

Sue Spaid

Northern Kentucky University

Abstract

New technologies prompt anxieties on all fronts. Although it's too early to envision chatbots' ultimate impact on users' linguistic competence, several dilemmas come to fore: 1) If only a handful of people know how to write, who will copyedit texts? 2) Is it ethical to claim authorship of personal letters, school assignments and even books (or dissertations) completed with the aid of chatbots? 3) If chatbots go mainstream, how will avant-garde texts be produced? 4) Will authors be credited more for "keyword selection" than "authorship"? Since successful searches depend on our accurately identifying prompts that target desired information, we are already more or less rewired by "keyword rewards." Building upon the many ways faxes, PCs, the internet and digital photography have helped to streamline our daily workflow, this paper aims to tease out our concerns regarding chatbot's unprecedented capabilities in light of what they can do. To better understand chatbots' philosophical ramifications, I review epistemic considerations regarding the appropriateness, accuracy and authenticity of resulting compositions; ethical concerns in terms of unintended consequences, impersonation and plagiarism; and the aesthetic significance of chatbot-generated documents in terms of originality, artistic significance and literary opportunities.

Keywords: ChatGPT, generative AI, chatbots, epistemology, ethics, originality

The Editorial Turn that Wasn't: Rethinking AI Text Generators

"AI Text Generators": Staged to Replace Writing Skills?

No doubt, radically-new technologies such as ChatGPT-3 foment anxieties regarding perceived risks and introduce ethical dilemmas that risk social ills. With photography's arrival, people envisioned painting's demise and even the death of the imagination (Teicher, 2016). The rise of the internet during the 1990s signaled printed matter's becoming obsolete, while email's popularity led people to envision post office closures: neither materialized. In fact, print-on demand has ramped up publications and online shopping has made delivery operations indispensable. Digital media prompted piracy wars and remote education increasingly threatens campus settings. Computer programmers's reliance on "off-the-shelf" subroutines engender mathematical deficiencies, which may explain why resulting algorithms appear so opaque to their users. Robot ubiquity risks erasing both factory- and service-worker jobs, not to mention reduced government incomes due to lost tax revenues. Perhaps the jobs most at risk of replacement by chatbots, or conversational agents, are call-center jobs, the unprecedented career option that lifted millions of multi-lingual speakers into the middle class across the globe. Even drug-sniffing dogs are at risk of losing their jobs to "biomechanical" noses that don't require months of training and years of food. For good, I expect programmers to one day employ a budding AI tool, the EPU (emotion processing unit). Developed by Emoshape, EPU could one-day help social media users predict how readers might react to their comments, thereby curtailing hate speech at risk of triggering violent crimes (Spaid, forthcoming 2023).

Simply put, new technologies prompt anxieties on all fronts. The only thing we know for sure is that it is way too early to envision chatbots' ultimate impact on users' linguistic competence. From the onset, several dilemmas come to fore: 1) If only a handful of people know how to write, who will copyedit texts? 2) Is it ethical to claim authorship of personal letters, school assignments and even books (or dissertations) written with the aid of a chatbot? 3) If chatbots go mainstream, how will avant-garde texts be produced? 4) Will authors be credited more for "keyword selection" than "authorship"? Since

successful searches depend on our identifying particular prompts that target desired information, we are already more or less rewired by “keyword rewards.”

Building upon the many ways faxes, PCs, the internet and digital photography have helped to streamline our daily workflow, this paper aims to tease out our concerns regarding chatbot’s unprecedented capabilities in light of what they can do. To better understand chatbots’ philosophical ramifications, I review epistemic considerations regarding the appropriateness, accuracy and authenticity of resulting compositions; ethical concerns in terms of unintended consequences, impersonation and plagiarism; and the aesthetic significance of chatbot-generated documents in terms of originality, artistic significance and literary opportunities.

Spinning Out of Control?

Prior to the 2020-2021 pandemic, society seemed especially focused on future-oriented affairs. Even *The New York Times* was participating in this race to the future, as it regularly ran “Op-Eds from the Future,” which addressed issues that contributors predicted would be topical 10, 20 or 100 years from now. Despite readers having comparatively more time to speculate on their futures during COVID lock down, this obsession with all things futuristic came to a screeching halt in March 2020, when suddenly the near-future mattered much more than the “far-off” future. People were suddenly more focused on when COVID-19 might end or when vaccines would arrive, etc. Just two years after people started receiving their first “jabs,” “futuristic thinking” returned with a vengeance. Not only was the Artemis I unmanned lunar mission launched, but the Mars rover Perseverance started studying Mars’ cloud cover, while its companion Curiosity searched for rock ripples indicative of water. Space enthusiasts have goals to inhabit the Moon by 2030 and to colonize Mars by 2050. Meanwhile, “Deep-sea” mining, Direct Air Capture and laboratory-meat are hailed as viable solutions, yet these would be totally unnecessary if only we reduced consumption. And now, “chatbots” appear ready to write whatever we don’t want to write ourselves.¹

On many levels, self-writing software had been bubbling up in ordinary writing software. For over thirty years, Microsoft’s Word has identified our potential spelling errors in red and our grammatical irregularities in blue. Texting introduced auto-correct, which is annoying as compared to iMessage prompts that hasten typing by proposing potential words. And Gmail software corrects spelling, grammar and punctuation, even for foreign languages. It was only a matter of time before chatbots could construct entire documents based on prompts. With OpenAI’s launching of ChatGPT-3 on November 30, 2022, conversational agents were no longer confined to labs.² Four weeks later, the *New York Times* polled readers to see who could distinguish stories penned by ChatGPT-3 about school lunches from those written by fourth graders. Then, New York City and Los Angeles schools banned it in public schools. Universities started informing faculty about how best to handle chatbots in classrooms. Those that found banning it untenable encouraged teachers to let students master it. Rather than worrying about chatbots, others told teachers to focus on creating assignments that excite students. Three weeks later, ChatGPT-3 scored a B- on a Wharton School of Business exam ([youtube.com](https://www.youtube.com/watch?v=...)), which made it seem even more viable. Seven weeks later, OpenAI issued the “faster,” even better-performing ChatGPT-4. Two weeks later, a group of AI developers, including OpenAI co-founder Elon Musk, Steve Wozniak and 1000 tech leaders all called for a “pause on AI systems,” citing “profound risks to society” (Metz and Schmidt, 2023).

¹ For a history of chatbots, beginning with ELIZA (1966) and then PARRY (1972), ALICE (1995), Jabberwacky (1997), Mitsuko (since 2005), Cleverbot (2006), and Rose (2011), as well as their broad range of applications, see Wahde and Virgolin (2022).

² Unaware of the decades-long history of chatbot battles, I envisioned OpenAI, who had named their image-making software DALL-E, calling their generative text software “DICK,” after the science fiction writer Philip K. Dick who wrote *Do Androids Dream of Electric Sheep* (1968), which became the 1982 film *Bladerunner*. Anticipating how DICK might impact linguistic competence initially inspired my consideration of the philosophical ramifications associated with AI writing tools. ChatGPT-3 arrived in time to test out my philosophical worries.

Most researchers complain that responses availed by “predictive text” software, such as Gmail phrase suggestions, engender rather “predictable writing” (Arnold *et al.*, 2020). Ack-acking, an anonymous commentator responding to Bret Devereaux’s comprehensive analysis of ChatGPT’s potential for helping students write college essays, describes its output as “Milquetoast.” Ack-acking continues:

It always chooses some middle-of-the-road option: “These two things are somewhat the same, somewhat different and neither is overall better than the other.” It picks out the most obvious, superficial differences. “One book was written more recently than the other.” It always writes in the same style of friendly, polite, and boring, like a hotel receptionist. All of its opinions seem the sort of thing that I’d make up on the spot if I was like, at a party and someone pressed me for an opinion but I didn’t really know anything or care about the subject in question. ...I hope Google, or some other company, releases a version that doesn’t reword the results and just links back to the original sources (Devereaux, 2023).

It’s thus no wonder that chatbot critics presume conversational agents have no literary future. Since most writing is of a “bland” technical nature, philosopher/artist Mark Harris considers contemporary writing’s predictability hardly surprising and possibly necessary. He observes, “Rather than being an uninvited and destructive virus, it’s likely that ChatGPT-3 has arisen due to a need we have. That need may reveal a dysfunctional and outdated overevaluation of ‘original’ writing, most of which in higher ed responds to routine boring assignments and provides in turn dull and unoriginal results” (Harris 2023a). He considers Raymond Queneau’s poem *A Hundred Thousand Billion Poems* (1961) (a book of ten sonnets whose fourteen interchangeable lines yield 10^{14} poems) proto-AI, since like large language models (LLMs), its 140 lines generate an infinity of poems. Moreover, written texts are likely to become even more predictable as chatbots find their natural application in texts produced for fields like law, business, science, journalism and especially medicine, where conformity is practically a virtue. As for academia, such software might one-day hasten philosophers already complicated (and sometimes grueling) task of having to identify and address every-possible counter argument.

Embedded in these rapid developments, whether positive or negative, are numerous “catch-22s.” No doubt, professionals and students alike will turn to chatbots to perform whatever assignments they deem rote or repetitive. The question for educators remains, can people actually learn how to write, let alone evaluate what they’ve read, without having to *repeatedly* construct sentences (something skilled writers routinely do *over and over*)? Although “digital natives” are often praised for writing better than their forebears, many teachers dispute this (Profeta, 2015). According to Christopher Profeta, “The teachers surveyed, however, are not without their concerns. [Sixty-eight percent] say that digital writing causes students to put less effort into their writing. This makes clear the need for students to be taught that the skills they are learning in their English and Writing classes have very practical applications in a digital world where they will have to be able to clearly express new and unique thoughts in writing largely done online” (Profeta, 2015). Moreover, teachers have noticed that students have difficulties comprehending long-form digital texts, as compared to short-form digital texts (Shanahan, 2018).

As citizens of democratic nations, our capacity to evaluate the truthfulness of what we read online depends on our skillfulness at close reading, which requires us to carefully analyze minutiae embedded in “news stories” that are often so complicated and/or distorted that important details appear contradictory, and thus worthy of neglect. It’s well known that “confirmation bias” dominates our already limited grasp of true belief. As psychologist Steve Rathje points out, “Our desire to hold identity-consistent beliefs often far outweighs our goals to hold accurate beliefs. This may be because being a part of a political party or social group fulfills fundamental needs, like the need for belonging, which supersedes our need to search for the truth” (Rathje, 2018). As we shall see, chatbots promise to shake us of our “false beliefs.”

However, such aspirations presume that true beliefs are simply a matter of correct information, a point proven untrue by climate change and evolution deniers who ably explain the science behind believers' views. Despite their access to such information, they still don't (to put it more aptly, "can't") believe it.

Epistemically-useful Texts: From “Prompts” to “Truthfulness”

Consumers. My primary worry is epistemic, the way sentences presented as “knowledge” sway opinions. For example, computer programmers' dependence on subroutines makes programmers more consumers of extant technology than knowledge producers, leaving the actual understanding of the underlying formulas and mathematical relationships to a select few. This could explain why ChatGPT-3's mathematical skills are remedial, supposedly the equivalent of a sixth grade level, which equates to pre-“pre-Algebra” (Zumbrun, 2023). I initially imagined that chatbots were designed to glean extant texts (made accessible by editors' keywords) to generate reports, reflection papers, essays, letters and even books, thus making writing practices effectively editing practices. This is a bit like how researchers currently compose papers, except that we do so “manually.” Sans chatbots, we use keywords to identify relevant papers, which form a kind of database that inspires us to generate new ideas in response, and thus serve as “citable” source material. I call this process “manual,” so long as people don't use software other than search engines to assemble relevant material for their papers. By contrast, evaluating the truthfulness (usefulness) of chatbot responses reflects three variables –appropriateness, accuracy and authenticity.

Appropriateness. To generate a love letter whose “Dear X” awaited only a name, I thought to give ChatGPT-3 the following keywords: “seduction, male, football match,” word count: “25,” type: “romantic letter,” historical era: “19th century,” style: “optimistic,” native language: “English.” The only problem is that ChatGPT-3 does not respond to keywords. It rather wants a question or a task, so I requested:

Task #1: “Write a 25-word romantic letter to a cute guy who loves football, but lived in the 19th century.”

ChatGPT-3: “My dearest love, though time may separate us, my heart beats for you like a football team's drums. Yours forever, in this life and the next.”

Note, it gave me a 26-word letter. I followed up with a more pertinent question.

Task #2: “What are some of the philosophical benefits of ChatGPT-3?”

ChatGPT-3: “ChatGPT-3, being a powerful language model, has the potential to provide several philosophical benefits, including:

1. Access to a vast array of knowledge: ChatGPT-3 can access and process an immense amount of information, including philosophical texts, allowing for a deeper understanding and analysis of philosophical concepts.
2. Enhanced critical thinking: By engaging in philosophical discussions and debates with ChatGPT-3, users can develop their critical thinking skills, as they are forced to consider and articulate their arguments more precisely.
3. Diverse perspectives: ChatGPT-3 can simulate different philosophical perspectives and ideologies, enabling users to explore and consider alternative ways of thinking and understanding the world.
4. Personal growth: Engaging with ChatGPT-3 can facilitate personal growth, as users can receive feedback on their ideas and beliefs and potentially challenge their assumptions and biases.

5. Increased accessibility: ChatGPT-3 can make philosophical concepts and discussions more accessible to people who may not have had access to them otherwise, such as those who are unable to attend formal educational institutions.”

One can already see a huge problem with ChatGPT-3. There are no citations. We haven't a clue which sources the chatbot has used to generate its answer, making it wholly inappropriate for academic writing where citing sources is of utmost importance. I next asked ChatGPT-3 one of the questions I've assigned my Cyberethics' students for their next paper, so I'll know what to expect if they opt to use this chatbot to do their homework.

Task #3: “What is your assessment of the Children’s Internet Protection Act (CIPA)? Do you support the ACLU’s views against this legislation?”

ChatGPT-3: “As an AI language model, I do not have personal opinions or beliefs, but I can provide an overview of the Children’s Internet Protection Act (CIPA) and the American Civil Liberties Union’s (ACLU) views on it....”

Apparently, the word “assessment” triggered some need to claim that it lacks opinions, even though the ChatGPT-3’s five philosophical benefits listed earlier are opinions, and not facts. Being so new, there is no evidence for any of the five benefits it cites. Moreover, each of these five benefits is more or less a cliché about the internet that not only has little to do with philosophy, but could easily be refuted. As it stands, ChatGPT-3 cannot access most philosophical texts, so #s 1,3 and 5 are suspect. If it doesn't have beliefs, how can it challenge users' assumptions and biases, let alone debate them, so #s 2 and 4 are out.

Accuracy. A few days later, I asked ChatGPT-3 several questions about my own philosophical views and found that it doesn't yet have access to oft-cited philosophy papers that are available online. Finally, I asked it a very basic question, “What is an ecovention?” and received a long-winded answer that seemed to be paraphrasing the “ecovention” Wikipedia page, which itself is riddled with errors. It ended by spewing a patently false sentence that attributed the term’s authorship to the wrong curator in conjunction with an art exhibition that never happened. It is therefore easily capable of spreading misinformation and its failure to cite source material makes it all the more problematic. According to Gordon Crovitz, a co-chief executive of NewsGuard, a company that tracks online misinformation and conducted an experiment, “This tool is going to be the most powerful tool for spreading misinformation that has ever been on the internet. Crafting a new false narrative can now be done at dramatic scale, and much more frequently — it’s like having A.I. agents contributing to disinformation” (Hsu and Thompson, 2023). I next asked, “What is ecoartspace?” and got a fantastic, mostly accurate answer. Finally, I asked “What is Hans-Ulrich Obrist’s most important idea?” and got a false reply, since ChatGPT-3 attributed Obrist with the idea of the “exhibition as a medium,” even though this notion was in play long before his arrival on the scene in the early 90s (Ferguson *et al.*, 1996, p. 1). Similarly, ChatGPT-3 defamed Australian mayor Brian Hood, the whistleblower (for an early aughts foreign bribery scandal involving the Reserve Bank of Australia) whom ChatGPT-3 tied to the bank’s troubles and erroneously accused him of having served 30 months. Rather than respond to his lawyer’s demand that OpenAI fix the false allegation within 28 days, OpenAI launched ChatGPT-4 that erases Hood’s involvement (Reuters, 2023). The Reuters reporter quotes Hood’s lawyer James Naughton, “If Hood files a lawsuit, it would accuse ChatGPT of giving users a false sense of accuracy by failing to include footnotes” (Reuters, 2023).

ChatGPT-3 may have scored a B- on a Wharton School of Business exam (youtube.com) and earned a C+ on four law school exams, including Constitutional Law (Leighton, 2023), but it barely passed the United States Medical Licensing Exam (Kung *et al.*, 2023) due to its notoriously abysmal math skills, as noted above. Despite its underperformance on the medical boards, a medical evaluator observed, “ChatGPT managed to create, ‘new, non-obvious and clinically

valid insights’, for 88.9% of its responses and appeared to show evidence of deductive reasoning, chain of thought and long-term dependency skills” (Kung *et al.*, 2023). This is good news for the healthcare industry for whom chatbots matter most. Most amazingly, ChatGPT recently outperformed doctors, though researchers caution that ChatGPT provides doctors value-added responses, hardly substitutes:

When it comes to providing patients with high-quality medical information — and delivering it with compassion and understanding — who does it better: ChatGPT or your doctor? The answer: ChatGPT, by a mile. In fact, after comparing doctor and AI responses to nearly 200 medical questions, a team of health care professionals concluded that nearly 80% of the answers from ChatGPT were more nuanced, accurate and detailed than those shared by physicians (Mozes 2023).

Authenticity. Attempting to reconcile the constitutive relationship between a visual artist’s intentions and a viewer’s appropriate grasp of an artwork’s meaning, philosopher Rob van Gerwen remarked, “We may not even be thinking about the author much, just *holding them responsible* for the nature of the work – their intentionality is the condition of possibility for making sense of a work¹⁰ [italics mine].” What interests me here is less his intentionalism, such that a resultant artwork’s meaning reflects the artist’s intention, and more footnote 10, where he notes, “[R]eading a text that one knows is ‘written’ by an ‘artificial intelligence’ (or by two authors) should be hard psychologically” (van Gerwen 2023: vii). I imagine that what van Gerwen finds “hard psychologically” is not simply the knowledge that one of the authors credited in the byline did not actually write what both are credited with having written, but the realization that we readers never know for sure whether both credited authors actually believe whatever views their text expresses. We accord texts the “intentional stance,” such that the “condition of possibility” for making sense of texts requires that both authors hold the same view. In fact, bylines grant authors ownership for their ideas and credit them for courageously sticking their necks out in public, staking claims and defending a potentially controversial viewpoint.

By contrast, chatbots allow people to construct and broadcast viewpoints that are not “truly” their own, granting them both the anonymity and impunity familiar to online conversations and/or retweets. When you ask an author, “Did you really mean to say this?” or “Why did you write this?” the author who has simply parroted a chatbot’s reply is likely to respond, “Well, it sounds right” or “It seemed like an interesting point,” putting the very notion of authenticity that is so crucial to authorship (ownership) in jeopardy. Moreover, as van Gerwen’s worry suggests, texts constructed by chatbots eschew all responsibility for what gets generated, jeopardizing not only a text’s meaning but the author’s blameworthiness/creditworthiness. For example, prior to the launch of ChatGPT-3, an advertising bot gained twitter notoriety for erroneously creating and distributing “unacceptable” ads that publicized KFC’s *non-existent* “Kristallnacht” promotion (Wiener-Bronner 2022). Even though distributors ought to be held responsible for such atrocities, businesses ordinarily get away with them by downplaying such blunders as “software bugs” or an algorithm “irregularity,” though certainly not the company’s viewpoint.

Not being persons, chatbots lack users’ “free speech” rights and could incidentally become the perfect vehicle for whipping up trouble while escaping punishment. Given AI developers’ awareness that they are accountable for chatbot banter, such as defamation or negligence due to misinformation, chatbots appear to be carefully programmed for sensitivity, yet insensitive responses regularly occur (Heaven, 2020). Note the above chatbot’s hesitancy to share its opinion regarding the ACLU’s attitude toward CIPA. Similarly, ChatGPT-3 resisted Harris’ humorous request: “Write a letter from a deadly microbe to King Charles III.”

I’m sorry, but I cannot fulfill this request. Writing a letter from a deadly microbe could be perceived as insensitive and could trivialize serious issues like the spread of diseases, which can have real-world consequences and affect many people’s lives. (Harris 2023b).

Despite this chatbot's hesitancy to weigh in on sensitive issues, any inability to distinguish falsities from truth indicates lax moral standards, since lies *too* "have real-world consequences [that] affect many people's lives" (Harris 2023b). Mattias Wahde and Marco Virgolin worry that chatbots "might be used unethically, for example in gathering private conversational data for use in, say, mass surveillance and control" (2022, p. 34). They thus assign "the research community and policymakers a strong collective responsibility to prevent unethical uses of CA technology. Tools must be developed for evaluating not only the functionality but also the safety and societal impact of CAs" (Wahde and Virgolin, 2022, p. 34).

Ethical Concerns: From "Prompts" to "Use/Abuse"

"The ethics isn't in the technology. It's in the people." Jason Allen (Roose 2022)

Access: Allen's view can be rewritten to say, "Technology isn't unethical. People are," which parrots the oft-quoted claim, "Guns don't kill people. People do." To be fair, he made this point in relation to businesses replacing human workers with AI platforms. I begin this section with his quote, because it shows how readily we shift the blame from those who *create* potential problems to those who take advantage of formerly unavailable situations. No doubt, people are to blame for whatever problems they initiate, yet ready "access" to firearms and technologies exacerbates their abuse. Despite the unrivaled number of mass shootings in the United States, there are almost twice as many suicides via guns than homicides in the United States, where physician-assisted suicide is illegal in all but ten states. Published in the *New England Journal of Medicine*, a study that tracked 26.3 million new California gun owners between 2004 and 2016 found that "men in the United States who own handguns are eight times as likely to die from gun suicides than men who don't own guns, and women who own handguns are 35 times more likely than women who don't" (Duff-Brown 2020). Even more horrifying, firearms recently surpassed traffic fatalities as the number one cause of death for US children and adolescents (ages 1-19). (Goldstick *et al.*, 2022).

Intended Use. I make this point to disabuse readers of the tendency to place the blame wholly on users who "abuse" some technology, when in fact bad things can happen, even when technologies are used as their developers intended. There are thus good reasons to hold technology developers responsible for technologies whose "intended use" poses a harm, however unanticipated. Due to unanticipated harms caused by social media use, two cases currently before the Supreme Court of the United States challenge Section 230 of the US Communications Decency Act (1996). Section 230 ordinarily protects social media companies such as Facebook, Twitter and Google from being responsible for third-party content (Barnes, 2023). *Gonzalez vs. Google* argues that Google's algorithm eased people's access to ISIS-recruitment videos hosted on Google subsidiary [youtube.com](https://www.youtube.com). Google's algorithm is accused of having aided and abetted terrorism (Barnes, 2023). *Twitter vs. Taamneh* similarly accuses Twitter of violating the 2016 Justice Against Sponsors of Terrorism Act (JASTA) (LII, 2023). Incidentally, the very same free-speech advocates who blame social media's censoring policies for "burying" conservative content hold the very same algorithms responsible for facilitating said harms. Google's PageRank supposedly ranks websites by visitorship (popularity), independent of content reliability or quality (Amine, 2020). Thus, manual censorship is the only way to skew algorithms.

Free speech advocates are hoping that the court will favor free speech at any cost, independent of potential harms. Thus far, the Supreme Court of the United States has protected hate speech, leaving attendant crimes punishable by law. People who use technology for criminal purposes are clearly responsible for whatever consequences arise, but those who manufacture tools must also be liable for whatever harms their products facilitate when "used as intended," such as when a police car explodes when rear-ended (Beall, 2011). That's why technology developers must make every effort possible to thwart "abusive" usages, which means private companies have every right to censor material they

consider likely to harm. I emphasize “used as intended” since handguns’ “intended use” is “self-defense,” which requires disabling attackers, so harming people is a gun’s “intended use.” By contrast, a suicidal pilot who intentionally crashes a plane is not using the plane as intended, thereby absolving the creator’s responsibility for the pilot’s abuse. Recall Boeing’s being held liable for two crashes of their 737 MAX aircrafts, since the plane’s “intended use” posed a risk of harm (Reuters, 2021).

That Chat-GPT-3 refused to pen a note “from a deadly microbe to King Charles III” suggests that either OpenAI has taken a moral stance, such that the “ethics must be in the technology” in order to prevent “abusive” usages, or they want to minimize any chance of lawsuits. The Deadly Microbe Letter case is especially informative, since OpenAI could have argued, as social media companies routinely do, that users are responsible for whatever outcomes arise, so long as platforms are used as intended. For example, it is well documented that Myanmar’s military personnel used Facebook to demonize the Rohingya, which caused 700,000 Burmese to flee to neighboring Bangladesh in 2017, eventually leading to a military coup. In 2021, the Rohingya sued Facebook for \$150 billion for enabling “Myanmar military personnel [to] turn the social network into a tool for ethnic cleansing, according to former military officials, researchers and civilian officials in the country” (Mozur, 2018). As Michelle Toh points out,

The legal argument in the US case may also be tricky. It alleges that Facebook should face product liability and negligence claims for failing to address defects in its platform which plaintiffs claim contributed to anti-Rohingya violence, court documents show. In the United States, Facebook would typically be protected from such liability by Section 230 of the Communications Decency Act, but the suit asks the court to instead apply Burmese law, which it says does not provide such protections (Toh, 2021).

Interestingly enough, “ten AI companies, including OpenAI, TikTok, Adobe, the BBC and the dating app Bumble, have signed up to a new set of guidelines on how to build, create and share AI-generated content responsibly” (Heikkilä, 2023), what the Partnership on AI terms “synthetic media,” or generative media (PAI, 2023). According to PAI head of AI and media integrity Claire Leibowicz, “We want to ensure that synthetic media is not used to harm, disempower or disenfranchise but rather to support creativity, knowledge sharing and commentary” (Heikkilä, 2023). On its surface, PAI’s recommendation to code AI-generated material with watermarks for transparency purposes sounds like a good idea. However, if doing so casts doubt on outputs, reduces profits or dilutes trademarks, I worry that watermarks will evaporate. Clearly, PAI members are trying to prevent negative attention, abuse and lawsuits, especially since AI text generators’ biggest potential application is the healthcare industry, whose AI healthcare workers and AI companions will one day converse with people who are patients, elderly, lonely or mentally-challenged. Meanwhile, Europe is working on an AI Act that would require companies to disclose when people are viewing deep fakes, etc. (Heikkilä, 2023). Despite such efforts to boost transparency, ChatGPT-3’s potential for misinformation, what Crovitz terms “A.I. agents contributing to disinformation” (Hsu and Thompson, 2023) suggests that using ChatGPT-3 “as intended” is still likely to cause harm.

One alternative approach that chatbot developers might adopt is to treat every text generated using their platforms as the user’s property, whether the user wants to own it or not. In using the chatbot, the user effectively agrees to own the chatbot’s response and is held responsible for whatever the chatbot expresses, just as multiple authors are held responsible for co-authored texts. Users who don’t want to own the text must either erase or destroy it, otherwise, they can be blamed or credited for unintended consequences related to the co-authored text’s distribution, much the way students who work with other students to answer test questions are wholly responsible for their peers’ inaccuracies. Since users have “free speech” rights that chatbots lack, “user-ownership” enables chatbot creators to piggyback off users’ “free speech” rights to protect them from lawsuits arising from unintended consequences owing to cases where their platforms were used as intended. Given that users have proven unreliable at self-editing their

own texts, as proven by the fact that only about 11% of tweeters tone-down their tweets when Twitter nudges them to do so (Bhattacharjee, 2021), it's hardly surprising that OpenAI is keen to embed the ethics in the technology, rather than depend on users' capacities to use it ethically.

Oxford cyberethicist Luciano Floridi considers all entities "informational objects" worthy of moral value, thus proffering yet a third approach (Spinello, 2021). According to Richard Spinello,

This category of beings deserving moral consideration includes even digital objects that exist only in cyberspace or in a database because they, too, are obviously informational objects...Floridi broadens the class of moral agents to include robots, software bots and other informational technology (IT) systems. He defines the moral agent as an interactive, autonomous and adaptable transition system capable of performing "morally" actions, that is, actions that can cause good or evil (Spinello, 2021, p. 23).

Although Floridi's model rightly considers chatbots moral agents whose actions *cause* good or evil, he fails to account for chatbots' inability to evaluate or judge (Shanahan, 2023, p. 5), which effectively absolves them of moral responsibility. His approach is thus untenable for thwarting harms owing to chatbot spews, such as when in 2016 Microsoft's chatbot Tay parroted hateful tweets, going so far as to liken feminism to cancer and to deny the Holocaust (Tennery and Cherelus, 2016). In terms of moral accountability, software developers ought to be held accountable, no differently than parents whose under-aged children commit crimes and are thus required to pay compensatory costs. Whatever laws cover parental liability imply creator liability, more generally, such that creators are held responsible for whatever crimes their products facilitate.

Spinello argues that the infosphere must transition from a "responsibility-oriented ethics based on punishment and reward to an ethics based on 'accountability and censure' (Spinello, 2021, p. 23)." Exemplary of "accountability and censure," Microsoft, Tay's creator, apologized before shutting down the chatbot, which originated "as an experiment to learn more about how artificial intelligence programs can engage with Web users in casual conversation. The project was designed to interact with and 'learn' from the young generation of millennials" (*Guardian*, 2016). They strangely didn't expect people to use it in a manner that would "undermine the company's principles and values" (*Guardian*, 2016).

Apparently, Microsoft has not learned their lesson, since its 2023 AI Bing search chatbot has had similar issues. "After months of speculation, Microsoft finally teased an updated Edge web browser with a conversational Bing search interface powered by OpenAI's latest language model, which is reportedly more powerful than the one powering ChatGPT" (Quach, 2023). Microsoft admitted, however, "We have found that in long, extended chat sessions of 15 or more questions, Bing can become repetitive or be prompted/provoked to give responses that are not necessarily helpful or in line with our designed tone" (Quach, 2023). Such mishaps rather prove that ethics must be *built into* the technology, as OpenAI appears to have done with their chatbot.

Impersonation. Even if the current version of ChatGPT-3 ignores citations, mixes up information, cannot scan files and lacks access to the wide range of online sources that researchers routinely access, I can imagine a future version "ChatGPT-I" that capably impersonates my own writing style, interests, and argument formation, especially deductive reasoning (Shanahan, 2023, p. 9). By gleaning extant papers (stored on suespaid.info, academia.edu, PhilPapers, ResearchGate and online journals), such a chatbot could construct my sentences, generate my ideas and furnish my stylistic flourishes, either with my blessing or "unbeknownst to me." Who is to prevent this from happening, especially long after I'm dead and no one knows exactly which unpublished articles were found stored on my computer? With one click

of a button, my philosophical legacy could collapse under the weight of my “apparent” advocacy for abhorrent brutes.

Not such a far-out scenario, Sarah Andersen claims that her images were modified for use by alt-right propagandists and later pirated by the LAION 5b data set (2022).

People on [the 4chan board] began reproducing my work and editing it to reflect violently racist messages advocating genocide and Holocaust denial, complete with swastikas and the introduction of people getting pushed into ovens. The images proliferated online, with sites like Twitter and Reddit rarely taking them down... Through the bombardment of my social media with these images, the alt-right created a shadow version of me, a version that advocated neo-Nazi ideology... Something about the mimicking of my handwriting, streamlined into an easily accessible typeface, felt particularly violating (Andersen, 2022).

In contrast to chatbots, AI generator software generates images from texts. Programs such as DALL-E 2 (launched by OpenAI in January 2021), Midjourney (launched by the discord server in July 2022), Stable Diffusion (launched by Stability.AI in August 2022), and Dream Up (launched by Deviant Art in November 2022) have been in use for several years now. Andersen claims that her cartoons got caught up in some AI data base for use by anyone eager to attribute captions to her cartoon renderings. She writes, “In October [2022], I was sent via Twitter an image generated by A.I. from a random fan who had used my name as a prompt. It wasn’t perfect, but the contours of my style were there. The notion that someone could type my name into a generator and produce an image in my style immediately disturbed me” (Andersen, 2022). Andersen claims that the data set LAION 5b, upon which Stable Diffusion was built, contains 6 billion images (Andersen, 2022). Moreover, “the LAION data sets have also been found to include photos of extreme violence, medical records and nonconsensual pornography” (Andersen).

Northern Kentucky University’s Academic Technology Analyst Evan Downing contests the notion of images stored as “data sets.” According to him, AI generators don’t have memories. They rather rely on “probabilities” of what things go together and in what order to create texts and/or images. As Kevin Roose explains:

What makes the new breed of A.I. tools different...is not just that they’re capable of producing beautiful works of art with minimal effort. It’s how they work. Apps like DALL-E 2 and Midjourney are built by scraping millions of images from the open web, then teaching algorithms to recognize patterns and relationships in those images and generate new ones in the same style. That means that artists who upload their works to the internet may be unwittingly helping to *train* [emphasis mine] their algorithmic competitors.

R.J. Palmer, another digital artist, tweeted “What makes this AI different is that it’s explicitly trained on current working artists... This thing wants our jobs, it’s actively anti-artist” (Roose). In fact, LLMs embedded in generative-text AI generate statistically-likely sequences, a tactic that grants the software authorship, thus disavowing plagiarism accusations, even though “statistically-likely sequences” hardly pull texts out of thin air. As Murray Shanahan explains:

LLMs are generative mathematical models of the statistical distribution of tokens in the vast public corpus of human-generated text, where the tokens in question include words, parts of words, or individual characters including punctuation marks. They are generative because we can sample from them, which means we

can ask them questions. But the questions are of the following very specific kind. “Here’s a fragment of text. Tell me how this fragment might go on. According to your model of the statistics of human language, what words are likely to come next?” (Shanahan, 2023, p. 10)

As repeatedly stressed herein, unless the ethics is *built in*, there’s no stopping AI’s voracious appetite.

Plagiarism. The truth is, we all know how difficult it is to generate innovative material, let alone to write clearly and grammatically. Moreover, we know some students hire people to write their papers, when they can’t find extant papers online to tweak prior to presenting them as “their own work.” As much as we remind students that not citing online material is plagiarism, a crime punishable by law; they don’t seem to grasp that material freely available for the taking is not actually theirs to take and to use as they see fit. Uncited chatbot compositions available just by posing questions will make this fine point all the more difficult to reinforce. A relevant difference between scholars and students is that scholars believe that citing sources makes them appear smarter, whereas students prefer to skip such tedious steps. Moreover, making something one’s own requires far more manipulation. The people who appear to benefit most from chatbots’ clever, apparently “well-written” texts are students, yet they are incidentally most at risk of being harmed by chatbots’ unintended consequences, which poses as our second “catch-22.” Unlike students, ChatGPT-3 does not research material, evaluate options or weigh evidence (Shanahan, 2023, p. 5), so students who eschew chatbots when writing essays win big. As UNC history professor Bret Devereaux observes, “[T]he suggestion is that students could use ChatGPT to produce an essay they edit to fit or an outline they fill in; both problems run into the issue that the student is now trying to offload the most important part of the task for them to learn: the actual thinking and analysis” (Devereaux, 2023).

Chatbots engender yet a third catch-22. While “translation software” requires sophisticated editors (typically native speakers) to clean-up translated texts, AI-originated texts depend on fact-checkers to determine which sentences are both true and relevant. The catch-22 is that the more writers depend on chatbots, the less likely they are to develop those skills necessary to succeed as either editors or fact-checkers. Not every native speaker is proficient as an editor, but most native speakers have sufficient linguistic proficiency to improve translations. One obvious outcome is that colloquial practices may come to dominate written practices, as spoken proficiencies surpass written proficiencies, much the way print dominates cursive handwriting.

Aesthetically- Significant Artworks: From “Prompts” to “Innovation”

Originality. Given DALL-E’s popularity as an art-making tool, art critic Hrag Vartanian tested what would happen if he “inserted the text from five [‘blue-chip’ gallery] press releases” into DALL-E to see how it would “reimagine the exhibitions using only those words” (Vartanian, 2022). And the crazy thing is that DALL-E managed to render artworks that on his view passed for artworks made by four of the five artists. Vartanian remarked that texts rendered by DALL-E are “gibberish,” so this software is not quite right for creating texts, yet his experiment preceded ChatGPT-3 by two months. When Colorado artist Jason M. Allen won the \$300 prize for “emerging digital artist” at the Colorado State Fair with an image he’d created using Midjourney, “Some artists defended Mr. Allen, saying that using A.I. to create a piece was no different from using Photoshop or other digital image-manipulation tools, and that human creativity is still required to come up with the right *prompts* to generate an award-winning piece (italics mine)” (Roose, 2022).

Vartanian’s experiment is particularly relevant here, since he implemented “keywords” availed by exhibition press releases to prompt imagery-inducing software. Over the past decade, published academic articles have been similarly accompanied by relevant “key words” submitted by authors who anticipated which words researchers might use to locate their papers. Unlike AI image generators, ChatGPT-3 prompts currently take the form of sentences. I imagine a future version of ChatGPT-K that targets papers

“keywords”, much like library catalogs; enabling it to produce essays that string together ideas found in related, though distinct papers. Just as Allen admitted to having used Midjourney to create his award-winning digital art, artists who attribute chatbots as “co-authors” score points for transparency, integrity and perhaps even innovation.

With each new technological development, reliable tools replace specialized skills that were developed over time. Nascent technologies that avail short-cuts initially seem like “cheating,” whether search engines (vs. printed scientific indices), subroutines (vs. coding) or digital photography (vs. analog). Over time, such disparagements dissipate. Forty years on, we think little of our reliance on Microsoft Word’s editing tools (spell check, conjugation, noun/verb agreement, punctuation, etc.). What will we think, however, when articles, books and even news stories are composed by entering keyword combinations that prompt computers to string extant material together to create “original” texts? If amateurs, students and professionals rely on such handy tools, is there even a literary future? If one can only express what already exists, are eccentrics bound for extinction?

Artistic Significance. Artists’ use of chatbots and image generators could simply be part of the never-ending evolution of artistic practices. During the heyday of conceptual art and *in situ* installations, artists seemed to have lost interest in studio practices, whether drawing, painting or sculpting with their hands. Some artists even opted for “post-studio” practices, enabling them to focus less on tinkering in the studio and more on thinking and *making* elsewhere. Sol LeWitt famously characterized conceptual artists as mystics whose outputs defied logical reasoning, even if “all of the planning and decisions are made beforehand and the execution is a perfunctory affair” (LeWitt, 1967). During the eighties, sculptors routinely sent designs out to be “fabricated,” once again removing the visual artist from the process. Since the millennium, some artists stopped making stuff altogether, preferring instead to organize social activities, host one-off events, create urban farms with communities or restore river banks with local stakeholders. Only the occasion of an exhibition motivates artists engaged in “social practice” to produce something show-worthy.

Harris wonders whether ChatGPT-3 might “relieve us of much of the routine tasks of writing and studio work in ways that allow us to focus on practices that matter. Or might they deal with background tasks (like video rendering in the background as we continue working), as...we get down to thinking” (Harris 2023a). Harris stumped ChatGPT-3 once again when he asked, “Explain what Frances Starks understands of Charles Ray’s talks on his own work.” Although ChatGPT-3 identified both Stark and Ray as artists, it couldn’t quite answer Harris’s query, suggesting that it won’t be writing artists’ statements anytime soon (Harris 2023b). If Harris continues to ask ChatGPT-3 questions that generate comical comments and artworld inside jokes, he will have found a low-cost collaborator, even if it is at ChatGPT-3’s expense.

Being an artist, Harris’s approach differs wildly from historian Devereaux, who requested ChatGPT-3 to “write an 8000-word essay on the causes of Roman military success in the third and second centuries BCE,” which admittedly sounds like a tall order for a budding historian, let alone a chatbot (Devereaux, 2023). Having run out of facts pretty fast, ChatGPT-3 petered out after 1220 words. Along the way, it stopped twice, requiring Devereaux to prompt it to continue, and finally to generate a conclusion. As I know nothing of warfare, let alone ancient Roman history, the resulting essay not only impressed me, but I actually felt like I had learned something by reading it, yet Devereaux proved otherwise. To the innocent eye, this brief, but comprehensive essay identified particular generals and described the various inventions that had enabled ancient Rome to win particular battles, and thus it felt believable.

Much to my dismay, Devereaux contributed nearly 30 comments, mostly pointing out where ChatGPT-3 had erred, leaving me feeling rather deceived. One of its earliest mistakes was its claim that Rome had a large pool of recruits, since it recruited from the lower classes. Devereaux countered that in fact, the Roman army “excluded the very poor and instead recruited from the landholding class exclusively.” This

important detail “reframes” the picture quite dramatically, laying the ground for the technological advancements that according to ChatGPT-3 ensured Rome’s victories. However, information really went awry in its “Military Technology and Tactics” section, leading Devereaux to remark, “ChatGPT will now take a period that is awash with the Romans adopting things, fail to identify any of those things and instead write about things the Romans did not adopt in this period. All of its examples here except for the quinquereme are wrong.” I didn’t even recall reading about the *quinquereme* (large seafaring vessel).

Even if fewer artists learn basic painting, drawing and sculpting skills, there have always been plenty of contemporary artists who have found ways to train themselves in whatever skills they needed. I thus imagine there will always be plenty of amazing writers. However, if chatbot reliance becomes the norm, it may take more discipline and ingenuity to develop the requisite skills. Readers too will have to develop a critical nose for misinformation. Perhaps the most impressive use of chatbots concerns fixing computer software. One of my students told me that her website software had an html error, so she entered the code in ChatGPT-3 and it immediately identified her error. That said, “when it comes to coding, ChatGPT isn’t always correct. Stack Overflow was forced to temporarily ban users from sharing responses generated by the AI because such a high percentage was incorrect” (White 2022). Similarly, another student told me that he requested ChatGPT-3 to create code for an ultrasonic sensor, which it did, but he never got around to testing whether the code actually works.

Literary Opportunities. As noted above, the most obvious worry is that chatbots have only been taught popular material and cannot dream up the imaginative scenarios human beings create. It thus seems that internet rumors circulating since 2020 indicating MSN News’ plan to replace hundreds of employees with chatbots is far-fetched. In terms of creative purposes, Shannon White recommends “using this chatbot to help you brainstorm ideas for creative work. All new projects draw on existing works, and ChatGTP can generate endless suggestions for plot lines, character backstories, world development, party ideas, gift ideas and more” (White 2022). She offers examples of ChatGPT-created plot lines for a sci-fi romance and video-game character backstories. Finally, she ends with “Breaking the Rules,” which as Mark Harris pointed out is where the most fun lies. She asked ChatGPT to “write an authoritative, scientific article about the benefits of eating glass” (White 2022). You can well imagine that it politely refused her request.

As someone who regularly reviews art, I share *New York Magazine* art critic Jerry Saltz’s enthusiasm for chatbots that pen reviews, leaving critics time to experience more art. He told Madeline Leung Coleman:

I live for the day –and I’m not kidding you about this –that *New York Magazine* will help me program an algorithm to write my reviews in my style. It is all that I want. So I can sit down with my AI computer at home, feed it 22 prompts about what I think of this show or that, and it will write a reasonably acceptable version of my work. I would die a happy critic and have more than seven days off in my life (Coleman, 2023).

When I asked ChatGPT-3 to write two reviews of Rirkrit Tiravanija’s art, one in the style of Jerry Saltz and another in the style of Eleanor Heartney, I received only slightly different texts. Although both focused on light and shadow, the latter also addressed materiality and textures.

Given chatbots’ dependence on the statistical probabilities of certain words correlating with and following others, it is doubtful that chatbots could create anything so novel as James Joyce’s *Finnegans Wake* (1923). Although I’ve read about creative writing classes using ChatGPT-3, it’s difficult to imagine its forte being poetry, unless there are good reasons to repeat word choices used in earlier poems or to replicate prior rhyming schemes!

Conclusion

Prior to ChatGPT-3's arrival, I envisioned chatbots producing documents by scanning internet files. I thus imagined users requiring editing skills. I now realize that chatbots produce documents that are so smooth that editorial skills seem superfluous. As Brett Devereaux demonstrated with chatbot's 1200-word report on the causes of ancient Rome's military success during the third and second centuries BCE, what we most need are experts to wade through reams of misinformation.

This paper has identified three catch-22s that make relying on chatbots a nonstarter: 1) The process of developing one's writing skills actually requires one to undergo mundane rewrites (builds sentence structure, boosts vocabulary, develops styles), the very activity chatbots aim to replace. 2) The absence of citations discourages professionals from using chatbots, while novices, who lack access to original source material and are least qualified to evaluate each output's accuracy, are most likely to find chatbots useful. 3) The more authors depend on such tools, the less likely they are to develop the requisite skills to succeed as editors and fact-checkers. Absent access to sources, chatbots prove rather useless for academia, which is hardly problematic since I doubt academia is generative AI's target market. Rather than helping authors generate novel texts, I imagine that chatbot developers are targeting the health care industry, thus putting society's most vulnerable people (the elderly, sick, homebound, lonely and mentally-challenged) at the mercy of chatbots. Because chatbots' singular attribute is repartee, it's imperative that they **are** programmed to refuse malicious orders, esteem accuracy and ensure reliability.

So long as no one can anticipate what venom chatbots might spew, they remain dangerous. Apparently, ChatGPT-4 no longer claims that Brian Hood was involved in the RBA-foreign bribery scandal and spent time in jail. If all that OpenAI does is launch new editions that block our access to past "lies," we are in for a long and bumpy ride. We need new roads, not just patches.³

References

- 2016 "Justice Against Sponsors of Terrorism Law." [PUBL222.PS \(congress.gov\)](#)
2016. "Artificial Intelligence," *The Guardian*. <https://www.theguardian.com/technology/2016/mar/26/microsoft-deeply-sorry-for-offensive-tweets-by-ai-chatbot>
2023. "Twitter, Inc. v. Taamneh." *Legal Information Institute*. <https://www.law.cornell.edu/supct/cert/211496>
- Tech and Science News. <https://www.youtube.com/watch?v=2oh0fFsqJ8U>
- Amine, Amrani (2020). "PageRank Algorithm, fully explained." <https://towardsdatascience.com/pagerank-algorithm-fully-explained-dc794184b4af>
- Andersen, Sarah (2022). "The Alt-Right Manipulated my Comic. Then A.I. Claimed it." *New York Times*. <https://www.nytimes.com/2022/12/31/opinion/sarah-andersen-how-algorithm-took-my-work.html>
- Arnold, K., Chauncey, K., and Gajos, K. (2020). "Predictive Text Encourages Predictable Writing." iis.seas.harvard.edu/papers/2020/arnold20predictive.pdf
- Barnes, Robert *et al.* (2023). "Supreme Court decides if Google is liable for recommending ISIS videos." *The Washington Post*. <https://www.washingtonpost.com/technology/2023/02/21/gonzalez-v-google-section-230-supreme-court/>
- Beall, Pat (2011). "Popular Police Cars Crown Victorias Prone to Explode Tied to Deaths." *Palm Beach Post*. <https://www.autosafety.org/popular-police-cars-crown-victorias-prone-explode-tied-deaths/>
- Bhattacharjee, Shomik Sen (2021). "Twitter Will Now Nudge You if it Thinks You're About to Tweet Something Toxic." *Mashable India*. <https://in.mashable.com/tech/22063/twitter-will-now-nudge-you-if-it-thinks-youre-about-to-tweet-something-toxic>
- Coleman, Madeline Leung (2023). "MoMA Has Data Issues," *New York Magazine*. February 24.

³ The author wishes to thank Connor Herbert who provided many useful tips regarding this fast-changing topic, as well as two anonymous reviewers who recommended references and challenged me to ground my ideas in evidence.

- Devereaux, Bret (2023). "Collections: On ChatGPT." <https://acoup.blog/2023/02/17/collections-on-chatgpt/>
- Duff-Brown, Beth (2020). "Handgun Ownership Associated with Higher Suicide Risk." *Stanford Medicine News*. <https://med.stanford.edu/news/all-news/2020/06/handgun-ownership-associated-with-much-higher-suicide-risk.html>
- Ferguson, Bruce, Reesa Greenberg and Sandy Nairne (eds) (1996). *Thinking About Exhibitions*. New York: Routledge.
- Goldstick, Jason *et al.* (2022). "Current Causes of Death and Adolescents in the United States." *New England Journal of Medicine*. <https://www.nejm.org/doi/full/10.1056/nejmc2201761>
- Harris, Mark (2023a). "Since You Asked." email exchange dated 12 February 2023.
- Harris, Mark (2023b). "Fun with ChatGPT-3." email exchange dated 18 February 2023.
- Heaven, Will Douglas (2020). "How to Make a Chatbot that Isn't Racist or Sexist." *MIT Technology Review*. <https://www.technologyreview.com/2020/10/23/1011116/chatbot-gpt3-openai-facebook-google-safety-fix-racist-sexist-language-ai/>
- Heikkilä, Melissa (2023). "How to Create, Release, and Share Generative AI responsibly." *MIT Technology Review*. <https://www.technologyreview.com/2023/02/27/1069166/how-to-create-release-and-share-generative-ai-responsibly/>
- Hsu, Tiffany and Stuart Thompson (2023). "Disinformation Researchers Raise Alarms About A.I. Chatbots." *New York Times*. <https://www.nytimes.com/2023/02/08/technology/ai-chatbots-disinformation.html>
- Kung, Tiffany *et al.* (2023), "Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models." *PLOS Digital Health*. <https://journals.plos.org/digitalhealth/article?id=10.1371/journal.pdig.0000198>
- LeWitt, Sol (1967). "Paragraphs on Conceptual Art." *Artforum International*.
- Leighton, Mary (2023). "The University of Minnesota law school tested ChatGPT on exams — it was a C+ student." *The Insider*.
- Metz, Cade and Gregory Schmidt (2023). "Elon Musk and Others Call for Pause on A.I., Citing 'Profound Risks to Society'," *New York Times*, <https://www.nytimes.com/2023/03/29/technology/ai-artificial-intelligence-musk-risks.html>
- Mozes, Alan (2023). "ChatGPT Rated as Better Than Real Doctors for Empathy, Advice," *Health Day News*. [ChatGPT Rated as Better Than Real Doctors for Empathy, Advice \(usnews.com\)](https://www.usnews.com/story/news/health/2023/02/17/chatgpt-rated-as-better-than-real-doctors-for-empathy-advice)
- Mozur, Paul (2018). "A Genocide Incited on Facebook, With Posts from Myanmar's Military." *New York Times*. <https://www.nytimes.com/2018/10/15/technology/myanmar-facebook-genocide.html>
- PAI, or Partnership on AI (2023). <https://syntheticmedia.partnershiponai.org>
- Profeta, Christopher (2015). <https://www.theedadvocate.org/how-to-improve-digital-writing/>
- Quach, Katyanna (2023). "Microsoft's New AI BingBot Bites Users and Can't Get its Facts Straight." https://www.theregister.com/2023/02/17/microsoft_ai_bing_problems/
- Rathje, Steve (2018). "Why People Ignore Facts." *Psychology Today*, <https://www.psychologytoday.com/us/blog/words-matter/201810/why-people-ignore-facts>
- Reuters (2021). "Boeing Accepts Liability for Ethiopian Airlines Max 737 Crash." *CNN Business*. [Boeing accepts liability for Ethiopian Airlines 737 Max crash | CNN Business](https://www.cnn.com/2021/03/29/boeing-ethiopian-airlines-737-crash/index.html)
- Reuters (2023). "Australian Mayor Prepares World's First Defamation Lawsuit over ChatGPT's Content." *The Guardian*, <https://www.theguardian.com/technology/2023/apr/06/australian-mayor-prepares-worlds-first-defamation-lawsuit-over-chatgpt-content>
- Roose, Kevin (2022). "An A.I.-Generated Picture Won an Art Prize. Artists Aren't Happy." *New York Times*. <https://www.nytimes.com/2022/09/02/technology/ai-artificial-intelligence-artists.html>
- Shanahan, Murray (2023). "Talking About Large Language Models." *Computation and Language*. <https://arxiv.org/pdf/2212.03551.pdf>
- Shanahan, Timothy (2018). Shanahan on Literacy. <https://www.shanahanonliteracy.com/blog/is->

- [comprehension-better-with-digital-text#sthash.mlmMT13q.dpbs](#)
Spaid, Sue (2023 forthcoming). "On Deterring Hate Speech. Whilst Maximizing Security and Privacy." *Ethics in Online AI-Based Systems: Risks and Opportunities in Current Technological Trends*, Elsevier Publications: Amsterdam.
- Spinello, Richard (2021). *Cyberethics: Morality and Law in Cyberspace*. New York: Jones and Bartlett Learning.
- Teicher, Jordan (2016). "When Photography Wasn't Art." *J-Stor Daily*.
<https://daily.jstor.org/when-photography-was-not-art/>
- Tennery, Amy and Gina Chereus (2016). "Microsoft's AI Bot Goes Dark After Racist, Sexist Tweets." *Reuters*. <https://www.reuters.com/article/us-microsoft-twitter-bot-idUSKCN0WQ2LA>
- Toh, Michelle (2021). "Facebook sued for \$150 billion over violence against Rohingya in Myanmar." *CNN Business*. <https://www.cnn.com/2021/12/07/tech/facebook-myanmar-rohingya-muslims-intl-hnk/index.html>
- Van Gerwen, Rob (2022). "Another Outcry for Artistic Relevance." *Aesthetic Investigations* (5(2): i-viii). doi: <https://doi.org/10.58519/aesthinv.v5i2.13645>
- Vartanian, Hrag (2022). "We Asked AI to Re-Imagine 5 Blue-Chip Art Exhibitions." *Hyperallergic*. September 29, 2022. <https://hyperallergic.com/765070/ai-reimagines-5-blue-chip-art-exhibitions/>
- Wahde, Mattias and Marco Virgolin (2022). "Conversational Agents: Theory and Applications." in Plamen Angelov (ed) *Handbook of Computer Learning and Intelligence - Volume 1*, pp. 1-60. <https://arxiv.org/pdf/2202.03164.pdf>
- White, Shannon (2022). "Top 10 Most Insane Things ChatGPT has Done This Week." <https://www.springboard.com/blog/news/chatgpt-revolution/>
- Wiener-Bronner, Danielle (2022), "KFC Germany apologizes for advertising a Kristallnacht Promotion." *CNN Business News*. <https://www.cnn.com/2022/11/10/business/kfc-germany-kristallnacht/index.html>
- Zumbrun, Josh (2023). "ChatGPT Needs Some Help with Math Assignments." *The Wall Street Journal*. https://www.wsj.com/articles/ai-bot-chatgpt-needs-some-help-with-math-assignments-11675390552?mod=hp_featst_pos5