The Use of Artificial Intelligence in Academic Research

Gordon M. Graham
E&Y Canada (retired partner)
gord.m.graham@ca.ey.com

Jennifer Wu Tucker
Fisher School of Accounting
Warrington College of Business
University of Florida
jenny.tucker@warrington.ufl.edu

First draft: October 9, 2025
This draft: November 19, 2025

Forthcoming in *Accounting Horizons*

Gord Graham is an editor of *Accounting Horizons*. Jenny Tucker is the senior editor of *Accounting Horizons*. The opinions expressed in this article represent the authors', not those of the Journal or its owner—the American Accounting Association (AAA). Any errors are ours. We have benefited from the discussions with Jackie Hammersley (AAA's Vice President for Research), Karen Sedatole (Chair of AAA's Research & Publications Committee), Senyo Tse (ad hoc editor of *Accounting Horizons*), and David Twiddy (AAA's Director of Publications). Our article has benefited from the Editors' Panel, "The rise of artificial intelligence: Implications for research and publishing," chaired by Ileana Steccolini with Amir Amelzadeh, John Dumay, Andrei Filip, Kathryn Kadous, Helen Tregidga, and Jenny Tucker as panelists at the 2025 European Accounting Association Congress in Rome, Italy. See https://aaahq.org/Research/Journals/Accounting-Horizons for information about *Accounting Horizons*. See https://aaahq.org/Research/Journals/Accounting-Horizons/Guide-for-Authors#use for information about AAA's publication ethics policies. Tucker thanks the J. Michael Cook/Deloitte Professorship for financial support and thanks her coauthors of AI-related research projects for the learning experience.

The Use of Artificial Intelligence in Academic Research

Abstract

This editorial is in response to the emerging issue of using artificial intelligence (AI) in academic research. We share our experience as editors in dealing with such an issue, provide a framework for understanding this issue, and encourage the academic community to ask questions about research, scholarship, and the role of humans in the world of increased availability of new technology. Our editorial should be of interest to authors, reviewers, editors, and readers of academic research as well as those who are involved in training future generations of researchers.

Keywords: artificial intelligence; AI; ChatGPT; technology; academic research.

Authors' Statement on the Use of Generative AI or AI-Assisted Technology:
The coauthors did not use Generative AI or AI-Assisted Technology at any stage of document preparation for this manuscript.

1. Introduction

The use of artificial intelligence (AI) in academic research has been evolving at an accelerated rate. Even though "artificial intelligence" was coined as a concept by a team of researchers 70 years ago, its earthshaking effect was not broadly felt until the introduction of ChatGPT to the public in November 2022. Since then, generative AI has not only been a buzzword but also permeated many aspects of our lives (Berger, Cai, Qiu, and Shen 2025). In corporate annual reports for 2023, for example, 12% of the U.S. publicly listed companies disclosed using AI in their operations and another 6% mentioned AI in other contexts (Tucker, Wang, and Zhao 2025). Based on those corporate reports, education was the industry with the highest percentage (45%) of its members disclosing AI use. As a general-purpose technology, AI use in academic research should not be a surprise. Yet, we were still surprised when we learned that the initial drafts of some manuscripts are generated by AI ("machine") with human-provided prompts. We have also observed or suspected the use of AI in preparing referee reports and conference discussion slides as well as non-existent references provided by authors or reviewers. These observations started us on a journey to learn how AI can be used in academic research, how to reconcile the roles of human vs. machine, and how researchers can really add value.

The first thing we learned in this journey is that we are not alone—the issue of AI use in academic research is not unique to our journal, the American Accounting Association (AAA), or the field of accounting. For example, Kobak, González-Márquez, Horvát, and Lause (2025) examine more than 15 million abstracts of biomedical publications in 2010-2024 and find a sudden increase in 2024 in the frequency of about 400 style words, 80% of which are verbs (e.g., "delve" and "underscore") or adjectives (e.g., "potential" and "significant"). The authors estimate that at least 14% of the abstracts published in 2024 were assisted by large language models (LLMs)—the models behind almost all popular generative AI tools today. Kwon (2025) surveyed over 5,000 researchers in March 2025 and reports that researchers are split about whether it is acceptable to generate text with AI for all or part of a paper. Regarding reviewing other researchers' work, Naddaf (2025b) reports that, based on a survey of about 5,000 researchers conducted by Wiley in February 2025, 19% of the respondents acknowledge using LLMs to increase the speed and ease of their review assignments. Moreover, AI is capable of producing complete research papers in a short order upon human prompts (Kolata 2025; Novy-Marx and Velikov 2025). What are the implications of AI for authors, reviewers, and editors?

In this editorial, we share our thoughts on the broad issue of AI use in academic research. In Section 2, we review what people might mean when they use the word "AI," what capabilities the technology can have, and what popular products are available to researchers who wish to embrace the technology. We then discuss existing policies by major journals for authors and reviewers on the use of AI. In Section 3, we provide a framework based on the type of tasks involved in an individual's academic research production. We discuss the extent to which those tasks can be performed by AI and, more importantly, the pros, pitfalls, and unintended consequences of using AI instead of humans for the tasks. We encourage researchers to consider three key questions when deciding whether to use AI for a given task. First, what would the use of AI mean to their research project and scholarship? Second, would AI improve their research efficiency and quality? Last, what might be the risks of using AI for a given type of task and how can researchers mitigate those risks? In the final subsection, we explain the factors we consider as editors in evaluating manuscripts that have used AI. In Section 4, we discuss some immediate questions faced by the research community and three additional challenges, including journal submissions of AI-generated *complete* papers under disguise, training future generations of researchers, and rediscovering the value of humans in research production. Section 5 concludes.

Our editorial intends to share our experience, open the dialogue, and encourage discussion about the use of AI in academic research. The editorial does not intend to pass judgment on the types of AI use or to interpret, defend, or criticize any AI policy. Academics have embraced technologies in the past, such as technology for preparing documents (e.g., MS Word), analyzing data (e.g., SAS and Stata), and copyediting (e.g., Grammarly). We believe that our editorial would be of interest to authors, reviewers, editors, and readers of academic research at the junction of AI injecting excitement, uncertainty, and confusion in the academic community.

2. AI and Existing AI Policies at Major Journals

2.1 What is AI?

There is no single agreed-upon definition of AI. In 1955, when a team of mathematicians, computer scientists, and neurologists came up with the terminology, they meant a type of intelligence that a machine can be made to simulate (McCarthy, Minsky, Rochester, and Shannon 2006). In announcing the Nobel Prizes in physics in 2024 to John J. Hopfield and Geoffrey Hinton for their contributions to AI, the Nobel Committee wrote, "When we talk about artificial intelligence, we often mean machine learning using artificial neural networks."[1] The concepts of machine learning and artificial neural networks (ANNs) are both decades old. Machine learning means automated detection of meaningful patterns in data by using algorithms on training data. ANNs use layers of nodes (called "artificial neurons"), which are connected by edges, whose weights are optimized via machine learning. When many hidden layers are used, that is "deep" learning, which became prominent since such a model won the annual ImageNet contest in 2012.

[1] See https://www.nobelprize.org/prizes/physics/2024/hopfield/facts/.

Almost all of today's LLMs are built on transformers—the most popular type of deep learning model since 2017. For example, GPT in ChatGPT stands for Generative Pretrained Transformers, where "generative" means the model can generate human-like content in response to user prompts. Even though the meaning of AI can be broad when organizations highlight their use of AI for promotion purposes, the issue of AI use in academic research is primarily about generative AI tools that are built on LLMs.[2]

Transformers have an attention mechanism, which allows the machine to "look around" the text and pay attention to more relevant context clues. Consequently, the model can learn contextual relations and understand a text "in a way that is similar to how humans analyze language" (Wang 2023). In addition, like all deep learning models, transformers can apply the knowledge learned from one task to another ("transfer learning") and therefore accumulate knowledge. The transfer learning ability also allows a model to be fine-tuned by researchers with the corpus of their interest. As a result, LLMs can classify, extract, summarize, and aggregate text. Today's AI tools can process and generate text, images, and videos.[3]

As the first generative AI application available to the public, ChatGPT, designed by OpenAI, is still the most commonly used AI tool.[4] Other products with similar functionalities include Gemini (by Google), Copilot (by Microsoft), Claude (by Anthropic), and Grok (by xAI). In addition, Gemini Deep Research and Perplexity.ai are newer tools that researchers increasingly use to gather information and synthesize literature on topics of interest. AI technology is still

[2] The concept of language model is decades old (Weizenbaum 1966). A language model can "predict what word comes next or more generally assign probability to a piece of text sequence" (Bochkay, Brown, Leone, and Tucker 2023, p.12). Thus, the essence of a language model is prediction, which is the main goal of machine learning. "Large" in "LLM" means that the language model is trained on large quantities of data to decipher complex patterns—the strength of deep learning. "Generative" means that the model can output the predicted text sequence. Bochkay et al. provide an overview of machine learning and discuss the evolution of deep learning.

[3] Some deep learning models, such as convolutional neural networks, were initially designed for processing images.

[4] About 81% of the respondents to Wiley's survey in February 2025 have used ChatGPT, and only a third of the respondents have heard of other generative-AI tools (Naddaf 2025a).

advancing. For example, reasoning models are designed to go through a sequence of logical steps and can address complex problems with greater depth and precision.[5] More advanced tools will become available to researchers. As technology becomes more powerful, it is urgent to ask how researchers can embrace technology to further their *human* contribution to research.

2.2 Existing AI policies at major publishing houses

Elsevier, Springer Nature, and Wiley are the three largest publishing houses for journals, with each house publishing about 2,000-3,000 journals. These publishing houses have their respective AI policies for authors and reviewers.[6] Elsevier allows authors to use generative AI or AI-assisted technologies in manuscript preparation but requires authors' oversight and disclosure, which includes the name of the AI tool used, the purpose of the use, and the extent of authors' oversight. To protect authors' confidentiality and proprietary rights, Elsevier does not allow reviewers to upload any part of manuscripts or referee reports into generative AI tools, even if the purpose is for improving language and readability.

Springer Nature does not prohibit authors' use of AI but requires proper documentation. The publishing house does not allow reviewers to upload manuscripts into generative AI tools. Wiley "embraces artificial intelligence (AI) as a transformative technology" and provides guidelines on authors' use of AI in writing.[7] It emphasizes that authors should preserve their authentic voice (i.e., using AI as a companion, not a replacement) and provide transparent disclosure (e.g., the purpose of AI use, the extent of human oversight, and whether key arguments

[5] See https://odsc.medium.com/the-rise-of-reasoning-models-unlocking-the-next-phase-of-ai-8d82683cb5ec.
[6] The policies were accessed on October 5, 2025. See details at https://www.elsevier.com/about/policies-and-standards/generative-ai-policies-for-journals, https://www.springernature.com/gp/policies/editorial-policies (and then click on the BMC link), and https://authorservices.wiley.com/ethics-guidelines/index.html#5.
[7] See https://www.wiley.com/en-us/terms-of-use/ai-principles and https://www.wiley.com/en-us/publish/book/ai-guidelines?utm_source=newsroom&utm_medium=pr&utm_campaign=AIAuthorGuidelines25&utm_content=presspost. Accessed on October 5, 2025. The latter also includes resources for authors, such as how to track AI use.

or conclusions are affected by AI use). Wiley also states that AI tools "cannot fulfil the role of, nor be listed as, an author of an article" and "The final decision about whether use of a GenAI tool is appropriate or permissible in the circumstances of a submitted manuscript or a published article lies with the journal's editor or other party responsible for the publication's editorial policy." Wiley does not allow reviewers to upload any part of a manuscript into a generative AI tool.

All three publishing houses note that they continue to monitor developments and will refine their policies as appropriate. In fact, Elsevier's policy has recently shifted. For example, its policy accessed as late as September 1, 2025 stated, "Where authors use generative AI and AI-assisted technologies in the writing process, these technologies should only be used to improve readability and language of the work." With Elsevier's policy shift, the AI polices at the three publishing houses are converging.[8]

2.3 Existing AI policies at major accounting journals

The Top 5 academic journals in accounting are *The Accounting Review* (*TAR*), *Contemporary Accounting Research* (*CAR*), *Journal of Accounting and Economics* (*JAE*), *Journal of Accounting Research* (*JAR*), and *Review of Accounting Studies* (*RAST*). *JAE* is published by Elsevier, *RAST* is published by Springer Nature, *JAR* and *CAR* are published by Wiley, and *TAR* belongs to the AAA.

JAE has recently updated its AI policy to mirror Elsevier's new policy.[9] *RAST*'s policy is identical to Springer Nature's. *JAR*'s policy follows Wiley's pivot on transparent disclosure of AI

[8] The AI policies generally prohibit using AI to generate images or alter existing copyrighted images.
[9] See https://www.sciencedirect.com/journal/journal-of-accounting-and-economics/publish/guide-for-authors for *JAE*, https://www.springer.com/gp/editorial-policies/artificial-intelligence--ai-/25428500 for *RAST*, https://onlinelibrary.wiley.com/page/journal/1475679x/homepage/ForAuthors.html#Submission10 for *JAR*, https://www.caaa.ca/car-editorial-policies for *CAR*, and https://aaahq.org/Research/Journals/The-Accounting-Review/Guide-for-Authors#use for *TAR*. Accessed on October 5, 2025. Note that the revised *TAR* policy was posted on November 14, 2025.

use and emphasizes that AI tools "cannot fulfill the role of, nor be listed as, an author of an article submitted and published in *JAR*" and that "The final decision about whether use of an AIGC tool is appropriate or permissible in the circumstances of a submitted manuscript or a published article lies with the editors," where "AIGC" stands for "AI generated content." *CAR*'s policy is similar to *JAR*'s and states that "The Editors will review and approve its use on a case-by-case basis."

On November 7, 2025, the AAA revised its AI policy for all its 17 journals, including *TAR*.[10] The policy does not explicitly prohibit the use of AI in academic research but requires authors to provide a disclosure statement explaining the tool used, the extent of use, and the reason for using the tool. Like the policies at *JAR* and *CAR*, the *TAR* policy states, "AI cannot fulfill the role of an author." The *TAR* policy further clarifies that "Authors should have contributed sufficiently to merit authorship status" and states that "The final decision about whether use of an AI tool is appropriate or permissible in the circumstances of a submitted manuscript lies with the editor(s)." In addition, the policy requires all submissions to be accompanied by an AI disclosure statement, specifying whether or not AI tools were used and, if so, how.

After these recent policy revisions, AI policies regarding manuscript preparation are converging at the Top 5 accounting journals. All five journals require authors to disclose their use of AI, and all exempt spelling and grammar checking (e.g., Grammarly) as AI use in the definitions. The location of the AI disclosure statement varies from journal to journal; authors

[10] The previous policy required authors to specify the tool(s) used, the extent of use, and the reason(s) for using the tool(s) and provided an example statement, "During the preparation of this work, the author(s) used [NAME TOOL/SERVICE] in order to [EXTENT/REASON]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication." As editors, we conditionally accepted a manuscript, which provided such a statement, "The authors used Open AI's ChatGPT 4o in order to prepare this manuscript. After using this tool/service, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication." Upon inquiry, we learned that the authors generated the initial draft by ChatGPT using a set of prompts. We rescinded the conditional acceptance decision and reevaluated the manuscript with the knowledge provided by the authors' expanded AI disclosure.

should consult each journal's policy. In addition, all these journals have policies for reviewers and editors. For example, to protect authors' confidentiality, the new *TAR* policy prohibits reviewers and editors from uploading any part of a manuscript into an AI tool and prohibits reviewers from uploading any portion of a referee report into an AI tool, even if it is just for the purpose of improving language and readability.

3. The Issue of AI Use at the Project Level

3.1 A framework

We use an archival project as an example to describe four types of activity involved in producing research. For each type, we explain how researchers typically assign credit to the individuals involved and whether and how AI may help. We then take the readers' perspective and discuss the information they might wish to know before consuming the research product.

3.1.1 Activity that is traditionally expected from authors ("Type A activity")

Authors are expected to identify the research question, decide what relations to examine, design how to examine those relations, make inferences from data analyses, and write up the research project for circulation and publication. Individuals materially involved in such activity earn credit as coauthors.

Authors could use AI in brainstorming research questions, critiquing the soundness of a research question, consulting on research design and statistical inferences, and generating sentences, paragraphs, or sections (including the abstract) of a paper.[11] Among these tasks, readers are most divided on the appropriateness of using AI to generate text. According to Kwon's (2025) survey of researchers, 35% would not be comfortable with using AI to generate content *under any*

[11] Carrillo, Stachl, and Talaifar (2024) propose a human-centered machine-assisted approach to generating research ideas in social psychology research.

circumstances, 23% would feel comfortable only if prompts are disclosed, 29% would be comfortable if AI use is disclosed, and 13% would be comfortable even with no disclosure of AI use or prompts.[12]

3.1.2 Activity that can be delegated to research assistants ("Type B activity")

A researcher can delegate some tasks to a research assistant (RA), such as gathering relevant information, reviewing the literature, collecting data, constructing variables, and analyzing data. The RA would serve as an agent to execute tasks under the researcher's instructions and supervision. The researcher assigns credit to human RAs by thanking them in the acknowledgement section of the paper.

AI can perform all the above tasks, but the quality of work may vary from model to model and depend on the complexity of a task and the quality of the prompts.[13] For example, AI can extract climate risk disclosure from a large corpus of corporate annual reports for researchers, classify text provided by researchers into predefined categories, and generate figures based on the provided data. de Kok (2025) shows that ChatGPT can detect nonanswers in 1.15 million Q&A pairs in conference calls with impressive accuracy and efficiency. If AI is used for Type B activity, readers would desire to know what prompts are used, whether a randomly selected machine output has been manually checked and validated, and what the machine performance statistics are. In our view, the researcher should treat AI just as a human RA and provide proper instructions and oversight.

3.1.3 Activity that is typically performed by a copy editor ("Type C activity")

---

[12] In their editorial for *Information Systems Research*, Susarla, Gopal, Thatcher, and Sarker (2023, p.403) state, "We believe that ceding control of the writing and knowledge-creation process to ChatGPT is a mistake."

[13] One of the authors of this editorial has compared sentence classification tasks performed by GPT-4o (API) vs. two human RAs and observed that the machine performed better than one human RA but much worse than the other human RA. In another task, the research team tried both zero-shot and few-shot prompts along with providing feedback to GPT-5, but the machine's performance remained unsatisfactory.

It is common, even among native speakers, to seek the service of a copy editor for checking grammar and typos and improving the readability of a paper. Researchers typically compensate the copy editor for such service. In addition to generative LLMs, some AI tools (e.g., Grammarly) are designed for copyediting. Using AI for such tasks is not only free but also convenient. Such AI use in academic research has faced the least tension.[14]

3.1.4 Activity that is typically performed by good colleagues ("Type D activity")

Conversations with colleagues may spark a research idea. A research workshop presentation or discussion may inspire one to pursue a new project. Researchers benefit from circulating their working papers and receiving feedback. Researchers can also consult their colleagues on specific technical issues. Good colleagues receive credit in the acknowledgement of a paper. With the increased capabilities of AI tools, researchers may consult with and benefit from "machine" colleagues. For example, one can use an AI tool called Paper-Wizard (https://paper-wizard.com/) to generate a referee report of their own work for the purpose of polishing their paper before submitting it to a journal.

3.2 Key questions to consider

Given the range of tasks that AI can perform, researchers might consider what research tasks should be given to AI and when AI use can be beneficial. We suggest that researchers ask three questions.

3.2.1 What would the AI use mean to the research project and the researcher's scholarship?

The goal of a research project is to explore and discover new knowledge to contribute to the body of research so that our understanding of a given phenomenon or topic is one step closer

[14] Kwon (2025) reports that about 10% of her survey respondents believe that using AI to edit a paper is not appropriate under any circumstances. Editing includes changing content, structure, and style, and is broader than copyediting.

to the truth. The key to good research is originality of thought. However, AI tools are trained to imitate humans with existing public knowledge and cannot be trusted with generating original thought.

Moreover, researchers accumulate knowledge, learn new skills, and sharpen their critical thinking with every research project they conduct. This is how scholarship is developed. Some uses of AI may assist scholarship development. For example, AI tools can help researchers brainstorm. This task would take advantage of the strength that LLMs are trained on billions or trillions of words sourced from the public domain so that the machine might bring out aspects that the researcher has not paid attention to.

However, other uses of AI might short circuit scholarship development. For example, academic writing is an important component of scholarship because research must be well communicated to make an impact. One of the best-known academic writers is Deirdre McCloskey, who lists the top two tips for economists as "writing is the economist's trade" and "writing is thinking" (McCloskey 1985). Moreover, writing is a process of deepening existing thought or cultivating original thought—some breakthroughs can come from such engagements. Using AI to generate text would deprive a researcher, especially a junior researcher, of an opportunity to craft their skills in writing and thinking as well as aha moments from close engagement with the subject matter.

3.2.2 Would AI improve the efficiency and quality of the research project?

According to the survey conducted by Kwon (2025), researchers whose first language is not English are more likely to use AI than other researchers. Her survey evidence was echoed at the Editors' Panel, "The rise of artificial intelligence: Implications for research and publishing," at the European Accounting Association Congress in May 2025. AI has been a boost to researchers whose first language is not English. They can use AI not only for copyediting or editing, but also for translating research material into the language which they can process more quickly or for summarizing literature quickly for researchers' initial screening.

Junior researchers may benefit more than senior researchers from using AI in data collection and analysis. Junior researchers typically carry most of the weight in data collection and analysis, which can be time-consuming, especially when samples cannot be collected or variables cannot be constructed from structured data. Using AI as a machine RA for those tasks can substantially increase junior researchers' efficiency and, furthermore, may reduce human error arising from fatigue or distraction.

However, one note of caution is that a researcher should always prepare an evaluation set of outcomes created by humans and be skeptical of the machine output until validation (de Kok 2025). Few research tasks are identical across projects because research questions, sample periods, or sample composition differ. A researcher should not use AI for a task without validation just because other researchers have performed and validated a similar task on their respective samples. Overall, the answer to the second question depends on the nature of the research project and the strengths and weaknesses of the researcher.

3.2.3 What might be the risks of using AI and how to mitigate them?

The best-known risk of using AI is hallucinations—the tendency of AI models to generate nonsensical or inaccurate output, including responses to text inquiries or data tasks. Researchers should check and verify, on a random basis if the sample is large, any machine output. Here, we focus on four less-well-known risks. The first risk is lack of replicability. If a researcher uses AI for data collection and variable construction, the collected or constructed data are not expected to be identical even if the same procedures are repeated on the same AI model. Except for the simplest models (e.g., linear or logistic regression), most machine learning models are stochastic in that the algorithm incorporates elements of randomness, such as random weight initialization and training data shuffling. In addition, a hyperparameter of an LLM is the temperature setting, which controls for the randomness in token selection for the output.[15] When the temperature is set to 0, the output is most deterministic. When the setting is 1 or higher, the output is more creative and spontaneous. To reduce the risk related to lack of replicability, researchers are advised to set the temperature low and save the LLM output as well as the prompts.

The second risk is losing a researcher's voice when generating text from AI using prompts. Even though academic writing is not a liberal art, a paper would be more persuasive to readers if it carries the researcher's voice. Even if we ignore its other issues, such as being overly confident and lacking detail and insight, AI-generated content is often bland. For example, one of the authors of this editorial recently assigned master's students to submit both their own essays and those generated by AI upon the same prompts. The instructor noticed that all the enlightening and enjoyable-to-read essays were written by students—a sentiment shared by the students. What distinguishes good human essays is the author's voice coming through from the writing. To find that voice, researchers can modify AI-generated content with detail, anecdotes, and style.

The third risk is copyright violation due to AI-generated content. When we write, we typically know where an idea, example, or argument comes from and would credit the source when appropriate. Moreover, we also know what ideas come from previous research and what are our new ideas; this distinction helps us identify and explain our own contribution. When AI generates text, it often does not know the source and could credit a source incorrectly because the model is trained on a gigantic amount of publicly available data and does not keep track of the sequence or

[15] See a good discussion of LLM temperature setting here: https://www.hopsworks.ai/dictionary/llm-temperature. Figure 1 of Cao, Long, Tucker, and Wan (2025) shows the sensitivity of machine output to temperature settings.

source of information. Relying upon AI-generated text, researchers may be exposed to the risk of using copyrighted information without due credit. Plagiarism software can reduce this risk to some extent, but the risk of copyright violation with AI-generated content remains much higher than that with human work.

The last risk is authorship identification. Authorship is an explicit way of assigning responsibility and giving credit for intellectual work.[16] When humans are involved in producing research, authorship is assigned to the individuals who have materially contributed to the work and those individuals, or the publishing house upon publication, may copyright that work. Legally, AI cannot be listed as a coauthor because it cannot take any responsibility as a coauthor. A related question is whether human coauthors can claim credit for 100% of the research product and receive copyright protection if a machine has contributed to the research production to the extent that a human who has done the same would have been credited as a coauthor. The U.S. Copyright Office ruled on February 22, 2023 that an image of a book cannot be copyrighted because it was generated by an AI tool based on the author's prompts (Brittain 2023). It is unclear whether the logic would extend to AI-generated text. Even if there is no legal hurdle, are readers comfortable about a research paper for which AI is a significant contributor, even if authors disclose the use of AI?[17] Perhaps this is a main reason why researchers are divided on the issue of using AI to generate text for all or part of a paper.

3.3 Factors we consider as editors of *Accounting Horizons*

First, as editors, we execute the existing AI policy, made by the AAA, instead of injecting our own preferences regarding which tasks AI should or should not be used for in research

[16] See https://hms.harvard.edu/sites/default/files/assets/Sites/Ombuds/files/AUTHORSHIP%20GUIDELINES.pdf.
[17] The case of omitted "machine writer" differs from the case of a ghost human writer, who is financially compensated and agrees to not take credit as a coauthor.

production. And we keep an open mind, allowing ourselves to learn and adapt to new situations. Second, we evaluate whether the research question addressed would be interesting to the journal's audience—accounting academics, preparers, practitioners, standard setters, and regulators who might be interested in the topic. Third, we examine whether the use of AI, according to the authors' disclosure, has undermined the academic integrity of a given manuscript. For example, what is the risk of the authors using inaccurate or uncredited information? Have the authors validated the output from their machine RA? Have the authors maintained control of the academic content? Does the study use logical arguments and offer fresh perspectives beyond the existing body of work on the topic? Does the study reflect a human voice as opposed to the machine's voice? Last, we keep an eye on signs of undisclosed AI use. On this respect, our reviewers have been very helpful in alerting us to suspected but undisclosed AI use. We occasionally inquire authors about their AI use in the absence of a disclosure or ask authors to expand their existing disclosure.

4. The Issue of AI Use at the Community Level

Academics have experienced and thrived with technological advances in the past, but the AI era seems to be different. For example, from handwriting to typewriting and to using document preparation tools such as MS Word, technology was merely assisting researchers but did not modify their content or generate new content. Even statistical software (e.g., SAS, Stata, and R) simply executes technical procedures determined and selected by researchers according to their research design and can be an open book for anyone who desires to understand the inner workings. In contrast, LLMs are a black box to users.

AI is a general-purpose technology in that it is widely used, is capable of ongoing technical improvement, and can enable innovation in sectors that apply the technology (Bresnahan and Trajtenberg 1995; Bresnahan 2010; Eloundou, Manning, Mishkin, and Rock 2024). According to a recent survey in June 2025, 19% of employees use AI a few times or more per week (Ip 2025).[18] The same article reports that "the pickup in productivity has been concentrated in technology and related sectors such as *scientific research*, engineering and consulting" (emphasis added). Next, we discuss some immediate questions faced by the academic community. We then discuss three additional challenges in this new, uncertain, and ever-evolving landscape of AI.

4.1 Immediate questions for the academic community

AI has stirred an unprecedented debate and confusion in the academic community, and the tension mostly resides in using AI for Type A activity, especially using AI to generate text. According to Kwon's (2025) survey, 4% of researchers have used AI to write the first draft of a paper without disclosing its use, 4% have done so while disclosing AI use. Of the remaining 92% of the respondents who report that they have not previously used AI, 29% would be willing to use AI to write the first draft and 63% would not consider it.

In addition to whether AI should be used in drafting academic papers, the research community could consider a few other questions. Should the use of AI be disclosed? Would readers find the disclosure helpful or important in their consumption of the research? If so, how? Would the disclosure result in reviewers and editors unintentionally applying different academic standards to manuscripts that are assisted with AI than those without? If that is true, would authors be incentivized to withhold information about their AI use? Would embracing AI in academic research lead to researchers' attention shifting from creating *new* knowledge to prompt-

[18] One externality of AI use is energy consumption and therefore burden to our environment. AI companies are reluctant to disclose the real cost incurred to deliver a response upon a user prompt. Stern (2025) reports that the estimated energy consumption is averaged at 0.34 watt-hours per prompt and that generating videos can consume as much energy as required for cooking a nice dinner. Thus, even "free" AI tools are not actually free to society.

engineering to extract and synthesize *existing* knowledge? If so, knowledge discovery might slow down or even dry up, and research might become an activity within an echo chamber.

4.2 AI-generated documents as journal submissions

According to a recent *New York Times* article, a deputy editor of a journal received a "letter to the editor" on a recent publication and later discovered that the authors had published many letters and comments in a variety of specialty journals all within a period of six months—an accomplishment that is infeasible for human authors (Kolata 2025). Novy-Marx and Velikov (2025) describe a process of using LLMs to automatically generate finance papers and show that AI can efficiently generate hundreds of papers that use existing public knowledge but might appear novel to unsuspecting readers. Journals need to implement a mechanism to distinguish AI-generated from AI-assisted submissions and avoid placing this burden on already overworked volunteer reviewers and editors. Furthermore, because AI can significantly shorten research production time, journals may expect growing submissions while their resources of human reviewers and editors are strained. Attracting and retaining quality reviewers and editors to identify truly innovative and impactful research would become more important than ever to a journal's success.

4.3 Training future generations of researchers

In a traditional Ph.D. program, students learn how to code, write papers, and critique others' research. The process is typically implemented through learning-by-doing. We have heard discussions at several Ph.D. programs regarding whether to allow, encourage, or forbid Ph.D. students from using AI during their training, especially in the early stages of training. As Finley (2025) points out, reliance on AI may deprive students of the opportunity to "learn how to think through, express or defend ideas." Even though senior researchers probably know how to evaluate the output of AI, deciding what to take in and what to discard, Ph.D. students likely have not yet gained the experience to exercise such judgment. This concern is consistent with the Dunning-Kruger effect, where junior researchers (including Ph.D. students) may become stalled in the phase of not knowing what they do not know and yet thinking they have mastery (Kruger and Dunning 1999).[19]

4.4 Rediscovering the value of humans in research production

The goal of research is to further the body of knowledge. Generations of researchers have achieved this goal while relying on agents to carry out facets of their work. AI can transfer many research activities from human agents to machine agents, but the transfer does not alter the primary need for researchers' human judgment to discern what research questions are important, how to most meaningfully answer them, and whether the research effort has contributed to a useful answer. As machines can execute many mundane tasks for researchers, the *thinker* behind the research project will become increasingly important.

For individual researchers, future competition might tilt more toward idea generation than task execution. For the research community, the focus might be on making fundamental advances in the discipline to remain relevant. For example, accounting is a means to answer the very human question about whether others entrusted with our economic resources have made good decisions and whether economic agents are better or worse off as a result (Tucker 2025). For the foreseeable future, only people can best judge what is important to society in this regard and accounting

[19] Consistent with this concern, in discussing how researchers can use AI in conducting literature reviews, Tingelhoff, Brugger, and Leimeister (2025, p.85) state that "it is imperative for researchers to recognize that the effective use of these tools necessitates a high level of skill and domain expertise. Researchers must be capable of independently performing the tasks they delegate to Gen.AI to critically evaluate and adapt the outputs generated. If a researcher lacks the necessary skills to carry out a task manually, it becomes challenging to assess the accuracy, relevance, and reliability of the Gen.AI output."

academics must focus on this human judgment question and what information would best accomplish this objective.

5. Conclusion

The various uses of AI in academic research will continue to grow and editorial policies on the use of AI in academic research may continue to evolve. We use our experience as editors of *Accounting Horizons* as a platform to have a broader conversation about the value and risks of the use of AI in academic research. We contribute to the debate by providing a framework for understanding the issue of AI use. We encourage the academic community to join the dialogue about how to use AI while mitigating its risks.

References

Berger, P. G., W. Cai, L. Qiu, and C. X. Shen. 2025. Employer and employee responses to generative AI: Early evidence. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4874061

Bochkay, K., S. V. Brown, A. J. Leone, and J. W. Tucker. 2023. Textual analysis in accounting: What's next? *Contemporary Accounting Research,* 40 (2): 765-805. https://doi.org/10.1111/1911-3846.12825

Bresnahan, T. 2010. General purpose technologies. Chapter 18 of the book "Handbooks in Economics," Volume 2. Elsevier B. V. https://www.sciencedirect.com/science/article/pii/S0169721810020022

Bresnahan, T. F. and M. Trajtenberg. 1995. General purpose technologies 'engines of growth'? *Journal of Econometrics*, 65: 83-108. https://www.sciencedirect.com/science/article/pii/030440769401598T

Brittain, B. 2023. AI-created images lose U.S. copyrights in test for new technology. *Reuters*. February 22. https://www.reuters.com/legal/ai-created-images-lose-us-copyrights-test-new-technology-2023-02-22/

Cao, Y, L. Chen, J. W. Tucker, and C. Wan. 2025. Can generative AI help identify peer firms? *Review of Accounting Studies*, 30: 3344-3386. https://doi.org/10.1007/s11142-025-09892-6.

Carrillo, A. H., C. Stachl, and S. Talaifar. 2024. A workflow for human-centered machine-assisted hypothesis generation: Commentary on Banker et al. (2024). *American Psychologist*, 79 (6): 800-802. https://doi.org/10.1037/amp0001256

de Kok, T. 2025. ChatGPT for textual analysis? How to use generative LLMs in accounting research. *Management Science*, 71 (9): 7888-7906. https://orcid.org/0000-0001-9577-4057

Eloundou, T., S. Manning, P. Mishkin, and D. Rock. 2024. GPTs are GPTs: Labor market impact potential of LLMs. *Science*, 384 (6702): 1306-1308. https://www.science.org/doi/10.1126/science.adj0998

Finley, A. 2025. AI's biggest threat: Young people who can't think. *Wall Street Journal*. June 23. A17.

Ip, G. 2025. Why the GDP-jobs split? Look to AI. *Wall Street Journal*. October 4. A2.

Kobak, D., R. González-Márquez, E. Horvát, and J. Lause. 2025. Delving into LLM-assisted writing in biomedical publications through excess vocabulary. *Science Advances*, 11 (27): 1-8. https://doi.org/10.1126/sciadv.adt3813

Kolata, G. 2025. 454 hints that a chatbot wrote part of a biomedical researcher's paper. *New York Times*, July 2. https://www.nytimes.com/2025/07/02/health/ai-chatgpt-research-papers.html?unlocked_article_code=1.TU8.Bz6M.lHpKTtkDye55&smid=url-share

Kruger, J. and D. Dunning. 1999. Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77 (6): 1121-1134. https://doi.org/10.1037/0022-3514.77.6.1121.

Kwon, D. 2025. Is it ok for AI to write science papers? *Nature* survey shows researchers are split. *Nature*, May 14. https://www.nature.com/articles/d41586-025-01463-8

McCarthy, J., M. L. Minsky, N. Rochester, and C. E. Shannon. 2006. A Proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27 (4): 12. https://doi.org/10.1609/aimag.v27i4.1904.

McCloskey D. 1985. Economical writing. *Economic Inquiry*, 24: 187-222.

Naddaf, M. 2025a. How are researchers using AI? Survey reveals pros and cons for science. *Nature*, February 4. https://www.nature.com/articles/d41586-025-00343-5

Naddaf, M. 2025b. AI is transforming peer review—and many scientists are worried. *Nature*, March 26. https://www.nature.com/articles/d41586-025-00894-7

Novy-Marx R. and M. Z. Velikov. 2025. AI-powered (finance) scholarship. Working paper. http://www.nber.org/papers/w33363

Susarla, A., R. Gopal, J. B. Thatcher, and S. Sarker. 2023. The Janus effect of generative AI: Charting the path for responsible conduct of scholarly activities in information systems. *Information Systems Research*, 34 (2): 399-408. https://doi.org/10.1287/isre.2023.ed.v34.n2

Stern, J. 2025. Tracing AI's insatiable appetite for energy. *Wall Street Journal*. July 1. A10.

Tingelhoff, F., M. Brugger, and J. M. Leimeister. 2025. A guide for structured literature reviews in business research: The state-of-the-art and how to integrate generative artificial intelligence. *Journal of Information Technology*, 40 (1): 77-99. https://doi.org/10.1177/02683962241304105

Tucker, J. W. 2025. Technology, accounting, and research. *Accounting Horizons*, Forthcoming. https://doi.org/10.2308/HORIZONS-2025-173

Tucker, J. W., J. Wang, and R. Zhao. 2025. Corporate use of artificial intelligence and subsequent product market performance. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=5223870

Wang, Z. 2023. Visualizing and explaining transformer models from the ground up. Available at: https://deepgram.com/learn/visualizing-and-explaining-transformer-models-from-the-ground-up

Weizenbaum, J. 1966. ELIZA—a computer program for the study of natural language communication between man and machine. *Communications of ACM,* 9 (1): 36-45.