the response of publishers such as Elsevier, which initially defended the term before correcting it, and mentions the use of automated tools like the Problematic Paper Screener, corroborating the discussion of challenges in correcting such AI-related scientific errors.
* <https://world.edu/a-weird-phrase-is-plaguing-scientific-papers-and-we-traced-it-back-to-a-glitch-in-ai-training-data/> - The piece highlights the broader implications for scientific publishing integrity and peer review processes due to AI-induced errors, supporting the article's point on the challenges and risks posed by erroneous AI-generated content in academia.
* <https://www.zmescience.com/science/news-science/ai-wrong-science-term/> - This source emphasizes the risk of permanent embedding of flawed concepts like 'vegetative electron microscopy' in AI knowledge bases, confirming the article's warning about long-term impacts on scientific knowledge and research integrity.