



Description

For the better part of a century, we have measured scholars by how much they publish. Generative AI has made publishing cheap, and in doing so it has quietly emptied the count of its meaning. The measure of merit now is not how many papers a researcher produces, but whether the work can be verified or whether it is trustworthy, accurate, and reliable.

“Publish or perish” has set the rhythm of academic life for generations. It decided who got hired, who got tenure, who got funded, and which institutions rose in the rankings and did real damage along the way, [driving burnout and distorted incentives through the early-career ranks](#). Its demerits were ignored all these years as the publication count was a usable proxy. It stood in for something we actually cared about: effort, expertise, contribution. However, that proxy is now broken by, and that the field needs to be honest about what should replace it.

The thing that should signal scholarly merit in 2026 and hereafter is the demonstrable trustworthiness of the work and the capacity, in a researcher and in a publishing system, to establish that trustworthiness rather than merely assert it.

Why the Count Ever Worked and Why AI has Hollowed it out

The publish-or-perish logic assumed scarcity. Writing a paper took months. Running an experiment took years. Producing volume was therefore a reasonable, if imperfect, proxy for effort and competence. The number was never the point — it was a stand-in for work that was hard to fake.

Generative AI has collapsed that assumption. Manuscripts, figures, methods sections, even reference lists can now be assembled in hours. Sakana AI’s “AI Scientist” has [demonstrated end-to-end automated paper generation](#), and computer-science conferences such as NeurIPS are already reporting submission volumes that reviewers cannot realistically absorb.

When the cost of producing the artifact falls to almost nothing, the count of artifacts stops carrying information about the person who produced them. The proxy has not become slightly noisier. It has detached from the thing it was supposed to measure. And you cannot repair a detached proxy by enforcing it more strictly — counting faster only accelerates the problem.

The Flood Is Not Just Larger. It Is Less Reliable

If the only change were volume, we might simply discount the count and move on. But AI has done something worse than make papers plentiful: it has made unreliable papers plentiful, and made them look exactly like sound ones.

Start with the citations. An early 2023 study found that a [large share of ChatGPT-generated references were fabricated or pointed to dead DOIs](#), and a parallel study in *Scientific Reports* found [GPT-4 inventing roughly 18% of the citations it produced](#). This did not resolve with better models. A 2025 experiment in *JMIR Mental Health* found that even GPT-4o [fabricated roughly one in five citations among the 141 non-fabricated ones, approximately 55% contained bibliographic errors, with incorrect DOIs being the most prevalent error type](#), and a 2026 synthesis now [treats hallucinated references as a standing threat to the integrity of the record](#). A fabricated citation is not a harmless slip. Once published, it is indexed, cited downstream, and ingested into the next model's training data, propagating a phantom through the literature.

Layer industrialized fraud on top of honest error and the reliability problem becomes systemic. [More than 10,000 retractions in 2023](#), and the [Retraction Watch Database now holds well over 55,000 entries](#). Wiley alone [retracted more than 11,300 articles from its Hindawi imprint and closed 19 journals](#). A machine-learning sweep of the cancer literature [flagged close to 10% of papers as carrying paper-mill signatures](#), and *Science* has documented an [explosion of low-quality papers strip-mining public datasets](#) with AI assistance.

This is the real indictment of the count. It was always a weak signal of merit; it is now a *contaminated* one. A higher number of publications can just as easily mark a polluted record as a productive scholar, and the number itself cannot tell the two apart.

The Honest Measure: Can the Work Be Trusted?

If anyone with an API key can generate a paper, then producing text is no longer where merit lives. Merit lives in whatever survives scrutiny, and what survives scrutiny is trustworthy, accurate, reliable work. That is the signal worth measuring, because it is precisely the signal AI cannot cheaply fake.

It helps to be concrete about what "verifiable" actually means, because the word is easy to wave at and hard to honor. A verifiable paper is one where the underlying data exist and can be re-analyzed; the methods are reported in enough detail to be repeated; every citation points to a real work that genuinely says what it is claimed to say; the figures and images are unmanipulated; any AI assistance is disclosed; and the authors are real, identifiable, and accountable for what appears under their names. Each of these is something a researcher can establish and a publisher can check, and each is something a language model can *imitate the surface of* but cannot actually deliver. A model can produce a citation that looks correct; it cannot make the cited paper say what it never said. The gap between looking sound and being sound is exactly where verification does its work.

This reframes what scholarly merit should reward. The most defensible thing a researcher can offer is no longer another entry on a publication list, but a paper whose every claim is auditable and holds up when someone checks. Verification, in other words, has become a core scholarly competency rather

than a clerical afterthought as both the discipline of producing work that can withstand scrutiny and the skill of scrutinizing others' work credibly. It is what authors owe when they check every AI-generated reference before submission and when they understand what AI-detection tools actually prove and where they fail.

What This Asks of the System

Verification is costly in exactly the way producing text no longer is, and that cost is the objection most likely to stall this shift. An overloaded peer-review system cannot simply will itself to check everything by hand. So, the honest answer is to put the routine, mechanical checks where machines belong and reserve human judgment for what only humans can do.

The mechanical layer should be infrastructure, not the burden of an individual author at midnight. This is already being built: The [STM Integrity Hub](#), launched in 2022, now [screens more than 125,000 manuscripts a month for over 35 participating publishers and intercepts roughly 1,000 suspected paper-mill submissions monthly](#). Springer Nature has [donated its in-house "Geppetto" AI-generated-text detector to the Hub](#) so smaller publishers can use it. Cross-publisher tools for image forensics, duplicate submission checks, "tortured phrase" detection, and irrelevant reference flagging are maturing; and they are increasingly being applied at *submission triage* rather than after publication.

A Closing Thought

AI is not essentially breaking science, nor would some tool would fix it. The claim is narrower and harder to dodge: producing a paper used to be hard, so counting papers told us something; producing a paper is now easy, so the count tells us almost nothing, and a worrying share of what it counts cannot be trusted at all.

The number of papers a scholar has published no longer reliably indicates what they know or what they have contributed. What does indicate it is whether the work holds up when a reviewer, a replicator, or a reader five years later actually checks — whether the data are real, the citations sound, the methods repeatable, the authorship accountable. That trustworthiness, and the capacity to establish it, is the scarce and valuable thing now.

If "publish or perish" was the slogan of the twentieth century, the twenty-first deserves something more honest. *Review and publish*. Or, with the warning left in: **verify, or perish**.

Disclosure: The image featured in this article was generated using NotebookLM for illustration purpose only. All the information in it is verified by a subject-matter expert.

Category

1. AI in Academia
2. Thought Leadership
3. trending

Date Created

2026/06/30

Author
anagha