statuSQuaestionis
language text culture

Ana Ilievska
Stanford University / University of Bonn

Frankenstein is about ChatGPT: Thinking and the Future of Literary Study in the Age of Generative AI

Abstract

This paper is a speculative inquiry into how generative Artificial Intelligence may affect our thinking abilities and what the role of the humanities and literary study will be in the process. Drawing on Hannah Arendt's Socratic notion of thinking as an internal dialogue as well as on various philosophical, psychological, and sociological perspectives on solitude and contemporary social character, I show how GenAIs such as ChatGPT present a dual challenge: the potential erosion of human cognitive autonomy and the need for a renewed focus on fostering critical thinking. To illustrate this challenge, I offer a fresh rereading of Mary Shelley's *Frankenstein* (1818) as a warning against thoughtlessness in the age of scientific progress by drawing a parallel between Victor Frankenstein's lack of solitude and friendship and our own contemporary attitudes towards technology, self, and others. I then ask where thinking can occur in contemporary society, where overstimulation, distraction, and a lack of solitude are widespread. Finally, I call for a renewed emphasis on thinking as a distinct human ability and underscore the importance of literature in preserving spaces for genuine thought and engagement with self and others.

"And even Socrates, so attracted by the marketplace, must go home
where he will be alone, in solitude, to meet the other fellow."
Hannah Arendt

"So rarely I saw him,
and always in haste.
Once, or so it seemed,
it was in one of the darkest
corners of a bar, at the port.
But was it me, was it him?"
Giorgio Caproni

1. *Introduction*

In 1961, while reporting for *The New Yorker* on the trial of Eichmann, one of the key figures behind the Holocaust, philosopher Hannah Arendt asked herself: how could great evil come from such an unremarkable person? Is thoughtlessness, rather than wickedness, the cause of evildoing? She went on to write a scandalous article where she coined the phrase "banality of evil," proposing a thesis that "the phenomenon of evil deeds, committed on a gigantic scale, [...] could not be traced to any particularity of wickedness, pathology or ideological conviction in the doer" but to "a curious, quite authentic inability to think" (Arendt 1971, 417). In other words, she suggested, evil acts on a mass scale are not the work of one, wicked, criminal mastermind, but are the product of a mass individual suspension of thought. Evil acts germinate in the soil of un-thinking, i.e., when a human being relies on "clichés, stock phrases" and "adheres to conventional, standardized codes of expression and conduct" (Ibid.: 418). These codes usually have the role of protecting us from "noise," that is, from overstimulation, stress, and from reality itself. There is no lack of terminology in the history of thought to designate the various modes of guidance on which humans rely when autonomy is impossible or undesirable. For Heidegger it was *das Man*. To exist inauthentically is to be lost in things and people, in the impersonal "man sagt, man hört, man ist dafür, man besorgt" ("one says, one hears, one is for, one gets;" Heidegger 2004, 113, my translation). Carlo Michelstaedter called it *retorica* ("rhetoric") and what he meant was the reliance on social rules, customs, and conventions which determine our lives and alleviate stress, but at the same time prevent us from walking the way of *persuasione* or authenticity (Michelstaedter 1982).[1] Sociologist David Riesman identifies three modes of reliance or of insuring conformity in the members of a given society: "tradition-direction," "inner-direction," and "other-direction." While these three types can coexist, contemporary Westerners are predominantly "other-directed," i.e., the globalized Western individual, far from being autonomous, relies on peer approval, the media, and influencers for guidance: "What is common to all the other-directed people is that their contemporaries are the source of direction for the individual –

1 For a provocative comparison between Heidegger and Michelstaedter, see Vašek 2019. On Michelstaedter, authenticity, and thought, see Harrison 2017 and Cangiano 2019.

either those known to him or those with whom he is indirectly acquainted, through friends and through the mass media" (Riesman et al. 2020, 20). Peer approval becomes a psychological need for the other-directed who invest all their and emotional reserves into "paying close attention to the signals from others" (Ibid.) in "a society increasingly dependent on manipulation of people" (Ibid.: 114).[2] Ultimately, such a cognitive stale ensues that, "while [an other-directed person] is willing to say what he likes, he cannot believe in himself enough to know what he wants" (Ibid.: 172).[3]

Needless to say, this form of overreliance on external guidance is the undoing of thought. But this undoing does not only characterize those eras and persons under the spell of great ideologies and charismatic leaders. Rather, unthought can most certainly be diagnosed in the most liberal, individualistic societies and persons as well. In other words, being an intellectual, a skilled professional, or a citizen of a liberal democracy is not a shield against thoughtlessness, just as oppression and the lack of formal education and power do not equal wisdom, solidarity, and innocence.

I would like to examine the implications of thoughtlessness in relation to the use of technology today in the attempt to offer some thoughts about what the use of generative AI (GenAI) could mean for human thought and autonomy. By taking over some of our cognitive tasks, will GenAI impact our ability to think and can "great evil" come from this human-machine cognitive distribution? Second, I want to understand what the role of the humanities and the study of literature will be in the process, the role, that is, of disciplines that historically have been primarily concerned with the pedagogy of language-based skills (reading, writing, listening, and speaking) through the translation and organization of thought into text (i.e., form). In this sense, GenAIs such as ChatGPT present us with a dual challenge: the potential erosion of human cognitive autonomy and the need for a renewed focus on fostering critical thinking. To illustrate this challenge, I draw on Mary Shelley's *Frankenstein*

2 Sincere thanks to the reviewer of this essay for bringing Riesman's book to my attention.
3 One only needs to think of the many forms that René Girard's idea of "mimetic desire" has taken on within marketing and tech companies who capitalize on this very need for peer approval (see Girard 1976). As Riesman et al. write, for other-directed persons, "the product approved by most of the others, or by a suitable testimonial from a peer consumer, becomes the 'best.' The most popular products, by this formula, are the products that happen to be used by the most popular" (Riesman et al. 2020, 69).

(1818) as an example of what happens when a knowledgeable person fails to engage in thinking. Usually read as a cautionary tale about technology running amok and turning against us, I put forward a reading of Shelley's novel as a warning against thoughtlessness. On one hand, it is a tale about a brilliant scientist who creates life out of death; on the other hand, it is about a deeply conflicted individual who is incapable of having an interaction with himself and does not know neither solitude nor company.

Thinking, as we will see later, is "a discourse that the mind carries on with itself about any subject it is considering. [...] When it reaches a decision – which may come slowly or in a sudden rush – when doubt is over and the two voices affirm the same thing, then we call that its 'judgment'" (*Theaetetus* 190a). The bane of Victor Frankenstein is not his superhuman murderous assemblage of a creature, but his own thoughtlessness. *Pace* Bostrom (2014) and many other contemporary tech-pessimists, our bane is not the advent of an artificial superintelligence, but our own inability to carve out spaces for thought and autonomy. In this sense, the present inquiry is not so much about the impact of technology on the literary imaginary but, conversely, on the impact that the literary imaginary can have on how we engage with technology. This calls for a reevaluation of our staple technological tales such as but not limited to the Frankenstein myth, issuing into a new culture of reading the classical stories that captivate the global imagination and furnished much of the imagery that we use to talk about technology today.

2. *Thinking and ChatGPT*

In November 2022, Open AI launched ChatGPT (generative pre-trained transformer), a chatbot that can do anything from answering any information-based question to writing a poem and composing a personal email in the style of Jorge Luis Borges. All this with minimal human input and in the matter of seconds. In addition to being an interactive version of Wikipedia, then, a GenAI such as ChatGPT can generate and organize text based on minimal instructions and keywords provided by a human. In February 2023, Jasper (a company that markets the homonymous software for businesses),[4] hosted in

4 Jasper's homonymous chatbot is a GPT-4 tool allowing businesses to generate brand-specific content efficiently and online.

San Francisco the first ever industry conference on GenAI. During her presentation, Meghan Keany Anderson, head of marketing at Jasper, stated that, rather than "let the best writers win," the new technology's headline is "let the best ideas win" (Anderson 2023). Bing Chat, BLOOM, Jasper, ChatGPT, or Sydney will help us save time on generating text so that we can focus on content and editing. So far so good. But if GenAI will provide sufficient leisure time for us to work on our ideas, then the question remains as to where these ideas will come from and whether we will use the free time to develop them or disengage from thinking activities altogether.

To further set-up the problem and how thinking relates to ChatGPT, two things need to be stated clearly. First, that humans are needy, fragile animals who avoid pain and seek pleasure. Freud had given this mechanism the name of "the pleasure principle." Later, he complemented it with the famous theory of the death drive, but the outcome is the same: both the life and death instincts in humans collaborate in bringing about the "return to the inanimate state," but on our own terms, making "ever more complicated *détours* before reaching [our] aim of death" (Freud 1989, 613).

Second, humans project themselves into the future and use technology to fulfill their needs. Let us recall Freud's three main sources of suffering in human life as he outlined them in his famous *Civilization and Its Discontents*: 1) our decaying bodies; 2) the external world with its rules and limitations; and 3) our relations to other human beings (Freud 1989[a], 729). While thinking can be perhaps subsumed under the first source – the body – and I do agree that thinking is embodied, for the sake of further focalizing our point of interest here, let us add a fourth source of potential suffering for humans: our minds. In the words of psychologist Julian Jaynes, whose theory of the bicameral mind informs the brilliant HBO series *Westworld* (2016), our mind is a "secret theater of speechless monologue and prevenient counsel, an invisible mansion of all moods, musings, and mysteries, an infinite resort of disappointments and discoveries. A whole kingdom where each of us reigns reclusively alone, questioning what we will, commanding what we can. A hidden hermitage where we may study out the troubled book of what we have done and yet may do" (Jaynes 1976, 1). This private mansion and its thinking processes are the focus of our inquiry.

Thinking (which involves discernment, decision-making, and planning) just like labor (which involves the exertion of physical force needed to make something) is strenuous. Unless done for leisure, both thinking and physical

labor can be outright painful. As John Guillory writes in his brilliant book on the history and future of the study of literature, "we think with our bodies" because "scholarly work is work with the eyes, ears, and hands," leading not just to professional deformation (i.e., specialization) but also to a physical one (the scholar is a hunchback).[5] As a society we have devised various ways of alleviating this labor.

It is then not surprising that one of the main goals of Western civilization until the twentieth century had been to reduce the physical strain on the body, i.e., to create technologies that will decrease or obviate the need for physical exertion. But since the twentieth century and the accompanying advancements in digital technology, we have shifted our attention to the mind. Now our foremost fixation seems to be with extending our mental capacities or, more correctly, with reducing the cognitive strain on our minds through the creation of technologies such as ChatGPT that will make it easier or outright unnecessary for us to think, plan, make decisions or even engage in creative acts. In scientific-engineering terms, this twofold endeavor to make life easier for our minds and bodies is called automation.

Speaking from a neuroscientific point of view, legal ethicist and bioscience professor Nita Farahany warns against the advent of an even more aggressive automation of our thinking processes. In a recent interview with Edward Helmore, she said that "the brain [...] is the one space we still have for reprieve and privacy, and where people can cultivate a true sense of self and where they can keep how they're feeling and their reactions to themselves. 'In the very near future that won't be possible.'"[6]

Thanks to automation, then, we have found some form of relief from the first three obstacles to human happiness as outlined by Freud: Our bodies thrive thanks to medicine, prostheses, and machines; the external world is manageable thanks to our dwellings, our families, and our social institutions; finally, other people are increasingly kept at bay thanks to the explosion of individu-

5 This is Guillory's reading of Nietzsche on the toils of being a scholar (Guillory 2022, 7f.). See also Nuttall 2011.

6 Helmore 2023: "It's been widely noted also that Elon Musk's Neuralink and Mark Zuckerberg's Meta are working on brain interfaces that can read thoughts directly. A new field of cognitive-enhancing drugs – called Nootropics – are being developed. Technology that allows people experiencing paralysis to control an artificial limb or write text on a screen just by Thinking it are in the works." See also Farahany 2023.

alism owed to the World Wide Web and, in particular, thanks to social media which enable interactions without face-to-face engagement. On the web, we are merely specters of ourselves. There is no obligation to interact and we in no way owe our presence and availability to the other. After each tweet, Facebook post, blog comment, or ChatGPT interaction we can retreat into the absolute privacy and security of our homes. All it takes to disengage is to turn off the computer, pull the plug, go out for a walk without my phone in my pocket. Everything else is compulsion. However, despite its promise to relieve us of our mental and physical troubles, automation does not only help humans to do repetitive tasks and heavy industrial work. It also creates unemployment, anxiety, and perhaps even depression. As Kevin Roose, the technology columnist for *The New York Times* writes, "there is no evidence that today's workers are happier than workers of previous generations. Overall rates of depression and anxiety are much higher in the United States today than they were thirty years ago, and self-reported workplace stress levels have been rising steadily for decades" (Roose 2021, 16). In other words, "we are not getting happier at work, despite the fact that our jobs are safer and less grueling than ever" (Ibid.).

How do things stand, then, when it comes to the fourth source of suffering, our own minds? The fundamental question both in logistical and psychological terms seems to be: How can we live with ourselves? In order to think, we need to not only carve out a solitary space for ourselves but also bear this solitude if we want to come anywhere close to David Riesman's call for autonomy: "[Riesman's] ideal of autonomy," Richard Sennett writes, "requires the capacity to take one's distance from one's surroundings. Which is to say, he wants people to dwell in solitude without feeling lonely" (Sennett 2020, xvi). But when we take away manual work, repetitive tasks, and routine, the mind tends to become self-referential. This is another point on which labor automation has failed to make people happier but has made them feel even more afloat and without 'things to do,' because we do not really know how to prevent the mind from cannibalizing itself. As Arendt writes, thinking has a high tendency towards self-destructiveness and "can seize and get hold of everything real" (Arendt 1977, 56). In order to 'distract' the mind from itself, we rely on psychoanalysis or therapy, medication, intoxication, friendships, sex, consumerism, and, of course, the arts, knowledge, and religion.

Now, if we are not getting happier, if machines can replace/supplement our bodies and AIs can replace/supplement our minds, purportedly doing

everything better and faster than we can, then what remains of what we've always thought of as authentically human? If the socialist utopia of the four-hour workday becomes possible, how shall we fill the remaining time? Will full automation first lead to even more negative effects on human beings and society before we figure out how to rewire an entire species (that for millennia has been functioning on survival mode) and learn how to create meaning for ourselves outside of labor and necessity? This is where Arendt's question returns to us with urgency, scaled down to our specific inquiry and *Zeitgeist*: can great evil come from cognitive automation and the technologies associated with it? Or more simply put, can great evil come from ordinary people using these technologies?

Again, all of these last points could be understood as problems concerning 'the body' and 'society.' But in order for us to get at thinking and remain there, we have to allow ourselves to dwell on the mind and, in particular, on thinking, as a problem in itself. This allows me to make the following claim: within the realm of thinking, as it takes place in the embodied mind, GenAI is the most recent, perhaps most impactful technology that changes the way we think and make decisions. It retroactively influences all the other three sources of potential suffering for humans, but it begins with the act of thinking itself. Thinking, let us say it again, like the plowing of the earth, is labor, it is work, it burns calories (but not enough and not visibly enough so that it can be marketable). It is time-consuming, inconvenient, it removes us from society, and is terribly inefficient. Thinking, in other words, goes entirely against the current optimization mind-set. Most importantly, it is done by humans, the most inefficient, unreliable, error-prone, (over)thinking machines. Thinking is the seat of rebellion. But it is also the most vulnerable of our cognitive faculties.

4. *Thinking and Literary Study*

Hannah Arendt's questions on thinking and moral considerations imposed themselves on me with urgency after attending two major conferences on AI: "AI in the Loop: Humans in Charge" organized by the Stanford Institute for Human-Centered AI (HAI) on 15 November 2022; and the aforementioned industry conference on GenAI hosted by Jasper. During both events, computer scientists, engineers, coders, marketers, investors, businesspeople, and

CEOs from Google, Microsoft, OpenAI, and Jasper – just to mention a few – brainstormed the future of GenAIs such as DALL-E, Jasper, and ChatGPT. The input from the humanities was from minimal and banal to non-existent although the physical space and walls of Jasper's GenAI conference in San Francisco were covered with quotes on creativity by Maya Angelou, Octavia E. Butler, and Alice Walker side by side with quotes by Steve Jobs and other important names in the tech industry.

On one hand, as someone who has made a life out of literature, it was impossible for me not to think about how literature is being made to serve industry and about how unlikely it is that any of these writers would have consented to their words being taken out of context and applied to a text-generating technology to call its output 'creativity.' But, as Socrates said, "once a thing is put in writing, the composition, whatever it may be, drifts all over the place, getting into the hands not only of those who understand it, but equally of those who have no business with it; it doesn't know how to address the right people, and not address the wrong. And when it is ill-treated and unfairly abused it always needs its parent to come to its help, being unable to defend or help itself" (*Phaedrus,* 275e). On the other hand, I also thought that the very foundation of our field, the very bread-and-butter of literature – the art of text-making – is being transformed right under our noses and we are not invited to the conversation. Both at the GenAI conferences and in all the recent best-selling publications on the topic, the task of thought and critique are relegated to New York Times columnists, Silicon Valley correspondents, marketers, professional AI speakers, and business consultants.

During one conversation at the conference in San Francisco, Nat Friedman, MIT graduate and former CEO of GitHub, said that GenAI will "rewrite civilization." While I strongly disagree with the grandiosity of similar predictions, I do admire the perspicuity of Friedman's word choice: GenAI is not going to *transform* or *disrupt* civilization, to use two favorite terms of Silicon Valley. It is going to *rewrite* it. What it is going to change is writing, one – at first glance insignificant but ultimately fundamental – prized feature of modern civilization. With writing comes education. For schools and universities are educational institutions that have elevated the teaching of writing skills to their primary concern, especially over the past decade. Indeed, no U.S. university of the several that I have attended lacked a writing center, writing tutors, composition courses for majors, and emphasized how much the role of the humanities is to trans-

mit the knowledge of form, i.e., to teach students how to write and evaluate final papers, dissertations, abstracts, CVs, cover letters, research statements, grant proposals, and articles. MFA programs have thrived in the past half a century, promising all but to teach the ultimate secret to writing best-selling novels and poetry. To put it less generously, humanities departments have become writing factories obsessed with the teaching of form and method, obsessed, that is, with the *how* of language. "The study of literature," Guillory writes, "is the site for development of modes of cognition specific to language use. These modes refer to potentialities in all the modes of language use – *listening, speaking, reading,* and *writing* – when these practices are wrought beyond their intuitive base" (Guillory 2022, 351f., cursive in the original).

However, today, very little effort goes into training students how to read, speak, or listen, and all our efforts go into teaching them how to write and 'brand' themselves and their research. Why? Because, with the teaching of such quantifiable skills, we enter the realm of professionalization which is based on pure *technē.* Plato scholar David Roochnik defines *technē* as follows: "a thorough, masterful knowledge of a specific field that typically issues in a useful result, can be taught to others, and can be recognized, certified, and rewarded" (Roochnik 1996, ix). It seems that if, like the sophists, we can show that the purpose of the humanities is to teach iterative skills that can be accredited and applauded, we gain public approval and legitimacy as a discipline, not to mention funding. This is how the study of literature was weened into becoming a 'profession' before it was even a discipline, Guillory aptly notes.

Now, as much as I agree with Guillory that these skills are specific to the study of literature, the question is why these cognitive abilities, as he calls them, are not already mastered at the high-school level only to be chiseled to perfection at the college level (and taken for granted at the graduate level). Why has the pedagogy of writing been imposed on underpaid graduate students who, often, are not good writers themselves, as well as on precarious postdoctoral scholars and overworked junior faculty? Making sure that persons gain high proficiency in these skills already during their formative teenage years would not only be the more democratic way to go about literacy (everyone goes to high school in the West while only very few can afford college degrees, especially in the U.S.); it will also leave space for the cultivation of the one cognitive skill or 'fifth element' which is the topic of this paper and which Guillory leaves out of his account: thinking. Thinking – like writing, speaking, listening, and

reading – is also a cognitive ability specific to language use. However, unlike other skills, thinking is difficult to quantify, evaluate, or market and does not immediately produce tangible results. At the same time, it also seems to be the one cognitive skill that artificial systems such as ChatGPT based on large language models just do not possess.

Without a doubt, postmodernism played a prominent role in this turn of literature departments towards method and skill and away from any claims to truth or morality, whereby McLuhan's idea of "the medium is the message" seems to have become the core of the humanistic endeavor: it is not necessarily about *what* we say, but about *how* we say it. This idea has seeped into recent scholarship on the nature of literary criticism and theory as craft, most prominently explored by Jonathan Kramnick in his much-discussed *Critical Inquiry* article on "Criticism and Truth." Kramnick has no doubts as to the *technē* nature of our profession: "writing criticism is knowing how to do something and the knowledge exhibited a kind of know-how" (Kramnick 2021, 225). In other words, literary criticism is "a most intensive handiwork" (Ibid.: 223) and the way in which literary scholars perform their craft, according to Kramnick, is through the method of close reading and in-sentence quotation. Those are expert practices "of writing prose and making text, of weaving one's own words with words that precede and shape them. This practice is craftwork in a literal sense. It is something one does or makes with one's hands" (Ibid.). Now the emergence of such a technology as ChatGPT (which is adroit at producing text and in-sentence quotation) casts a dense, dark shadow onto similar arguments about the particularity of literary study. GenAI exposes with an even greater force the shift of Western civilization and educational institutions from rhetoric to writing or, what Walter J. Ong characterizes as the gradual migration "from the oral to the chirographic world" (Ong 1982, 116). For Guillory "the postrhetorical condition" is the result of this transition that ushered into a "writing-centered pedagogy" in education (Guillory 2022, Ch. 5).

The question then begs itself as to what remains of our field and pedagogical aims when every student, professor, and academic can simply instruct a GenAI to translate their research keywords into the form of an abstract or a final paper, intermingled with quotations from other thinkers and writers. Indeed, as Roose writes, "[f]or decades, most automation was focused on repetitive manual tasks [...] but today [it] involves lots of tasks like planning, prediction, and process optimization. As it turns out, these are exactly the kinds of

things AI does well. In fact, white-collar workers may actually be *more* likely to be automated out of a job than blue-collar workers" (Roose 2021, 28). If the particularity of literary criticism and theory lies in teaching optimization, planning, and method, and if this method is reproducible by an artificial intelligence – technology engaging in *technē* – then what is the point of literature departments and the humanities at large? What is the future of what we call 'theory'? Here is where Arendt and the notion of thinking as fostering internal dialogue and moral considerations become relevant.

Already Socrates taught us that thinking and writing are not the same thing. This is an important distinction that can allow us to understand that, while ChatGPT can produce high-quality text, there is no critical thought behind its generative processes, but merely likelihood and statistics. Input from GenAIs can surely be helpful to structure text and get us started on the writing process, as Anderson emphasized. But it should not be mistaken for thinking. Those who rely on writing, Socrates says in the *Phaedrus*, will no longer call "things to remembrance […] from within themselves, but by means of external marks" (275a). Sharing information with persons and students in this sense, without teaching them, "will make them seem to know much, while for the most part they know nothing, and as men filled, not with wisdom, but with the conceit of wisdom, they will be a burden to their fellows" (275b). He then goes on to compare writing with painting: "the painter's products stand before us as though they were alive, but if you question them, they maintain a most majestic silence. It is the same with written words; they seem to talk to you as though they were intelligent, but if you ask them anything about what they say, from a desire to be instructed, they go on telling you just the same thing forever" (275d). This latter description of written texts sounds uncannily similar to how users (and some of my former Stanford students whom I asked to experiment with ChatGPT for their final papers) describe their interactions with large language models. No thought comes out of them, no rationale, but an endless repetition of the same (which some users anthropomorphize as 'obsessiveness').

At the end of the day, generative AI is a tool, one that is much better at 'making text' than we are just as calculators are much better at calculating than we are. But it does not think and, as Chomsky and his colleagues write, it reasons in a fundamentally different way: "The crux of machine learning is description and prediction; it does not posit any causal mechanisms or physical laws. Of course, any human-style explanation is not necessarily correct; we are

fallible. But this is part of what it means to think: To be right, it must be possible to be wrong" (Chomsky et al., 2023). To ventriloquize Socrates's words, GenAI seems "to talk to you as though [it] were intelligent, but if you ask [it] anything about what [it] say[s], from a desire to be instructed, [it] go[es] on telling you just the same thing forever."

Of course, we have Plato to thank for having disagreed on this issue with Socrates and then preserved his teacher's thoughts on paper for posterity. But an important point is made here about thinking and knowledge acquisition, and we can find in Socrates's words a suggestion about what exactly the role of the academy, i.e., of the university should be today. Hannah Arendt takes Socrates as the prime example of what thinking is about: "The need to think," she writes, "can be satisfied only through thinking" (Arendt 1971, 422). Relying on Kant's distinction between *Vernunft* (reason) and *Verstand* (intellect), she posits that reason, or thinking, requires constant practice. It requires an ongoing internal dialogue whereby a temporary separation from society and doing is necessary. Since Arendt was famously a political thinker, it is important to note that this thinking process should not be an end goal in itself, but has to be followed by a confrontation with other people's thoughts within the public domain so that judgment can be practiced (Rodowick 2021). Knowledge (or the domain of intellect) involves the acquisition and retention of facts. However, without thinking (or the domain of reason), knowledge can become rigid and ideological. This is how a society can have knowledgeable engineers, knowledgeable coders, and knowledgeable scholars who, nevertheless, all have the potential to become versions of Eichmann.

The good news is that thinking is not only innate to all human beings, but it can be actively fostered and practiced: "Thinking [...] is not a prerogative of the few but an everpresent faculty of everybody" just as "the inability to think is not the 'prerogative' of those many who lack brain power but the everpresent possibility for everybody – scientists, scholars and other specialists in mental enterprises not excluded – to shun that intercourse with oneself whose possibility and importance Socrates first discovered" (Arendt 1971, 445). This gives the humanities, liberal arts education, and the study of literature enormous leeway as to their role in fostering critical thinking in the age of GenAI and perhaps even co-determining how interactions with GPTs can be used in productive ways. One such noteworthy achievement is the Socratic Artificial Mind (SAM) developed by a group of researchers at the University of

Catania in Sicily who want to explore "the 'comprehension' abilities of GPT-3 and [its] capability to generate more natural and diverse responses, including asking questions, making statements, or provide counterexamples" (Caneva et al. 2023, 414). Now if the Socratic dialogue can provide meaningful ways of interaction with generative AIs, why should we not envision similar models stemming from the literary imaginary? Can literature and literary study become a locus for the cultivation of critical thought and the practice of moral considerations rather than the practice of form and writing which are now in the hands of generative AIs?

To achieve this, a move away from form and method is necessary in addition to a transition from "paranoid theory" (Guillory 2022, 86) towards the practice of moral judgment through literature. Guillory rightly notes "the difficulty of *teaching* judgment" but he agrees that "a reassertion of judgment as the legitimate practice of all readers of literature" is necessary (Ibid.: 377). This is not to say that students and readers should be taught to distinguish between "good" and "bad" novels or poems. Rather, the type of moral judgment that I have in mind is a process or *praxis*, something that we cannot learn from machines nor they from us because it involves the judgment of human nature as "a constantly moving historical target" (Ibid.: 381). The latter quote happens to be Guillory's definition of "literature." For, to be human is to "constantly work out a historically shifting and culturally varying account of what exactly it takes to be the kind of minded animal we are," philosopher Markus Gabriel writes (Gabriel 2015, 2). And this work, I suggest, alongside the internal dialogue necessary to engage in a robust process of thinking, can best be cultivated through the study of literature because, as the ancient Greeks knew well, Guillory reminds us, "action capable of being judged as right or wrong has the status of a manifest *content* in most narrative writing" (Guillory 2022, 356). It is in literature that complex, diverging perspectives, not bound by a single system of thought, as well as human fears, anxieties, hopes, experiences, and dreams are sedimented. It is only through literature that we can talk about ourselves, other people, and their actions without really talking about ourselves, other people, and their actions. If anything, literary works are the greatest archive of internal dialogues put to paper, as Bakhtin recognized when he wrote about the dialogic imagination.

Furthermore, thinking for Arendt (via Socrates and Plato) is first and foremost a "soundless dialogue (*eme emautô*) between me and myself, the two-in-

one" that is 'I' (Arendt 1971, 442).[7] In the stranger's words in Plato's *Sophist*, "thinking is, precisely, the inward dialogue carried on by the mind with itself without spoken sound" (263e). Let us set aside the 'soundless' part in the definition of thinking as internal dialogue and think about the split within the self that appears to be a precondition for thinking. Arendt names Spinoza when she writes that "every determination [...] is negation" (Ibid.: 440) and we can also think of Hegel's dialectic in this sense. Freud calls each part in this split of the self respectively the ego and the superego. But Arendt has something less daunting, more mundane than this in mind. The thinking partner of the 'I' is "another fellow" who waits for us at home at the end of the day. This fellow shuns the marketplace and, as the epigraph to this paper anticipates, only appears in solitude. One contemporary implication of this definition of thinking as the 'I' confronting itself in the privacy of our solitude is that thinking cannot take place on platforms such as Twitter or Facebook. It can ultimately also not take place between myself and any other technology, no matter how much I consider it a prosthetic extension of myself. Thinking is absolutely private but, in this absolute privacy, we either encounter our best friend or our biggest enemy: ourselves. Therefore, as Arendt writes, "if you want to think you must see to it that the two who carry on the thinking dialogue be in good shape, that the partners be friends" (Ibid.: 442). Thinking should be this amicable encounter of the self with the self, the true locus of the three Delphic maxims which for millennia have constituted the pillars of autonomy and the good life: know thyself (*gnothi seauton*), nothing in excess (*mēden agan*), and certainty brings ruin (*engya para d'atē*).

The question of how we can live with ourselves is one of the most interesting philosophical, literary, and psychological questions ever asked. Here too literature has been that realm of human creativity where this inner conflict has been best explored throughout the centuries – that fear of living with ourselves and how it impacts our living with others, and that endless dialogue that arises in our minds when we are in conflict with ourselves (Hamlet's "to be or not to be" being perhaps the most famous exemplar of this struggle ever put down on paper). Our modern go-to solution to this problem has been *not* to make friends with that "other fellow" but to silence him or replace him

7 German philosopher Markus Gabriel proposes, via Arendt, a notion of thinking as a "sixth sense" (Gabriel 2020).

with more external stimuli. We seek to silence the internal dialogue by taking recourse to distraction and entertainment, whether in couple life, friendship, videogames, medication, alcohol, the gym, social media, or shopping sprees. With this, we are back to Freud's "deflections," "intoxicating substances," and "substitutive satisfactions." But what does this have to do with GenAI? Or rather: What happens when, after having spent the day on the physical or virtual marketplace, we go home and instead of finding ourselves in solitude we find a machine who takes over our mental labor? Are we happier? Do we have more time for leisure? But isn't leisure the absence of thought anyways? If we spend our entire days not thinking but working, and if we spend our entire time off not thinking but chatting/posting/viewing, then when is that internal dialogue supposed to develop and take place? In bed, at night, as we start drifting away into unconsciousness and the realm of dreams? Or in the aftermath of our actions? Or, as in Eichmann's case, *never*?

5. *Thinking and the Case of Frankenstein*

Western literature of the nineteenth and early-twentieth centuries was fascinated with the problem of the split self and the double (*Doppelgänger*). We only need to think of E.T.A. Hoffmann's *The Sandman* (1816), Shelley's *Frankenstein* (1818), Stevenson's *Strange Case of Dr. Jekyll and Mr. Hyde* (1886), and Oscar Wilde's *The Picture of Dorian Gray* (1890) all the way to Luigi Pirandello's *The Late Mattia Pascal* (1903), and Fernando Pessoa's heteronyms. Is it a coincidence that this interest culminated precisely during the two industrial revolutions when, incidentally, Hegel outlined his dialectics of consciousness in the *Phenomenology of Spirit* (1807) and Freud discovered the unconscious? In other words, why do the split self and the *Doppelgänger* become a topic of such fascination for European writers and audiences just as technology was exploding onto the European continent? I believe that the theme of the fragmentation of the self (Taylor 2001; Ricoeur 1990) in literature at this time did not only occur because of a destabilization of values and hierarchies in the aftermath of the French Revolution and the drastic socio-economic changes brought about by capitalism in the nineteenth century. It was also a response to the overwhelming mental stimuli that came with the proliferation of machines, noises, images, and information (Danius 2002; Crary 1992 and 1999).

More concretely, at least in the case of the industrialized part of the European continent at the time, humans were promised the freedom of social, economic, and physical movement. But this also brought about the capitalist multiplication of 'choices,' the amassment of products, venues, devices, and humans to choose from, leading to an increased difficulty in making choices that, in previous times, were relegated to religious or state authorities. We could say that the nineteenth century witnessed the birth of stress. As Jaynes writes, "decision-making [...] is precisely what stress is" (Jaynes 1976, 94) and what better way is there to avoid stress and decision-making than through distraction and non-thinking? A fresh reading of Shelley's *Frankenstein*, for instance, can offer some answers.

From a Faustian warning about the excesses of knowledge to an allegory for race relations, queerness, and the dangers of technology, Shelley's book is now a staple of the global imaginary, the book that gave us the archetype of the mad scientist obsessed with knowledge and glory, and the invention that turns against him. Since the publication of the novel, the creature has taken on the name of its creator, Victor Frankenstein, so that when we speak of Frankenstein today, we rarely mean the scientist but his creation. I would like to pause here immediately and draw attention to this productive case of metonymy. Shelley never named the creature. This allows readers to confuse creator with creature throughout the narrative, making us wonder at the end of the book as to who really is the monster of the story and who the human. I propose that this ambiguity is fundamental to understanding what I see as the core problem of *Frankenstein:* the shutdown of internal dialogue. Victor Frankenstein has *Verstand* (cognition, intellect, knowledge) but he in no way engages in *Vernunft* (reason, thinking). His lack of moral considerations when he decides to make the creature and his subsequent rejection of his creation are a direct consequence of his inability to, first, double himself and, second, face that double, his inability to first engender and then have that internal dialogue between I-and-I, to make friends with himself. How did this come about? There are two seemingly contradictory reasons: a lack of confrontation with peers and a lack of solitude.

Already in the beginning of the novel we see that Victor was a highly educated person brought up in a well-off, loving family. He lacked neither companionship nor education nor entertainment nor opportunities to travel. All the preconditions for a sophisticated, friendly, internal dialogue are thus in

place. And yet, Victor rarely thinks. Rather, he is invested in knowing, i.e., in the accumulation of knowledge that can be systematized, stored for later use, verified, and employed in the creation of a product. He first reads Paracelsus, Albertus Magnus, and Cornelius Agrippa, and later transitions to the study of chemistry and modern natural philosophy. As a child and a young man, Victor's "hours were fully employed in acquiring and maintaining a knowledge of [...] various literature" (Shelley 2011, 24). Indeed, even after his mother passes away, we do not encounter the verb "to think" nor instances of his thinking in the narrative until after the creation of the monster. However, it is not just relentless knowledge acquisition that is detrimental to Victor's internal dialogue. In a first moment, the desire for knowledge is his passion and later serves as a coping mechanism for his mother's death. What leads to Victor's suspension of thought and eventual moral degradation is that, even as an adult, he amasses knowledge *without confrontation and discussion*. He does not submit his ideas to public judgment (or what we call peer-review).

Upon his arrival at Ingolstadt, he makes the following observation: "I was now alone. In the university, whither I was going, I must form my own friends, and be my own protector. But I believed myself totally unfitted for the company of strangers. Such were my reflections as I commenced my journey; [...] I ardently desired the acquisition of knowledge" (ibid.). Consequently, he sets up his "workshop of filthy creation [...] in a solitary chamber, or rather cell, at the top of the house, and separated from all the other apartments by a gallery and staircase" (Ibid.: 34). Even Kant, the famous "Königsberg clock" who loved his daily, solitary walks, knew that society is healthy for the mind and discouraged thinkers from even eating alone. In *Anthropology*, he writes: "Eating alone *(solipsismus convictorii*) is unhealthy for a scholar who philosophizes; it is not restauration but exhaustion [...]. The *savouring* human being who weakens himself in thought during his solitary meal gradually loses his sprightliness which, on the other hand, he would have gained if a table companion with alternative ideas had offered stimulation through new material which he himself had not been able to track down" (Kant 2006, 180f., italics in the original).

It is perhaps not far-fetched then to read Shelley's novel as being framed in terms of friendship, both with oneself and with another, or the desire for and lack thereof. Already on the first few pages, we read Walton's letter to his sister where he Socratically laments the shortcomings of writing and articulates his desire for a friend:

I have no friend, Margaret: [...] I shall commit my thoughts to paper, it is true; but that is a poor medium for the communication of feeling. I desire the company of a man [...] to approve or amend my plans. [...] It is true that I have thought more, and that my day dreams are more extended and magnificent; but they want [...] *keeping*; and I greatly need a friend who would have sense enough not to despise me as a romantic, and affection enough for me to endeavour to regulate my mind. (Shelley 2011, 10, italics in the original)

The term "keeping," borrowed from painting, is key here: without a friend's counsel, without peer review, and collective discussion there can be no depth to one's knowledge, no contrast. And, as we saw, contrast or the dialectical confrontation with oneself and with others is generative of human thought and creativity. This need for confrontation is mirrored in the entire structure of *Frankenstein* itself, which is one of communal entrustment, of one character's narrative being nested in that of another: the outer frame of Walton's letters in the hands of his sister Margaret; the inner frame of Victor's oral narrative in the ears of Walton; and within it the third frame of the creature's narrative to Victor and various letters from Victor's family to him. Finally, we as readers hold the stitched-together parts of the novel as a whole in our hands and minds. Timothy Morton has excellently fleshed out the implications of this structure of the novel as a form of "environmentality" (Morton 2016).[8] But Victor neglects his courses at the university, his family and friends, and his own body in order to create a product, a humanoid assemblage, which he then also abandons: "A restless, and almost frantic impulse, urged me forward;" he tells Walton, "I seemed to have lost all soul or sensation but for this one pursuit" as if in a state of "trance" (Shelley 2011, 33). Indeed, at no point in the novel until after the creation of the monster does Victor pause and reflect on his actions or discuss them with others.

Now even according to twenty-first century ideas of intelligence and education, the amassment of knowledge, its translation into practice, and the freedom to perform work without distraction should engender that internal dialogue which Socrates and Arendt considered necessary for thinking. Yet, this does not come about for Victor. Why? The answer is contained in a

8 Environmentality, a concept originally developed by Heidegger, denotes a mode of coexistence that presupposes uncertainty and a mixing of worlds (human, technological, natural), promoting care and cohabitation instead of an existence based on power relations, hierarchies, and orders.

phenomenon widely observed in our present times: Victor, although alone, does not know solitude. Both as a child in the company of his family and, later, when he lives alone in Ingolstadt, Victor continuously substitutes one theoretical framework with another while avoiding any confrontation either with himself or with others: "I could not entirely throw [Agrippa and Albertus Magnus] aside, before some other system should occupy their place in my mind" so that "[m]y dreams," he tells Walton, "were [...] undisturbed by reality" (Ibid.: 23). After his mother passes away and he goes to live alone at the university, rather than mourn, he turns all his efforts towards the goal of "bestowing animation upon lifeless matter" (Ibid.: 32). "The other fellow," a confrontation with whom is necessary for healing, decision-making, and moral behavior, is locked away, like Dorian Gray's portrait in the attic. Victor now lives, as we might call it today, on autopilot. In the aftermath of trauma, he has numbed his mind or "the offending part," in McLuhan's words, in an act of "self-amputation" which again "forbids self-recognition" (McLuhan 1964, 42f.) – in this case, the recognition of a mourning self for the first time confronted with existence outside of the family, by himself, and in a foreign country.

The creation of the monster, in addition to simulating Victor's taking control over the creation of life to compensate for the lack of control over his mother's death, serves the purpose of a neurological extension of his mind. The act of creation is a "strategy of equilibrium," to quote McLuhan again (Ibid.), because Victor's mind could not serve as a protective buffer against trauma and his helplessness. Unlike Dante's Ulysses and Goethe's Faust, Victor Frankenstein's tragic flaw does not lie so much in the fact that he "dares" to know too much, meddling with God's secrets, but in the fact that he 1) does not think, and 2) is an epistemological capitalist, thirsty for cognition and the amassment of knowledge that will usher into a tangible, useful, marketable, final product. Victor is the archetype of the modern technologist who engages in knowledge extraction so that he can order it "to stand by, to be immediately on hand, indeed to stand there just so that it may be on call for a further ordering" (Heidegger 1954, 322). This is why he imagines the creature as his progeny, a future "thankful" species at his beck-and-call because Victor exemplifies that "combination of narcissism with helplessness," of which Martha Nussbaum has written, "a helplessness that is resented and repudiated" and which makes one want to subordinate another to one's own

needs (Nussbaum 2013, 172). But, as we know, the creature will have none of it. Rather, he develops his own personality and desires, and throws back at Victor the ugly picture of Victor's own moral degradation.

As soon as the creature then opens his eyes, Victor immediately projects onto him all the shortcomings of his own behavior up to that fateful night. "Disgust fill[s his] heart," and he is "unable to compose [his] mind to sleep [...] endeavouring a few moments of forgetfulness" (Shelley 2011, 36). Victor can feel nothing but disgust for the creature that he had so carefully assembled and brought to life. Nussbaum's analysis of the emotion of disgust clarifies this reaction of the creator. What elicits disgust, according to Nussbaum, are "reminders of helplessness" whereby a subordinate being, in this case the monster, "is stigmatized [...] in order to serve the inner need of the dominant group for a surrogate for its own animality" (Nussbaum 2013, 183f.). It is not surprising, then, that Victor constantly refers to the creature as a "monster," "beast," "wretch," "thing," "ogre," and "daemon" even before it commits any crimes. Rather than turning the creature into that "other fellow" who is putting him to task, i.e., into a friend, Victor – full of disgust before the fruits of his own thoughts and labor – turns him into an enemy. "What is the cause of all this?" his friend Clerval asks, upon finding Victor half-dead and scared out of his mind. "Do not ask me," he replies, "*he* can tell – Oh save me! Save me!" (Shelley 2011, 39, italics in the original).

What is it exactly that the newly-come-to-life creature can *tell* about Victor? What exactly is Victor so terrified of than hearing his own story told out loud? Like the monster, ChatGPT and other GenAIs similarly throw back at us our ugliest prejudices and narcissistic helplessness. They quite literally contain within themselves our own stories (Frankenstein's monster, let us remember, reads Milton's *Paradise Lost* and Goethe's *Werther* to learn the ways of humans). After all, who educates Victor's creation and who educates our machines if not we, their creators through our own stories? Whose prejudices and animality do they cast back at us if not our own, all the knowledge but also hate and bias that is in our archives and our minds? When we ask to be saved from technology and its "threats," who are we asking to be saved from if not from ourselves?

Without a doubt we are dealing here with some of the most insidious problems of Western society today that have grave epistemological and ethical consequences: 1) our inability to be *together* with ourselves, whence fol-

low 2) the confusion of social distancing with solitude, 2) the confusion of information with knowledge, and finally 3) the confusion of knowledge with thinking. Norwegian philosopher Lars Svendsen frames the question of solitude and technology as follows, in what most of us would find counter-intuitive: "the main problem we are facing today is not that loneliness is on the rise, but rather that solitude is too scarce" (Svendsen 2017, 17). According to Svendsen, rather than suffering from loneliness, we suffer from "hypersociality" (Ibid.: 106) due to the omnipresence of communication devices and social media in our lives so that, even when we return home in the evening after work or school, and even if we live alone, we are online, perpetually entangled in other people's stories, words, and images. Rather than building a robust internal dialogue, we engage in infinite chat. This is an interesting twist on Sherry Turkle's idea according to which, in the age of mass communication, we are "alone together" (Turkle 2011).

Svendsen defines solitude as an "indefinite openness to a variety of experiences, thoughts and emotions" (Svendsen 2017, 108) which, for Ralph Waldo Emerson, should ideally even exclude reading and writing. Furthermore, solitude is the freedom from the gaze of others: "In solitude," Svendsen writes, "I achieve a more direct relationship to myself because it is not mediated by others' gaze. In solitude we escape the experience of being an object for another person" (Ibid.: 123). However, Svendsen omits the obverse – solitude should also involve a freedom from our gaze at others, a freedom from our own voyeurism. Unlike Victor Frankenstein whose work primarily involves the handling of human material, thought's object of reflection should not be immediately present before our eyes and other senses. Hence the need for solitude. This is what Arendt means when she writes that "an object of thought is always a re-presentation" for "thinking always deals with objects that are absent, removed from direct sense perception" (Arendt 1971, 423). Thinking is inward-looking. Technology, in this sense, the omnipresence of screens and distractions, is not the origin of thought but works as a doubling of the same. Svendsen insightfully points out that the very meaning of the word 'distract' is 'to be drawn apart' whereby "[w]hat [we] are drawn apart from is ourselves" (Svendsen 2017, 125). "In loneliness," he continues, "one is *alone* with oneself, whereas in solitude, one is *together* with oneself." (Ibid.: 126, my emphasis).

6. *Conclusion*

To bring our reflections to a conclusion: It is from overreliance on external guidance, including that of GenAIs, that "great evil" can arise and not from the use of technology itself. The "great evil" is this abandonment of ourselves, the relegation of all that we refuse to confront, of all that we dislike about ourselves, of all we want to forget and repress, to our technologies. For, at the end of the day, ChatGPT is yet another eloquent but unseemly (because it also reflects the worst image of ourselves) Frankensteinian creature, another portrait of Dorian Gray in the attic. And what will happen when this portrait takes on a life of its own and begins thinking and acting for itself? It works with everything that we do not want to think about and memorize, everything that we have relegated to online forums, clouds, and platforms. ChatGPT is Socrates's written text that "repeats to us all over again," like Frankenstein's creature to Victor, what we do not want to hear. It spotlights an absence, the absence of our own selves, and continuously points back at our modern inability to be together with ourselves. "Compose yourself!" Clerval admonishes a distraught Victor in the aftermath of the creature's awakening (Shelley 2011, 40). But Victor is incapable of doing just that and instead sets out to destroy his creation which he calls his "enemy," his "own vampire, [his] own spirit let loose from the grave" (Ibid.: 51). But, as Heidegger famously thought with Hölderlin, "where danger is, also grows / the saving power" ["Wo aber Gefahr ist, wächst das Rettende auch"]. If we live in a world of *dis-traction* whereby, as Svendsen writes, we are further "drawn apart from ourselves" by technology and the lack of solitude that it enhances, how can we heed Clerval's admonition and *compose ourselves* so that thinking can flourish?

By bringing together ChatGPT and *Frankenstein* under the umbrella of thinking in this somewhat experimental paper, I hope to have shown that what large language models do is first and foremost to remind us of how absent we are indeed from ourselves and how irrelevant writing and information are for thought. *Frankenstein* is obviously not about ChatGPT. But it is about the thoughtlessness which could ensue both from the lack of a robust internal dialogue and from the lack of meaningful peer-confrontation. For, in all its dangerous implications, ChatGPT points out to the one thing that it cannot, but humans can do: thinking. This is perhaps the nature of its "saving power." GenAIs are also a reminder of one unmarketable truth: that the humanities deal

in thought and thought does not require – although it can work with – technology. But what happens when the I-and-I must face one another in solitude, now *that* takes place in the most sacred of spaces, in the last bastion of privacy and moral considerations: the mind. The humanities classroom is where the mind is then put to task, asked to compose itself in the free, embodied confrontation with others. "Great evil" raises its banal head where these two fundamental human activities (solitary thinking and communal confrontation) are not cultivated and given their space and time. In this sense, ChatGPT isn't the monster, we are; just as the creature isn't the monster in *Frankenstein*, but Victor is. Perhaps the future of thought lies right there, in the communion between the two.[9]

9 I thank Prof. Robert Pogue Harrison (Stanford) for our stimulating conversations on thinking and, recently, Artificial Intelligence. A heartfelt thanks also goes to the students in my winter 2023 course at Stanford, "Deep Tech at the Intersection of Literature, Philosophy, and Society," who allowed me to think with them through some of the questions discussed in this paper and offered many insightful suggestions (Alexander Huang, Bella Raja, Anita Too, Benjamin Perez). This paper was discussed at the research workshop in Literature, Philosophy, and the Arts, Stanford University (18 April 2023), and some of its ideas were presented on two podcasts and at the World Summit AI 2023 in Amsterdam: Harrison, Robert, host. "Humanities in the Age of Artificial Intelligence with Ana Ilievska." *Entitled Opinions*, KZSU Stanford, 12 April 2023, https://entitled-opinions.com/2023/04/12/humanities-in-the-age-of-artificial-intelligence-with-ana-ilievska/; Lyden, Christopher, host. *Radio Open Source*, 5 May 2023, https://radioopensource.org/failing-intelligence/; Ilievska, Ana. "Exploring ChatGPT's Cognitive Abilities: The Prospect of Critical Thinking and the Humanities in the Era of Generative AI." World Summit AI, Amsterdam, 11–12 October 2023, https://www.youtube.com/watch?v=o-zotcaq43Q. Finally, sincere thanks to the two anonymous reviewers of this paper who made excellent reading recommendations and offered constructive criticism which I endeavored to incorporate to the best of my abilities.

Bibliography

Anderson, Meghan Keany. "A Marketing Turning Point." *The GenAI Conference,* San Francisco, February 14, 2023. Conference presentation.

Arendt, Hannah. 1971. "Thinking and Moral Considerations." *Social Research* 38, no. 3: 417-446.

—. 1977. *The Life of the Mind.* New York: Harcourt.

Bostrom, Nick. 2014. *Superintelligence: Paths, Dangers, Strategies.* Oxford: Oxford University Press.

Caneva, Gianfranco, Calcagno, Salvatore, and Giordano, Daniela. 2023. "Socratic Artificial Mind (SAM): Lessons from the Evaluation of a GPT-3 Based Chatbot for Socratic Dialogue." *Sistemi intelligenti* XXXV, no. 2: 413-434.

Cangiano, Mimmo. 2019. *The Wreckage of Philosophy: Carlo Michelstaedter and the Limits of Bourgeois Thought.* Toronto: University of Toronto Press.

Chomsky, Noam, Roberts, Ian, and Watumull, Jeffrey. 2023. "Noam Chomsky: The False Promise of ChatGPT." *The New York Times*, March 8. https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html (accessed: 30/03/2023).

Crary, Jonathan. 1999. *Suspensions of Perception: Attention, Spectacle, and Modern Culture.* Cambridge, MA: MIT Press.

—. 1992. *Techniques of the Observer: On Vision and Modernity in the Nineteenth Century.* Cambridge, MA: MIT Press.

Danius, Sara. 2002. *The Senses of Modernism: Technology, Perception, and Aesthetics.* Ithaca: Cornell University Press.

Farahany, Nita. 2023. *The Battle for Your Brain: Defending the Right to Think Freely in the Age of Neurotechnology.* New York: St. Martin's Press.

Freud, Sigmund. 1989. "Beyond the Pleasure Principle." In *The Freud Reader,* edited by Peter Gay, 594-626. New York: W.W. Norton.

—. 1989[a]. "Civilization and its Discontents." In *The Freud Reader,* edited by Peter Gay, 722-771. New York: W.W. Norton.

Gabriel, Markus. 2020. *The Meaning of Thought*. Translated by Alex Englander and Markus Gabriel. Cambridge, UK: Polity.

Girard, René. 1976. *Deceit, Desire, and the Novel: Self and Other in Literary Structure.* Baltimore and London: The Johns Hopkins University Press.

Guillory, John. 2022. *Professing Criticism: Essays on the Organization of Literary Study.* Chicago: The University of Chicago Press.

Harrison, Thomas. 2017. "Michelstaedter and Existential Authenticity." In *Storia e storiografia di Carlo Michelstaedter*, edited by Valerio Capozzo, 25-46. Jackson, MS: University Press of Mississippi.

Heidegger, Martin. 2008. "The Question Concerning Technology." In *Basic Writings*, edited by David Farrell Krell, 311-341. New York: Harper Perennial.

—. 2004. *Der Begriff der Zeit.* In *Gesamtausgabe*, Band 64, edited by Friedrich-Wilhelm v. Herrmann. Frankfurt am Main: Vittorio Klostermann.

Helmore, Edward. 2023. "The Professor Trying to Protect Our Private Thoughts from Technology." *The Guardian,* March 26, https://www.theguardian.com/science/2023/mar/26/nita-farahany-the-battle-for-your-brain-neurotechnology (accessed: 28/03/2023).

Ilievska, Ana. 2023. "Exploring ChatGPT's Cognitive Abilities: The Prospect of Critical Thinking and the Humanities in the Era of Generative AI." World Summit AI, Amsterdam, October 11-12, https://www.youtube.com/watch?v=0-zotcaq43Q.

Ilievska, Ana, Harrison, Robert Pogue. 2023. "Failing Intelligence." *Radio Open Source*. Christopher Lyden, host. May 5, https://radioopensource.org/failing-intelligence/.

Ilievska, Ana. 2023. "Humanities in the Age of Artificial Intelligence." *Entitled Opinions*. Robert Pogue Harrison, host. KZSU Stanford, April 12, https://entitled-opinions.com/2023/04/12/humanities-in-the-age-of-artificial-intelligence-with-ana-ilievska/.

Jasper. 2023. "Meet Jasper." Last modified March 25, https://www.jasper.ai/.

Jaynes, Julian. 1976. *The Origins of Consciousness in the Break-Down of the Bicameral Mind.* New York and Boston: Houghton Mifflin.

Kant, Immanuel. 2006. *Anthropology from a Pragmatic Point of View*. Translated by Robert B. Loudon and edited by Manfred Kuehn. Cambridge, UK: Cambridge University Press.

Kramnick, Jonathan. 2021. "Criticism and Truth." *Critical Inquiry* 47: 218-240.

McLuhan, Marshall. 1964. *Understanding Media: The Extensions of Man*. Cambridge, MA: MIT Press.

Michelstaedter, Carlo. 1982. *La persuasione e la rettorica*. Milano: Adelphi.

Morton, Timothy. 2016. "*Frankenstein* and Ecocriticism." In *The Cambridge Companion to* Frankenstein, edited by Andrew Smith, 143-157. Cambridge: Cambridge University Press.

Nussbaum, Martha C. 2013. *Political Emotions: Why Love Matters for Justice*. Cambridge, MA: Belknap Press.

Nuttall, Anthony. 2011. *Dead from the Waist Down: Scholars and Scholarship in Literature and the Popular Imagination*. New Haven: Yale University Press.

Ong, Walter J. 1982. *Orality and Literacy: The Technologizing of the Word*. London and New York: Routledge.

Plato. 1961. "Phaedrus." In *The Collected Dialogues,* edited by Edith Hamilton and Huntington Cairns, 475-525. Princeton: Princeton University Press, Bollingen Series LXXI.

—. 1961[a]. "Theaetetus." In *The Collected Dialogues,* edited by Edith Hamilton and Huntington Cairns, 845-919. Princeton: Princeton University Press, Bollingen Series LXXI.

—. 1961[b]. "Sophist." In *The Collected Dialogues,* edited by Edith Hamilton and Huntington Cairns, 957-1017. Princeton: Princeton University Press, Bollingen Series LXXI.

Ricoeur, Paul. 1992. *Oneself as Another*. Chicago: The University of Chicago Press.

Riesman, David, Glazer, Nathan, and Denney, Reuel. 2020. *The Lonely Crowd: A Study of the Changing American Character*. New Haven and London: Yale University Press.

Rodowick. David N. 2021. *An Education in Judgment: Hannah Arendt and the Humanities.* Chicago: The University of Chicago Press.

Roochnik, David. 1996. *Of Art and Wisdom: Plato's Understanding of Techne.* University Park, PA: The Penn State University Press.

Roose, Kevin. 2021. *Futureproof: 9 Rules for Humans in the Age of Automation.* New York: Random House.

Sennett, Richard. 2020. Introduction to *The Lonely Crowd*, by David Riesman, xi-xvii. New Haven and London: Yale University Press.

Shelley, Mary. 2011. *Frankenstein.* Edited by J. Paul Hunter. New York: W.W. Norton.

Svendsen, Lars. 2017. *A Philosophy of Loneliness*. Translated by Kerri Pierce. Chicago: The University of Chicago Press.

Taylor, Charles. 2001. *Sources of the Self: The Making of Modern Identity.* Cambridge, MA: Harvard University Press.

Turkle, Sherry. 2011. *Alone Together: Why We Expect More from Technology and Less From Each Other.* New York: Basic Books.

Vašek, Thomas. 2019. *Schein und Zeit: Martin Heidegger und Carlo Michelstaedter – auf den Spuren einer Enteignung.* Berlin: Matthes & Seitz.

Dr. Ana Ilievska is Senior Research Fellow on the joint project "Desirable Digitalisation: Rethinking AI for Just and Sustainable Futures" between the Universities of Bonn and Cambridge funded by Stiftung Mercator. Before joining the Bonn team, she served as an Andrew W. Mellon Postdoctoral Fellow at the Stanford Humanities Center and Lecturer in the department of French and Italian at Stanford University. She has studied at Yale, Lisbon, Tübingen, and in Sicily, and holds a PhD from the University of Chicago (2020). A comparatist specializing in Italian, Lusophone, and Balkan Studies, her research and teaching focus on the relationship between literature, philosophy, and technology from a Southern European and Mediterranean perspective. Dr. Ilievska has published articles on Luigi Pirandello, Eça de Queirós, Fernando Pessoa, and Luso-African writers, in addition to numerous translations, reviews, and public scholarship on modernism, poetry, noise, the public humanities, and the ethos of Silicon Valley.