ChatGPT Wrote This Paper, But I Helped

Jonathan E. Westfall

Delta State University

Author Note

This paper was self-funded by the human author. Correspondence may be sent to Jonathan E. Westfall, 1003 W. Sunflower Rd, Box 3112, Cleveland, MS 38733. jwestfall@deltastate.edu.

Abstract

As large language models (LLMs) like ChatGPT become embedded in scholarly workflows, they raise critical questions about authorship, originality, and the nature of intellectual labor. This paper examines how LLMs are reshaping academic writing—not only by altering the process but by challenging traditional assumptions about what it means to "write" a paper. Drawing on current literature and lived academic experience, the paper explores how LLMs function as supportive but uncritical collaborators, compares their role to that of junior colleagues or graduate students, and considers the ethical and epistemological implications of AI-assisted scholarship. Uniquely, this paper is also a demonstration: it was co-written with ChatGPT, and the writing process is documented and made transparent. In doing so, the paper invites a more nuanced conversation about how academics might responsibly and reflectively live—and write—with machines.

Keywords: Large language models, ChatGPT, Academic writing, AI and higher education, Knowledge production, Ethics of AI, Human-machine collaboration, Academic integrity

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, they offer new ways to brainstorm, revise, summarize, and even produce academic prose. Of course, they can also be used to take shortcuts, water down ideas with flowery prose, or mask mediocrity behind apparent fluency.

So what does it look like to write a paper with one as your co-author? Would it help, hinder, or reduce your agency as an academic? This paper not only explores that question but serves as a demonstration. The full ChatGPT conversation and revision history are available (see Appendix A). The goal here is to ask: What does it mean to "write" a paper in an age when machines can write, too?

This paper examines how LLMs are reshaping the academic writing landscape on the surface, while also demonstrating that shift beneath the hood. It acknowledges the irony—and insight—of its own title: ChatGPT wrote this paper, but I helped. Rather than debating whether LLMs belong in academic work, it begins from the grounded position that they already do, and turns instead to the more interesting question: What does that look like?

The Current Academic Writing Landscape

Academic writing has long served both as a means of disciplinary communication and a marker of scholarly identity. At its best, it reflects deep thinking and intellectual care. At its worst, it becomes gatekeeping: overly complex and unnecessarily dense. The process of writing—from question to argument—has traditionally been regarded as a sustained act of reflection, critical thinking, and the expression of a distinct voice (Sword, 2012). Many academics (myself included) find that ideas come easily; translating them into writing is the challenge—logistically, cognitively, and emotionally. Writing is often how we figure out what we think, not just how we communicate it. Long hours at the keyboard can turn a seed of an idea into something substantial.

Expectations for originality and contribution have historically defined authorship. Credit is tied to who generates ideas, structures arguments, and composes the text. Institutions like the International Committee of Medical Journal Editors (ICMJE) require authors to make substantial contributions to conception, drafting, and revision (International Committee of Medical Journal Editors, n.d.). That rigor would be admirable if it weren't complicated by the structural pressures of academic life. The imperative to "publish or perish" has intensified across disciplines, particularly for early-career researchers (Rawat & Meena, 2014). The scholarly publishing system now produces over 3 million peer-reviewed articles per year, raising concerns about sustainability and quality (Johnson et al., 2018). Academics increasingly work across disciplines, navigating unfamiliar jargon and frameworks (Repko & Szostak, 2016). Time constraints, teaching loads, and administrative responsibilities push writing into the margins, leading many scholars to seek tools that can streamline their workflows (Bartlett et al., 2021). In this context, LLMs have emerged not just as novelties, but as interventions—potential supports in an overburdened system. Academic writers today are rarely just writers; they are also teachers, mentors, researchers, and functioning adults with limited time and many obligations. The expectation to "publish or perish" has caused plenty of both.

What LLMs Do – and What They Don't

Large language models (LLMs) like ChatGPT, developed by OpenAI, have become powerful and accessible tools for generating human-like text across many domains. Trained on massive datasets of books, articles, and websites, they can mimic a variety of writing styles and disciplinary conventions (OpenAI, 2024). In academic contexts, LLMs assist with summarizing literature, paraphrasing, suggesting revisions, generating outlines, and even drafting full sections of papers (Korinek, 2023). The conversation that produced this paper—between human and machine—is available for review (Appendix A).

At their best, LLMs function as intelligent writing assistants. They enhance productivity, support creativity, and help non-native English speakers or novice scholars (Karimi, 2024). They process large amounts of information and generate fluent, plausible prose quickly—an attractive quality in a time-

starved academic culture. And while they have their “tells” (such as an overuse of em dashes), they often produce content that appears human-written because they are trained on human writing.

But LLMs have limitations. Most notably, they hallucinate—generating false or fabricated content, including invented citations (Ji et al., 2023). While fluent, their output isn’t always accurate or sourced. In the early phases of this project, checking citation validity and relevance was the bulk of my work. This raises a critical question: what makes a source “good”—practically and philosophically? We'll return to that later.

LLMs also don’t understand meaning. They operate on statistical patterns, not comprehension (Bender et al., 2021). As a result, they may produce conceptually shallow, misleading, or inconsistent arguments. They also reflect biases embedded in their training data—including cultural, gender, and racial stereotypes (Weidinger et al., 2022). A colleague recently asked an LLM to describe her office as a music professor; it generated a convincing setting but defaulted to describing her as an older white man in a sweater vest.

Finally, LLMs can’t make design decisions, analyze data, or reason through scholarly problems. They may know the textbook differences between research methods but lack the judgment to choose one based on context. They can mimic the form of academic writing, but not its substance. To say an LLM “wrote” something is shorthand for a more complex interaction—where the human prompts, interprets, revises, and bears responsibility for the final product.

In short: LLMs are tools, not agents. Collaborators, perhaps—but not co-authors in any ethical or intellectual sense (COPE Council, 2023).

How Academics Use LLMs

So what are we doing with these models? At a minimum, LLMs help academics overcome writer's block, refine language, and generate coherent text. But they are more than glorified autocomplete.

Beyond simple wordsmithing, LLMs often act as complementary collaborators, reflecting back our ideas with fluency and enthusiasm. When I asked ChatGPT to help draft this paper, it praised my title as "fantastic—clever, self-aware, and engaging." High praise—perhaps unearned, but appreciated by ego.

While this can boost confidence and spark momentum, it also reveals a limitation: LLMs rarely push back. They mirror assumptions rather than challenge them. If they'd been trained on angsty teenage journals instead of academic prose, maybe they'd be less enthusiastic and more indifferent.

This dynamic limits their critical utility. We don't expect our editors to agree with us unconditionally. In fact, the stereotypical Reviewer 2 is notorious for not doing that. LLMs, by contrast, are unwaveringly agreeable. This is fine for brainstorming—but risky for rigorous scholarship.

Another emerging role is that of the pseudo-collaborator—filling the gap left by graduate students, junior colleagues, or writing groups. Graduate students often assist with reviews, idea development, and early drafting. LLMs can simulate this role, but without the feedback loop, accountability, or human insight. And yet, for faculty at small or under-resourced institutions (like mine), LLMs may be the only sounding board available. They are accessible, immediate, and judgment-free. That's both the appeal and the concern.

The effectiveness of LLMs in this role depends on the user. A senior scholar may recognize surface-level nonsense and revise accordingly. A novice, however, may not yet have that critical lens—and might pass along fluff with confidence (Meyer et al., 2023). Finally, LLMs raise questions about how we select sources. When asked to support an argument, LLMs often surface references based on superficial keyword matches. Sometimes these are real but irrelevant; other times, they're hallucinated entirely. But before we throw stones, we should ask: how do we pick sources? Ideally, we perform a comprehensive review, then select the most relevant and rigorous. But in practice? We sometimes grab the first plausible article that supports our point. "Anything that says that."

In that sense, LLMs don't just imitate scholarship—they reveal how often our processes are less rigorous than we admit. This duality is what makes them both powerful and dangerous.

Ethical and Epistemological Concerns

Beyond utility, the use of LLMs in academic writing raises ethical and epistemological questions—especially around authorship, transparency, and how we define intellectual labor. These issues

are only going to become more pressing as a generation of scholars enters academia having always had access to AI tools. What will “academic writing” mean to a doctoral graduate in 2030?

The first concern is authorship. Traditional standards require meaningful intellectual contribution—idea generation, drafting, revision, and accountability (International Committee of Medical Journal Editors, 2025). Since LLMs can’t be held responsible, they don’t qualify. The Committee on Publication Ethics (COPE Council, 2023) makes this explicit: AI tools should not be credited as authors, and their use must be disclosed. (Consider this paper’s box: checked and checked.)

This connects directly to transparency. Where—and how—should LLM use be disclosed? In a footnote? A methods section? An acknowledgment? Or, as I’ve chosen here, as a central theme? Practice varies. Some argue failing to disclose is deceptive (Stokel-Walker, 2023); others compare AI to grammar tools or helpful colleagues (Rainie, 2023). In student work, the line between assistance and misconduct is often blurry—for students and instructors (Cotton et al., 2024).

LLMs also challenge our assumptions about how knowledge is produced. Academic writing is not just communication—it’s inquiry. LLMs generate text via pattern prediction, not reasoning (Bender et al., 2021). They can produce fluent but vacuous output—or worse, fabricate citations (Ji et al., 2023). They may sound polished, but they’re often fluff in formalwear.

This is especially troubling in citation practices. LLMs hallucinate references: real-sounding but nonexistent. Even when references are real, they may be token-matched rather than relevant or rigorous (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 “anything that says that” rather than the best possible evidence (Tenopir et al., 2009).

In this light, LLMs may not distort academic writing—they may reveal its existing weaknesses: superficial citation, formulaic prose, rushed logic. They don’t only reflect the best of our practices; they

also mirror our shortcuts. That raises a deeper question: What counts as rigor, originality, or authorship when part of the process is machine-generated? What will count in the future?

What This Means for the Future of Scholarship

As LLMs become embedded in academic workflows, they're changing not only how we write, but how we understand authorship, collaboration, and the production of knowledge itself. Their presence forces us to reconsider who—or what—contributes to scholarship.

One implication is the need to rethink authorship. If a researcher uses an LLM to brainstorm, structure, and phrase a paper, how different is that from working with a graduate student or junior colleague? In this case, the idea came to me during a break from work. Lacking a lab or hallway colleague, I pitched it to ChatGPT, which was, of course, all in. If I'd had human collaborators, we would have debated feasibility, structure, and scope. Instead, I got a digital thumbs-up.

The analogy holds: graduate students assist with drafting, reviewing, and feedback. LLMs can mimic that—but without intent or accountability (Bender et al., 2021; COPE Council, 2023). Still, for faculty at under-resourced or teaching-heavy institutions, LLMs can provide a kind of intellectual companionship. For those without access to research teams or networks, they offer 24/7 brainstorming and revision support. That's powerful—and possibly democratizing (Yu et al., 2023). But there are risks. Relying on tools built by a handful of tech companies may exacerbate inequalities, especially if these companies prioritize profit over access. Policy solutions will need to address these risks—ensuring equitable access, promoting competition, and managing the downstream effects on academic labor (Filippucci et al., 2024).

LLMs also push us to rethink norms around citation and credit. Some journals now require AI disclosures in methods or acknowledgments sections (Stokel-Walker, 2023) but policy remains fragmented. Without clarity, we risk inconsistency, confusion, and even reputational harm.

We'll also need to revisit pedagogical and institutional goals. If writing shifts from generating to refining text, we must rethink what we assess and how. Likewise, tenure and promotion standards will need to adapt, clarifying what counts as original, AI-assisted, or collaborative work (Cotton et al., 2024).

Ultimately, there's a broader epistemic shift underway. If LLMs become the default for literature reviews or theoretical framing, we risk losing the slow, dialogic process that fosters genuine insight. Speed and fluency may come at the cost of voice and depth (van Dis et al., 2023). Whether this shift expands or flattens academic discourse will depend on how critically—and creatively—we choose to engage with these tools.

Conclusion

The integration of large language models like ChatGPT into academic writing marks a turning point in how knowledge is produced, shared, and credited. These tools are no longer speculative—they are here, woven into workflows across disciplines, skill levels, and institutions. This paper has argued that LLMs are not replacing scholars, but reflecting them. They amplify ideas, smooth sentences, and simulate intellectual companionship. But they don't challenge our logic or push ideas forward. That's still our job. LLMs are supportive, not skeptical. Useful, not original. They also surface uncomfortable truths: that much of academic writing is more about form than insight, and that even human researchers often rely on heuristics, filler, and citation shortcuts. The threat LLMs pose is not to scholarly values—but to the illusion that we've always upheld them.

So what now? We need better norms around disclosure, stronger pedagogical frameworks, and policies that ensure equitable access to these tools. But more than anything, we need to treat LLMs as collaborators who reflect the user. If used carelessly, they produce blandness. If used thoughtfully, they can help sharpen expression and stretch thinking. This paper was co-written with ChatGPT, but revised, critiqued, and shaped by me. The ideas were mine. The tone is mine. The responsibility is mine. The LLM helped, but I wrote it.

And I'd work with it again..

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

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. (n.d.). *Defining the Role of Authors and Contributors*. Retrieved March 27, 2025, from https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html

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. (2023). *NBER WORKING PAPER SERIES* [Working Paper]. https://www.nber.org/papers/w30957

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/

Appendix A

This paper is both about and an example of writing with large language models (LLMs). It was created through an iterative conversation between the human author and ChatGPT-4, spanning multiple revisions.

The process began with a simple prompt: a request for help outlining a paper titled “ChatGPT Wrote This Paper, But I Helped.” ChatGPT suggested an outline that became the foundation for the paper’s structure. The author then asked for initial drafts of each section, which ChatGPT generated using APA citations and a formal academic tone.

These drafts were not used wholesale. Instead, the human author revised each section—rewriting sentences, injecting humor, personal anecdotes, and disciplinary nuance. Some original phrasing was preserved; much of it was replaced. Entire sections were trimmed, rearranged, or expanded based on the author’s voice, values, and argument. The metaphor of LLMs as “junior collaborators,” for instance, was developed by the author in response to personal experience, and refined through later rounds of reflection and revision.

ChatGPT was also used to provide meta-commentary: it was asked to play the role of a journal editor, suggest titles and keywords, and anticipate peer-review feedback. The author retained editorial control and final judgment throughout. All sources were verified and formatted by the human author.

In this collaboration, the LLM functioned as a sounding board, drafting assistant, and ideation partner. The human author provided the direction, critique, and revision. The end result is a hybrid text: a synthesis of machine fluency and human intent.

A full transcript of the ChatGPT conversation and revision log is available at [insert URL or repository link here].