KULA KNOWLEDGE CREATION, DISSEMINATION, AND PRESERVATION STUDIES

Blechinger, Joel. 2026. Probabilistic Obliteration and Formulaic Fabrication: Citational (In)justice in the Age of Generative Artificial Intelligence. *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 9(1). https://doi.org/10.18357.kula.310

RESEARCH ARTICLE

Probabilistic Obliteration and Formulaic Fabrication: Citational (In)justice in the Age of Generative Artificial Intelligence

Joel Blechinger
University of Calgary

What does it mean to cite generative AI (GenAI) tools—both in an instrumental, information retrieval sense, and in a symbolic sense that has more to do with recognition? What does it mean that GenAI tools are also able to produce plausible-looking yet false citations? This theoretical article critically explores the possibility of citational justice in the GenAI era through an analysis of two sets of examples: (1) the existing citation guidance to GenAI output as articulated by the major style guides, and (2) the issue of "hallucinated" (or fabricated) citations produced by GenAI large language model (LLM) chatbots like ChatGPT. Using ideas from Robert K. Merton, Eugene Garfield, Emily M. Bender, Robert J. Connors, and Sam Popowich, I argue that GenAI, across both sets of examples, is antithetical to citational justice. In the first set of examples, I make the case that the human authors and their works, the real source of GenAI tools' textual output, have been "obliterated"—a term that I borrow from Merton—as part of LLM training. This renders the official style guides' recommendations for how to cite GenAI tools—particularly APA's guidance—deeply inadequate. In the second set of examples, I see the fabricated citations produced by GenAI as antithetical to citational justice because they decentre the human. Crucially, though, these fabricated citations are actually perfectly suited to the irrelational context of contemporary higher education transformed by neoliberalism, where commodified student outputs are made to stand as evidence of students' internal transformation. As a closing gesture, I contend that the issues brought to the fore by GenAI and citation could present a pedagogical opportunity to radically reconceive of library instruction, focusing it on the importance of attribution and relationality in academic work if we so chose, and I offer some questions to guide that reimagined pedagogy.

Keywords: generative artificial intelligence; GenAI; large language models; ChatGPT; citational justice; hallucination; citation styles; APA; MLA; Chicago

Introduction

Robert K. Merton, an early sociologist of science, published two articles that, in their own tenuous way, indirectly laid some of the groundwork for the present-day citational justice movement. Across "The Matthew Effect in Science" (1968) and "The Matthew Effect in Science, II" (1988), Merton theorizes the titular, biblically named effect, which he describes in his earlier article (1968, 58) as "the accruing of greater increments of recognition for particular scientific contributions to scientists of considerable repute and the withholding of such recognition from scientists who have not yet made their mark." Citational justice, associated with "the deliberate act of citing more work authored by people from marginalized communities in society and in academia" (Coalter 2023, 62), essentially attempts to counteract the Matthew Effect's logic by encouraging scholars to cite more intentionally and equitably.

One cannot credibly claim that Merton's theory was social justice–oriented in intention, particularly in comparison to the citational justice movement. Nevertheless, if only by negative example, the Matthew Effect serves as a useful and unfortunately still timely shorthand for a citationally unjust scholarly reality where

recognition begets more recognition for a select few privileged scholars to the detriment of—as Merton identified—less notable co-authors, less senior colleagues, and—as figures associated with the citational justice movement have identified (Smith et al. 2021)—racialized, marginalized, non-normative identities in academia. Merton's approach was not at all intersectional, but it signaled a self-reflexivity around scientific practice and knowledge production that reflects concerns similar to those found in movements like citational justice.

Since 2022, the explosion of interest in generative artificial intelligence (GenAI) technologies has severely complicated the citational justice project. Put simply, I believe that the use of GenAI to do citational work is an unjust, irrelational style of citation. To illustrate this claim in this paper, I will examine issues central to citational justice and GenAI across two sets of (related) examples: (1) guidance issued by the official style guides about citing text outputted by GenAI tools, and (2) fabricated (or "hallucinated") citations generated by GenAI. Through my examination of these two sets of examples, I conclude that GenAI is antithetical to citational justice. In my first set of examples, I argue that authors and their works are—to use a concept from Merton—"obliterated" as part of large language model (LLM) training. As a result, and most markedly in the American Psychological Association (APA) case, official citation guidance from a style guide only serves to reify the corporate enclosure of the language commons via an overly simplistic citation to the model creator as the "author" of the GenAI tool's textual output. My second set of examples, the fabricated citations generated by GenAI, similarly illustrates the technologies' decentring of the human by revealing how formulaic reference forms have become: That they can be deployed without any meaningful connection to existing works or people illustrates that citation, in contemporary academic contexts, has become completely irrelational. Moreover, I argue that LLMs—as irrelational technologies that have no connection to knowledge about the world and only model the linguistic distribution of word forms—are a technology perfectly positioned to exploit the irrelational reality of contemporary higher education materially transformed by neoliberalism, where commodified student outputs are forced to stand as surrogates for inferred internal academic transformation.

Part I: Merton's Dual Function Framework and Citational Justice

In "The Matthew Effect in Science, II," Merton (1988, 622) provides an elegant explanation of the dual function of a reference:

> The bibliographic note, the reference to a source, is not merely a grace note, affixed by way of erudite ornamentation. (That it can be so used, or abused, does not of course negate its core uses.) The reference serves both instrumental and symbolic functions in the transmission and enlargement of knowledge. Instrumentally, it tells us of work we may not have known before, some of which may hold further interest for us; symbolically, it registers in the enduring archives the intellectual property of the acknowledged source by providing a pellet of peer recognition of the knowledge claim, accepted or expressly rejected, that was made in that source.

The instrumental and symbolic functions of a reference[1] will be intuitive to many instructional librarians who have led undergraduate students in exercises going over the reasons why scholars cite in academia. Librarians frequently offer workshops on citation as part of instruction programs on information literacy and academic integrity. These workshops can include content about the symbolic function of citation and the ethics of giving credit for someone else's ideas, but more often they focus on the instrumental function of the reference, on how to cite correctly according to different style guides (e.g., whether to include the author's full name or a surname and initial, where to put the year, how to format the volume and issue numbers of journals, etc.). Students learn primarily that the reason for citing (to avoid plagiarism) and the correct format of citation are of paramount importance. Certainly, a host of other reasons for citing can be added to Merton's two functions, including the many that Eugene Garfield (1965, 189) identifies, such as correcting one's own and others' work, alerting readers to forthcoming work, and identifying methodology.[2]

[1] Merton's instrumental and symbolic functions of a reference conveniently map onto competing discourses within librarianship and library and information science. The instrumental function reflects the commitments of positivist information science (or information retrieval), while paying greater attention to the reference's symbolic function could be said to be more of the province of critical librarianship, as reflected in the citational justice movement.

[2] However, as Blaise Cronin (1984, 30) articulated decades ago—and others following him have debated at great length (Nicolaisen 2007; Small 2016; Wouters 2016)—"in some cases it may well be possible to adduce the motive [for why authors cite others' work], but this is an attributive exercise, and cannot make the author's intentions explicit. There will invariably be a gap between why the author cited and why we *think* the author cited." This fundamental unknowability of why an author cited another is intellectually rich for scholars in scientometrics and citation analysis, but, in library instructional practice, less relevant, particularly when the limitations of the one-shot instructional session (Nicholson 2016; Pagowsky 2021; Almeida 2022) continuously impede deeper exploration of the issues at hand.

The current context of undergraduate writing is fundamentally surveillant (Watters 2020). Students are routinely expected to submit their work to text-matching (or "plagiarism detection") software like Turnitin and Ouriginal, and they experience significant stress in doing so (Goddiksen et al. 2024). The threat of an academic misconduct accusation looms behind poorly understood and frequently misinterpreted quantitative metrics like the Turnitin "similarity score" (Canzonetta 2018, 2021)—the percentage of writing in a submission that matches to another document in the company's proprietary database, which includes previously submitted student writing, crawled webpages, and scholarly journal content (Lee 2019)—and its newer tool for detecting the percentage of AI-generated text in a work, which has been shown to be less than reliable (Weber-Wulff et al. 2023).[3] Set against this context, citation pedagogy that is directed at students "from a place of compliance and fear" (Chenevey 2023, 152) is, unfortunately, commonplace in my experience. That pedagogical approach, in turn, inculcates in students an orientation towards citation where the reference as mere grace note becomes, essentially, its primary usage, deflated and self-protective.

In stark contrast to the deflated use of citation as self-protective grace note stand the citation practices envisioned by the citational justice movement. I first became aware of the citational justice movement early on in my MLIS degree program. Having previously studied at the graduate level in English, I was equipped with the theoretical vocabulary to think critically about the politics of citation when I arrived in my new discipline of library and information science. Many overviews of citational justice—for example, the ones that one might read on a standard LibGuide—will cite Sara Ahmed's blog post "Making Feminist Points," where she claims that citation is "a rather successful reproductive technology, a way of reproducing the world around certain bodies" (2013, para. 3). According to Ahmed, "the reproduction of a discipline can be the reproduction of *these techniques of selection* [like citation], ways of making certain bodies and thematics core to the discipline, and others not even part" (2013, para. 4). In addition to Ahmed's writing, Christen Smith's work with the Cite Black Women (CBW) movement has proven to be influential beyond the bounds of her home discipline of feminist anthropology and has diffused across librarianship. Essentially, Smith and the CBW movement sharpen Ahmed's points about citational politics by combining them with an explicitly anti-racist, decolonial set of political commitments: "Cite Black Women is a Black feminist intellectual project, praxis, and global movement to decolonize the practice of citation by redressing the epistemic erasure of Black women from the literal and figurative bibliographies of the world" (Smith et al. 2021, 12). It is important to note, however, that the goals of citational justice movements like the CBW extend far beyond cynical, tokenistic referencing of scholars of colour. Instead, CBW highlights citation understood "as a practice of *relation*" (Shange 2022) that signifies deeper engagement with Black women's ideas and works (Craven 2021) and that recognizes their agency beyond the page (Smith et al. 2021).

Conceptually, citational justice has clear connections to library instruction (Coalter 2023), even though the perpetual constraints of the standard one-shot instructional session (Nicholson 2016; Pagowsky 2021; Almeida 2022) may hamper one's ability to integrate citational justice content as thoroughly as desired. Librarians are optimally positioned to help students reflect critically on the voices they choose to amplify in their scholarly work, consider the importance of consulting a diversity of perspectives on a given research topic, and even explore different source types beyond traditional scholarly publishing formats (Thomas 2024).

Part II: Citational Injustice: AI Obliteration and Fabrication

Over the past several years, the popularization of citational justice discourse within librarianship during the 2010s has co-existed uncomfortably in my mind with the more recent frenzy of excitement around GenAI technologies. It is this sense of dissonance that I wish to interrogate in this piece by juxtaposing GenAI and citational justice, and I see this work as complementing a growing body of literature that examines the uncritical embrace of GenAI technologies across information work and the information professions.[4]

As articulated above, I see Merton's work as continuous with present-day citational justice movements like CBW. I also think his concepts are particularly useful in thinking through the implications of GenAI citation for citational justice and, ultimately, for illustrating that GenAI citation is unjust and irrelational. In what follows, I first make the case that human authors and their works, which are the real source of GenAI tools' textual output, have been "obliterated" (in Merton's words) as part of LLM training. Citational justice movements rely on the stability and accuracy of references (i.e., one must know the author of a work and the

[3] Though, in the Turnitin AI writing detector's case, this percentage is not even to a direct database match; rather, it "indicates the qualifying text . . . within the submitted document that Turnitin's AI writing detection model determines could be generated by AI or could be generated by AI and could be further modified using an AI paraphraser tool or an AI bypasser tool" (Turnitin 2025).

[4] See, for example, Andrea Baer (2025a, 2025b); Blechinger (2024a); Violet B. Fox (2024); Norah Mazel (2025); Nicole Lucero Murph (2025); Matthew Pierce (2025); Sam Popowich (2024a); and Kay Slater (2025).

work itself to be able to cite them), and the erasure of authorship that occurs in AI-generated text makes attribution of underrepresented scholars that much more difficult, obscuring already obscured voices. Second, I argue that the fabricated citations produced by GenAI are antithetical to citational justice because they decentre the human; however, in decentring the human, they align with the absence of relationality in the current paradigm of contemporary neoliberal higher education. The structural issues brought to the fore by GenAI and citation could, therefore, present a pedagogical opportunity to radically reconceive of library citation instruction, focusing it on the importance of attribution and relationality in academic work in the spirit of promoting citational justice.

Example I: Style Guide Citations of GenAI

Across the 2023–2024 academic year, I co-presented workshops at Mount Royal University with a colleague from the institution's Academic Development Centre. The workshops were entitled "Considering How Citation Guidance for AI Can Support Student Learning."[5] Playing a traditional librarian role in the sessions, I guided attendees through how each style guide had addressed (or not addressed) the citation of GenAI tools' outputs.

The key difference that I highlighted in the sessions had to do with how each style guide approached the question of GenAI authorship. The Modern Language Association's (MLA) (2023) GenAI citation guidance leaves the "author" element blank, taking a firm stance—also applied in the policy of its own journal, *PMLA*—that AI tools cannot be considered for authorship. Instead, the MLA centres the prompt in works cited list entries:

> "Describe the symbolism of the green light in the book *The Great Gatsby* by F. Scott Fitzgerald" prompt. *ChatGPT*, 13 Feb. version, OpenAI, 8 Mar. 2023, chat.openai.com/chat.

The APA, which published its guidance less than a month after the MLA, takes a different approach. Timothy McAdoo (2024, under "Creating a Reference to ChatGPT or Other AI Models and Software"), writing for the APA,[6] claims that the author element in a GenAI citation should be filled by "the author of the model," which, in the ChatGPT example he discusses, is OpenAI:

> OpenAI. (2023). *ChatGPT* (Mar 14 version) [Large language model]. https://chat.openai.com/chat

The University of Chicago Press's *Chicago Manual of Style* (*CMOS*) (2023) splits the difference between the MLA and APA, allowing a GenAI tool like ChatGPT to "stand in" as the author of generated output and relegating the developer to "publisher or sponsor" in both notes-bibliography and author-date styles.[7] Due to the irretrievability of GenAI output, *CMOS* also stipulates that authors not include GenAI content in a bibliography or reference list unless they are using a stable link.

The sessions were well received, but I had misgivings about presenting "rules" for citing GenAI output as if it were a simple problem that had already been solved. Across the sessions, I had a nagging feeling that the templates I provided were, at best, representative of a naive approach to the real, material issues at hand, and, at worst, illustrative of how the academic or cultural capital of librarianship was being hijacked by technology companies that were themselves indifferent—if not outright hostile—to authorship and scholarship itself. In examining the MLA guidance's comment section, I came across a commenter, the writing instructor Anna Mills, who proposed a revised template that crystallized my discomfort around the erasure of authorship in the official style guide templates. According to Mills's comment (on April 11, 2023) on MLA (2023), the way to "foreground the human critics who are the real sources of the ideas about the symbolism of the green light in the . . . [F. Scott] Fitzgerald example" that the MLA provided is to reinsert the critics back into the GenAI citation's author element. Mills therefore suggests a different template, which I later decided to include in a version of the AI citation guidance workshop tailored to an arts faculty audience:

[5] Slides for one version of this workshop can be viewed at this URL: https://tinyurl.com/y9nnkmd4.

[6] Over the course of the editing of this article, the APA updated its official GenAI citation guidance by releasing a newer, three-part series of blog entries on the subject. The central issue that I discuss in this section of the company behind the GenAI model occupying the "author" element in APA's official reference templates, however, persists in its updated guidance. See McAdoo, Samantha Denneny, and Chelsea Lee (2025) for even more examples of technology companies occupying the "author" element in sample GenAI references.

[7] Like the APA, *CMOS* updated their GenAI citation guidance over the course of this article's editing. In *CMOS*'s new eighteenth edition in-text and notes GenAI citation examples, the University of Chicago Press (2025, para. 1) hedges on GenAI authorship by stating, "ChatGPT is the author of the content (though not in the traditional sense)." For a reference to AI-generated content in a reference list or bibliography, though, the University of Chicago Press (2025) then stipulates that the name of the "publisher or developer" (for example, Google or OpenAI) should be cited in the author element.

> Unknown human authors statistically remixed by *ChatGPT*, 13 Feb. version, OpenAI, 8 Mar. 2023, chat.openai.com/chat. "Describe the symbolism of the green light in the book *The Great Gatsby* by F. Scott Fitzgerald" prompt.

Though one can certainly dispute whether Mills's word choice of "statistically remixing" is an accurate description of how a GenAI tool like ChatGPT functions, what her citation is doing is attempting to restore Merton's symbolic function to GenAI reference. Instrumentally, however, Mills's citation is completely useless. The proposed template exhibits a marked tension between its instrumental uselessness but striking symbolic power.

The APA's assertion that the model's "author"—itself an awkward formulation—should assume de facto authorship of its generated textual output is deeply ideological. It is ideological because it reifies the tech corporation as the source of many humans' intellectual labour by naming the corporation—and, therefore, obfuscating the human(s) in attribution—and mystifying human written work as "data" to be unscrupulously scraped, trained on, and monetized. This is a tension that is keenly felt by (real, human) authors when—at the moment of this article's composition—there are ongoing lawsuits alleging that the LLMs developed by OpenAI, Anthropic, Meta, and NVIDIA have been trained on massive corpora of pirated ebooks (Hawley 2023; Albanese 2024; Belanger 2023, 2025; Van der Sar 2024; Reisner 2025).

Papering over the host of copyright-related issues associated with the training of LLMs on large corpora of copyrighted texts, the arrival of GenAI citation guidance from the style guides—specifically the APA's guidance, with its deceptively simple advice to cite the model author—has buttressed (a thoroughly individually responsibilized)[8] AI literacy (Hervieux and Wheatley 2024; LaFlamme 2025; Lo 2025a, 2025b). AI literacy, as a nascent project for libraries and librarians to pursue, is at least in part an effort to remain relevant (Blechinger 2024a). Essentially, librarianship—experiencing its continuous, technologically deterministic identity crisis in which the profession's focus is pressured repeatedly to change in order to keep up with, and still remain useful in the wake of, evolving technologies—has now clearly moved on from its late 2010s investment in media and information literacy to AI literacy.

The fact that the MLA, APA, and *CMOS* all issued their respective GenAI citation guidance within six months of ChatGPT's debut at the end of November 2022 is also important to analyze critically. ChatGPT's release in the waning weeks of the fall 2022 academic semester was cause for significant consternation and stress for educators (Herman 2022; Scott 2023; Weissman 2023), already burnt out and overburdened from the COVID-19 pandemic's fallout (McMurtrie 2020; Gewin 2021). Due to mounting anxiety around GenAI tools in early 2023, the official style guides and other institutions intervened to codify directives—including authorized GenAI citation guidance—and to issue statements[9] that were intended to assuage educators' fears. One effect of these interventions was to effectively sidestep the ethically complicated dimension of what it even means to cite GenAI tools. What it means, I contend, is the erasure of the human author and individual voices.

Paying close attention to the throughline that exists between shadow libraries—illicit online text-sharing communities like Anna's Archive, Library Genesis (or LibGen), and Sci-Hub that flourished over the past decade, particularly in areas like the Global South with less well-resourced academic institutions (Karaganis 2018)—and LLM training corpora may appear gauche in library circles that are invested (as contemporary librarianship is) in copyleft initiatives and maximal openness as tenets of a kind of modern, technologized library faith.[10] One need not even be overly invested in the maintenance of traditional copyright regimes, however, to be troubled by LLMs' relation to the texts used in their training. Matthew Kirschenbaum and Rita Raley (2024), for example, wrote an article that touched upon Alex Reisner's (2023) publication of the Books3 dataset for authors to check whether their works were used in LLM training. Though they feel the need to highlight "the reinvestment in origins, the earnest searching for sources that attends the new copyright fundamentalism"—as represented by authors expressing legitimate concern over the unauthorized reproduction and use of their works—even Kirschenbaum and Raley admit that for-profit LLMs training on massive text corpora represents "an enclosure of the language commons" (2024, 510).

What does all of this focus on LLM training corpora and GenAI citation have to do with citational justice? As a movement, citational justice is fundamentally *dependent* on the stability of the reference form. One way

[8] AI literacy, understood to be individually responsibilized, is continuous with previous individually responsibilized projects of media and information literacy (Haider and Sundin 2022).

[9] One can interpret the Library Copyright Alliance's "Principles for Copyright and Artificial Intelligence" (2023) as a related effort to, essentially, pre-emptively authorize GenAI use in academia by categorically asserting that "the ingestion of copyrighted works to create large language models or other AI training databases generally is a fair use."

[10] I feel obligated to note here how amenable these tenets of the library faith are to exploitation by cyberlibertarian political actors. For more detail on this, see David Golumbia (2024, 21–5).

that the movement's dependence on the reference form's stability has been recognized has, somewhat ironically, been through the creation of procedural diversity audit tools like the Gender Citation Balance Indicesalyzer, originally developed by Brad Postle and Jacqueline Fulvio (2024) for authors submitting manuscripts to the *Journal of Cognitive Neuroscience*, and Dale Zhou et al.'s (2022) cleanBib code base, which probabilistically analyzes references for race and gender. These tools are misguided in their attempt to automate the deeper, more relational work of citational justice by tokenistically auditing one's reference list. The importance of the reference form's stability remains true, however, even as citational justice figures like CBW's Christen Smith emphasize that "just changing the proportion of authors from minoritized groups in reference lists is not enough" and advocate for more sustained engagement with and acknowledgement of diverse scholars' ideas (Kwon 2022, 571): One cannot enter into relation with a work or author without being able to reliably identify that work or author.

GenAI tools, however, completely destabilize the reference form, rendering it—in the citation guidance templates provided by the style guides—close to if not completely meaningless (Pierce 2025). As Avery Slater (2024, 223) declares, "at the level of their engineering, LLMs offer an altered relation to the extant network of ideas and proper names, and thus to the ethics of accreditation." Even calls for greater diversity in GenAI training data (Marwala 2024) do not improve the situation from the citational justice standpoint because such changes will materially amount to technology companies ultimately getting credit for marginalized individuals' ideas.

Related to this destabilization of the reference form, another term from Merton has taken on new, unexpected relevance again with GenAI's advent: the "anatopic or palimpsestic syndrome" of attribution, otherwise known as the "obliteration effect." Merton (1985, 218) describes obliteration as the phenomenon where "in the transmission of ideas[,] each succeeding repetition *tends to erase* all but one antecedent version." He elaborates in a footnote:

> Naturally enough, most of us tend to attribute a striking idea or formulation to the author who first introduced us to it. But often, that author has simply adopted or revived a formulation which he [*sic*] (and others versed in the same tradition) know to have been created by another. The transmitters may be so familiar with its origins that they mistakenly assume these to be well-known. Preferring not to insult their readers' knowledgeability, they do not cite the original source or even refer to it. And so it turns out that the altogether innocent transmitter becomes identified as the originator of the idea when his [*sic*] merit lies only in having kept it alive, or in having brought it back to life after it had long lain dormant or perhaps in having put it to new and instructive use. (Merton 1985, 218–19n)

Followers of Merton, like Garfield (1975), have framed the obliteration effect positively. Says Garfield (1975, 398), "Obliteration—perhaps even more than an astronomical citation rate—is one of the highest compliments the community of scientists can pay to the author." Using the example of Archimedes's constant, Garfield (1975, 398) enthuses, "If Archimedes were alive today, he could take comfort in the fact that his primordial paper on pi had been obliterated. It would mean that his contribution was so basic, so vital, and so well-known that scientists everywhere simply take it for granted. He would have been obliterated into immortality!"

LLMs engage in a kind of dynamic and probabilistic obliteration effect,[11] though neither Merton nor Garfield could have foreseen how presciently this concept would anticipate our present moment. It is strange, however, to now read Garfield's remarks, and to note how optimistic he is about the phenomenon. Rather than obliteration signaling "one of the highest compliments the community of scientists can pay to the author," present technological obliteration—as I have argued elsewhere (Blechinger 2024b)—liquidates authorship and attribution, stripping authors of even the paltry symbolic remuneration accorded them through citational recognition, not to mention any possible monetary compensation. Garfield, in his moment, was able to celebrate the teleological progress of science and the registering of a name like Archimedes within the enduring intellectual archive of scientific history. What has changed now is that the same archive has been enclosed, datafied, and is now used for relentless for-profit value extraction. This is certainly a "new and instructive use," but one with darker undertones than Merton and Garfield could have predicted.

[11] Compare LLM text generation's obliteration effect with Stephanie Decker's (2025, para. 8) idea of "citation laundering" that she uses in a blog post about GenAI model training on open access content: "AI outputs, especially when subsequently cited by human researchers, become a form of 'citation laundering,' where original sources are hidden or misattributed through AI generation. Not only will this make it impossible to trace the intellectual lineages of ideas, but the researchers who built foundational knowledge will no longer receive credit for their innovation."

Example II: Fabricated Citations Generated by GenAI

In the months after ChatGPT's November 2022 release, the chatbot's proclivity for generating erroneous but convincing information became apparent. This phenomenon was widely reported in the press, both popular (Klein 2023; Metz 2023; O'Brien 2023) and educational (Hicks 2023; Young 2023). The term "hallucination" was used frequently in much of this reporting,[12] and the claim that ChatGPT was "hallucinating" has been critiqued for anthropomorphizing GenAI, making it seem like it is conscious (Bender 2022; Barrow 2024; Hasan 2024). Many non-anthropomorphizing candidate terms have since been proposed to more precisely describe AI-generated falsehoods, including "bullshit" (Hicks, Humphries, and Slater 2024) and "mirages" (Mills and Angell 2025).[13] Bracketing whether this is an example of "criti-hype"—criticism that is "parasitic upon and even inflates hype" (Vinsel 2021, para. 2)—or not, OpenAI's (2023, 51) own documentation for GPT-4 claimed that "the profusion of false information from LLMs . . . has the potential to cast doubt on the whole information environment, threatening our ability to distinguish fact from fiction."

The term "hallucination" also suggests that the chatbot somehow "knows"—except when it is hallucinating—what information is "correct" in response to a prompt, and what information is "incorrect." Perhaps because of this misunderstanding of chatbot capabilities, some users prompt chatbot interfaces under the erroneous belief that they possess information retrieval capabilities.[14] This is a fundamental misunderstanding of how an LLM chatbot like ChatGPT generates text. LLMs like ChatGPT are not retrieving information in response to a question, and do not fact-check information against source materials. Instead, they generate a "reasonable continuation" of the preceding text, token by token (Wolfram 2023). Put slightly differently, "language models are prone to making stuff up . . . because they are not designed to express some underlying set of information in natural language; they are only manipulating the form of language" (Shah and Bender 2022, 222). Their output has nothing to do with "correct" and "incorrect," and everything to do with pure linguistic probabilities.[15]

Supporters of using GenAI as a component of information retrieval systems may interject at this stage in my argument and point to retrieval-augmented generation (RAG) as a solution to the fabrication issues that I have detailed. RAG is "an AI framework for improving the quality of LLM-generated responses by grounding the model on external sources of knowledge to supplement the LLM's internal representation of information" (Martineau 2023, para. 2). Admittedly, much of my thinking in this article related to GenAI fabricated citations was developed when earlier general-purpose models without RAG functionality, like ChatGPT and Gemini, dominated the GenAI conversation. Regardless, however, issues highly relevant to citational justice persist and, if anything, become even more complicated in the RAG context. First, a general point to remember: Text output by an LLM—whether or not enhanced with RAG—is still synthetic and therefore prone to error (Bender 2024a). Second, citational justice issues become even more complicated in the RAG context because complex numeric calculations undergirding processes like the creation of embedding models and vector similarity searches are now being asked to represent both users' search queries and authors' ideas expressed in a corpus of documents. Two concepts from the RAG literature illustrate some of the profound difficulties inherent to these information architectures most significantly: "citation faithfulness" and "negative rejection." Citation faithfulness is a RAG-specific term that, essentially, concerns how a RAG system will "'misinterprete' [*sic*] the item that is cited" (Tay 2024). Negative rejection describes when "a LLM should reject to answer the question when the required knowledge is not present in any retrieved document" (Chen et al. 2024, 17755), but sometimes the system does not do so and attempts to answer anyway because, ultimately—like its general-purpose, sycophantic chatbot counterparts—it aims to please the user (Lotz 2025).

Setting RAG-specific citational justice issues aside,[16] I want to look at one example of LLM-produced falsehoods that is particularly salient to the library instructional or reference interview context: fabricated citations being generated on demand for users (Simpson 2023; Tucci 2023; Welborn 2023). One of the most infamous and public examples of this phenomenon and its consequences was likely when the lawyers Steven Schwartz and Peter LoDuca were excoriated for submitting fabricated legal citations and made-up quotes in court in New York (Brodkin 2023). This later happened in Canada to BC lawyer Chong Ke too

[12] Though the term had a different prior meaning in the field of computer vision, where hallucination was seen as a valuable asset (Maleki, Padmanabhan, and Dutta 2024). See, for example, Simon Baker and Takeo Kanade's (2000) documentation of their "resolution enhancement algorithm" that had the ability to "hallucinate" additional pixels in a low-resolution image to improve overall image quality.

[13] As part of endeavouring to be careful with the language that I have used to describe false information, I have chosen the word "fabrication" in my own GenAI-related library instruction, and this is the term that I will use for the remainder of this paper.

[14] For a rigorous analysis of the (in)applicability of GenAI to information access systems, see Chirag Shah and Emily M. Bender (2024).

[15] There is also a risk that fabrication rates may *increase* with more advanced, so-called "reasoning" models. See, for example, Roland Moore-Colyer's (2025) reporting on a study released by OpenAI.

[16] Further engagement with RAG-specific citational justice issues could be a way to extend the work that I am attempting to do in this piece in a subsequent article.

(Proctor 2024).[17] Across 2025, fabricated citations made the news in several other, even higher-profile examples. May saw the Make America Healthy Again Commission release its Make Our Children Healthy Again Assessment (otherwise known as the MAHA Report), which contained fabricated citations to seven nonexistent studies among many other errors (Kennard and Manto 2025). In August, the academic Chris Rudge found numerous fabrication errors in a Deloitte report commissioned by the Australian government, eventually resulting in the firm partially refunding its fee back to the government (Karp 2025). Finally, fabricated citations surfaced in the Canadian context in September when an Education Accord NL final report was found to contain more than fifteen fabricated citations (Butler 2025), and—also in Newfoundland—a government-commissioned health human resource plan created by Deloitte was found to contain four fake citations (Whitten 2025).

Many scholarly attempts to explore the possible consequences of hallucinated or fabricated citations outputted by GenAI tools (Day 2023; Walters and Wilder 2023; Giray 2024; Watson 2024) have, thus far, been methodologically similar to what Fabian Offert and Ranjodh Singh Dhaliwal have identified as "red teaming": "a practice, in the fields of security and cybersecurity, wherein a group takes on (by prior planning and often consent) the role of an enemy and tries to infiltrate, attack, or harm in other ways, the entity/organization that finally benefits from knowing how its defenses could be breached so as to endeavor improving them" (2024, 2). In the most basic terms, this style of analysis essentially amounts to the author (or authors) prompting a GenAI tool for references and then scrutinizing them for the rate of fabrication. For example, Alex P. Watson (2024) finds that of the forty-six suspected fabricated citations he analyzed, which were either submitted by University of Mississippi instructors to him or by students to his library's chat service in spring 2023, two were wholly fabricated, whereas the other forty-four had at least one real citation element that was combined with fabricated elements.

These fabrication-audit-style articles satisfy a positivist discipline like library and information science (Popowich 2024b, 104) because they contain measurable data that can be reported back, laid out in tables, and broken down by source type. This approach fits right in with what Offert and Dhaliwal (2024, 2) characterize as a red-teaming methodology: "Specific prompts that generate specific texts, images, or sounds" being asked "to stand in for a universal critique of the abilities and possibilities offered by generative AI." As Offert and Dhaliwal explain, however, this approach has natural limitations: It is "a 'method' which actually threatens to become increasingly less successful as models continuously improve" (2024, 2). Offert and Dhaliwal put forward that it is unclear what the goals of fabrication-audit-style articles are when applied to continuously changing, probabilistic systems beyond the authors' amassing of point-in-time, unreplicable data for data's sake.

A more trenchant critique of this type of analysis, however, can be found in Emily Bender's (2023) work. She argues that the application of LLM technology to the search use case is fundamentally ill-conceived, and, therefore, that a red-teaming methodology of LLM critique is misguided: "If [LLM chatbots] get tuned to be wrong less of the time, they're going to become more dangerous because people are going to be more trustful of them" (Bender 2023, at 36:49). Bender's critique illuminates a more profound issue with fabrication-audit-style analyses. Specific to the library and information science context, how low would LLM chatbots' fabrication rates have to be for information professionals to feel comfortable authorizing their use for search tasks? Bender's insight suggests that the answer is not zero; it is simply low enough for us to feel like we can stop checking their work.

I think Bender's point and its implications should concern us in librarianship more than LLM chatbots' rates of fabricated references. I am much more interested in examining what AI-generated citations—and our interest in lowering their fabrication rates—reveal about citation in contemporary teaching, learning, and knowledge production contexts. These references—more precisely, their circulation in academic texts over the last several years—make clear that citation, at its most basic level, has become a formulaic rhetorical gesture that can be imitated in formal terms but is fundamentally evacuated of meaning.[18] Citation coming to operate in this way poses significant challenges not only for the project of citational justice but also, more fundamentally, for citation as "the scaffolding of scholarship" (O'Sullivan 2025, para. 2).

Returning to Merton's dual function of the reference with which I began Part I of this article, we can say that an AI-generated fabricated citation *clearly* has no instrumental function in that it does not lead to a real source that one could consult. The symbolic function of AI-generated fabricated citations is, however, more

[17] Fabricated references in the legal context are now so numerous that they are being tracked in several different databases online (Charlotin n.d.; Lee n.d.; Wondracek 2025).

[18] Even pre-GenAI, the formulaic and purely symbolic nature of citational metadata like DOIs was noted by figures working in library technology like Geoffrey Bilder (2016), who identified what he termed "DOI-like strings." Presciently, like Merton, Bilder's "DOI-like strings" resurface in the GenAI fabrication context, cropping up as plausible-looking DOIs in the lists of references to non-existent works that circulated virally across academic social networks in the fall of 2022 and winter of 2023.

complicated to parse, in particular regarding the case of a fabricated source attributed to a real scholar. In this instance, one could perhaps argue that—despite the non-existence of the actual source—the first symbolic function (a perverted "pellet of peer recognition," to use Merton's words) is being provided by the LLM chatbot, though not for a specific knowledge claim traceable to an actually existing source. Rather, the reference to a real scholar perhaps merely evinces the statistical notability of that author's name within a textual corpus.[19] Distressingly, more than being largely symbolically empty, this notion of statistical notability within a textual corpus also reenacts many of the systemic problems that the citational justice movement was founded to work against because statistical notability within a corpus is, essentially, prevalence of citation to a particular scholar. Whereas citational justice advocates like Carrie Mott and Daniel Cockayne implore scholars to "understand citation as a performative practice of conscientious engagement" (2017, 963) with others' ideas, GenAI citation, in this example, reveals itself to be "a technology for reproducing sameness and excluding difference" (2017, 960).

An AI-generated fabricated citation also has another symbolic function: that of a *purely* symbolic "grace note." Merton (1988, 622), though he would not have been contemplating AI at the time, briefly touches upon this function too: "The bibliographic note, the reference to a source, is not merely a grace note, affixed by way of erudite ornamentation. (That it can be so used, or abused, does not of course negate its core uses)." The abuse of the bibliographic note as "grace note," an occluded second symbolic function even at the time of Merton's writing, is less respectable than the first symbolic function, and fundamentally more formulaic and self-protective.[20] It is a symbolic function operative within and perfectly suited to a thoroughly commodified higher educational context materially transformed by neoliberalism.

Sam Popowich (2024a, 2024b) has written compellingly about how GenAI technologies have arrived to "disrupt" knowledge work and higher education set against this context. For Popowich (2024b), if we focus solely on the sensational "threat" that GenAI technologies pose to learning, we overlook how higher education has been transformed as the result of managerialism, corporatization, and commodification in such a way as to make it possible for synthetic, generated outputs to be mistaken for proxies of a student's internal academic transformation in the first place. Crucially, though, as Popowich (2024b, 207–8) articulates,

> Large Language Models like the GPTs . . . challenge a major assumption of capitalist education. . . . This assumption is that there is in fact a relationship between a student's output and their learning, that an output (essay, exam, etc.) directly reflects their learning, their internal experience with knowledge and language . . . ChatGPT may have the signal result of decoupling that particular signified/signifier pair and putting paid to a cardinal tenet of academic ideology leftover from pre-corporate days: that education is about internal transformation of subjectivity rather than about the generation of texts.

Before November 2022, learning in higher educational contexts was already transactional, output-focused, and—despite important efforts to mitigate this tendency (Fitzpatrick 2019; Littletree, Andrews, and Loyer 2023)—irrelational. GenAI arrived perfectly suited to this context. As Bender (2024b, 116) states, "the 'knowing' that we program into 'AI' is . . . *irrelational*, that is, ostensibly abstracted from the web of relations within which we have all of these experiences [of ourselves, our lives, and our world]." Rampant GenAI use in higher education irrevocably alters any extant meaningful connection between learning and output that undergirds neoliberal assessment practices. More profoundly, however—and more to Bender's point—irrelationality is *fundamental* to GenAI technologies' architecture, and we risk making a significant category error if we mistake "information about the distribution of word forms" (Shah and Bender 2024, 9) for grounded knowledge about the world.

The fact that authors' names could even be (quite literally) *tokenized* as part of fabricated GenAI citations is the logical endpoint of the evolution of the citation styles that are now dominant in contemporary academia. As Robert J. Connors (1998, 1999) shows across two pieces historicizing and unpacking what he terms the "rhetoric of citation systems," parenthetical citation systems like MLA and APA effectively reduced authors' names to deracinated data points prior to GenAI's advent. For Connors (1999, 238–39), "parenthetical citations gave new importance to the date of published works as they diminished the importance of the author. Rather than being 'Samuel Johnson,' the author becomes '(Johnson 1755)' in the text and 'Johnson, S. (1755)' in the List of Works Cited. Rhetorically, authors lose agency here, as their surnames become nametags

[19] David Smerdon (2023) performed an analysis on Twitter (now X) of a fabricated reference provided by ChatGPT in response to the prompt "What is the most cited economics paper of all time?" wherein he noted that Douglass North was fabricated as the paper's "most probable" author likely for this reason.

[20] This formulaic, self-protective use of the bibliographic note is similar to uses of citation that Bruno Latour (1987) identified when he examined the primacy of citation to rhetorical persuasion and authority claim in scientific writing.

for works." Scientific disciplines' adoption of parenthetical citation systems and those systems' central focus on dates—and their consequent de-emphasis on authors and authorial agency—is reflective of a significant shift in those disciplines' "attitude toward cumulation of knowledge and supersession of outdated knowledge in a field" (Connors 1999, 240). Connors (1999) also contrasts parenthetical citation systems with footnote systems in terms of how the latter—as reflective of humanities disciplines—are much more concerned with being able to quote a specific author's exact words. In contrast, in scientific disciplines that came to favour parenthetical citation systems, summary and paraphrase as opposed to direct quotation predominate, and "citations came more and more to be of a whole article rather than of specific pages" (Connors 1999, 239).

Like Merton, Connors is not speaking to the GenAI context in these articles, but it is striking how much his analysis of parenthetical citation systems presages GenAI, particularly the AI-generated fabricated citations that I have discussed in this section. He says, tongue firmly in cheek, "If authors' first names are reduced to initials, they should get over it. If 'the research' is made to have more reality than the persons who conducted it, so much the better. . . . What is important is that the march of science not be impeded by any issues that would cloud the referential efficiency of what is, after all, a professional literature" (Connors 1999, 242). Indeed, research being "made to have more reality than the persons who conducted it" is, perversely, an accurate description of the present AI-powered academic conjuncture.

It is also worth highlighting here the note that Connors sounds about "the referential efficiency . . . of a professional literature." GenAI is the apotheosis of this referential efficiency where even (now negligible) details like proper attribution and the actual existence of specific scholars or works can be disregarded in service to the hyper-accelerated text generation required of both students (producing the assignment commodities they desire to exchange for grades, degrees, jobs, and, ostensibly, a higher quality of life) and scholars (subject to the competitive dictates of "publish or perish") alike. Human, relational processes are decentred and marginalized in each case. According to Popowich (2024a, 14), "only if we ignore the hidden, obscured human processes that we participate in when we write (or perform or paint or teach), only as long as we focus solely on the output, can we mistake LLM-generated text for the human process of creative production." To attempt to speak of citational justice—let alone material justice—in light of such profoundly irrelational technologies leaves one uncertain where to even begin. What can citational justice even mean if we are not centring the people themselves who we are citing, and, presumably, for whom we want that justice? How can we credibly speak of anything related to citational justice if we are not firmly opposed to the reduction of an author's name to a linguistic token that can be probabilistically recombined with other tokens, irrespective of that token's relation to the person it "points to" out in the world and their body of work and embeddedness in a network of relations with others? These issues highlight that, as with citation guidance from the official style guides that uncomplicatedly advise referencing GenAI tools, AI-generated fabricated citations are also antithetical to citational justice.

Conclusion

To part on a more hopeful note, however, I would like to provide some provocations that may point toward what a reimagined library instruction around GenAI citation *could* be if we actually chose to engage with some of the issues at hand that I have identified. To articulate this takes me back, full circle, to the initial discomfort that I felt when presenting GenAI citation workshops to faculty in workshops in 2023–2024. During those sessions—as I parroted GenAI citation guidance from the official style guides—I felt that the real opportunity for meaningful learning about and *against* GenAI was submerged beneath the surface of the official guidance, barely touched by the conversation happening at the pedantic level of citation elements and templates.

To rectify this, I propose the following five questions that could prompt reflection on what a re-engaged, relational citation pedagogy in the age of GenAI could be:

- How can we take the opportunity that GenAI presents for our pedagogy to radically reimagine library instruction around the importance of attribution and relationality in academic work?
- What if, instead of teaching the mechanics of citation from a place of either rote, templated compliance or honour-code-scolding fear, we approached citation instruction with the explicit aim of humanizing authors, intentionally restoring some of their agency through acknowledgement, and, in turn, working to bolster the agency of aspiring student authors?
- If we took seriously legitimate claims that the writing process is also an embodied process of thinking and of feeling (Menary 2007; Warner 2025), how might this change our teaching of citation practices?

- How could we productively combine analyses of GenAI's material harms—such as the significant climate impact of LLM training and use (Kneese and Young 2024; Luccioni, Trevelin, and Mitchell 2024) or the exploitative labour practices involved in data cleaning and annotation (Dzieza 2023; Perrigo 2023; Muldoon, Graham, and Cant 2024)—with critical analyses of GenAI's discursive or citational harms done to authors and other content creators?
- What would it mean to teach against the grain of hegemonic style guides like the APA, and to disagree productively—emphatically and explicitly—with their official guidance in our own pedagogy and in our interactions with students and faculty?[21]

Of course, these avenues for pedagogical exploration of GenAI citation issues would not singlehandedly bring about citational justice in the GenAI context, but they would at least clear space for us to begin to think through it again in the wake of increased use of these fundamentally irrelational technologies.

Competing Interests

The author declares no conflict of interest regarding this manuscript.

References

Ahmed, Sara. 2013. "Making Feminist Points." *feministkilljoys* (blog), September 11. https://feministkilljoys.com/2013/09/11/making-feminist-points/.

Albanese, Andrew. 2024. "Authors Sue AI Firm Anthropic for Copyright Infringement." *Publishers Weekly*, August 20. https://www.publishersweekly.com/pw/by-topic/industry-news/publisher-news/article/95768-authors-sue-ai-firm-anthropic-for-copyright-infringement.html.

Almeida, Nora. 2022. "Library Tautology: A Reenactment of the One-Shot." *College & Research Libraries* 83 (5): 833–36. https://doi.org/10.5860/crl.83.5.833.

Baer, Andrea. 2025a. "Investigating the 'Feeling Rules' of Generative AI and Imagining Alternative Futures." *In the Library with the Lead Pipe*. https://www.inthelibrarywiththeleadpipe.org/2025/ai-feeling-rules/.

Baer, Andrea. 2025b. "Unpacking Predominant Narratives about Generative AI and Education: A Starting Point for Teaching Critical AI Literacy and Imagining Better Futures." *Library Trends* 73 (3): 141–59. https://dx.doi.org/10.1353/lib.2025.a961189.

Baker, Simon, and Takeo Kanade. 2000. "Hallucinating Faces." In *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*. Institute of Electrical and Electronics Engineers. https://doi.org/10.1109/AFGR.2000.840616.

Barrow, Nicholas. 2024. "Anthropomorphism and AI Hype." *AI and Ethics* 4: 707–11. https://doi.org/10.1007/s43681-024-00454-1.

Belanger, Ashley. 2023. "Sarah Silverman Sues OpenAI, Meta for Being 'Industrial-Strength Plagiarists.'" *Ars Technica*, July 10. https://arstechnica.com/information-technology/2023/07/book-authors-sue-openai-and-meta-over-text-used-to-train-ai/.

Belanger, Ashley. 2025. "Authors Celebrate 'Historic' Settlement Coming Soon in Anthropic Class Action." *Ars Technica*, August 26. https://arstechnica.com/tech-policy/2025/08/authors-celebrate-historic-settlement-coming-soon-in-anthropic-class-action/.

Bender, Emily (@emilymbender@dair-community.social). 2022. "And let's reflect for a moment on how they phrased their disclaimer, shall we? 'Hallucinate' is a terrible word choice." Mastodon, November 16. https://dair-community.social/@emilymbender/109355539906866849.

Bender, Emily. 2023. "ChatGP-Why: When, if Ever, Is Synthetic Text Safe, Appropriate, and Desirable?" Virtual lecture, August 8. Posted August 12, 2023, by Global Research Alliance for AI in Learning and Education. YouTube, 59:16. https://youtu.be/qpE40jwMilU?si=2wiibpdmsSk-0OD3.

Bender, Emily. 2024a. "Information Literacy and Chatbots as Search." *Mystery AI Hype Theater 3000: The Newsletter*, November 5. https://buttondown.com/maiht3k/archive/information-literacy-and-chatbots-as-search/.

Bender, Emily M. 2024b. "Resisting Dehumanization in the Age of 'AI.'" *Current Directions in Psychological Science* 33 (2): 114–20. https://doi.org/10.1177/09637214231217286.

Bilder, Geoffrey. 2016. "DOI-like Strings and Fake DOIs." *Crossref Blog*, June 29. https://doi.org/10.64000/8mj8h-2t809.

[21] Perhaps a gesture like this would also go some way towards legitimizing librarians as educators in the eyes of (other) faculty, as librarians would work against an image of themselves as service providers or neutral vessels for official style guidance and instead appear actively engaged in their own dynamic field of study.

Blechinger, Joel. 2024a. "Insist on Sources: Wikipedia, LLMs, and the Limits of Information Literacy Instruction." Paper presented at CAPAL24: Sustainability and Shared Futures in Academic Libraries Conference, virtual, June 7. https://doi.org/10.60770/36Y6-3562.

Blechinger, Joel. 2024b. Review of *Verified: How to Think Straight, Get Duped Less, and Make Better Decisions about What to Believe Online*, by M. Caulfield & S. Wineburg. *Journal of Radical Librarianship* 10: 8–11. https://journal.radicallibrarianship.org/index.php/journal/article/view/108.

Brodkin, Jon. 2023. "Lawyers Have Real Bad Day in Court After Citing Fake Cases Made Up by ChatGPT." *Ars Technica*, June 23. https://arstechnica.com/tech-policy/2023/06/lawyers-have-real-bad-day-in-court-after-citing-fake-cases-made-up-by-chatgpt/.

Butler, Patrick. 2025. "N.L.'s 10-Year Education Action Plan Cites Sources That Don't Exist." *CBC News*, September 12. https://www.cbc.ca/news/canada/newfoundland-labrador/education-accord-nl-sources-dont-exist-1.7631364.

Canzonetta, Jordan. 2018. "Can We Use Plagiarism Detection Services Responsibly?" In *Student Plagiarism in Higher Education*, edited by Diane Pecorari and Philip Shaw. Routledge. https://doi.org/10.4324/9781315166148.

Canzonetta, Jordan. 2021. "Repurposing Plagiarism Detection Services for Responsible Pedagogical Application and (In)Formative Assessment of Source Attribution Practices." *Assessing Writing* 50: Article 100563. https://doi.org/10.1016/j.asw.2021.100563.

Charlotin, Damien. n. d. "AI Hallucination Cases." *Damien Charlotin*. https://www.damiencharlotin.com/hallucinations/.

Chen, Jiawei, Hongyu Lin, Xianpei Han, and Le Sun. 2024. "Benchmarking Large Language Models in Retrieval-Augmented Generation." *Proceedings of the AAAI Conference on Artificial Intelligence* 38 (16): 17754–62. https://doi.org/10.1609/aaai.v38i16.29728.

Chenevey, Liz. 2023. "Teaching the Politics of Citation: Challenging Students' Perceptions." *College & Research Libraries News* 84 (5): 152–57. https://doi.org/10.5860/crln.84.5.152.

Coalter, Jodi H. 2023. "Citation Power: Overcoming Marginalization One Citation at a Time." In *Perspectives on Justice, Equity, Diversity, and Inclusion in Libraries*, edited by Nandita S. Mani, Michaelle A. Cawley, and Emily P. Jones. IGI Global. https://doi.org/10.4018/978-1-6684-7255-2.ch004.

Connors, Robert J. 1998. "The Rhetoric of Citation Systems—Part I: The Development of Annotation Structures from the Renaissance to 1900*." *Rhetoric Review* 17 (1): 6–48. https://doi.org/10.1080/07350199809359230.

Connors, Robert J. 1999. "The Rhetoric of Citation Systems—Part II: Competing Epistemic Values in Citation." *Rhetoric Review* 17 (2): 219–45. https://doi.org/10.1080/07350199909359242.

Craven, Christa. 2021. "Teaching Antiracist Citational Politics as a Project of Transformation: Lessons from the Cite Black Women Movement for White Feminist Anthropologists." *Feminist Anthropology* 2 (1): 120–29. https://doi.org/10.1002/fea2.12036.

Cronin, Blaise. 1984. *The Citation Process: The Role and Significance of Citations in Scientific Communication*. Taylor Graham. https://garfield.library.upenn.edu/cronin/citationprocess.pdf.

Day, Terence. 2023. "A Preliminary Investigation of Fake Peer-Reviewed Citations and References Generated by ChatGPT." *The Professional Geographer* 75 (6): 1024–27. https://doi.org/10.1080/00330124.2023.2190373.

Decker, Stephanie. 2025. "Guest Post — The Open Access–AI Conundrum: Does Free to Read Mean Free to Train?" *The Scholarly Kitchen*, April 15. https://scholarlykitchen.sspnet.org/2025/04/15/guest-post-the-open-access-ai-conundrum-does-free-to-read-mean-free-to-train/.

Dzieza, Josh. 2023. "AI Is a Lot of Work." *The Verge*, June 20. https://www.theverge.com/features/23764584/ai-artificial-intelligence-data-notation-labor-scale-surge-remotasks-openai-chatbots.

Fitzpatrick, Kathleen. 2019. *Generous Thinking: A Radical Approach to Saving the University*. Johns Hopkins University Press.

Fox, Violet B. 2024. *A Librarian Against AI or I Think AI Should Leave*. https://violetbfox.info/against-ai/.

Garfield, Eugene. 1965. "Can Citation Indexing be Automated?" In *Statistical Association Methods for Mechanized Documentation: Symposium Proceedings, Washington 1964*, edited by Mary Elizabeth Stevens, Vincent E. Giuliano, and Laurence B. Heilprin. National Bureau of Standards. https://nvlpubs.nist.gov/nistpubs/Legacy/MP/nbsmiscellaneouspub269.pdf.

Garfield, Eugene. 1975. "The 'Obliteration Phenomenon' in Science—and the Advantage of Being Obliterated!" *Current Comments* 51/52: 396–98. https://garfield.library.upenn.edu/essays/v2p396y1974-76.pdf.

Gewin, Virginia. 2021. "Pandemic Burnout Is Rampant in Academia." *Nature* 591: 489–91. https://doi.org/10.1038/d41586-021-00663-2.

Giray, Louie. 2024. "ChatGPT References Unveiled: Distinguishing the Reliable from the Fake." *Internet Reference Services Quarterly* 28 (1): 9–18. https://doi.org/10.1080/10875301.2023.2265369.

Goddiksen, Mads Paludan, Mikkel Willum Johansen, Anna Catharina Vieira Armond, Mateja Centa, Christine Clavien, Eugenijus Gefenas, Nóra Kovács, Marcus Tang Merit, I. Anna S. Olsson, Margarita Poškutė, Júlio

Borlido Santos, Rita Santos, Vojko Strahovnik, Orsolya Varga, P. J. Wall, Peter Sandøe, and Thomas Bøker Lund. 2024. "The Dark Side of Text-Matching Software: Worries and Counterproductive Behaviour Among European Upper Secondary School and Bachelor Students." *International Journal of Educational Integrity* 20: Article 15. https://doi.org/10.1007/s40979-024-00162-7.

Golumbia, David. 2024. *Cyberlibertarianism: The Right-Wing Politics of Digital Technology*. University of Minnesota Press.

Haider, Jutta, and Olof Sundin. 2022. *Paradoxes of Media and Information Literacy: The Crisis of Information*. Routledge. https://doi.org/10.4324/9781003163237.

Hasan, Ali. 2024. "Are You Anthropomorphizing AI?" *Blog of the American Philosophical Association*, August 20. https://blog.apaonline.org/2024/08/20/are-you-anthropomorphizing-ai-2/.

Hawley, Steven C. 2023. "OpenAI Trains Its GPT Model Using Pirated E-books, Contends Authors' Lawsuit." Piracy Monitor. June 30. https://piracymonitor.org/chat-gpt-trained-using-pirated-e-books/.

Herman, Daniel. 2022. "The End of High-School English." *The Atlantic*, December 9. https://www.theatlantic.com/technology/archive/2022/12/openai-chatgpt-writing-high-school-english-essay/672412/.

Hervieux, Sandy, and Amanda Wheatley. 2024. "Building an AI Literacy Framework: Perspectives from Instruction Librarians and Current Information Literacy Tools." https://www.choice360.org/wp-content/uploads/2024/08/TaylorFrancis_whitepaper_08.28.24_final.pdf.

Hicks, Maggie. 2023. "No, ChatGPT Can't Be Your New Research Assistant." *The Chronicle of Higher Education*, August 23. https://www.chronicle.com/article/no-chatgpt-cant-be-your-new-research-assistant.

Hicks, Michael Townsen, James Humphries, and Joe Slater. 2024. "ChatGPT Is Bullshit." *Ethics and Information Technology* 26: Article 38. https://doi.org/10.1007/s10676-024-09775-5.

Karaganis, Joe, ed. 2018. *Shadow Libraries: Access to Knowledge in Global Higher Education*. The MIT Press. https://doi.org/10.7551/mitpress/11339.001.0001.

Karp, Paul. 2025. "AI-Tainted Deloitte Report Was Worse Than Previously Thought." *Financial Review*, November 6. https://www.afr.com/politics/ai-tainted-deloitte-report-was-worse-than-previously-thought-20251106-p5n863.

Kennard, Emily, and Margaret Manto. 2025. "The MAHA Report Cites Studies That Don't Exist." *NOTUS*, May 29. http://notus.org/health-science/make-america-healthy-again-report-citation-errors.

Kirschenbaum, Matthew, and Rita Raley. 2024. "AI and the University as a Service." *PMLA* 139 (3): 504–15. https://doi.org/10.1632/S003081292400052X.

Klein, Naomi. 2023. "AI Machines Aren't 'Hallucinating.' But Their Makers Are." *The Guardian*, May 8. https://www.theguardian.com/commentisfree/2023/may/08/ai-machines-hallucinating-naomi-klein.

Kneese, Tamara, and Meg Young. 2024. "Carbon Emissions in the Tailpipe of Generative AI." *Harvard Data Science Review* (Special Issue 5). https://doi.org/10.1162/99608f92.fbdf6128.

Kwon, Diana. 2022. "The Rise of Citational Justice: How Scholars Are Making References Fairer." *Nature* 603: 568–72. https://doi.org/10.1038/d41586-022-00793-1.

LaFlamme, Katherine A. 2025. "Scaffolding AI Literacy: An Instructional Model for Academic Librarianship." *The Journal of Academic Librarianship* 51 (3): 103041. https://doi.org/10.1016/j.acalib.2025.103041.

Latour, Bruno. 1987. *Science in Action: How to Follow Scientists and Engineers Through Society*. Harvard University Press.

Lee, Christine. 2019. "How Much Do You Know About Turnitin's Content?" *Turnitin Blog*, May 30. https://www.turnitin.com/blog/how-much-do-you-know-about-turnitins-content.

Lee, Matthew. n. d. "AI Hallucination Cases Tracker." *Natural and Artificial Intelligence in Law*. https://naturalandartificiallaw.com/ai-hallucination-cases-tracker/.

Library Copyright Alliance. 2023. "Library Copyright Alliance Principles for Copyright and Artificial Intelligence." Copyright and Artificial Intelligence, issued July 10. https://www.librarycopyrightalliance.org/wp-content/uploads/2023/06/AI-principles.pdf.

Littletree, Sandra, Nicola Andrews, and Jessie Loyer. 2023. "Information as a Relation: Defining Indigenous Information Literacy." *Journal of Information Literacy* 17 (2). https://doi.org/10.11645/17.2.8.

Lo, Leo S. 2025a. "AI Literacy: A Guide for Academic Libraries." *College & Research Libraries News* 86 (3): 120–22. https://doi.org/10.5860/crln.86.3.120.

Lo, Leo S. 2025b. "AI Literacy for All: A Universal Framework." Preprint, University of New Mexico Digital Repository. https://digitalrepository.unm.edu/ulls_fsp/213/.

Lotz, Avery. 2025. "AI Sycophancy: The Downside of a Digital Yes-Man." *Axios*, July 7. https://www.axios.com/2025/07/07/ai-sycophancy-chatbots-mental-health.

Luccioni, Sasha, Bruna Trevelin, and Margaret Mitchell. 2024. "The Environmental Impacts of AI – Primer." *Hugging Face*, September 3. https://huggingface.co/blog/sasha/ai-environment-primer.

Maleki, Negar, Balaji Padmanabhan, and Kaushik Dutta. 2024. "AI Hallucinations: A Misnomer Worth Clarifying." Preprint, arXiv, January 9. https://doi.org/10.48550/arXiv.2401.06796.

Martineau, Kim. 2023. "What Is Retrieval-Augmented Generation?" *IBM Research*, August 22. https://research.ibm.com/blog/retrieval-augmented-generation-RAG.

Marwala, Tshilidzi. 2024. "The Commercial Importance of Diversity in Generative AI." *Daily Maverick*, August 21. https://www.dailymaverick.co.za/opinionista/2024-08-21-the-commercial-importance-of-diversity-in-generative-ai/.

Mazel, Norah. 2025. "Generative AI and the Reinforcement of Information Privilege: Implications for Information Literacy Instruction." *Library Trends* 73 (3): 323–43. https://dx.doi.org/10.1353/lib.2025.a961197.

McAdoo, Timothy. 2024. "How to Cite ChatGPT." *APA Style Blog*, February 23. Last updated September 15, 2025. https://apastyle.apa.org/blog/how-to-cite-chatgpt.

McAdoo, Timothy, Samantha Denneny, and Chelsea Lee. 2025. "Citing Generative AI in APA Style: Part 1—Reference Formats." *APA Style Blog*, September 9. https://apastyle.apa.org/blog/cite-generative-ai-references.

McMurtrie, Beth. 2020. "The Pandemic Is Dragging On. Professors Are Burning Out." *The Chronicle of Higher Education*, November 5. https://www.chronicle.com/article/the-pandemic-is-dragging-on-professors-are-burning-out.

Menary, Richard. 2007. "Writing as Thinking." *Language Sciences* 29 (5): 621–32. https://doi.org/10.1016/j.langsci.2007.01.005.

Merton, Robert K. 1968. "The Matthew Effect in Science." *Science* 159 (3810): 56–63. https://doi.org/10.1126/science.159.3810.56.

Merton, Robert K. 1985. *On the Shoulders of Giants: A Shandean Postscript.* Vicennial ed. Harcourt Brace Jovanovich.

Merton, Robert K. 1988. "The Matthew Effect in Science, II: Cumulative Advantage and the Symbolism of Intellectual Property." *Isis* 79 (4): 606–23. https://doi.org/10.1086/354848.

Metz, Cade. 2023. "What Makes A.I. Chatbots Go Wrong?" *The New York Times*, April 4. https://www.nytimes.com/2023/03/29/technology/ai-chatbots-hallucinations.html.

Mills, Anna, and Nate Angell. 2025. "Are We Tripping? The Mirage of AI Hallucinations." Preprint, SSRN, February 13. http://dx.doi.org/10.2139/ssrn.5127162.

Modern Language Association. 2023. "How Do I Cite Generative AI in MLA Style?" *MLA Style Center*, March 17. https://style.mla.org/citing-generative-ai/.

Moore-Colyer, Roland. 2025. "AI Hallucinates More Frequently as It Gets More Advanced—Is There Any Way to Stop It from Happening, and Should We Even Try?" *Live Science*, June 21. https://www.livescience.com/technology/artificial-intelligence/ai-hallucinates-more-frequently-as-it-gets-more-advanced-is-there-any-way-to-stop-it-from-happening-and-should-we-even-try.

Mott, Carrie, and Daniel Cockayne. 2017. "Citation Matters: Mobilizing the Politics of Citation Toward a Practice of 'Conscientious Engagement.'" *Gender, Place & Culture* 24 (7): 954–73. http://dx.doi.org/10.1080/0966369X.2017.1339022.

Muldoon, James, Mark Graham, and Callum Cant. 2024. *Feeding the Machine: The Hidden Human Labor Powering A.I.* Canongate Books.

Murph, Nicole Lucero. 2025. "Beyond Information Literacy: Exploring AI's Impact on Labor in Academic Libraries." *Library Trends* 73 (4): 609–27. https://dx.doi.org/10.1353/lib.2025.a968498.

Nicholson, Karen P. 2016. "'Taking Back' Information Literacy: Time and the One-Shot in the Neoliberal University." In *Critical Library Pedagogy Handbook*, edited by Nicole Pagowsky and Kelly McElroy. ACRL. https://hdl.handle.net/20.500.14721/29062.

Nicolaisen, Jeppe. 2007. "Citation Analysis." *Annual Review of Information Science and Technology* 41 (1): 609–41. https://doi.org/10.1002/aris.2007.1440410120.

O'Brien, Matt. 2023. "Chatbots Sometimes Make Things Up. Is AI's Hallucination Problem Fixable?" *The Associated Press*, August 1. https://apnews.com/article/artificial-intelligence-hallucination-chatbots-chatgpt-falsehoods-ac4672c5b06e6f91050aa46ee731bcf4.

Offert, Fabian, and Ranjodh Singh Dhaliwal. 2024. "The Method of Critical AI Studies, A Propaedeutic." Preprint, arXiv, December 10. https://doi.org/10.48550/arXiv.2411.18833.

OpenAI. 2023. *GPT-4 System Card.* https://cdn.openai.com/papers/gpt-4-system-card.pdf.

O'Sullivan, James. 2025. "Gen AI and the Citation Problem." *Text, Culture, Algorithms* Substack, February 28. https://jamescosullivan.substack.com/p/gen-ai-and-the-citation-problem.

Pagowsky, Nicole. 2021. "The Contested One-Shot: Deconstructing Power Structures to Imagine New Futures." *College & Research Libraries* 82 (3): 300–9. https://doi.org/10.5860/crl.82.3.300.

Perrigo, Billy. 2023. "OpenAI Used Kenyan Workers on Less Than $2 per Hour to Make ChatGPT Less Toxic." *Time*, January 18. https://time.com/6247678/openai-chatgpt-kenya-workers/.

Pierce, Matthew. 2025. "Academic Librarians, Information Literacy, and ChatGPT: Sounding the Alarm on a New Type of Misinformation." *College & Research Libraries News* 86 (2): 68–70. https://doi.org/10.5860/crln.86.2.68.

Popowich, Sam. 2024a. "ChatGPT and the Commodification of Knowledge Work." *CAUT Journal*. https://doi.org/10.63409/2024.37.

Popowich, Sam. 2024b. *Solving Names: Worldliness and Metaphysics in Librarianship.* Litwin Books.

Postle, Brad, and Jacqueline Fulvio. 2024. "GCBI-alyzer." University of Wisconsin-Madison, Department of Psychology. Last updated September 17, 2025. https://psych.wisc.edu/gcbialyzer/.

Proctor, Jason. 2024. "B.C. Lawyer Reprimanded for Citing Fake Cases Invented by ChatGPT." *CBC News*, February 26. https://www.cbc.ca/news/canada/british-columbia/lawyer-chatgpt-fake-precedent-1.7126393.

Reisner, Alex. 2023. "These 183,000 Books Are Fueling the Biggest Fight in Publishing and Tech." *The Atlantic*, September 25. https://www.theatlantic.com/technology/archive/2023/09/books3-database-generative-ai-training-copyright-infringement/675363.

Reisner, Alex. 2025. "The Unbelievable Scale of AI's Pirated-Books Problem." *The Atlantic*, March 20. https://www.theatlantic.com/technology/archive/2025/03/libgen-meta-openai/682093/.

Scott, Inara. 2023. "Yes, We Are in a (ChatGPT) Crisis." *Inside Higher Ed*, April 18. https://www.insidehighered.com/opinion/views/2023/04/18/yes-we-are-chatgpt-crisis.

Shah, Chirag, and Emily M. Bender. 2022. "Situating Search." In *CHIIR '22: Proceedings of the 2022 Conference on Human Information Interaction and Retrieval*, edited by David Elsweiler, Udo Kruschwitz, and Bernd Ludwig. Association for Computing Machinery. https://doi.org/10.1145/3498366.3505816.

Shah, Chirag, and Emily M. Bender. 2024. "Envisioning Information Access Systems: What Makes for Good Tools and a Healthy Web?" *ACM Transactions on the Web* 18 (3): Article 33. https://doi.org/10.1145/3649468.

Shange, Savannah. 2022. "Citation as Ceremony: #SayHerName, #CiteBlackWomen, and the Practice of Reparative Enunciation." *Cultural Anthropology* 37 (2): 191–98. https://doi.org/10.14506/ca37.2.03.

Simpson, Colin. 2023. "ChatGPT - When Good Citations Go Bad." *Screenface* (blog), February 24. https://www.screenface.net/chatgpt-when-good-citations-go-bad.

Slater, Avery. 2024. "Phantoms of Citation: AI and the Death of the Author-Function." *Poetics Today* 45 (2): 223–31. https://doi.org/10.1215/03335372-11092818.

Slater, Kay. 2025. "Against AI: Critical Refusal in the Library." *Library Trends* 73 (4): 588–608. https://doi.org/10.1353/lib.2025.a968497.

Small, Henry. 2016. "Referencing as Cooperation or Competition." In *Theories of Informetrics and Scholarly Communication*, edited by Cassidy R. Sugimoto. De Gruyter. http://dx.doi.org/10.1515/9783110308464-006.

Smerdon, David (@dsmerdon). 2023. "Now we have the title of our fake paper: 'A Theory of Economic History'. We need the most probable author." Twitter (now X), January 26. https://x.com/dsmerdon/status/1618817753514938368.

Smith, Christen A., Erica L. Williams, Imani A. Wadud, and Whitney N. L. Pirtle, and The Cite Black Women Collective. 2021. "Cite Black Women: A Critical Praxis (A Statement)." *Feminist Anthropology* 2 (1): 10–17. https://doi.org/10.1002/fea2.12040.

Tay, Aaron. 2024. "Primo Research Assistant Launches - A First Look and Some Things You Should Know." *Aaron Tay's Musings About Librarianship*, September 15. https://musingsaboutlibrarianship.blogspot.com/2024/09/primo-research-assistant-launches-first.html.

Thomas, Jasmine. 2024. "Teaching Citation Justice." *WOC + Lib*, September 25. https://www.wocandlib.org/features/2024/9/24/teaching-citation-justice.

Tucci, Ryan. 2023. "When Chat GPT Falters: Implications for Reference Services in Academic Libraries." *8 Bit Librarian* (blog), March 10. https://www.8bitlibrarian.com/when-chat-gpt-falters-implications-for-reference-services-in-academic-libraries/.

Turnitin. 2025. "AI Writing Detection in the New, Enhanced Similarity Report." Turnitin Guides. August 27. https://guides.turnitin.com/hc/en-us/articles/22774058814093-AI-writing-detection-in-the-new-enhanced-Similarity-Report.

University of Chicago Press. 2023. "Citation, Documentation of Sources: FAQ 422." *The Chicago Manual of Style Online*. 17th ed. https://www.chicagomanualofstyle.org/qanda/data/faq/topics/Documentation/faq0422.html.

University of Chicago Press. 2025. "14.112: Citing AI-Generated Content." *The Chicago Manual of Style Online*. 18th ed. https://www.chicagomanualofstyle.org/book/ed18/part3/ch14/psec112.html.

Van der Sar, Ernesto. 2024. "NVIDIA: Copyrighted Books Are Just Statistical Correlations to Our AI Models." *TorrentFreak*, August 17. https://torrentfreak.com/nvidia-copyrighted-books-are-just-statistical-correlations-to-our-ai-models-240617/.

Vinsel, Lee. 2021. "You're Doing It Wrong: Notes on Criticism and Technology Hype." *Medium*, February 1. https://sts-news.medium.com/youre-doing-it-wrong-notes-on-criticism-and-technology-hype-18b08b4307e5.

Walters, William H., and Esther Isabelle Wilder. 2023. "Fabrication and Errors in the Bibliographic Citations Generated by ChatGPT." *Scientific Reports* 13: 14045. https://doi.org/10.1038/s41598-023-41032-5.

Warner, John. 2025. *More Than Words: How to Think About Writing in the Age of AI.* Basic Books.

Watson, Alex P. 2024. "Hallucinated Citation Analysis: Delving into Student-Submitted AI-Generated Sources at the University of Mississippi." *The Serials Librarian* 85 (5–6): 172–80. https://doi.org/10.1080/0361526X.2024.2433640.

Watters, Audrey. 2020. "School Work and Surveillance." *Hack Education* (blog), April 30. https://hackeducation.com/2020/04/30/surveillance.

Weber-Wulff, Debora, Alla Anohina-Naumeca, Sonja Bjelobaba, Tomáš Foltýnek, Jean Guerrero-Dib, Olumide Popoola, Petr Šigut, and Lorna Waddington. 2023. "Testing of Detection Tools for AI-Generated Text." *International Journal for Educational Integrity* 19: Article 26. https://doi.org/10.1007/s40979-023-00146-z.

Weissman, Jeremy. 2023. "ChatGPT Is a Plague Upon Education." *Inside Higher Ed*, February 8. https://www.insidehighered.com/views/2023/02/09/chatgpt-plague-upon-education-opinion.

Welborn, Aaron. 2023. "ChatGPT and Fake Citations." *Duke University Libraries Blog*, March 9. https://blogs.library.duke.edu/blog/2023/03/09/chatgpt-and-fake-citations/.

Whitten, Elizabeth. 2025. "N.L. Asks Deloitte to Carry Out Review After 'Incorrect' Citations Found in $1.6M Provincial Health Plan." *CBC News*, November 24. https://www.cbc.ca/news/canada/newfoundland-labrador/nl-deloitte-citations-9.6990216.

Wolfram, Stephen. 2023. "What Is ChatGPT Doing . . . and Why Does It Work?" *Writings* (blog), February 14. https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work.

Wondracek, Jenny. 2025. "Coming Soon: The Interactive GenAI Legal Hallucination Tracker— Sneak Peek Today!" *AI Law Librarians* (blog), August 10. https://www.ailawlibrarians.com/2025/08/10/coming-soon-the-interactive-genai-legal-hallucination-tracker-sneak-peek-today/.

Wouters, Paul. 2016. "Semiotics and Citations." In *Theories of Informetrics and Scholarly Communication*, edited by Cassidy R. Sugimoto. De Gruyter. https://doi.org/10.1515/9783110308464-007.

Young, Jeffrey R. 2023. "Georgia Tech Is Trying to Keep a ChatGPT-Powered Teaching Assistant From 'Hallucinating.'" *EdSurge*, June 22. https://www.edsurge.com/news/2023-06-22-georgia-tech-is-trying-to-keep-a-chatgpt-powered-teaching-assistant-from-hallucinating.

Zhou, Dale, Jennifer Stiso, Max Bertolero, Eli Cornblath, Erin Teich, Ann Sizemore Blevins, Kieran Murphy, Kendra Oudyk, Cleanthis Michael, Anne Urai, Jordan Matelsky, Virtualmario, Christopher Camp, Raul Alacantra Castillo, Rebecca Saxe, Jordan Dworkin, and Dani Bassett. 2022. "dalejn/cleanBib." Zenodo. Last modified December 13. https://doi.org/10.5281/zenodo.3672109.

How to cite this article: Blechinger, Joel. 2026. Probabilistic Obliteration and Formulaic Fabrication: Citational (In)justice in the Age of Generative Artificial Intelligence. *KULA: Knowledge Creation, Dissemination, and Preservation Studies* 9(1). https://doi.org/10.18357.kula.310

Submitted: 28 March 2025 **Accepted:** 14 November 2025 **Published:** 19 January 2026

KULA: Knowledge Creation, Dissemination, and Preservation Studies is a peer-reviewed open access journal published by University of Victoria Libraries.

OPEN ACCESS