

The AI Revolution: Speculations on Authorship, Pedagogy, and the Future of the Profession

Nir Evron
Tel Aviv University

Roi Tartakovsky
Tel Aviv University

In one moment of Edith Wharton's *Age of Innocence* ([2008] 1920), a casual conversation between three of the characters turns to “the fantastic possibility that they might one day actually converse with each other from street to street, or even—incredible dream!—from one town to another.” The scene is set in the 1880s, just a few years after Alexander Graham Bell patented the telephone. The idea of a device that can carry the human voice across vast distances still seems to the characters like something out of Jules Verne. We are not made privy to the details of their discussion, but the narrator informs us that they quickly fall into “such platitudes as naturally rise to the lips of the most intelligent when they are talking against time.” The scene is rich with dramatic irony. Wharton's characters cannot see what her narrator and readers know full well: that they are living on borrowed time, that Old New York—the cultural world that sustains their identities and lends meaning to their lives—is about to change in ways they cannot yet foresee or prepare for.

These days, we are all talking against time. That artificial intelligence (AI) will play an outsized role in shaping our collective near-term future is a

statement hardly anyone in 2024 would care to deny. What this future will hold for our profession and for the cultural world we call the university is still very much an open question. Moreover, unlike Wharton's narrator, we do not enjoy the privilege of hindsight. The historical change is happening—or will happen—to us. All we can do, given this situatedness, is speculate about what's coming in an effort to better manage the transition. Such speculation is the purpose of this issue.

In May 2023 we approached scholars in the fields of literature and literary theory, broadly conceived, and invited them to muse about the meaning and potential implications of large language models (LLMs) for our understanding of authorship, for the future of academic writing and pedagogy, and for the university at large. Long-form articles, we felt, would not be appropriate for engaging with developments that are still so nascent and open-ended. So we asked participants to submit shorter, more speculative pieces that unfold their claims in a subjunctive rather than indicative mood.

Our species' peculiar penchant for speculation takes on added significance in the context of the coming AI revolution. For, as it is often conceived, speculation involves taking a mental leap beyond the known, gesturing toward the still unarticulated, inventing a hitherto unimagined future. If we moderns have abandoned the idea that history is an endlessly recurring cycle in favor of a linear view, in which human time unfolds as a non-repeatable sequence of departures, then it seems to follow that novelty, born of speculation, experimentation, and chance, is a genuine possibility. And isn't such novelty precisely what LLMs, constrained as they are by the texts they were fed or trained on, cannot (in principle, by definition and design) achieve?

Then again, can human acts of innovation and imagination ever be said to be genuinely *sui generis*? After all, as many have argued, novelty is always relative rather than absolute, a reconfiguration of already existing materials rather than the deliverance of some transcendent creative faculty. Writers and poets often say that constraints are a prerequisite for creativity, that artistic freedom is intimately related to rules if not to algorithms.

In more ways than one, then, the encounter with AI is bound up with questions of originality, authorship, creativity, and freedom that have long occupied literary critics and theorists. To the extent that ChatGPT and its kind can be said to author texts, including original (are they original?) poems, short stories, and entire novels, its brand of authorship would seem to align with the poststructuralist, post- or non-Romantic conceptions of authorship that question the individuality (perhaps humanness) of the author. Indeed, quite a few of the contributors to this issue take as their

starting point Roland Barthes’s seminal “The Death of the Author” and Michel Foucault’s “What Is an Author?”—texts that famously conceptualize the flesh-and-blood author as either “dead” or relegated to the position of a discursive function—and ask what ChatGPT has done to these conceptions. Has it confirmed or even literalized them? Has it reminded us of what we already “know” but may have secretly been unable or unwilling to internalize? Is an artificial author the same as a dead author, an anonymous author, or a collection of authors?

Another direction undertaken by some of our contributors is to use ChatGPT as an opportunity to rethink human authors and their (allegedly) unique and irreplicable capabilities. Some of these essays include an experimental component: the writer prompts ChatGPT to produce a poem or a short story, specifying desired features (a sonnet, an unreliable narrator), and then evaluates the result, highlighting its deficiencies when compared with texts produced by humans. Of course, these discovered deficiencies may only be temporary setbacks to be overcome by the next, even more scaled-up iteration of the technology. Still, while acknowledging this possibility, some of our contributors strive to articulate those human capacities that LLMs may never, in principle, be able to emulate or indeed surpass.

A complementary direction represented in the short essays collected here asks not what AI is unable to do but, rather, what it can. For one, the rapidity with which it can produce passable and sometimes surprisingly effective texts, offers some dizzying promises for literary criticism. Imagine, for example, prompting ChatGPT to produce a fictional text in the style of, or emanating from, all the known works of a specific historical author. Or feeding it with all the known works in a particular genre and then asking for more fictional works that fit that generic repertoire. What, these essays ask, is the status and, more importantly, the utility of such *fictional* fictional works, works of fiction never in fact written by human authors but ones that *could* have been written?

Like other technological breakthroughs in the past, but perhaps more dramatically than most, AI is poised to disrupt many aspects of our lives. Its implications for the domains of employment, law, commerce, medicine, design, journalism, and many other fields are expected to be far-reaching. Many of the pieces collected here show the potential of AI to change the ways we see authorship, the ways we think about writing and reading. It stands to reason, too, that an AI-saturated future will have a dramatic effect on our very profession: on how we write, teach, and conduct research. Some of the essays try to sketch out these possible effects. How, they ask, might pedagogy adapt to a reality in which students can easily “outsource” their writing assignments? How will the AI revolution

change the way we go about doing research and even producing our own academic publications? How would universities reorganize themselves in the years and decades to come, as this technology becomes still more powerful and ubiquitous?

Our hope when we began thinking about this special issue was that it would supply its readers with provocative and potentially useful ways to think about how LLMs will affect our theoretical commitments, pedagogic practices, profession, and the university at large. Thanks to the enthusiastic cooperation of our contributors, we feel that that hope has been more than fulfilled.

The following is an overview of some of the themes discussed in the issue's essays.

James Phelan describes a challenge that he posed to ChatGPT, aimed at testing its storytelling prowess: to produce a fictional narrative that features unreliable narration. Phelan uses ChatGPT's failure to produce such a narrative to illustrate some of the technology's current limitations as well as to make a claim about the deeply ingrained rhetorical assumptions we bring—as lay readers and as narratologists—to our reading of fictional texts.

Charles Altieri likewise puts ChatGPT's alleged creative abilities to the test, but to evaluate how well it does as a poet. He prompts the LLM to produce two texts in the styles of Shakespeare's Sonnets 128 and 143. The results, argues Altieri, are "less than mediocre." His subsequent comparative close readings of ChatGPT's and Shakespeare's sonnets are geared toward demonstrating what he describes as the impoverished nature of LLM intelligence, "woefully lacking in the ability to provide models of careful sympathy or dynamic empathy."

Eamon Duede and Richard Jean So take a more favorable view of LLMs, arguing that they offer unique affordances for humanist inquiry. Striking a middle path between structuralist and humanist approaches to literary interpretation, they suggest that LLMs' ability to construct new fictional texts from well-defined though lacking corpora—their example: Asian American fiction before 1940—may be used to enhance our ability to see more deeply into the "underlying cultural structure" that informed such bodies of literature in ways that would otherwise be impossible.

Just like ChatGPT can produce fictive texts on command, it is also notorious for lacking the capacity to restrict itself to truths, a fact that Avery Slater explores in her contribution. Focusing on what is sometimes called AI's hallucinations, Slater homes in on the way ChatGPT will often cite sources that simply do not exist. Are ChatGPT's confident yet false citations of nonexistent research "merely a fixable bug or tied to a deeper, foun-

dational feature,” Slater asks, and she proceeds to show that they are key to the way ChatGPT may alter the poststructuralist version of authorship.

Radhika Koul asks not about AI as author but as reader or critic. Literary critics, she writes, routinely reason about others’ mental states (theory of mind) and make conjectures about the relation between the fictional text and the “real world” (high-level analogical reasoning). Are LLMs capable of similar feats of empathic reasoning? Koul reports on studies showing that LLMs perform decently well on some tasks aimed at measuring both capabilities and ends with her own experiment with GPT-3.5, asking it to assess the emotional appeal on readers of different stories, and demonstrating its brand of “affective awareness.”

While many of our contributors focus on how AI impacts our understanding of the author, Kurt Beals invites us to think about the position of the reader. “Who,” he asks, “will read these AI-generated texts, and why?” Using a series of thought experiments featuring a hypothetical LLM trained on Franz Kafka’s oeuvre (“KafkAI”), Beals argues that we remain attached, as readers, to the idea that some person is speaking to us through the text. The main readership for AI-produced fiction, he concludes, will likely be other computers.

For Ed Finn, the uncanny experience of interacting with ChatGPT leads to a meditation on the mysterious process of (human) writing and the shaping power of the imagination. Like Duede and So, he sees the advent of LLMs as an opportunity rather than a threat: “Instead of erasing human poiesis, these systems could help us rediscover and empower it.”

Striking a similarly positive note, N. Katherine Hayles discusses the potential pedagogical benefits of LLMs. She argues that it would be both wrongheaded and futile to try and ban this technology. Instead, we should embrace LLMs “as useful tools to accelerate student learning.” After a preliminary discussion of the concept of plagiarism, she offers some practical suggestions for what the college assignment might look like in an LLM-saturated world.

LLMs, argues Katherine Elkins, are much more than stochastic parrots. While conscious of the threat this technology might pose to our fields, she marvels at what it can already achieve and expects still more impressive achievements to come. In her piece, she reflects primarily on the implications of LLMs for linguistics, claiming that their ability to arrive at coherent utterances through a process in which grammar is secondary rather than primary poses a serious challenge to views like Noam Chomsky’s, which hold up grammar as the universal baseline of language.

Eric Hayot’s essay takes up three interrelated questions in turn: Can or should we speak of LLMs as quasi-sentient beings? What does the anx-

ious response to this technology by many humanist scholars reveal about the broader underlying tension between our poststructuralist assumptions and lingering Romantic attachments? And lastly, Hayot asks what should humanist critics *do* with AI-produced literature—in what sense, if any, might such texts function as a form of evidence?

Rita Raley and Russell Samolsky use Jorge Luis Borges's parable, "Borges and I," to explore the eerie awareness that every word, phrase, and argument they write will likely be ingested and appropriated as training data for LLMs at some point, occluding their authorship and perhaps eventually rendering human-produced humanist scholarship obsolete. Like Hayles, they believe that humanist pedagogy will have to adapt to the "new techno-linguistic order." However, they are less sanguine about the relative gains and losses of this transition.

Leah Henrickson's provocation seeks to undercut the idea that literature is about human connection between reader and writer. Drawing on her own research, in which she "sought to identify differences between reading human-written and computer-generated texts—only to reach the realization that they are negligible," Henrickson argues that reading literature is not so much about entering into another person's world as much as it is about engaging more deeply with ourselves. Our anxiety toward AI-produced literature, she suggests, is symptomatic of a broader societal crisis of involuntary loneliness.

Alexandre Gefen begins by taking us into the long history of imagining speaking and reasoning machines, from Albert the Great's fourteenth-century "Universal Doctor" to Borges's 1941 "Library of Babel." ChatGPT, he writes, takes us from the imaginings of writers to "the possibility for everyone to experiment with forms of assisted creation." And yet, in the two experiments he reports with two versions of ChatGPT, he finds that the earlier version actually produces a more creative and delightful text, while the later version already begins to acquire that ChatGPT signature blandness which seems to curtail any thoughts of unfettered creativity.

Hoyt Long's essay takes on the topic of LLMs and translation studies. Like many of our contributors, he does not believe that the attempts to ignore the rapidly evolving AI-based translation technology is a viable strategy. Nor, however, should we buy into the transhumanist fantasy, that this technology will enable us to "transcend the contextual specificity of language and thought in pursuit of pure informational exchange." After making a case against this either-or approach, Long proceeds to outline two alternative paths for coping with the growing incursion of LLMs into the realm of translation: "coordinated friction" and "playful experimentation."

Nir Evron broaches the question of how the advent of LLMs will affect our institutional home, the university. His speculative piece anticipates a future in which AI-driven tutoring systems, “able to teach any topic, on any level, and in a self-correcting pace, style, and language tailor-made for its student,” are ubiquitous. This development, he argues, would spell the end of the post–World War II era of mass higher education. The university as a site of teaching and learning will likely survive the AI revolution, but as a much smaller and more culturally marginal institution.

Matthew Kirschenbaum cautions us against the comforts of historicizing AI as simply one more technological innovation in a series that leads from the ballpoint pen to Autocomplete. LLMs, he suggests, are “qualitatively different” from such tools, and their impact, “not just on writing instruction but on writing itself as a category of human expression,” will likely be profound. To make good on this claim, Kirschenbaum offers a speculative scenario that spells out what the experience of writing a scholarly article—say, for *Poetics Today*—might be like in a few years’ time, after LLM technology has been fully integrated into our word processors.

Reference

Wharton, Edith. (1920) 2008. *Age of Innocence*. Project Gutenberg, posted August 12. <https://www.gutenberg.org/files/541/541-h/541-h.htm>.