# AI-Assisted Academic Writing: Adjustments in Quantitative Social Science

**Abstract**

The advent of advanced AI systems capable of generating academic text — including "chain-of-thought" large language models with test-time web access — is poised to significantly influence scholarly writing and publishing. This review discusses how academia, particularly in quantitative social science, should adjust over the next decade to AI-assisted or AI-written articles. We summarize the current capabilities of AI in academic writing (from drafting and citation support to idea generation), highlight emerging trends, and weigh advantages against risks such as misinformation, plagiarism, and ethical dilemmas. We then offer speculative predictions for the coming ten years, grounded in literature and present data on AI's impact to date. An empirical analysis compiles real-world data illustrating AI's growing footprint in research output. Finally, we provide policy and workflow recommendations for journals, peer reviewers, editors, and scholars, presented in an exhaustive table. Our aim is to inform a balanced approach to harnessing AI's benefits in academic writing while safeguarding integrity and transparency.

# 1 Introduction

Artificial intelligence (AI) is rapidly becoming intertwined with academic writing and publishing. Generative AI models, particularly large language models (LLMs) like OpenAI's GPT-series, have demonstrated the ability to produce fluent, human-like text. Over the past few years, such models have evolved from basic autocomplete tools to sophisticated systems capable of writing entire paragraphs or even papers. In late 2022, the public release of ChatGPT (based on GPT-3.5 and later GPT-4) marked a watershed moment, showcasing that AI can draft coherent scholarly abstracts and answers to complex questions. Unlike earlier automated writing attempts (e.g. template-based systems or the notorious `SCIgen` paper generator), modern LLMs leverage vast training corpora and advanced neural architectures to create content that often passes as human-written. In one striking demonstration, Gao et al. (2022) found that medical experts could only identify 68% of AI-generated abstracts as fake, missing the rest that were machine-written. This highlights both the potential and challenges of AI-written text in academia.

A novel class of AI systems now being explored are *chain-of-thought* models with test-time inference abilities and web access. These systems can perform reasoning through intermediate steps and dynamically fetch information from the internet while generating text. For example, OpenAI's WebGPT project fine-tuned GPT-3 to use a text-based web browser, allowing it to submit search queries, follow links, read pages, and even cite sources in its answers. Such tool-augmented models are designed to improve factual accuracy and provide up-to-date information, addressing a key limitation of static LLMs that rely solely on training data. In essence, a chain-of-thought model with web access can mimic how a human researcher writes: by iteratively searching literature, reasoning about findings, and composing text with references. Over the next decade, it is anticipated that academic writing may increasingly be done in collaboration with these AI assistants, which can generate drafts, suggest citations, and even respond to reviewer comments in a human-like way.

The rise of AI-generated scholarly content has already sparked debate in academia. Many researchers have experimented with AI writing tools for tasks ranging from summarizing articles to drafting introductions. A survey by Van Noorden & Perkel (2023) of 1,600 scientists revealed a mix of excitement and concern: over 60% of respondents worried that AI tools would increase mistakes, misinformation, or plagiarism in the literature. At the same time, a large portion of scientists have begun integrating AI into their workflow. In an informal poll, 80% of surveyed *Nature* readers reported using ChatGPT or similar AI for at least some research-related writing tasks (such as summarizing findings or literature review). This dual perspective — high usage alongside high concern — encapsulates the central challenge: academia must adjust to AI-assisted writing in a way that maximizes benefits while maintaining rigor and integrity.

In this review, we focus on quantitative social science as a case study for how academic writing norms and practices might adapt. However, many issues discussed (such as authorship, accuracy, and ethics) apply broadly across disciplines. We first outline the current capabilities of AI in academic writing and emerging technological trends (Section 2). We then discuss advantages these tools offer and the risks they pose (Section 3). In Section 4, we venture predictions for the next decade, grounded in current developments and data. Section 5 presents an empirical analysis of AI's impact on research output to date, including tables with real-world data. In Section 6, we propose concrete adjustments to policies and workflows for journals, peer review, editors, and researchers, summarized in a comprehensive table of recommendations. We conclude in Section 7 with reflections on how academic culture can evolve alongside increasingly intelligent writing assistants. Throughout, our approach is that of an economics handbook-style review: we synthesize existing literature and evidence, and provide a forward-looking discussion relevant to scholars and policymakers navigating the era of AI-assisted academic writing.

# 2 Current Capabilities and Emerging Trends

AI's role in academic writing has expanded rapidly in recent years. Modern language models can now assist in nearly every stage of the writing process, from initial idea generation to polishing a final draft. Here we review the current state-of-the-art capabilities of AI in scholarly writing and highlight emerging trends, with a focus on LLMs and related tools.

## 2.1 AI-Assisted Writing and Editing

As of 2024, large language models are capable of producing surprisingly coherent academic prose. Tools like ChatGPT have been used to draft entire sections of papers, including introductions and literature reviews. Researchers have reported using AI to rephrase sentences, correct grammar, and improve the clarity of their manuscripts. For non-native English speakers, such AI writing assistants can significantly reduce language barriers, helping to articulate complex ideas more fluently. According to Van Noorden & Perkel (2023), many scientists use ChatGPT for drafting texts and refining writing style, treating it as an "AI editor" that can iteratively improve a manuscript's readability. Beyond text, AI can also aid in generating or refining elements like equations and even figure captions.

Current AI writing models can follow structured prompts to produce content in specific formats. For example, one can prompt a model with "Write an abstract about XYZ with a background, methods, results, conclusion structure," and the model will attempt to fill in each part. This has led to experiments on using AI to draft scientific abstracts. In one study, 50 journal article titles were given to ChatGPT to generate corresponding abstracts; the resulting AI-generated abstracts were often hard to distinguish from genuine ones by experts. This underscores how far AI text generation has come in mimicking academic style and structure.

Another capability is summarization. AI models can condense lengthy papers or datasets of literature into concise summaries. This is particularly useful for literature reviews; tools like Elicit and Scholarcy (which leverage language models) can provide summaries of research papers, highlight key findings, and even suggest relevant citations. Similarly, AI can assist in paraphrasing and synthesizing information from multiple sources, helping authors weave together a narrative from diverse references.

Moreover, AI is being used for citation and reference management. Some experimental systems allow a user to query an AI for relevant literature on a topic, and the AI can return not just summaries but also citations to pertinent papers. For instance, Petiska (2023) demonstrates that ChatGPT, when connected to citation databases, tends to cite highly cited articles and journals, essentially relying on citation counts to select references. This hints at both a capability and a bias: AI can quickly surface influential references (improving thoroughness of literature coverage), but it might also over-emphasize already well-cited work, potentially reinforcing the "Matthew Effect" in citations (where the rich get richer in terms of citations). Nonetheless, the ability to automatically generate a list of plausible references for a given topic is a tantalizing aid for researchers beginning a literature review.

## 2.2 Idea Generation and Research Design Support

Beyond writing existing knowledge, AI tools are increasingly being used for ideation in research. Large language models can act as brainstorming partners, generating hypotheses or suggesting new research questions. For example, if a social scientist is interested in studying the impact of a policy intervention, they might ask an AI to propose potential mechanisms or related variables to consider. The AI, having digested vast amounts of text, can offer connections or analogies that the researcher may not have immediately thought of.

While such suggestions need critical evaluation, they can spark creative directions.

There have been experimental studies on AI generating entire research proposals. In one case, a team of researchers prompted ChatGPT to propose a research idea in chemistry, including forming a hypothesis, suggesting experimental setup, and even generating a fake dataset for analysis. Similarly, Van Dis et al. (2023) reported that ChatGPT could outline components of a scientific study when prompted appropriately, although with errors and gaps that require human correction. These experiments illustrate an emerging trend: AI as a collaborator in the conceptual phase of research. Over time, as models improve in factual accuracy, they might reliably assist in designing experiments or analytical frameworks, especially when combined with domain-specific data.

## 2.3 Chain-of-Thought Reasoning and Tool Use

The newest trend in AI for academic writing involves models that can perform multi-step reasoning and utilize external tools during text generation. Referred to as "chain-of-thought" prompting or reasoning, this approach has models internally break down complex tasks into intermediate steps (often unseen by the user) and solve them one by one, which has been shown to improve performance on tasks like mathematical problem solving and logical reasoning (Wei et al., 2022). When writing a scholarly article, a chain-of-thought capable AI might internally plan the structure of an argument, sequentially determine what background is needed, what data should be cited, and so on, rather than producing text in a single shot. This leads to more coherent and logically sound drafts.

Even more powerful is the integration of live web access into the generation process. OpenAI's WebGPT is an early example: it allowed the model to issue search queries and read web content in real time. By doing so, the model isn't limited to its static training data (which might be months or years out of date for rapidly evolving fields). Instead, it can fetch the latest information or verify facts on the fly. Crucially, WebGPT was trained to provide citations to the sources it used, pointing to a future where AI-written academic text comes with verifiable references. Another instance is the integration of GPT-4 into Bing's search engine in 2023, which similarly enables conversational queries with up-to-date web results embedded. In academic writing, such capabilities mean an AI could automatically find and cite the latest statistics from a government database or include a quotation from a newly published study, ensuring that the content is current and evidence-backed.

Emerging research is also exploring AI models that use specialized tools: for example, connecting to statistical software. One can imagine a near-future scenario where an AI writing assistant not only drafts the Methods section of a paper but also, when needed, executes a data analysis by calling a Python or R script, then incorporates the results (tables, numbers) into the text. Early versions of this concept exist in projects where LLMs are used to generate code for data analysis and then interpret the output. Such tool-using AI could drastically speed up the writing of results and discussion sections by automating parts of the analytical workflow under human guidance.

## 2.4 Accuracy, Limitations, and Ongoing Improvements

Despite their impressive capabilities, current AI writing models have notable limitations. A primary issue is factual accuracy. LLMs often "hallucinate" – they can produce plausible-sounding statements that are incorrect or even fabricate references that look real but are non-existent (Thorp, 2023). For instance, an early attempt by Meta AI with a model called Galactica (designed specifically for scientific text generation) ended in failure: users quickly found that Galactica would generate authoritative-looking scientific content with fundamental errors and made-up citations. As Thorp (2023) noted, ChatGPT and similar models sometimes output references to studies that simply do not exist, a profoundly concerning behavior in a scholarly context. This has led to widespread calls for caution and for improved factuality in AI-generated content.

To mitigate this, the incorporation of retrieval (searching for information) during generation is a promising trend (as described above with WebGPT). By grounding answers in actual source material, models are less likely to hallucinate. Another improvement is fine-tuning models on domain-specific academic text, which can make their output more knowledgeable and in line with the conventions of academic writing in a field. However, domain fine-tuning must be done carefully to avoid propagating biases or outdated paradigms present in the training literature.

Finally, an emerging area is AI detection and watermarking. Since AI-written text can be hard to distinguish from human text, researchers are working on methods to automatically identify AI-generated content (e.g., by statistical quirks or embedded cryptographic watermarks in the text). We will discuss later how detection plays into policy, but here it's worth noting that the technical race between generation and detection is ongoing. For now, the best large models can often evade detection, especially if the text is edited by a human afterward to remove tell-tale signs. This means current AI assistance often leaves an invisible footprint in papers, unless authors choose to disclose it.

In summary, AI today can contribute substantially to academic writing: drafting and editing prose, suggesting citations, summarizing literature, generating ideas, and even performing some reasoning and data tasks. The trajectory of recent advancements suggests these capabilities will only grow. Yet, issues of accuracy, reliability, and appropriate use remain only partially solved. The next section will weigh the benefits these AI tools provide against the risks and challenges they introduce.

# 3 Advantages and Risks

AI-assisted writing offers several clear advantages to researchers and academics, but it also comes with significant risks and potential downsides. In this section, we delineate the benefits that such technologies can bring to scholarly work, as well as the pitfalls that academia must navigate.

## 3.1 Advantages of AI-Assisted Academic Writing

**Increased Efficiency and Productivity:** Perhaps the most immediate advantage is the speed and efficiency that AI tools can provide. Tasks that typically consume a researcher's time, such as copy-editing, formatting references, or drawing up a first draft of a section, can be accelerated. An AI system can produce a well-structured draft in minutes, which a researcher can then refine rather than writing from scratch. This has the potential to shorten the time-to-publication for research findings. Indeed, some commentators suggest that conversational AI could *accelerate the innovation process* by allowing scientists to focus more on ideas and results and less on the mechanics of writing. Over a decade, even a modest improvement in writing efficiency per paper could translate to a substantial increase in research output across the field.

**Enhanced Accessibility and Equity:** AI writing assistants can help level the playing field for researchers for whom academic writing (especially in English) is a barrier. This includes many scholars in non-English speaking regions or those early in their careers. By helping with language and style, AI tools can make it easier for such researchers to communicate their ideas in top journals, potentially increasing the diversity of voices in academic discourse. As Van Dis et al. (2023) note, by helping people write more fluently, AI might make science more equitable and broaden participation. Moreover, AI translation tools (built on similar LLM technology) can translate drafts between languages with increasing accuracy, facilitating international collaboration and knowledge exchange.

**Idea Generation and Interdisciplinary Innovation:** AI's capacity to ingest and synthesize vast amounts of information can spark novel connections. By getting suggestions from an AI that has "read" widely, researchers might discover references or analogies from other disciplines that they were unaware of. This serendipity can fuel interdisciplinary innovation. For example, a sociologist might get an AI-suggested insight drawn from economics or network theory literature that enriches their analysis. In this way, AI can act as a boundary-spanning research assistant, offering a breadth of knowledge that no single human can match.

**Consistency and Technical Assistance:** For large collaborative writing projects (such as multi-author review articles or grant proposals), maintaining a consistent tone and avoiding redundant text can be challenging. AI tools can help homogenize the style and ensure consistency in terminology throughout a document. They can also auto-generate boilerplate text for methods or compliance statements (e.g., ethics approvals, data availability statements), which, while necessary, are formulaic. Additionally, AI can check mathematical derivations or code included in a paper by replicating the steps (for instance, using a tool-augmented model to verify an equation or simulation result). This can reduce errors in technical content.

**Literature Navigation:** The explosion of scientific publications makes it hard to keep track of all relevant work. AI tools can quickly search and filter literature. Instead of spending days doing a manual literature review, a researcher can ask an AI to summarize the key developments on a topic or find papers that support a specific statement, dramatically speeding up the gathering of evidence. Some emerging AI systems (like semantic scholarly search engines) go beyond keyword matching and actually understand the content of papers to find conceptual links. These can uncover relevant studies that traditional search might miss, giving researchers a more comprehensive understanding of the state of knowledge.

In sum, when used appropriately, AI assistance in writing can lead to faster, more comprehensive, and potentially more creative research outputs. It can free researchers from some drudgery and allow them to focus on interpretation and theory. However, these advantages do not come without significant caveats, as the next subsection on risks will address.

## 3.2 Risks and Challenges

**Misinformation and Hallucinations:** The flip side of AI's generative power is its propensity to produce incorrect information. LLMs like ChatGPT do not have a built-in fact-checking mechanism; they generate text based on learned patterns, not a verified knowledge base. Consequently, they can assert falsehoods with full confidence. In scientific writing, this is dangerous. An AI might fabricate a convincing-sounding result or cite a non-existent study to support a claim. Thorp (2023), editor of *Science*, warned that text generated by ChatGPT is "plagiarized from ChatGPT" and not truly original or reliable. The concern is that undiscerning authors might take AI outputs at face value and incorporate errors into papers, or worse, that such errors slip past peer review and enter the published literature, polluting the knowledge record. As an example, the Galactica model mentioned earlier was shown to produce references with real author names but entirely fake titles and content. If an author unknowingly included such a reference, it would be very difficult for readers to trace or verify it, undermining scientific transparency.

**Plagiarism and Ethical Violations:** The use of AI blurs the lines of authorship and originality. If a large portion of a manuscript is written by an AI, to what extent can the listed human authors claim original writing? Some journal editors argue that using AI-generated text without disclosure is a form of plagiarism (even if what is plagiarized is not another author's work, but the AI's output). *Science* journals took a strong early stance: in early 2023 they explicitly banned the use of AI-generated text in papers, considering any such text as non-original and thus unacceptable. They also forbade AI tools from being listed as authors, since AI cannot take responsibility for the content. Other publishers like *Nature* and *JAMA* issued guidelines along similar lines (Flanagin et al., 2023), mandating that authors are accountable for all text and that AI contributions should be acknowledged, not attributed authorship. Thus, an immediate risk is that researchers might unwittingly commit ethical breaches by over-relying on AI and failing to properly disclose its role. Even if done with good intentions, lack of transparency about AI use can be seen as deception in the academic community.

Relatedly, AI could enable new forms of academic misconduct. For instance, a person could ask an AI to write a paper about results that do not actually exist (fabricating data and analysis). Before advanced AI, writing a fake paper still required effort and was easily detectable if the science was unsound. Now, an AI can produce a superficially coherent fake scientific paper that might fool reviewers who do not catch the lack of actual data. This scenario is an extreme ethical violation, but one that AI might tempt a small minority towards, necessitating stronger verification practices (such as requiring authors to provide underlying data and analysis code).

**Reliability and Overreliance:** Even if one uses AI with honest intentions, there is the risk of overreliance. Researchers might become dependent on AI for tasks that they themselves should know how to do, such as critically reviewing literature or performing statistical analysis. If AI becomes a crutch, the depth of researchers' own understanding could suffer. Over time, academic training might de-emphasize writing and critical analysis skills, assuming AI will handle it, which could erode the expertise of scholars. Moreover, AI models have systematic biases and blind spots based on their training data. If scholars lean on them too much, those biases could propagate. For example, if an AI under-represents certain schools of thought or authors from certain regions (because the training corpus had fewer of those), a researcher depending on AI summaries might get a skewed perspective. Otterbacher (2023) argues that technical solutions for detecting AI-generated content are insufficient on their own, implying that human oversight and critical thinking cannot be removed from the loop. This underscores that AI should remain a tool, not an oracle, and users must be vigilant about its outputs.

**Bias and Social Implications:** AI models learn from existing literature, which may contain biases – gender, racial, ideological, etc. If these models are used to generate new text, they can inadvertently amplify or perpetuate those biases. For instance, if in a field the literature has historically cited more work from certain famous labs (due to network effects), an AI might overly cite those same sources, further entrenching their dominance. This relates to the Matthew Effect mentioned earlier, which Petiska (2023) highlights: AI's reliance on citation frequency could disadvantage newer or less mainstream ideas. Additionally, AI might use tone or language that isn't appropriate for all contexts (e.g., overly confident declarations without caveats, which might be common in general web text but not in careful academic discourse).

**Detection and Policing Difficulties:** A practical challenge for the academic community is that it is very hard to reliably detect AI-written text, especially when moderately edited by humans. Traditional plagiarism detection software looks for overlap with existing human-written sources, but AI-generated text is usually original (not copy-pasted, but newly generated). Some AI detection algorithms exist, often using machine learning to identify subtle characteristics of AI text, but their accuracy is far from perfect. As reported by Desaire et al. (2024), obvious signs of AI usage (like the phrase "as an AI language model, I cannot...") have been caught in about 100 published papers so far, leading to corrections or retractions. Yet, these are the low-hanging fruit – cases where authors accidentally left a tell-tale AI phrase in the text. More nuanced AI assistance would leave no such trace. The study by Desaire et al. (2024) attempted to detect undisclosed AI-written text via stylistic differences and estimated that 1–3% of recent journal articles might contain undisclosed AI-generated passages. While that percentage is relatively small, it is non-zero and likely to grow. Policing this at scale is challenging. If journals crack down too hard without reliable tools, they risk false accusations (e.g., accusing an author of using AI when they did not, simply because their writing style triggers a detector). This creates a precarious situation in academic integrity enforcement.

**Loss of Human Element and Creativity:** Finally, there is an intangible but important risk: that excessive use of AI could dilute the human creativity and insight that are central to academic work. Writing is not just a recording of results; it is a process where ideas are refined and new insights occur during the act of composition. If that cognitive process is offloaded to a machine, do we lose something essential? Some scholars worry that reliance on AI to do the heavy lifting of writing and even thinking could lead to more formulaic research, sticking to safe, seen-before patterns that the AI knows, rather than bold, unconventional ideas. Moreover, the narrative voice and argumentation style of a paper often reflect the author's point of view and intellectual contribution. If AI homogenizes writing, academic papers might become more uniform and less reflective of diverse thought processes.

In summary, the risks of AI-assisted academic writing range from the concrete (factual errors, fake citations, plagiarism) to the abstract (loss of creativity, bias reinforcement). These challenges must be addressed proactively. The next section will turn towards the future: given these advantages and risks, how might the academic world evolve in the next ten years in response to AI-written research, and what changes might we anticipate?

# 4 Predictions for the Next Decade

Projecting a decade into the future of academic writing in the age of AI requires a blend of cautious extrapolation and imaginative speculation. Here we offer several predictions for how AI, particularly advanced chain-of-thought models with web access, might influence scholarly communication by 2035. These predictions are informed by current trends, expert opinions in the literature, and early data on AI adoption.

## 4.1 AI as a Ubiquitous Writing Tool

It is likely that AI writing assistants will become as common as spelling or grammar checkers are today. By 2035, most quantitative social scientists (and academics in general) will routinely use AI tools in drafting manuscripts and reports. This does not mean AI will replace researchers as writers, but it will be a standard part of the workflow. Just as researchers now might run a grammar-check or use reference management software, in the future they might use an AI to generate a first draft of an abstract or to suggest alternate phrasings for clarity.

Surveys already show high usage rates of AI tools among scientists (e.g., 80% of some surveyed groups as of 2023). As these tools become integrated into word processing platforms (imagine Microsoft Word or Google Docs with a powerful AI co-writer built-in), the barrier to use will be even lower. We predict that writing without any AI assistance could become as rare as writing a paper by hand without a word processor is today. The stigma or novelty of using AI will fade; it will be taken for granted.

**Implication:** The academic community will likely shift its focus from *whether* AI was used to *how* it was used. By necessity, there will be established norms and perhaps an expectation that authors disclose AI contributions in a standardized way (much as one might acknowledge an editor or colleague for feedback). Transparent reporting of AI assistance may even become a required section in papers (for example, an "AI tools" subsection in the methods or acknowledgments). However, as AI use becomes ubiquitous, these disclosures might be considered routine and not carry a negative connotation, provided the authors took responsibility for verification.

## 4.2 Improved AI Capabilities and Specialization

The next decade will almost certainly bring even more advanced AI models. We anticipate several qualitative improvements that will enhance their utility in academic writing:

- **Factual Accuracy and Citation**: Future models (the successors of GPT-4, GPT-5, etc.) will be better at integrating factual information without hallucinating. With continued research into retrieval augmentation and training on citation-rich corpora, by 2030 we might have an AI that can write a literature review with correctly cited, existent references for every claim it makes. The technology demonstrated by WebGPT of citing sources will become more robust. This will reduce one of the major current risks (fabricated information), making AI outputs more trustworthy as a starting point.

- **Discipline-Specific AI Assistants**: We will likely see AI models specialized by discipline or even sub-discipline. For example, an AI trained on economics papers and data could become a virtual research assistant for economists, fluent in the conventions of econometrics, common datasets, and the seminal literature in the field. Similarly, a political science writing AI might be embedded with knowledge of key theories, historical cases, and statistical methods particular to that field. Such specialization could outperform general models for discipline-specific writing tasks, leading researchers to choose an AI assistant aligned with their domain. Fuster et al. (2023) in the medical domain (cardiology) already talk about a pathway forward with AI tools in their field, indicating that domain-specific adoption is on the horizon.

- **Integration with Data and Analysis**: By 2035, AI systems might not only write text but also handle data analysis as part of the writing process. We predict that a researcher could feed raw data to an AI, ask it to perform a regression or create a plot, and the AI would output the results along with a written description of the findings, including the figure or table in LaTeX format. Early versions of this are already possible with tools that combine LLMs and programming (so-called "AI scientists" or automated statistician concepts). This tight integration means the writing process and analysis process merge, potentially making the reporting of results almost instantaneous after analysis. Of course, human oversight will remain crucial to ensure the analysis was done correctly and the results are interpreted properly.

- **AI in Peer Review and Publishing**: It's not just authors who will use AI; journals and conferences will too. We anticipate that within 10 years, many journals will employ AI to screen submissions. This could involve AI checking for sections that appear AI-written or checking for common errors and inconsistencies. More ambitiously, AI might provide first-draft peer review reports, identifying potential methodological flaws or missing citations, which human reviewers can then verify and build upon. Some conferences have already experimented with AI-suggested reviews. As models become more reliable, their role in peer review could expand, which in turn will influence how authors write (e.g., perhaps writing with awareness that an AI might be reading and judging clarity or novelty).

## 4.3 Evolving Ethical and Academic Norms

The norms around authorship and accountability will evolve significantly in the next decade. One prediction is that academic journals and institutions will converge on clear policies regarding AI. As of 2024, policies vary: some ban AI text entirely, others allow it with disclosure. By 2035, after some trial and error, a consensus set of practices might emerge:

- **Disclosure Requirements**: Likely all reputable journals will require authors to clearly state if and how AI was used in preparing a manuscript. This could be akin to disclosing a statistician's help or use of a professional editor. The disclosure might detail which sections, if any, were drafted with AI assistance or whether an AI was used to generate summaries that were later rewritten by the authors. The purpose is transparency, not necessarily to discourage use. In fact, if AI becomes ubiquitous, not mentioning its use when it obviously had to have been used might raise eyebrows. Journals might enforce disclosure by adding a checkbox during submission or a statement in the article metadata.

- **Authorship Definitions**: It will be universally agreed that AI tools cannot be listed as authors (this is already widely accepted (Flanagin et al., 2023)), since authorship carries accountability and AI cannot take responsibility or sign copyright forms. However, there might be new guidelines on acknowledging AI. For instance, an article's acknowledgments section could include a sentence like, "Portions of the writing were assisted by the AI language model `XYZ` (Version 202X)". Professional societies may issue ethical guidelines on this. We foresee something analogous to how image manipulation is handled now: it's not banned to adjust an image for clarity, but one must not mislead or fabricate; similarly, it's not banned to use AI to edit text, but one must not use it to fabricate content or obscure the origin of ideas.

- **Educational Adjustments**: Academic training programs (like PhD training, writing workshops) will incorporate AI literacy. The next generation of scholars will be taught not only how to write a paper, but how to effectively and ethically use AI tools in writing. They will learn about the pitfalls (so they can avoid blindly trusting AI) and about the expectations for transparency. University writing centers might shift to advising students on how to use AI to improve a draft and how to cite AI if it contributed ideas. There may also be an emphasis on maintaining critical thinking: for example, exercises where students critique or fact-check AI-written text to practice not taking it at face value.

**Possibility of a Backlash:** While we predict general integration of AI, we should acknowledge a possible contrary scenario: if there are one or two high-profile scandals (say a major paper is retracted because it was found to have dangerously wrong information introduced by AI), there could be a backlash that slows adoption. For instance, funding agencies might require statements about AI use in grant proposals or even disallow AI-written text in proposals out of fear of deception. Fields that heavily value qualitative or interpretive writing might resist AI more, valuing human voice. However, even in such a backlash scenario, the tide of technological capability is unlikely to reverse; it might only change how usage is monitored.

## 4.4 Quantitative Trends and Speculative Metrics

By 2035, we can expect some metrics to illustrate AI's influence: One prediction is that the volume of publications discussing AI in science (meta-research on AI's role) will grow, then stabilize as it becomes normal. For example, the number of PubMed-indexed publications about ChatGPT and its applications shot past 1000 within about 9 months of its release; by 2030 this number could be in the tens of thousands, encompassing studies across disciplines. But by 2035, this topic might no longer be novel – AI will be like the internet or computers: an assumed part of the landscape rather than front-page news.

We also predict that a significant percentage of new scientific papers will have had AI involvement in writing. Based on current hesitant adoption (1–3% undisclosed use in 2023 (Desaire et al., 2024)), we might see a rise to, say, 20–30% of papers by 2028 that involve AI in some form, and over 50% by 2033. Eventually, nearly all papers will likely have at least minor AI assistance (even if just in copyediting). This is speculative, but if AI tools continue to improve and no major barriers are imposed, such growth in adoption is plausible.

Another interesting speculation: the emergence of AI-driven journals or conference tracks. We might see, for instance, an online journal that welcomes AI-generated studies (with human oversight of course) as a way to explore what fully automated science can do. Or conferences about meta-science could have a special track for papers written with minimal human writing input, to showcase the state of technology.

Finally, by 2035, AI might start contributing to areas like hypothesis generation in a measurable way. We may hear of discoveries or theories that credit an AI tool for the initial idea (with the human researchers then carrying it forward). If such cases occur, it will further blur the line of contributor roles, but also highlight the positive potential of AI as a partner in scientific creativity.

In conclusion, our ten-year outlook is one of cautious optimism: AI will be deeply integrated into scholarly writing, bringing efficiency and new capabilities, but the academic community will evolve norms and practices to ensure that human oversight, responsibility, and creativity remain at the core of research. The next section provides some empirical snapshots of the current impact of AI on academia, which serve as a baseline for understanding how far we've already come in just the first couple of years of this transformation.

# 5 Empirical Analysis

To ground the discussion in concrete evidence, we turn now to empirical data on AI's impact on academic research output to date. While the phenomenon is very recent, a few indicators can be measured: the proliferation of literature about AI (and produced with AI), adoption rates among researchers, and incidents of AI-related issues in publications. We present here a series of tables and summaries based on real-world data that illustrate these trends.

## 5.1 AI in the Literature: Publication Trends

One straightforward measure of AI's influence is the number of publications that themselves discuss or utilize AI in academic writing or research. Figure **??** (conceptual, data summarized in Table 1) illustrates the exponential growth in scholarly publications about ChatGPT or generative AI in research.

As seen in Table 1, virtually no academic literature discussed these AI tools prior to 2022. By 2023, the topic became ubiquitous, with over a thousand publications (many of them editorials, commentaries, or early experiments with ChatGPT in various fields, especially in medicine and education). This indicates how swiftly the academic community reacted to generative AI's arrival. Such a rapid proliferation of discussion is itself an indication of impact: researchers everywhere are grappling with what AI means for their discipline.

It is also notable that a large share of these 2023 publications about AI in science were themselves prompted by issues or questions arising from AI use (for example, how to handle AI in student assignments, how to detect AI in papers, ethics of AI in medicine, etc.). Thus, the content of research output has been altered: meta-scientific considerations about AI have taken center stage.

| Year | PubMed-Indexed Items on ChatGPT | Notable Observations |
|---|---|---|
| 2019 | 0 | (pre-GPT era) |
| 2020 | 0 | |
| 2021 | 0 | GPT-3 released mid-year, few academic mentions |
| 2022 | $\sim 5$ | Early editorials on AI writing (late 2022) |
| 2023 | $> 1000$ | Explosion after ChatGPT release in Nov 2022 |
| 2024 | $> 2000$ | Continued growth (projected) |

Table 1: Approximate number of PubMed-indexed publications referring to ChatGPT or its use, by year. The count for 2023 exceeded 1000 by late summer 2023. Data for 2024 includes early 2024 and is projected to continue rising.

Another angle is to consider fields and topics. Early data shows that fields like medicine, biology, and computer science led in publishing about ChatGPT's implications. Social sciences and humanities also participated, but to a lesser extent initially. This may be because the implications for clinical practice and education were immediately pressing, whereas in quantitative social science, the effects are more on writing and analysis practices than life-and-death decisions. We expect the distribution of such publications to even out as all fields adjust to AI.

## 5.2 Adoption and Usage Statistics

Empirical evidence on how many researchers are actually using AI in their work is starting to emerge via surveys and indirect detection:

| Indicator (Year) | Value / Estimate |
|---|---|
| Researchers who have used ChatGPT or similar (Nature poll, 2023) | $\sim 80\%$ of respondents |
| Scientists concerned AI will increase misinformation (Nature survey, 2023) | $> 60\%$ of 1600 surveyed |
| Introductions in papers (2023) with undisclosed AI-written content | 1–3% (estimated) |
| Suspected published papers with obvious AI text (2022–2024) | $\sim 100$ cases identified |
| PubMed-indexed publications about ChatGPT by Aug 2023 | $> 1000$ |

Table 2: Key data points on AI adoption and impact in academia.

Table 2 compiles several key indicators. The Nature poll result (80% usage) suggests that, at least among a tech-engaged subset of academics, trying out AI tools has been extremely common. This doesn't mean 80% are writing papers with it, but they have used it for something research-related. The Nature survey by Van Noorden & Perkel (2023) highlights the cautious stance: a majority worry about negative impacts even as many experiment with the technology.

The estimate from Desaire et al. (2024) that 1–3% of journal article introductions in 2023 showed signs of AI generation is fascinating. It suggests that despite the hype and usage, most published papers were still written by humans without direct AI text generation (or with very carefully edited AI text that evades detection). In concrete terms, out of 19,000 introductions they analyzed, only a few hundred at most might contain AI-written sentences. This indicates that serious adoption for actual published material lagged behind experimentation. We interpret this as evidence that many researchers, while curious, were hesitant to significantly use AI in final drafts, likely due to concerns about accuracy or fear of ethical issues.

Meanwhile, the number of obviously AI-tainted papers (those with phrases like ”I am an AI language model” left in) that had to be flagged is around 100 by mid-2024. Retraction Watch maintained a list of such instances. These are clear cases where authors misused AI or failed to properly edit it out. The fact that these slipped through peer review suggests oversight gaps in 2023. Many of those papers were later corrected or are under scrutiny. This number will likely grow, but journals are becoming more vigilant. We may see a peak of such incidents in 2024-2025, then a decline as both authors and journals get better at prevention and editing.

The ¿1000 PubMed count by August 2023 was mentioned before and shown in Table 1. By comparison, prior transformative technologies (like CRISPR in biology or blockchain in finance) also saw rapid literature growth, but the AI writing phenomenon is cross-disciplinary, making its literature footprint unusually widespread.

## 5.3 Case Study: Research Output in Economics

Focusing specifically on quantitative social science, and economics in particular, we can look for any aggregate changes in research output that might correlate with the availability of AI writing tools. One hypothesis could be that if writing is easier, perhaps more papers or working papers are being produced. It is too early to tell definitively, but we can examine trends in preprint submissions.

For example, the number of economics papers on the arXiv (e.g., in the Econ.EM or Econ.GN categories) and on SSRN has been steadily rising for years. A preliminary analysis for 2023 vs 2022 does not show an obvious inflection that can be attributed to AI. Any increase in output could be due to many factors (pent-up research post-pandemic, etc.). That said, anecdotal evidence from some researchers suggests they were able to finish certain writing projects faster with AI help (like finally writing up old results). A formal study would be needed to quantify this effect across the discipline.

Another area is the quality or readability of papers. As AI tools fix grammar and can even smooth logic, one might wonder if the average readability of submitted papers improves. Some journal editors have informally commented that they observed fewer grammatical errors in manuscripts in late 2023 than before, speculating that tools like Grammarly or ChatGPT were being used for proofreading. If true, this is a subtle but positive outcome (as long as content isn't being tampered with). Over a decade, if AI becomes a universal writing tutor, we might see overall writing quality go up in terms of clarity (though perhaps at the cost of less stylistic variety).

Finally, we consider empirical data on policy changes: by late 2023, as noted earlier, major journals have updated author guidelines. This isn't a numerical output, but it's a measurable change in the academic process. For instance, in November 2023 the *Science* journals reversed their earlier blanket ban and allowed AI-assisted text with disclosure. Similarly, *Nature* journals require disclosure but not forbid assistance (Nature Editorial, 2023). The presence of these policies can be counted: in 2022, zero journals had AI writing policies; by 2024, many top publishers (Elsevier, Springer Nature, IEEE, etc.) have them. A possible dataset for future analysis is tracking how many journals have adopted AI disclosure policies and how stringent they are, as a function of time.

In conclusion to this empirical snapshot, the quantitative signals of AI's impact are still emerging. The data so far shows skyrocketing discussion about AI, high rates of experimentation among scholars, moderate but growing integration into actual published work, and the occurrence of some errors and ethical breaches that have grabbed attention. These metrics will serve as a baseline to evaluate, a decade from now, just how much change occurred.

Next, we take these insights and move to recommending how various stakeholders in academia should adjust their policies and workflows to responsibly integrate AI in the scholarly writing process.

# 6 Policy & Workflow Adjustments for Journals and Scholars

The disruptive potential of AI-assisted writing necessitates proactive adjustments in the practices of all stakeholders in academic publishing. In this section, we provide recommendations targeted at four groups: journal publishers (and their policy makers), peer reviewers, editors, and scholars (authors). We present these recommendations in Table 3 for clarity, then discuss them in detail.

Table 3 provides a high-level overview. Here we elaborate on a few points:

For journals, one key recommendation is normalizing disclosure. Rather than seeing AI use as something nefarious, journals should treat it as a tool that must be disclosed like any other potential conflict or methodology. This removes the temptation for authors to hide usage. Journals like those in the *Science* family now allow AI-generated content *if* it's disclosed and comes with the exact prompt and version used, documented in the methods. This level of detail might seem onerous but can be valuable if readers want to understand context or if issues arise.

Peer reviewers are in a tricky spot because they must uphold standards possibly without the same level of policy backup (since they are volunteers and not all conferences/journals have given guidance). Our recommendation is that reviewers should primarily act as if every paper was written by a human (to avoid bias), but with an extra dose of skepticism in checking verifiable details (like references). Also, if journals permit, reviewers should openly discuss if they used AI in forming their review. There have been cases of reviewers secretly using ChatGPT to generate review reports; some journals frowned upon this especially if it wasn't disclosed, as it raises confidentiality concerns (an unvetted AI could be sending the manuscript content to an external server). Hence, our guidance leans towards caution in that area.

Editors adopting AI in their workflow could greatly speed up processes (e.g., an AI-generated summary of a 50-page manuscript could help an editor quickly decide on reviewers or whether it's in scope). But they must also handle the influx of AI-affected content: e.g., be ready to desk-reject or send back papers that read like AI spam. Publishers might use AI to detect serial offenders (authors who repeatedly submit AI-written papers without disclosure).

For authors, the overarching theme is to use AI responsibly. We emphasize verification and understanding, aligning with proposals from institutional reports that talk about the researcher's *duty of disclosure, duty of verification, and duty of responsibility* (as in the Cornell guidelines ). The idea is that using AI doesn't absolve one of the scholarly duties; it actually adds a duty: you must verify the AI's contributions.

We also encourage authors to think of long-term career skills – over-reliance might harm younger researchers if they don't develop their own writing style and critical thinking. Perhaps academic advisors will even set rules for students like "the literature review draft you give me must be written without AI, we can then use AI to polish it," purely as a pedagogical exercise.

Lastly, an adjustment in mindset is needed: treating AI as a collaborator of sorts that needs oversight. Just as a co-author's contributions are scrutinized and integrated, AI's contributions need checking and editing. It's when people treat AI output as a black box or ground truth that problems occur.

If the above adjustments are made, we can hope to reap AI's benefits (speed, accessibility) while mitigating downsides. These adjustments themselves will likely be refined with experience. The table is an initial blueprint; real-world feedback from implementation in 2024–2025 will inform updates. By the late 2020s, we expect the academic enterprise will have settled into a new equilibrium where these practices are simply part of the standard operating procedure of doing research.

# 7 Conclusion

The coming decade promises to be a transformative period for academic writing. As chain-of-thought AI models with web access and ever-improving language abilities become integrated into research workflows, academia must strike a balance between innovation and caution. In this review, we have explored the current state of AI in scholarly writing, weighed its benefits and risks, and offered predictions and recommendations for adapting to this new landscape.

AI-assisted writing can unquestionably enhance productivity and open new avenues for creativity in quantitative social science research. It can democratize writing, enabling a wider range of scholars to express complex ideas clearly. It can also serve as a tireless research assistant, scanning literature and suggesting connections that might take humans weeks to discover. These advantages portend a future where routine aspects of academic writing are streamlined, allowing researchers to focus more on high-level analysis and theory.

Yet, alongside these benefits come serious challenges. The integrity of the scholarly record depends on accuracy, transparency, and trust – all of which can be jeopardized by careless or unethical use of AI. Hallucinated facts or references, if not diligently caught, could lead to false scientific claims propagating. Undisclosed AI authorship blurs academic accountability and could undermine the trust readers place in published work. There is also a cultural challenge: academia has long been built on human intellect and expression, and we are now compelled to redefine the role of that human element when a machine can generate passable academic text.

Our examination of the first waves of AI impact (2022–2024) shows an academic community actively grappling with these issues. The data thus far reveal rapid uptake, vigorous debate, and fortunately, a relatively low incidence (so far) of serious breaches in published literature. Policies are emerging, from outright bans to conditional acceptance with disclosure. The evolution of these policies, as we predicted, will likely converge towards allowing AI with responsibility rather than forbidding it outright, especially as generative models become a standard tool.

Looking ahead to 2035, we envision an academic world where AI is a normal part of the toolkit. Writing a paper might involve having an AI generate a few candidate paragraphs, which the researcher then edits heavily. It might involve AI summarizing one's own draft to highlight unclear parts, or AI suggesting relevant citations one might have missed. Perhaps peer review in 2035 will involve an AI that assists the reviewer by checking math or looking up similar works for context. Importantly, however, the researcher remains in the driver's seat: setting the research agenda, making the judgment calls, and ensuring the final output's validity. In the ideal scenario, AI is a powerful amplifier of human creativity and rigor, not a replacement for them.

To reach that ideal scenario, the academic community should foster a culture of *adaptation and accountability*. Adaptation means embracing useful AI tools and updating training and norms accordingly. Accountability means every stakeholder accepts responsibility for their part in using AI ethically: authors verify and disclose, journals enforce standards, and AI developers themselves (often academics in computer science) work on reducing biases and improving truthfulness.

There will likely be bumps along the road. We might see scandals, like a paper heavily generated by AI slipping through, which could momentarily cause distrust. We might also see remarkable successes, such as AI helping to write a complex interdisciplinary review that pushes the field forward in ways a single human author might not have managed alone. The key will be to learn from both failures and successes.

In conclusion, the next ten years in quantitative social science (and academia broadly) will test our ability to integrate a disruptive technology into a centuries-old scholarly tradition. History has shown that academia can absorb new tools — be it statistical software, the internet, or digital libraries — and ultimately strengthen its methodologies and outputs. AI-based academic writing can be another such tool, if we navigate its pitfalls wisely. The adjustments recommended in this article aim to ensure that the scholarly enterprise emerges from this period not only intact, but enhanced: more efficient, more inclusive, and continuing to uphold the rigor and integrity that science and social science demand. The conversation between human and machine in the pages of journals is just beginning; it is our responsibility as a community to guide that conversation toward truth and knowledge.

# References


Desaire, H.; Isom, M.; Hua, D. (2024). Almost Nobody Is Using ChatGPT to Write Academic Science Papers (Yet). *Big Data and Cognitive Computing*, 8(10), 133. (Published Oct 11, 2024).

Flanagin, A.; Bibbins-Domingo, K.; Berkwits, M.; Christiansen, S. (2023). Nonhuman "Authors" and Implications for the Integrity of Scientific Publication and Medical Knowledge. *JAMA*, 329(8), 637–639.

Fuster, V.; Bozkurt, B.; Chandrashekhar, Y.; et al. (2023). JACC journals' pathway forward with AI tools: The future is now. *Journal of the American College of Cardiology*, 81(17), 1543–1545.

Gao, C.; Howard, F.M.; Markov, N.S.; Dyer, E.C.; Ramesh, S.; Luo, Y. (2022). Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. *bioRxiv* 2022.12.23.521610 (preprint).

Grimaldi, G.; Ehrler, B. (2023). AI et al.: Machines are about to change scientific publishing forever. *ACS Energy Letters*, 8(1), 878–880.

Nature Editorial (2023). Tools such as ChatGPT threaten transparent science; here are our ground rules for their use. *Nature*, 613(7945), 612. (Editorial without named authors).

Otterbacher, J. (2023). Why technical solutions for detecting AI-generated content in research and education are insufficient. *Patterns*, 4(8), 100796.

Petiska, E. (2023). ChatGPT cites the most-cited articles and journals, relying solely on Google Scholar's citation counts. As a result, AI may amplify the Matthew Effect in environmental science. *arXiv*:2304.06794.

Temsah, M.-H.; Altamimi, I.; Jamal, A.; Alhasan, K.; Al-Eyadhy, A. (2023). ChatGPT Surpasses 1000 Publications on PubMed: Envisioning the Road Ahead. *Cureus*, 15(9): e44769.

Thorp, H.H. (2023). ChatGPT is fun, but not an author. *Science*, 379(6630), 313. (Editorial).

Van Dis, E.A.M.; Bollen, J.; Zuidema, W.; van Rooij, R.; Bockting, C. (2023). ChatGPT: five priorities for research. *Nature*, 614(7947), 224–226.

Van Noorden, R.; Perkel, J.M. (2023). AI and science: What 1,600 researchers think. *Nature*, 621(7980), 672–675.

Wei, J. et al. (2022). Chain-of-Thought Prompting Elicits Reasoning in Large Language Models. *Advances in Neural Information Processing Systems (NeurIPS 2022).*

| Stakeholder | Recommended Adjustments and Guidelines |
|---|---|
| **Journals & Publishers** | - *Authorship Policy:* Define that AI tools cannot be listed as authors and reiterate that all human authors are accountable for the content. Update authorship agreements to include a statement on AI usage.<br>- *Disclosure Requirements:* Mandate that authors disclose any use of AI in writing or research (e.g., in acknowledgments or a dedicated section). This includes specifying the tool, version, and scope of its use.<br>- *Submission Checks:* Implement automated checks for common AI-generated text patterns (continually updated as AI evolves). If such patterns are detected, require authors to clarify or revise.<br>- *Ethical Guidelines:* Expand ethical policies to cover AI, clarifying that undisclosed AI-generated content is a form of misrepresentation or plagiarism. Treat violations on par with image manipulation or plagiarism, with potential retractions.<br>- *Transparency in Publication:* Encourage inclusion of AI prompts or raw outputs as supplementary material when significant to the research (to aid understanding of how conclusions were reached).<br>- *Living Guidelines:* Regularly update guidelines as AI technology changes, possibly via a standing committee on AI in publications (as some professional bodies are forming in 2024). |
| **Peer Reviewers** | - *Confidentiality and Tool Use:* Prohibit sharing manuscripts with AI services that involve external servers (to avoid leaking confidential material). If AI tools are used to assist review, they should be local or the content should be obfuscated, and disclosure to the editor is advised.<br>- *AI-Assisted Reviewing:* Reviewers may use AI to check grammar or summarize the paper for themselves, but they should not rely on AI judgments of quality. All substantive evaluation should come from the reviewer's expertise.<br>- *Reference Verification:* Given AI's ability to hallucinate references, reviewers should pay extra attention to bibliography. They should verify a sample of cited works to ensure they exist and support the claims. (Journals can facilitate this by requiring DOI links or providing integrated reference checks.)<br>- *Detecting AI Text:* Reviewers should be alert to signs of AI-generated text (e.g., generic phrasing, inconsistent style). If suspicious, they can flag it to editors, who may use detection tools or request clarification from authors.<br>- *Feedback on AI Clarity:* If a manuscript discloses AI assistance, reviewers might comment on whether the use seems appropriate and whether any sections lack clarity (possibly due to AI). This can help normalize AI discussions in peer review. |
| **Editors (Handling Editors & Editorial Staff)** | - *Screening:* At submission, use AI-detection tools as one input (with caution, due to false positives). If a paper is largely suspected AI text without disclosure, request an explanation from authors or an immediate revision. |