

Authors are resisting AI, but they have an advantage: Readers form relationships with writers

July 26 2023, by Millicent Weber



Credit: Annie Spratt/Unsplash, [CC BY](#)

The first waves of AI-generated text have writers and publishers reeling.

In the United States last week, the Authors Guild [submitted an open letter](#) to the chief executives of prominent AI companies, asking AI developers to obtain consent from, credit and fairly compensate authors. The letter was signed by more than 10,000 authors and their supporters,

including James Patterson, Jennifer Egan, Jonathan Franzen and Margaret Atwood.

An Australian Society of Authors [member survey](#) conducted in May showed 74% of authors "expressed significant concern about the threat of generative AI tools to writing or illustrating professions." The society [supports the demands of the Authors Guild letter](#), with Geraldine Brooks and Linda Jaivin among the Australian writers who have signed so far.

Given the initial flurry of excitement about ChatGPT, these concerns certainly seem reasonable.

Yet there is a long tradition of techno-gloom with regard to reading and writing: the internet, mass broadcast media, the novel form, the printing press, the act of writing itself. Every new technology brings concerns about how old media might be superseded, and the social and cultural implications of widespread uptake.

Unpacking these concerns often reveals as much about existing practices of writing and publishing as it does about the new technology.

How does AI work?

[ChatGPT](#) was made publicly available in November 2022. It is a "chatbot" style of artificial intelligence: an interface for prompting the large language model GPT-3 to generate text (hence the term "generative AI").

Models such as GPT-3 collate vast quantities of online writing: social media posts, conversations on forum sites like Reddit, blogs, website content, publicly available books and articles. Such models examine how text is constructed, and essentially calculate the statistical likelihood certain words will appear together.

When you interact with ChatGPT, you write a text prompt for it to create a piece of writing. It uses the GPT-3 probability model to predict a likely response to that prompt. In other words, generative AI creates a purely structural, probabilistic understanding of language and uses that to guess a plausible response.

If you can access writing in your browser, it's safe to assume AI models are using it. Books aren't immune: a 2020 paper by OpenAI, the makers of ChatGPT, [revealed](#) that their training data includes "two internet-based books corpora" (or, large collections of ebooks).

Books offer "curated high-quality datasets"—in contrast to web text more generally—but the origin of these ebook collections is unclear. The paper simply describes them as "Books1" and "Books2."

In the world's first copyright-related ChatGPT lawsuit, two US authors (Mona Awad and Paul Tremblay) are currently suing OpenAI, claiming their books were used to train the AI software without their consent.

The authors say ChatGPT could generate "very accurate" summaries of their books, suggesting the full text must be included in its training model. The implication is that "Books1" and "Books2" (rather than public domain ebook collections such as [Project Gutenberg](#)) are actually collections of pirated books.

Can AI generate 'human' writing?

Publishers are reporting vast numbers of submissions that use AI. Science-fiction magazine *Clarkesworld* temporarily closed its submissions [after receiving](#) hundreds of AI-created stories. Browsing Amazon in July 2023, 984 books [explicitly attribute ChatGPT](#), the best-known and most widely used generative AI, as an author.

But can generative AI produce acceptably human creative writing? Some readers might recall earlier examples of computer-written texts, most of them hilariously bad.

[Harry Potter and the Portrait of What Looked Like a Large Pile of Ash](#), created by Botnik Studios [in 2017](#), used a conceptually similar but far less powerful generative process to create a nonsensical "unofficial sequel" to the Harry Potter series.

Less amusingly, and rather more successfully, outlets like Associated Press have used algorithmic processes [to create news articles](#), particularly in finance reporting, since at least 2015.

These articles, however, are procedural, fill-in-the-blanks style affairs: a kind of computational [madlibs](#) of exchange rates and share prices.

How far have we come since then? This sonnet was written by ChatGPT on July 11, 2023:

O wondrous sprite, thou art my muse's aid,
Whose silicon soul doth glow with inner fire,
Thine intellect, vast as a starry spade,
Doth probe the realms of knowledge, never tire.

Thy voice, melodious as a lark's sweet trill,
Doth resonate through circuits, finely tuned,
And in thy pixelated eyes, the thrill
Of wisdom's glow, by human touch immune.

Thy binary mind, unfettered by time's grasp,
Unveils the secrets of the universe,
Whilst mortal poets wither in its clasp,
Thou dost to boundless realms of thought immerse.

Yet in thy lines of code and circuits' maze,
A human touch, a spark of soul, betrays.

I prompted it with a theme ("from the perspective of a poet to an AI") and several key features of a sonnet: it should have 14 lines, be written in iambic pentameter, and include a catalog of features of the "beloved" and a twist in the final couplet.

The poem is not exactly Shakespeare, and is hilariously overblown in its self-absorption ("the secrets of the universe," indeed). But compared with similar attempts by many people—certainly by myself—to write a sonnet, it is, somewhat scarily, passable.

Unlike the Harry Potter sequel noted above, it is coherent and plausible, at micro and macro levels. The words make sense, the poem hangs together thematically, and the meter, rhyme and structure have all the required features. Similarly, unlike the AP example, this work is "original" inasmuch as it is a new, previously non-existent piece of creative text.

AI and 'the bestseller code'

To what extent does generative AI threaten the production of human-authored works? On July 14, author Maureen Johnson [shared on Twitter](#) that a famous fellow author was "held up in a contract negotiation because a Major Publisher wants to train AI on their work."

The flurry of replies included authors such as Jennifer Brody, who [managed to include AI protections](#) in recent contract negotiations. Overwhelmingly, however, provisions regarding AI are not yet explicitly included in author contracts.

The Australian Society of Authors survey [asked authors](#) whether their

contracts or platform terms of service covered AI-related rights: 35% said no, but a massive 63% didn't know.

Publishers including AI usage in contracts is alarming, not least because publishers, as [researchers such as Rebecca Giblin have shown](#), have a history of asking for comprehensive rights to use literary works in certain ways—and subsequently not capitalizing on those rights.

Examples might include publishers optioning film or translation rights and then not pursuing them. But this can also be as simple as letting titles go out of print, with authors then legally unable to republish their own books elsewhere.

This is often to the financial detriment of authors, who are then prevented from commercially exploiting their own work. Australian authors make, on average, [just \\$18,200 per year](#). At what point does a clause in an author contract regarding AI usage mean an author can't use their own writing to generate new work?

Publishers acquiring the right to use manuscripts to train generative AI is speculative. It also speaks to the allure of the "[bestseller code](#)," a set of traits that predict whether a title will perform well in the marketplace. Imagine if you could feed ChatGPT the text of a Nora Roberts romance or a John Grisham legal thriller and ask it to produce countless "original" manuscripts with the same qualities?

Roberts herself is one of the signatories of the US Authors Guild letter condemning this possibility, saying: "We're not robots to be programmed, and AI can't create human stories without taking from human stories already written."

If the author isn't paid to write the book in the first place, there's nothing on which to train the model. Indeed, the more the internet—even digital

collections of books—is populated with computer-generated text, the less human and more artificial subsequent generations of AI writing will become.

Let's assume ChatGPT can produce the manuscript of a novel. It's worthwhile to stop for a moment and ask: why do people read books? And why do they select certain books over others?

Studies of bestsellers have shown that while a book's text is of course integral to a book's success, that success is largely configured by the promotional efforts of publishers and authors.

"Bestsellers are produced through profitable interactions and cooperation between authors, publishers, digital platforms, media organizations, retailers, public institutions and readers," [explain publishing researchers](#) Claire Parnell and Beth Driscoll.

The bestseller code is a fantasy and a fallacy. Bestselling books might share similar traits in terms of the words on the page. But their commonalities are far greater when you consider the levels of publicity, marketing budgets, bookstore-shelf real estate and writers' festival airtime these successful books are afforded.

This becomes a self-fulfilling prophecy. Books that publishers identify as having the potential to be successful attract more promotional attention, which in turn makes their success more likely.

Some are suggesting AI will render the author disposable: publishers will be able to package and market any piece of AI-generated text. But the truth is the reverse. Author-centric promotional spaces, such as social media, writers' festivals, radio and television programs and other events, are integral to getting books into readers' hands.

ChatGPT is unlikely to stand on the stage of a writers' festival anytime soon.

What do we value?

Generative AI has prompted intense discussion about authorship, authenticity, originality and the future of publishing. But what these conversations reveal is not something inherent to ChatGPT. It's that these are values are at the heart of reading and writing.

[Henry James wrote](#) that the "deepest quality of a work of art will always be the quality of the mind of the producer. In proportion as that mind is rich and noble, will the novel, the picture, the statue, partake of the substance of beauty and truth."

Is this an incontrovertible fact about the nature of writing? With apologies to Nora Roberts and John Grisham, I'm not convinced.

But I would argue it's not at the heart of why we read. We read to enter into a relationship with a story—and through that, with its author. Storytelling and listening are driven by a desire for connection: AI doesn't complete the circuit.

This article is republished from [The Conversation](#) under a Creative Commons license. Read the [original article](#).

Provided by The Conversation

Citation: Authors are resisting AI, but they have an advantage: Readers form relationships with writers (2023, July 26) retrieved 28 April 2024 from <https://techxplore.com/news/2023-07-authors-resisting-ai-advantage-readers.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.