ChatGPT Wrote This Paper, But I Helped

Jonathan E. Westfall

Delta State University

Author Note

Abstract

To be added

Keywords: Add keywords here.

ChatGPT Wrote This Paper, But I Helped

Today, large language models (LLMs) like ChatGPT are no longer fringe novelties—they are collaborators, assistants, and in some cases, co-authors. Trained on millions of examples, these AI systems can generate text that is coherent, responsive, and stylistically versatile (to a point). For researchers, students, and educators alike, they can offer new ways to brainstorm, revise, summarize, and even produce academic prose. Of course they also can be used to take shortcuts, potentially waterdown important concepts with flowery prose, and masquerade mediocrity behind apparent talent. Undoubtably, these innovations come with complex questions: Who gets credit for a paragraph written by an algorithm? How does outsourcing aspects of writing change the intellectual labor of scholarship? What does it mean to "write" a paper in an age when machines can write, too? This paper explores how LLMs are reshaping the academic writing landscape. It considers not just what these tools are capable of, but how they are currently used, the ethical and epistemological challenges they raise, and the ways they may prompt a rethinking of core academic values such as authorship, originality, and intellectual contribution. In doing so, it acknowledges the irony—and the insight—of its own title: *ChatGPT wrote this paper, but I helped.* Rather than asking whether LLMs should be part of the academic process, it begins from the more grounded position that they already are, and turns instead to the harder question: What do we do now?

The Current Academic Writing Landscape

Academic writing has long served as both a means of communicating within a discipline and a core marker of scholarly identity, for better or worse. The process of crafting a paper—from formulating a question to articulating an argument, has traditionally been regarded as an intellectual endeavor requiring sustained reflection, critical thinking, and a distinct authorial voice (Sword, 2012). Indeed, many academics today (myself included) can come up with ideas easily, but translating them from an concept to paper is the challenge, mostly in terms of logistics. Writing is not merely a way to communicate ideas; it is often how scholars come to understand their ideas in the first place, taking that small idea and fleshing it out further and further via long hours at a word processor. Expectations for originality and individual contribution have historically defined academic authorship. In most disciplines, credit is closely tied to

who generates ideas, structures arguments, and composes text. Institutions like the International Committee of Medical Journal Editors (ICMJE) set clear guidelines for authorship, emphasizing the necessity of substantial contributions to the conception or design of the work, as well as drafting or revising it critically (*Defining the Role of Authors and Contributors*, n.d.). This might be a fine system in its rigor, if at the same time, structural pressures within academia have reshaped how writing is approached. The imperative to "publish or perish" has intensified across disciplines, particularly for early-career researchers seeking tenure or grant funding (Rawat & Meena, 2014). The growing volume of scholarly publication—estimated at over 3 million peer-reviewed articles per year—has raised concerns about both the sustainability and quality of academic writing (Johnson et al., 2018). Moreover, the increasing interdisciplinary nature of scholarship requires academics to navigate unfamiliar fields and vocabularies, further complicating the writing process (Repko & Szostak, 2016). Time constraints, teaching loads, and administrative responsibilities often push writing into the margins of academic life, leading many scholars to seek tools that can streamline or support their writing workflows (Bartlett et al., 2021). Writing has become essential to the modern day academic, however it isn't always approached in the spirit of academia – it's a task that one must do in order to gain, maintain, and grow in their job, along with dozens of other things. In this context, LLMs like ChatGPT have emerged not simply as technological novelties, but as potential interventions in an overburdened system. Their integration into the academic landscape should be understood against this backdrop of heightened expectations, constrained resources, and shifting norms around authorship and collaboration.

What LLMs Do (And Don't Do)

Large language models (LLMs) like ChatGPT, developed by OpenAI, have rapidly become powerful and accessible tools for generating human-like text across a wide range of domains. These models are trained on massive datasets that include books, articles, websites, and other digital content, enabling them to produce text that mimics a variety of styles, tones, and disciplinary conventions (OpenAI, 2024). In academic contexts, LLMs have been adopted for tasks such as summarizing literature, paraphrasing text, suggesting revisions, generating outlines, and even drafting entire sections of papers

(Korinek, n.d.). The entire process for this paper, in fact, will be available for those interested in how LLMs might be used in writing.

At their best, LLMs function as intelligent writing assistants—capable of enhancing productivity, sparking creativity, and supporting non-native English speakers or novice scholars (Karimi, 2024). They can synthesize large amounts of information quickly and generate plausible prose on demand. This makes them appealing tools in an environment where time and cognitive bandwidth are scarce resources.

However, LLMs also have critical limitations. One of the most well-known issues is **hallucination**—the tendency of LLMs to produce false or fabricated information, including invented citations or misrepresented facts (Ji et al., 2023). While the generated text may appear fluent and authoritative, it often lacks grounding in verified sources. This poses particular risks in academic writing, where evidence and citation integrity are paramount. Thus far in the creation of this paper, finding if citations are real or false, and are appropriate, has been the bulk of the human author's work!

Moreover, LLMs do not understand meaning in a human sense. They operate based on statistical patterns in language rather than comprehension of ideas, arguments, or disciplinary logic (Bender et al., 2021). As a result, they may generate text that is grammatically correct but conceptually shallow, misleading, or logically inconsistent. Their outputs reflect the biases embedded in their training data, including cultural, gender, and racial biases (Weidinger et al., 2022).

Importantly, LLMs cannot make decisions about research design, analyze data, or formulate arguments in ways that reflect genuine critical thinking or scholarly intent. While they can simulate the *form* of academic writing, they do not participate in the *substance* of scholarship. To say that an LLM "wrote" something is, therefore, a shorthand for a much more complex human-machine interaction—one in which the human user prompts, interprets, revises, and ultimately takes responsibility for the final product.

This distinction is central to understanding the role of LLMs in academic work. They are tools, not agents; collaborators, perhaps, but not co-authors in any meaningful ethical or intellectual sense (COPE Council, 2023).

How Academics Are Using LLMs

Large language models (LLMs) like ChatGPT have become integral to various stages of academic writing, assisting scholars in tasks ranging from drafting and summarizing to editing and proofreading. These AI tools help researchers overcome writer's block, refine language, and generate coherent text, thereby enhancing productivity and facilitating the writing process.

In many instances, LLMs function as complementary collaborators, responding to user prompts by generating text that aligns with the user's initial ideas. For example, when I asked ChatGPT to help draft this paper, it told me "That title is fantastic—clever, self-aware, and immediately sets the tone for an engaging and thought-provoking paper." High praise that I don't think was warranted by merit but definitely appreciated by ego. This supportive role can be beneficial during the drafting phase, providing users with alternative phrasings and elaborations. However, LLMs typically lack the capacity to critically evaluate or challenge the content they generate, as they are designed to produce text based on patterns in the data they were trained on, rather than through critical analysis. One wonders if LLMs had been primarily trained upon the angsty writings of teenagers if they wouldn't be as apt to provide praise as they would to feign indifference.

Another use of the LLM by the academic is as a pseudo-collaborators or graduate students. In academic settings, early-career researchers often assist with literature reviews, idea generation, and initial drafting, while also providing critical feedback and questioning assumptions. In contrast, LLMs, although capable of generating text and suggesting revisions, do not possess the ability to engage in critical discourse or offer evaluative feedback, as they operate without understanding or awareness. One wonders, though, how long it takes for a seasoned writer to feel comfortable challenging a senior's judgment, so in that case, LLMs might not be any less critical than a first-year assistant professor.

For faculty at smaller institutions or in under-resourced settings, where access to graduate students or research assistants may be limited, LLMs may serve as accessible tools for brainstorming and drafting. Their availability (much greater than that of a junior faculty member) and responsiveness make them appealing options for scholars seeking immediate assistance in developing ideas and structuring

their writing. However, the effectiveness of LLMs in this role depends on the user's ability to critically assess and refine the AI-generated content, ensuring that the final work maintains scholarly rigor and originality (Meyer et al., 2023).

Finally, a notable concern in utilizing LLMs for academic writing is their approach to source selection. When tasked with supporting an argument, LLMs may retrieve references based on keyword matching without evaluating the credibility, relevance, or recency of the sources.

While scholars ideally conduct comprehensive literature reviews to identify and select the most pertinent and rigorous sources, practical constraints such as time limitations often lead to more expedient methods, such as seeking any source that supports a given argument. This pragmatic approach, although sometimes necessary, underscores the importance of maintaining critical engagement with source material, a quality that LLMs currently lack. Consequently, while LLMs can mimic certain aspects of scholarly research practices, they do not replace the nuanced judgment and critical appraisal that human researchers bring to source selection and evaluation.

In summary, LLMs are increasingly being integrated into academic writing as tools that can enhance productivity and assist in various stages of the writing process. However, their use necessitates careful oversight and critical engagement from scholars to ensure that the resulting work adheres to academic standards of quality, originality, and integrity.

Ethical and Epistemological Concerns

The rise of large language models (LLMs) in academic writing raises several ethical and epistemological concerns, particularly around authorship, transparency, integrity, and the nature of knowledge production. While LLMs can be powerful tools for augmenting scholarly productivity, their use challenges long-held norms about what it means to create, critique, and take responsibility for intellectual work.

One of the most immediate concerns is the question of authorship. Traditional academic authorship requires substantial intellectual contribution, including idea generation, drafting, revision, and accountability for content (International Committee of Medical Journal Editors, 2025). Since LLMs

cannot be held accountable, they do not meet the basic criteria for authorship, even if their textual contributions are substantial. The Committee on Publication Ethics (COPE Council, 2023) has emphasized that AI tools should not be listed as authors, and any use of such tools must be clearly disclosed in the manuscript (for this paper: done and done).

This issue is closely tied to the broader problem of transparency. When scholars use LLMs to generate or revise text, should that be disclosed in a footnote, a methods section, or an acknowledgment? Practices vary widely, and most journals have yet to establish consistent policies. Some researchers argue that failing to disclose AI involvement misrepresents the origin of the work (Stokel-Walker, 2023), while others believe that AI-generated assistance is no different from using a grammar-checking tool or consulting a colleague (Rainie, 2023). This ambiguity creates ethical gray zones, especially in student work, where the line between legitimate assistance and academic misconduct is often unclear (Cotton et al., 2024).

Beyond questions of credit and disclosure, LLMs raise epistemological concerns about how knowledge is constructed and validated. Academic writing is not just a means of communication—it is a mode of inquiry, where argumentation, evidence, and critical engagement are essential. LLMs, however, generate text based on probabilistic associations in training data, not through any understanding of logic, truth, or evidence (Bender et al., 2021). This means they may produce fluent but vacuous text, or worse, propagate misinformation or fabricated references (Ji et al., 2023).

This is especially troubling in the context of citation integrity. As mentioned, LLMs have been shown to hallucinate references—creating plausible-looking but nonexistent citations (Touvron et al., 2023) . Even when real references are suggested, they are often chosen based on surface-level token matches rather than relevance, credibility, or scholarly rigor. This stands in contrast to the ideal of academic research, where sources are carefully vetted and contextualized. Yet, as discussed previously, real-world academic practice is not always ideal. Time pressures, search fatigue, and publication deadlines can lead even experienced scholars to adopt heuristic approaches to sourcing—choosing "something that says that" rather than the best possible evidence (Tenopir et al., 2009).

In this way, LLMs may reflect and reinforce problematic tendencies that already exist in the academy. Rather than transforming academic writing, they may expose its weaknesses—its dependence on stylistic fluency, its tolerance for shallow citation, and its sometimes arbitrary approach to evidence. This invites a deeper question: What counts as rigor, originality, or authorship when machines are part of the process?

Implications for the Future of Scholarship

As large language models (LLMs) become more deeply embedded in academic workflows, they are not only changing how scholars write, but also reshaping how we conceive of authorship, collaboration, and the production of knowledge itself. Their influence prompts a reconsideration of long-held assumptions in academic culture—particularly regarding who (or what) contributes to scholarship, and how.

One emerging implication is the need to reimagine academic authorship. If a researcher uses an LLM to help brainstorm, structure, and phrase a paper, how different is that from working with a graduate assistant or junior colleague? The analogy is increasingly apt. Just as early-career researchers often contribute by offering feedback, drafting sections, or synthesizing sources, LLMs can perform many of the same functions (Korinek, n.d.). However, unlike human collaborators, LLMs lack accountability, intent, and the capacity for original insight—attributes many consider central to scholarly contribution (Bender et al., 2021; COPE Council, 2023).

Still, in institutions with limited access to graduate support or collaborative research networks, LLMs may offer a form of intellectual companionship that democratizes participation in scholarly discourse. For faculty at teaching-intensive institutions or in the global south, who may have fewer resources for research infrastructure, LLMs can provide affordable, 24/7 feedback and writing assistance (Yu et al., 2023). This could reduce disparities in academic productivity and access to publication, though

it may also reinforce dependence on tools developed and controlled by a small number of private tech companies. This concentration of AI development within major firms can exacerbate existing inequalities, as these entities may prioritize commercial interests over equitable access to AI resources. Addressing these challenges requires a comprehensive policy approach to ensure AI's beneficial development and diffusion, including measures to promote competition, enhance accessibility, and address job displacement and inequality (Filippucci et al., 2024).

Another implication lies in the development of new norms and citation practices. As more scholars use LLMs to generate content, review literature, or assist with synthesis, the academic community will need to develop clearer standards for how to disclose and credit such use. Some journals have begun implementing policies requiring authors to declare AI involvement in methods or acknowledgments sections (Stokel-Walker, 2023). However, much of this policy-making is reactive and fragmented. Without a coordinated response, inconsistencies may lead to confusion, mistrust, or even reputational damage.

The integration of LLMs also invites a reconsideration of pedagogical and institutional goals. If writing becomes less about constructing prose and more about guiding or refining generated text, educators will need to rethink how they assess writing skills and what learning outcomes they prioritize (Cotton et al., 2024). Similarly, institutions may need to adapt promotion and tenure guidelines to clarify expectations about originality, contribution, and acceptable use of AI tools in scholarly output.

Finally, there is a broader epistemic shift at stake. If LLMs are regularly used to produce literature reviews, abstracts, and even theoretical frameworks, we risk eroding the reflective, dialogic process that underpins knowledge creation. The automation of scholarly writing may speed up production, but it also threatens to flatten originality and reduce the diversity of academic voice (van Dis et al., 2023). Whether this transformation leads to a flourishing of ideas or a homogenization of discourse will depend on how scholars choose to engage with these tools—critically, creatively, and ethically.

Conclusion

The integration of large language models like ChatGPT into academic writing marks a profound shift in how knowledge is produced, shared, and credited. These tools are not merely technological novelties; they are already active participants in scholarly workflows—generating prose, shaping arguments, and even simulating intellectual partnership. As this paper has argued, their rise compels scholars to grapple with complex ethical, epistemological, and practical questions: What counts as authorship when machines can write? How do we preserve academic rigor when generative fluency can mask conceptual shallowness? And how do we ensure that these tools empower rather than homogenize or marginalize scholarly voices?

LLMs are, by design, enthusiastic collaborators. They amplify ideas, smooth sentences, and offer endless variations on a theme. But they rarely challenge assumptions or provide meaningful critique (although the human author here has tried to push it to do so). In this way, they function less like peer reviewers and more like agreeable junior colleagues—useful, yes, but potentially uncritical. This raises an important challenge for future development: can we build LLMs that not only support writing but sharpen thinking?

Moreover, as the academic community negotiates how LLMs are used, cited, and disclosed, we must also confront a more uncomfortable truth: many of the habits that LLMs expose—surface-level citation, formulaic prose, reliance on heuristics—are already common in scholarly practice. Rather than treating these tools as external threats to academic integrity, it may be more productive to see them as mirrors—reflecting both the strengths and the weaknesses of how we write, teach, and evaluate knowledge today.

Ultimately, this paper has not argued for or against the use of LLMs in academia. Instead, it has suggested that they are already here, already working alongside us, and has shown a process of using them. You may view the original ChatGPT thread that drafted this paper by going to ==URL==. The real question, then, is not whether we should let them in, but how we live and write with them—responsibly, reflectively, and with a renewed attention to what matters most in scholarly work: clarity, rigor, creativity, and care.

References

Bartlett, M. J., Arslan, F. N., Bankston, A., & Sarabipour, S. (2021). Ten simple rules to improve

academic work–life balance. *PLoS Computational Biology*, *17*(7), e1009124.

https://doi.org/10.1371/journal.pcbi.1009124

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the Dangers of Stochastic

Parrots: Can Language Models Be Too Big? 🦜 . *Proceedings of the 2021 ACM Conference on

Fairness, Accountability, and Transparency*, 610–623. https://doi.org/10.1145/3442188.3445922

COPE Council. (2023, February 13). *Authorship and AI tools*. COPE Position - Authorship and AI -

English. https://doi.org/10.24318/cCVRZBms

Cotton, D. R. E., Cotton ,Peter A., & and Shipway, J. R. (2024). Chatting and cheating: Ensuring

academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*,

61(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Defining the Role of Authors and Contributors. (n.d.). Retrieved March 27, 2025, from

https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-

authors-and-contributors.html

Filippucci, F., Gal, P., Jona-Lasinio, C., Leandro, A., & Nicoletti, G. (2024). *The impact of Artificial

Intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and

policy challenges* (OECD Artificial Intelligence Papers No. 15; OECD Artificial Intelligence

Papers, Vol. 15). https://doi.org/10.1787/8d900037-en

International Committee of Medical Journal Editors. (2025). *Recommendations for the Conduct,

Reporting, Editing, and Publication of Scholarly Work in Medical Journals*.

https://www.icmje.org/recommendations/

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, D., Xu, Y., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023).

Survey of Hallucination in Natural Language Generation. *ACM Comput. Surv.*, *55*(12), 248:1-

248:38. https://doi.org/10.1145/3571730

Johnson, R., Watkinson, A., & Mabe, M. (2018). *The STM Report, 5th edition: An overview of scientific and scholarly publishing*.

https://policycommons.net/artifacts/1575771/2018_10_04_stm_report_2018/2265545/

Karimi, H. (2024). *AI in Education: Friend or foe? A researcher's perspective*.

https://www.bera.ac.uk/blog/ai-in-education-friend-or-foe-a-researchers-perspective

Korinek, A. (n.d.). *NBER WORKING PAPER SERIES*.

Meyer, J. G., Urbanowicz, R. J., Martin, P. C. N., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N., Won, K. J., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, *16*(1), 20. https://doi.org/10.1186/s13040-023-00339-9

OpenAI. (2024, January 12). *GPT-4*. https://openai.com/index/gpt-4-research/

Rainie, J. A. and L. (2023, February 24). The Future of Human Agency. *Pew Research Center*. https://www.pewresearch.org/internet/2023/02/24/the-future-of-human-agency/

Rawat, S., & Meena, S. (2014). Publish or perish: Where are we heading? *Journal of Research in Medical Sciences : The Official Journal of Isfahan University of Medical Sciences*, *19*(2), 87–89.

Repko, A. F., & Szostak, R. (2016). *Interdisciplinary Research: Process and Theory*. SAGE Publications.

Stokel-Walker, C. (2023). ChatGPT listed as author on research papers: Many scientists disapprove. *Nature*, *613*(7945), 620–621. https://doi.org/10.1038/d41586-023-00107-z

Sword, H. (2012). *Stylish Academic Writing*. Harvard University Press. https://www.jstor.org/stable/j.ctt2jbw8b

Tenopir, C., King, D. W., Edwards, S., & Wu, L. (2009). Electronic journals and changes in scholarly article seeking and reading patterns. *Aslib Proceedings*, *61*(1), 5–32. https://doi.org/10.1108/00012530910932267

Touvron, H., Lavril, T., Izacard, G., Martinet, X., Lachaux, M.-A., Lacroix, T., Rozière, B., Goyal, N., Hambro, E., Azhar, F., Rodriguez, A., Joulin, A., Grave, E., & Lample, G. (2023). *LLaMA: Open*

and Efficient Foundation Language Models (No. arXiv:2302.13971). arXiv.

      https://doi.org/10.48550/arXiv.2302.13971

van Dis, E. A. M., Bollen, J., Zuidema, W., van Rooij, R., & Bockting, C. L. (2023). ChatGPT: Five

      priorities for research. *Nature*, *614*(7947), 224–226. https://doi.org/10.1038/d41586-023-00288-7

Weidinger, L., Uesato, J., Rauh, M., Griffin, C., Huang, P.-S., Mellor, J., Glaese, A., Cheng, M., Balle,

      B., Kasirzadeh, A., Biles, C., Brown, S., Kenton, Z., Hawkins, W., Stepleton, T., Birhane, A.,

      Hendricks, L. A., Rimell, L., Isaac, W., … Gabriel, I. (2022). Taxonomy of Risks posed by

      Language Models. *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and

      Transparency*, 214–229. https://doi.org/10.1145/3531146.3533088

Yu, D., Rosenfeld, H., & Gupta, A. (2023, January 16). *The 'AI divide' between the Global North and

      Global South*. World Economic Forum. https://www.weforum.org/stories/2023/01/davos23-ai-

      divide-global-north-global-south/