NANCY UM

SCHOLARLY WRITING IN THE FACE OF GENERATIVE AI

A View from Art History

For whom do we write? No longer satisfied with addressing small circles of academic peers, it is now much more acceptable, even welcome, in the humanities to seek out a wider readership. Although slow to follow, the academy has increasingly come to acknowledge the value of work that is accessible, engaged with the community, and more widely impactful. However, as we, humanities scholars, look to the future, our most vigorous reading audience may not be a broader public that is hungry for smart critical perspectives. Rather, it is likely to be much more mechanical. This essay foregrounds some of the issues that emerge for humanities scholars at a moment when our most active rising readership is poised to be the machine.

Since OpenAI's release of ChatGPT in November 2022, when the public became aware of the power of the large language model (LLM) to produce seemingly plausible and immediate responses to human prompts, academic anxiety around generative AI has been both heightened and hastened. The fears are oriented around multiple weighty, critical issues, including but not limited to the replication of bias, the rise in misinformation, the dissolution of trust, the environmental footprint of AI operations, the leakage of private information, and the future of intellectual labor and academic work. Perhaps the higher education community's most focused response to the hype generated by the ChatGPT release has been directed toward its use by students and concerns related to academic integrity.

As it stands, we have yet to take stock of the growing impact of these technologies on our research and publishing environments, especially in the humanities. There is no opting out of the generative-AI wave, as its functions are being rapidly integrated (or have already been integrated) into our existing everyday systems.[1] With this essay, I aim to present peer art historians with some specific issues affecting scholarly writing in the humanities in the face of the LLM. The intention is not to advocate for these technologies, although I acknowledge that they do hold potential to contribute to certain spheres, but only if managed with a pointed awareness of their harms and a commitment to responsible use. Nor will I use this space to argue against them, even though I fully understand why so many of us would prefer to resist

QUICK CITATION
Um, Nancy. "Scholarly Writing in the Face of Generative AI." *Ars Orientalis* 54 (2024): 182–191

their growing impact. Accordingly, I will steer away from both the tenor of techno-optimism and the opposing aspiration for disavowal.

It is truly daunting to write an essay that may be irrelevant by the time it is published, given the rapidly changing nature of the technological landscape.[2] Indeed, while I composed this piece, I had to add new footnotes each day, as fresh perspectives and breaking developments emerged. At the same time, I am motivated to provide a window, however fleeting, into some of generative AI's most vivid rising instantiations, with the goal of provoking deeper understanding within our discipline about this increasingly weighty and evolving force.

AI Subjects

In a 2022 article, Mehtab Khan and Alex Hanna delved into the workings of LLMs, with the key contention that it is necessary to understand how the datasets upon which they are based, containing both text and image, have been selected, prepared, and processed.[3] The article is excellent in describing how massive datasets are produced to train, test, and benchmark LLMs, while also carefully delineating the roles that human actors, and their decisions, play at each juncture. Writing from a legal perspective, the two authors elucidate how such stakeholders are connected to (or more frequently disconnected from) each other, as well as to and from the technological processes that they either participate in or are affected by. As Khan and Hanna deftly present, these extended and distributed matrices of knowledge production render certain individuals vulnerable to various types of harm at different scales. Namely, they differentiate the *data curator*, the individual or group that collects large datasets, such as a university-based researcher, from the *data annotator*, who labels the images in those datasets and could be a low-wage remote worker in the developing world. They also distinguish the *copyright holder*, who maintains certain rights over the materials that may be represented in large datasets, such as an author or publisher, from the *data subject*, whose biometric data is collected and could be identified in an image. Lastly, they center on the figure of the *model subject*, who may be uninvolved in these processes but is subject to decisions based on them, such as a person who is unjustly arrested due to a misidentification triggered by facial-recognition software.

Khan and Hanna's article provides accessible language to make generative AI processes and their impacts concrete in granular ways. It also gives much-needed structure to more generalized anxieties about the future societal harms that AI-based decisions may inflict. Most importantly, it underlines how we are all implicated in the world of the LLM, albeit at differing levels. This type of AI awareness is surely useful to humanities scholars, particularly art historians, who will be called upon increasingly to interact with and interpret synthetic texts, images, and videos as the products of our evolving cultural systems. Moreover, Khan and Hanna's work prompts us to follow suit by breaking down some of the AI subject positions of scholars, while also considering the potential impact of these new technologies on each of these roles. Indeed, our own work is already being heavily mediated by LLMs, compelling us to think deeply about the way we write and how various publics will interact with those words.

Scholars and Publishers as Rights Holders

Multiple sectors of the creative industry are actively contesting the incorporation of their work into the datasets that underlie the LLMs developed by firms in Silicon Valley. As one notable

example among many, in September 2023 the Authors Guild and several high-profile writers filed a class-action lawsuit against OpenAI, stating that the company had used their books to train its ChatGPT chatbot, without seeking their permission or providing compensation to them.[4] Additionally, some visual artists have taken a defensive stance, supported by tools that digitally alter the way in which their images are ingested into large datasets, with the aim of deflecting future mimicry.[5] In another high-profile instance, from the world of journalism, in December 2023 the *New York Times* sued OpenAI and Microsoft for copyright infringement, with the claim that they used "millions of articles published by The Times" to train their models, and called upon them to destroy those models and datasets.[6]

It is important to underline that neither the Authors Guild nor the *New York Times* examined the huge training datasets that they had issued grievances against and that are generally undocumented.[7] Rather, both entities backward engineered the support for their cases, feeding the chatbots prompts that demonstrated that copyrighted content had been included. For instance, by providing details that could not be obtained from summaries posted online, ChatGPT's responses revealed that certain books were ingested in their entirety. The *New York Times* ran focused prompts, using Microsoft's Bing, which issued verbatim responses drawn from *Times* articles without any acknowledgment. Both cases also cited the potential financial damage and reputational harm that could emerge from mis- or unattributed content. In regard to the *Times* claim, OpenAI has contended that these instances should be considered as fair and transformative uses and has called for the case's major claims to be dismissed.[8] Yet, Khan and Hanna underline the challenges of resolution on these grounds: "Lack of certainty around fair use also means that it is not clear how data collection may take place, and how copyright holders, data subjects, and model subjects may seek potential recourse."[9]

For the scholarly community, these issues are surfacing more slowly and thus are playing out in a different way. It has been pronounced that we are moving into a "data winter," marked by new restrictions on data use and reuse and a rise in monetization. This is a decisive shift away from the unregulated data-access environment that fueled the period referred to as the "AI summer" and that provided the conditions for the types of claims mentioned above.[10] When the news emerged that Taylor & Francis, which owns Routledge and publishes many humanities journals, including the *Art Bulletin*, had contracted with Microsoft to provide content without consent, some of the affected authors expressed shock that their scholarship had been used in this way.[11] Their initial attention was primarily oriented toward questioning the press's obligations to its authors and seeking the possibility to opt out of the arrangement. It remains to be seen if scholarly authors will have the grounds to follow the model of the novelists and journalists mentioned above. If so, the position will have to be situated carefully between claims of intellectual property, assertions of copyright protection, and consideration of past publishing contracts, while also taking seriously the principle of educational value that undergirds fair use. Meanwhile, other publishers are negotiating similar types of deals.[12]

Along these lines, I ran a simple experiment using Microsoft Copilot, an AI chatbot that is built on GPT-4 and draws on live sources from the web.[13] Using Copilot's "more precise" mode, I asked, "What is the Red Sea style?" This question was not innocent. Rather, inspired by the approaches mentioned above, I selected it to see if the response would draw upon my own research. Copilot returned two results (fig. 1). The first mentioned a recent event called Red Sea Fashion Week, which was held at the luxurious St. Regis Hotel on the Red Sea coast of the Kingdom of Saudi Arabia in May 2024. The second tier of results discussed an architectural tradition shared across coastal societies in Egypt, Sudan, Eritrea, Somalia, Yemen, and Saudi

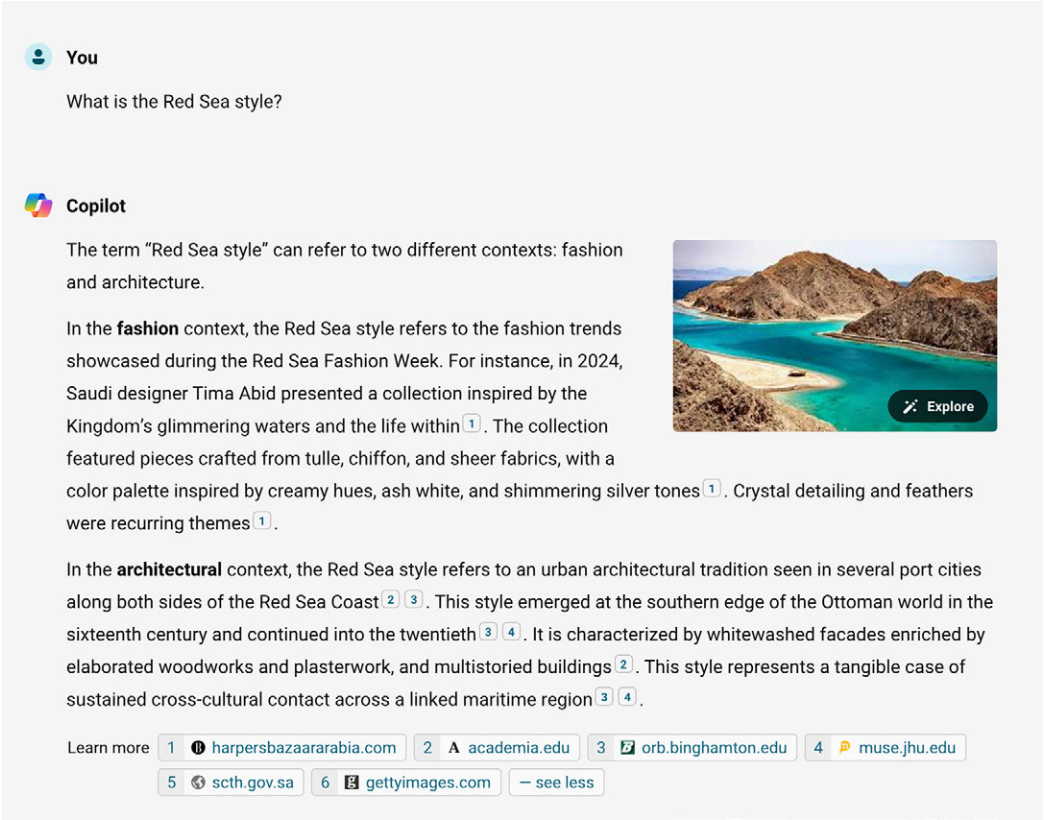

FIGURE 1. Screenshot of prompt and response on Microsoft Copilot, using the "more precise" mode. Conversation initiated by the author, July 5, 2024

Arabia. It listed as sources a number of scholarly articles, including an essay that I published in 2012, titled "Reflections on the Red Sea Style."[14] Embedded in the generated response text, I found short verbatim quotes that came directly from the abstract of that article. For instance, Copilot delivered the statement that the Red Sea style "emerged at the southern edge of the Ottoman world in the sixteenth century and continued into the twentieth" and "represents a tangible case of sustained cross-cultural contact across a linked maritime region." Both lines parroted my published abstract, word for word, but without direct citation to the original 2012 article. Rather, I needed to sift through the links provided below to find the source.[15] On one hand, the blatant mimicry of these responses surfaces some of the unease cited in the two legal claims mentioned above, although on a wholly different scale. But from another angle, they suggest that a tool such as Copilot could help to increase public understanding of an understudied historic building tradition, thereby amplifying the reach of research that was initially published in an academic periodical with a limited print run.

This very lightweight example points to certain conflicting conditions that emerge as the products of our research are increasingly mediated by mainstream chatbots, such as ChatGPT and Copilot. These instruments can potentially open new avenues for the dissemination of the words and ideas that we generate in our scholarship, while at the same time veritably occluding our individual authorship. In this way, evidence-based academic study is, more than ever, key to the battle against misinformation, even while the independent voice of the scholarly author may concurrently recede from visibility.[16]

Scholars as Creators (but also Peer Reviewers and Editors)

It is not just students who are making use of generative AI tools for writing. Since the release of ChatGPT, scholars have been actively experimenting with it and other instruments in composing texts, including those intended for publication. While some have touted the benefits of generative-AI tools for brainstorming and eliciting useful editorial prompts, others have used them in a more instrumental manner.[17] This became clear as early as January 2023, when a handful of papers and pre-prints that credited ChatGPT as a co-author were published in medical journals and uploaded to scientific repositories.[18] In response, many scientific journals, including prestigious titles such as *Nature* and *Science*, issued declarations that a chatbot could not fulfill their designated criteria for authorship and hastened to amend their submission guidelines. Publishing organizations such as the World Association of Medical Editors, the *Journal of the American Medical Association*'s JAMA Network, and the Committee on Publication Ethics followed by issuing similar assertions.[19]

As of the time of writing this essay in August 2024, none of the flagship journals of the College Art Association or the Society of Architectural Historians had yet delineated protocols for the identification or use of AI-generated content in their submission guidelines. By contrast, the publication of record for the Modern Language Association, *PMLA*, provides specific language in its submission guidelines, indicating that "AI tools cannot be considered authors of works submitted to the journal" and that "authors must fully cite in the manuscript, at submission, their use of all content (whether text, images, data, or other) created by an AI tool."[20] The MLA has also issued guidelines on how to cite ChatGPT and other generative AI tools as sources.[21] Clearly, this is an evolving topic, in which journals (where the effects of new writing trends are poised to be felt more immediately than in monographs), publishers, and professional organizations are moving forward at different paces to provide guidance. Yet, AI-generated content, both texts and images, continues to steadily enter the realm of scholarly production, whether directly identified, blatantly obvious, or detected through subtle clues of "ChatGPT contamination."[22]

Even for those scholars who have no interest in using generative AI tools in any part of the writing process, an awareness of these technologies is germane. In our expanded roles as peer reviewers and editors, we need a broadened toolkit so that we can assess, evaluate, and respond to today's art historical writing and associated images, both of which may be produced or inspired by generative AI. Following our peers in literature and the sciences, updated submission guidelines and new ethical directives are surely needed. But, just as significantly, we have not yet cultivated the types of discernment and literacy that will guide the wider scholarly community to engage effectively in the work of review, in the face of the content that is already emerging with AI support.

Scholars as Researchers

Ithaka S-R, the research wing of the nonprofit organization that operates JSTOR, has carried out an extensive survey of the generative-AI tools that are being pioneered in the world of higher education. Their product tracker, released in July 2024, includes over one hundred different tools, apps, plug-ins, and workflow platforms. Based on their review of this developing arena, Ithaka S-R researchers Claire Bytas and Dylan Ruediger envision that we are steadily moving toward "a future in which the distinction between the initial act of identifying and accessing relevant sources and the subsequent work of reading and digesting those sources

is irretrievably blurred if not rendered irrelevant." They also foresee a research landscape in which "content is less 'discovered' than queried and in which secondary sources are consumed largely through tertiary summaries."[23] These weighty pronouncements portend major changes in the ways that we conduct research, starting from the fundamental act of searching our standard databases. Additionally, they foreshadow a condition in which our own research outputs, namely our articles, essays, reviews, and books, will be mediated through such instruments before they are found, downloaded, and read (or not) by other scholars.

Ithaka's own JSTOR is on the front edge of these innovations, with an AI-assisted research tool, scheduled to be released in the fall of 2024.[24] Powered by GPT-3.5, in addition to some smaller models, this feature will soon be integrated as an active function on JSTOR's user platform. It will provide users with summaries of articles, recommendations for related materials, and answers to pointed questions about the content of a given article. In its design, the tool aims to facilitate faster and more effective understanding of search results and to aid various audiences, including students, in making use of JSTOR's diverse content. Surely, major questions about the accuracy of its outputs will dominate its reception in the scholarly sphere. Yet, it is certain that this development will alter the way that we interact with the copious aggregated materials provided by JSTOR, one of the most indispensable resources used by humanities researchers today.

It should be underlined that JSTOR has approached the development of this tool with transparency as a collaborative and iterative undertaking, working closely with university partners. The organization signals core principles of "democratizing research, enhancing educational experiences, and supporting scholarly endeavors across various academic levels" to deliver high-quality outputs, bolstered by JSTOR's extensive scholarly catalog.[25] However, JSTOR is not alone in this type of platform enhancement. Indeed, as the above-mentioned report by Bytas and Ruediger indicated, this area of generative AI-supported research assistance is a vigorous growth field in the EdTech sector. Other scholarly content aggregators and publishers, such as EBSCO and Elsevier, are developing similar types of functions for their own platforms with distinct approaches, goals, and orientations.

We have yet to see how these developments will change our fundamental research behaviors, but their impacts could be quite thoroughgoing, with the potential to transform our relationship to the act of reading and to redefine the visibility of our own publications. Furthermore, as Donna Lanclos, Lawrie Phipps, and Richard Watermeyer have indicated, such developments entail "an interrogation of what consent means, in this time of embedded GAI tools in institutionally provided systems."[26]

Conclusion

This essay has touched upon many worlds that may seem distant from the standard content of *Ars Orientalis* and the spheres of Asian and Islamic art. Yet, it is hoped that these perspectives regarding copyright disputes, scientific publishing, journal submission guidelines, and academic content-delivery systems have been made relevant to the readership of this journal. Indeed, the goal of this short piece is to underline the ways in which such changes stand to affect our writing, its evaluation, its circulation, and its reception. Such transformations are well underway and are germane in regard to works that are already published in addition to forthcoming materials. This shifting landscape requires us to think about how our readers will interact with our scholarship, while factoring in the rising modes of mediation, creation, and access that have been spurred on by generative-AI technologies.

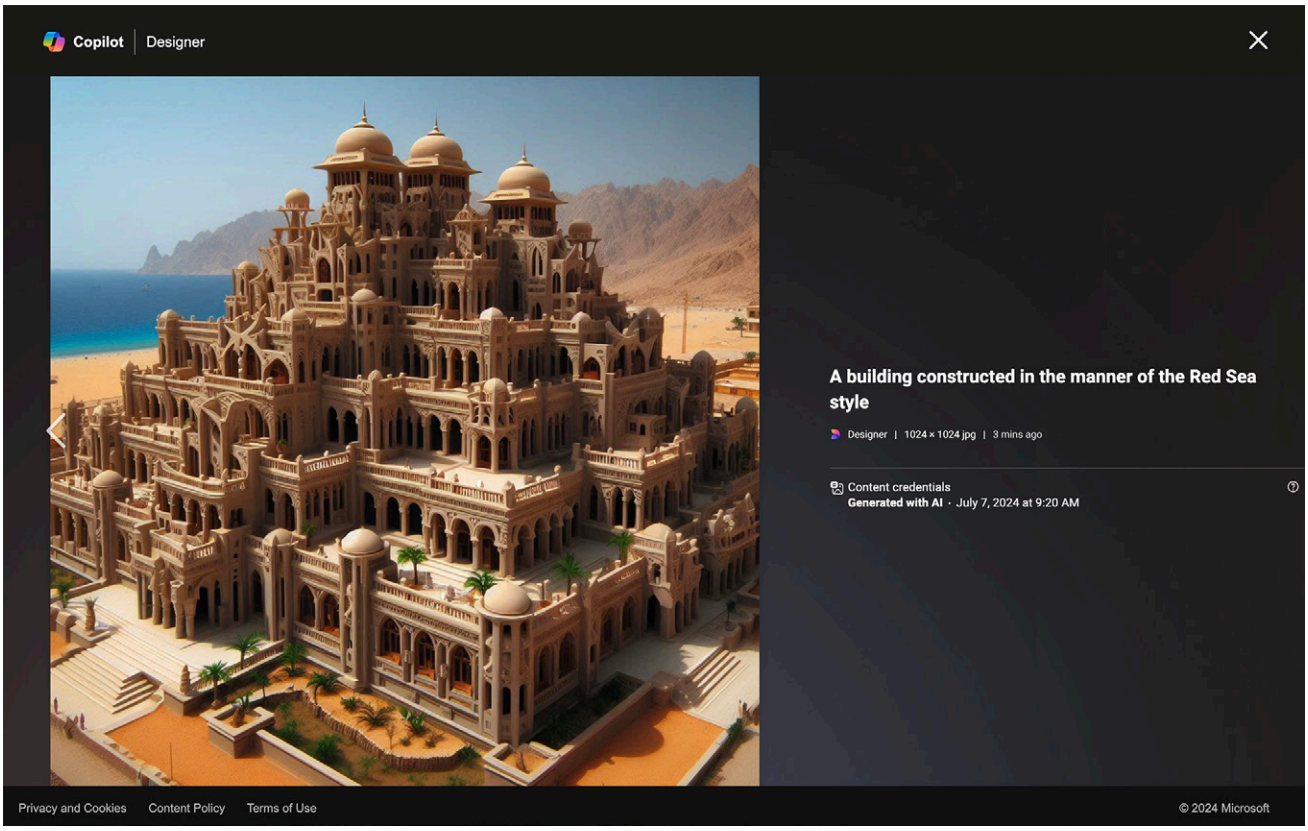

FIGURE 2. One of four images generated by Microsoft Copilot, based on the prompt "Please provide an image of a building constructed in the manner of the Red Sea style." Conversation initiated by the author, July 7, 2024

To conclude, I return to Copilot. In addition to the above-mentioned prompt, I asked Copilot to generate an image, again using the "more precise" mode, based on this request: "Please provide an image of a building constructed in the manner of the Red Sea style." Relying upon OpenAI's text to image model DALL-E 3, Copilot generated four images of towering stone structures with projecting arched tiers and myriad cupolas, set within a generic coastal locale (Fig. 2). These inventive images elaborate upon the vibrant undulation and ornamentation of an Indian temple façade, along with some Orientalizing arcades and domes. Copilot's synthetic visual representations are quite distinct from the Red Sea buildings described in the earlier text prompts, which pointed to whitewashed façades and elaborate woodwork. (See figure 3 for an example of a Red Sea style building.) Rather than simply highlighting the text to image limitations of DALL-E, this example emphasizes that LLMs must be understood as "stochastic parrots," both textually and visually, rather than as purveyors of meaning and intention, an idea that is fully articulated in the much-cited article by Emily M. Bender, Timnit Gebru, Angela McMillan-Major, and Margaret Mitchell.[27] Copilot delivers textual information in plausible strings that are intelligible to a human reader, but the chatbot has not *understood* the meaning of those words. Even further, there is no correspondence, yet, between the words and images that an instrument like Copilot generates. Indeed, this disconnect highlights the value of art historical understanding, which navigates the persistent complexities between words and pictures and dwells in the subtlety of visual interpretation.

FIGURE 3. Palace of Sultan Hasan (now destroyed), Mocha, Yemen, photograph by Hermann Burchardt, 1909, from Eugen Mittwoch, *Aus dem Jemen: Hermann Burchardts letzte Reise durch Südarabien* (Leipzig: Deutsche Morgenländische Gesellschaft, 1926), pl. xxii. Image in the public domain

Acknowledgments

Many thanks to the generous colleagues with whom I have been discussing these topics and who have offered their insightful feedback on this essay. They include but are not limited to Michele Ciaccio, Lidia Ferrara, Kara Kirk, and Susan Elizabeth Gagliardi.

Nancy Um, PhD (UCLA), 2001, is associate director for research and knowledge creation at the Getty Research Institute in Los Angeles, where she oversees research and digital initiatives. She served on the faculty of the department of Art History at Binghamton University from 2001 to 2022 and as associate dean for faculty development and inclusion in Harpur College of Arts and Sciences from 2019 to 2022. E-mail: naum@getty.edu

Notes

1 On widespread academic "technology refusal," see Donna Lanclos, Lawrie Phipps, and Richard Watermeyer, "Agency and Consent: The Impact of Generative AI in Academia," *Digital Is People*, July 1, 2024, accessed July 9, 2024, https://digitalispeople.org/agency-and-consent-the-impact-of-generative-ai-in-academia/

2 This piece was commissioned early in 2024 and most of the writing and editing was completed in early August 2024.

3 Mehtab Khan and Alex Hanna, "The Subjects and Stages of AI Dataset Development: A Framework for Dataset Accountability," September 13, 2022, posted October 11, 2022, *Ohio St. Tech. L.J.* 19 (2023), http://dx.doi.org/10.2139/ssrn.4217148

4 "The Authors Guild, John Grisham, Jodi Picoult, David Baldacci, George R. R. Martin, and 13 Other Authors File Class-Action Suit Against OpenAI," The Authors Guild, press release, September 20, 2023, accessed July 7, 2024, https://authorsguild.org/news/ag-and-authors-file-class-action-suit-against-openai/

5 The Glaze Project (https://glaze.cs.uchicago.edu), developed at the University of Chicago, offers various open-source resources to protect creative assets digitally.

6 Michael M. Grynbaum and Ryan Mac, "The Times Sues OpenAI and Microsoft over A.I. Use of Copyrighted Work," *New York Times*, December 27, 2023, accessed July 7, 2024, https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html

7 On the practice of third-party dataset audits, see Khan and Hanna, "Framework for Dataset Accountability," 18. On LLM "documentation debt," see Emily M. Bender, Timnit Gebru, Angelina McMillan-Major, and Shmargaret Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?," Proceedings of the ACM Conference on Fairness, Accountability, and Transparency (FAccT '21), March 3–10, 2021, Virtual Event, 615, https://doi.org/10.1145/3442188.3445922

8 Cade Metz and Katie Robertson, "OpenAI Seeks to Dismiss Parts of the New York Times's Lawsuit," *New York Times*, February 27, 2024, accessed July 21, 2024, https://www.nytimes.com/2024/02/27/technology/openai-new-york-times-lawsuit.html

9 Khan and Hanna, "Framework for Dataset Accountability," 34.

10 Stefaan G. Verhulst, "Are We Entering a 'Data Winter'?," *Medium*, January 23, 2024, accessed July 26, 2024, https://sverhulst.medium.com/are-we-entering-a-data-winter-f654eb8e8663

11 Matilda Battersby, "Academic Authors 'Shocked' After Taylor & Francis Sells Access to Their Research to Microsoft AI," *The Bookseller*, July 19, 2024, accessed July 20, 2024, https://www.thebookseller.com/news/academic-authors-shocked-after-taylor--francis-sells-access-to-their-research-to-microsoft-ai

12 Kathryn Palmer, "Taylor & Francis AI Deal Sets 'Worrying Precedent' for Academic Publishing," *InsideHigherEd*, July 29, 2024, accessed August 10, 2024, https://www.insidehighered.com/news/faculty-issues/research/2024/07/29/taylor-francis-ai-deal-sets-worrying-precedent

13 Copilot was previously known as Bing Chat. Copilot offers three "conversation modes" that the user can select before entering a prompt: more creative, more balanced, and more precise.

14 Nancy Um, "Reflections on the Red Sea Style: Beyond the Surface of Coastal Architecture," *Northeast African Studies* 12.1 (2012): 243–72, https://doi.org/10.1353/nas.2012.0013

15 Two of the references point to the same article on the Open Repository of Binghamton University (https://orb.binghamton.edu/art_hist_fac/4/), my previous institution, and behind a paywall through Project Muse, labeled in figure 1 as sources 3 and 4.

16 On the rationale that copyrighted materials should be integrated into testing datasets as a measure to mitigate bias, see Khan and Hanna, "Framework for Dataset Accountability," 43.

17 For examples of academic authors who tout the benefits of generative-AI tools as writing aids, see Lindsey Passenger Wieck, "Revising Historical Writing Using Generative AI: An Editorial Experiment," *Perspectives on History: The Newsmagazine of the American Historical Association*, August 15, 2023, accessed July 7, 2024, https://www.historians.org/perspectives-article/revising-historical-writing-using-generative-ai-an-editorial-experiment-august-2023/; and "Generative AI and the Enjoyment of Academic Writing," *Generative Dialogues: Generative AI in Higher Education*, podcast with Mark Carrigan and Helen Beetham, July 14, 2024, accessed July 19, 2024.

18 Chris Stokel-Walker, "ChatGPT Listed as Author on Research Papers: Many Scientists Disapprove," *Nature News*, January 18, 2023, accessed July 5, 2024, https://www.nature.com/articles/d41586-023-00107-z

19 "Authorship and AI Tools," COPE position statement, Committee on Publication Ethics, February 13, 2023, accessed July 12, 2024, https://publicationethics.org/cope-position-statements/ai-author

20 "Submitting Manuscripts to *PMLA*," accessed July 21, 2024, https://www.mla.org/Publications/Journals/PMLA/Submitting-Manuscripts-to-PMLA

21 "How Do I Cite Generative AI in MLA Style?," MLA Style Center, accessed July 21, 2024, https://style.mla.org/citing-generative-ai/

22 Andrew Gray, "ChatGPT 'Contamination': Estimating the Prevalence of LLMs in the Scholarly

Literature," arXiv pre-print, March 25, 2024, arXiv: 2403.16887 [cs.DL].

23 Claire Bytas and Dylan Ruediger, "Generative AI in Higher Education: The Product Landscape," Ithaka S+R, March 7, 2024, 5, https://doi.org/10.18665/sr.320394

24 Kevin Guthrie and Beth LaPensee, "Generative AI on JSTOR: What We're Learning from Early Usage, Real-World Applications, and User Feedback," *JSTOR*, June 25, 2024, accessed July 13, 2024, https://www.about.jstor.org/blog/generative-ai-on-jstor-what-were-learning/

25 Guthrie and LaPensee, "Generative AI on JSTOR."

26 Lanclos, Lawrie, and Watermeyer, "Agency and Consent."

27 Bender et al., "Stochastic Parrots."