Nietzsche, AI Writing, and Authorship as Directing

What a bedridden philosopher in 1878 can teach us about authorship in 2026

Tobias Brücker 10 min read · 22 hours ago



A device made so that a hand could write. The question of where the writing then comes from — hand, instrument, or the arrangement of both — is older than any machine. An image from G.-B. Duchenne, "Physiologie des mouvements" (1867). Credit: Wellcome Collection, London. CC BY 4.0.

Late in his life, Nietzsche told a revealing story about how he wrote one of his books. He had been ill and so his most loyal assistant, Peter Gast, took dictation. Gast wrote down the words. Gast corrected them. And Nietzsche, surveying the result, delivered a verdict that should give every user of today's language models pause mid-prompt:

> *I dictated, my head bandaged up and in pain; he copied out and made corrections, too — basically he was the actual writer, while I was just the author. (*Ecce Homo*, "Why I Write Such Good Books," HA Section 5, translation by Duncan Large, Oxford World's Classics)*

Keep that distinction in mind; we will return to it shortly. Once you notice that the man who gave us *Thus Spoke Zarathustra* cheerfully admitted he sometimes didn't do the actual writing of his own books, the contemporary

panic about machines doing our writing for us starts to look less like a revolution and more like a very old argument with a new face.

The myth we keep telling ourselves

Ask most people what writing is, and if you scratch the surface of their answer, you'll find a stubborn romantic fantasy: a lightning bolt. Inspiration descends, the genius receives, the masterpiece arrives more or less intact. We cling to this picture because it flatters us — it makes the writer a kind of priest rather than a mere laborer.

The Nietzsche of the late 1870s — the cooler, more skeptical, aphoristic Nietzsche of *Human, All Too Human* — found this fantasy not just false but a bit of a racket. One that artists, he observed, have a professional interest in our believing. In an aphorism bluntly titled "Belief in Inspiration," he punctures the whole thing.

> *In truth, the imagination of a good artist or thinker continually produces good, mediocre, and bad things, but his power of judgment, highly sharpened and practiced, rejects, selects, ties together; thus, we now see from Beethoven's notebooks that he gradually gathered together the finest melodies and selected them, as it were, out of multiple beginnings. (*Human, All Too Human I*, Aphorism 155, this and subsequent quotations from this work trans. Gary Handwerk, Stanford University Press*)

What distinguishes the artist is not production but judgment: a faculty, sharpened by practice, that rejects, selects, and ties together. He points to Beethoven's notebooks, where the most sublime melodies turn out to have been assembled gradually, picked out of many discarded attempts.

His conclusion is one of the great deflations in the history of aesthetics:

> *All the great artists were great workers, tireless not only in inventing, but also in rejecting, sifting, reshaping, ordering. (*Human, All Too Human I*, Aphorism 155)*

Read that list again — rejecting, sifting, reshaping, ordering — and notice that it describes almost perfectly what you do when you sit with a language model's output and decide what to keep and what not.

Writing was always a bundle of practices

Here is the move that makes Nietzsche so useful for thinking about AI. He refuses to treat writing as a single magical act. For him, "writing" is an umbrella term covering a whole bundle of practices, most of which have nothing to do with the romantic moment of composition. The notebook stands for gathering, the rewrite for selecting, the crossed-out line for rejecting. Judgment, crucially, is not purely a mental event that happens before the hand moves. It happens in the material itself — in choosing a word, killing a phrase, recasting a thought.

Once you see writing this way, the actual act of putting down letters — typing, scribbling, or, for a language model, generating — shrinks to a single component in a much larger operation. And nearly everything around that component still requires human decisions: the idea, the sources, the audience, the length, the structure, the stylistic texture, the venue, the promotion. None of that vanishes when a machine drafts the prose.

Nietzsche himself sometimes delegated. Fair copies, transcriptions, editorial cleanup — some of it went to Gast. What we now call 'prompting' he might have recognized at once: the craft of handing a task over in a precise manner, with the right information attached. The managing of a writing process — the directing of it — was always there. AI merely makes it impossible to ignore.

What actually changes (and what doesn't)

If you take Nietzsche's view seriously, you stop expecting AI to abolish writing altogether and start looking out for something subtler: a redistribution. Editing, correcting, rewriting were always the heart of the work. What's new is only that some of these tasks can now be handed over to a machine rather than a grad student. The activity doesn't transform; the division of labor between human and technical hands does. A few consequences follow.

Judgment, not generation, is the scarce resource. The bottleneck was never the number of ideas — a model can flood you with those. It's the faculty that tests them, picks the good ones, throws out the rest, and rearranges what survives. Choosing the prompt, the fine-tuning, the data; deciding whether and where to publish the result — these are all acts of judgment, and they remain stubbornly human. This is precisely where the current discourse around AI tends to go quiet.

Judgment has to be trained. In his aphorism "The seriousness of craft," Nietzsche sketches an imaginary novelist and insists on the heap of unglamorous tasks we habitually look past (*Human, All Too Human I*, Aphorism 163). What separates the real novelist from the amateur is the craftsman's seriousness: learning the work through patient, repetitive practice. The same is true, unromantically, of prompting, fine-tuning, and curating data. These are skills, and skills are earned.

This is a cultural anxiety, not a technical one. Strip away the noise and the debate about AI-generated text is really a debate about a deeply rooted model of authorship — one still entangled with genius, talent, sudden visitations. Adopt instead the *Human, All Too Human* conception, where creativity is a long sequence of decisions, manual tasks, and delegable labor, and AI-assisted writing becomes merely a changed writing situation. Nietzsche's idea of authorship was, in this sense, already open to the participation of instruments, materials, and the ordinary circumstances of life — something I traced manuscript by manuscript in a book on how he composed *The Wanderer and His Shadow* on his walks through the Swiss Alpine valley of the Engadine.

Authorship as a spectrum

Authorship was never a clear thing. It was always a label — usually a proper noun — stretched across a wide spectrum of social, cultural, and technical realities: noting and fair-copying, pens and desks, self-staging and reviews and the politics of reputation. Into this crowded scene, language models arrive less as a rupture than as a shift in emphasis — much as the virtual band Gorillaz succeeded not by inventing a new kind of music but by rearranging the familiar interplay of musicians, producers, and media staging.

And if we were ever truly tracking demonstrably human writing, that ship has sailed. The tools are now embedded in the word processors, email clients, and browsers through which most prose passes; the line between human and machine text has dissolved not through some dramatic rupture but through quiet, ambient assistance. This is, by every indication, only the beginning. As the tools grow cheaper almost by the month, the question of whether human authorship matters will be posed more and more radically — and the real question underlying this is: how much of literature's value lies in the functional production of good sentences, and how much in our expectation of encountering another human consciousness speaking to us through the text?

Why AI writing doesn't land yet — and why that's temporary

Here is a claim worth defending: AI-based authorship is entirely possible. What's missing is not quality. What's missing is the cultural scaffolding.

To borrow and adapt a term from the literary scholar Steffen Martus, Werkpolitik, or the politics of the work, is the practice by which authors, critics, and audiences invest a text with authority and force, positioning readers so that they feel they are hearing the text's appeal, its claim on them. A dozen philosophical-sounding sentences from a chatbot command no deep attention. The reason is not their shallowness, nor even their machine origin; it is that they have no format in which such a claim to meaning could be ascribed to them.

Think about what you actually do with a text that matters to you. You mark up a passage, cite it, interpret it, review it, situate it against the author's earlier work or against rival authors. All of this requires stable anchors — a name, an intent to publish, a context. Author-centered authorship functions as a guarantee of sense, an ordering authority against an overwhelming flood of text. Without a publication and an author's name, a generated text is almost impossible to quote, interpret, or discuss. The trouble isn't the sentences. It's that nobody has been positioned to take them seriously.

Which leads to a conclusion that is more radical than it first appears. If authorship is essentially an interpretive function of texts — a set of circumstances under which we grant a text the presumption of meaning — then AI texts can, under the right cultural conditions, fulfill that function too. They will need to earn it the way texts always have: through publication, recognizability, referenceability, and a critical mass of texts coherently

related to one another. History is not on the skeptics' side here. The pseudo-Aristotelian writings, the long tradition of pseudepigrapha, the anonymously published texts of the Enlightenment: texts have repeatedly acquired authority without a fixed, verifiable human behind them. It seems only a matter of time before generated philosophical or literary texts find their foothold — most likely, at first, nestled within the context of established human authorship.

Authorship as directing

So what is the author, in a world where sentences can be outsourced?

The most useful word comes from the visual arts, where this problem was observed early. Back in 2017, Blaise Agüera y Arcas — then head of Google's Machine Intelligence group — put it memorably:

> *As the tools become more powerful, what we have thought of as creation becomes the end of a giant lever. Rather than worrying about every work or every brushstroke, you can produce a lot by picking and curating. (Blaise Agüera y Arcas)*

Translate that to writing and the trajectory is clear. The more capable the language models become, the more the author's work shifts from formulating to selecting — from writing to editing proposals, from drafting to evaluating alternatives. The word that fits this best is 'directing': the craft of controlling how a performance develops out of material composed or assembled by others — with all the emphasis on evolution, on a product that takes shape only in the making. The film director decides on vision, casting, set, tone, while delegating the detailed execution. Detach 'directing' from the institutional job title and treat it as an activity, and you get artistic directing: the active, ongoing steering and shaping of a creative process. Nietzsche's own list — rejecting, sifting, reshaping, ordering — is a catalogue of actions that could be considered directing.

For writing with AI, then: the author sets the direction, chooses, decides on tone and structure and effect, and hands the detailed execution to another actor — human or machine. Each of us, writing this way, becomes a small organization in which we act as the artistic director, drawing on a team of people, agents, data, models, software, and tools.

And now we can redeem the joke from the beginning. Nietzsche said of Gast: he was really the writer, while I was merely the author. The more honest formulation, for our moment, inverts his ranking. In a process driven by genuine human judgment, you don't say "the machine was the real author." You say: the AI is merely the writer, while I was the actual artistic director. The status doesn't flow to whoever produced the words. It flows to whoever exercised the judgment.

A closing caveat, in fairness

It would be too tidy to end there. But this division of labor can shift: as fine-tuning and autonomous agents take on more of the directing itself

(structural and stylistic decisions, not just sentences), the human role may shrink from directing to mere framing, and honesty might then demand a more modest title than “artistic director.” But even this is a Nietzschean point, not a refutation of one. Every new technology foregrounds some aspects of writing and pushes others into the margins. The lesson of Nietzsche was never that the human author is eternal and untouchable. It was that “the author” was always a label stretched over a shifting bundle of practices, instruments, and circumstances — and that what deserves our attention is not the romantic act of creation but the unglamorous, profoundly human work of judgment.

Consider how this essay came to be. It began as a side thread in a German article written for Nietzsche scholarship. My thinking on authorship as directing would reach only a small German-speaking community and go no further. Three colleagues read the draft and found the reflections on authorship its most compelling thread. That nagged at me. So I lifted one strand out to let it stand on its own.

I translated the relevant passages with an AI language model, in a back-and-forth: drafts proposed, passages weighed, paragraphs rejected, a few sifted, reshaped, and ordered into what survived. The text changed a great deal along the way. Then I changed it again, working the machine’s translations into my own. Finally I sent it to a friend, a bilingual artist and English teacher, whose careful corrections and comments I largely accepted, rewriting several passages accordingly.

And yet, when a text is carried into another language this way — drafted with colleagues, translated and revised in dialogue with a machine, with oneself, and with a friend — what remains of the demand that it be genuinely authored, carefully reviewed, owned by a single hand? Doesn’t that make this piece, a little too neatly, an instance of its own argument? Was I the writer here, or the artistic director?