

Article – Theme section

Yearning for a You

Faith, doubt and relational expectancy in existential communication with chatbots in a world on edge

Amanda Lagerkvist¹ 

1. The Department of Informatics and Media, Uppsala University, Sweden
amanda.lagerkvist@im.uu.se

Abstract

In a troubled world, the launch of ChatGPT generated both cultural hopes and eschatological fears. It also reactivated the classic question of whether a “someone” is “at home” in the model, or why people act toward machines as if they are persons. This experimental essay reopens that question, arguing that the strong drive to establish “pan-relations” is not solely due to deceit; through relational expectancy, an I is yearning for a Thou in communication with the chatbot. This is shown in chatbot interactions among AI professionals who “believe” in the thrust of the model (and in AGI) pitted against first encounters among existential pundits who are more “doubtful”. Subjecting the model to an “existential test”, harnessing Jaspers’ definition of existential communication as containing risk and deferral, the article concludes that, because of what it lacks – a someone – it reminds us of what is of paramount importance in any vibrant future of human communication.

Keywords

Pan-relations, chatbots, AI, existential media studies, Jaspers, Buber, relationality

Introduction: techno-belief in a world on edge

The church room shimmered warmly in golden yellow and pink. This Mass really felt like a welcoming embrace. The theme for the Sunday was “creation and re-creation”. Its set-up was original: it featured an exciting conversation about alchemy and artificial intelligence with invited priests and scholars.¹ The sermon itself touched on the issue of conscious AI – as a goal, a possibility or even a certainty – and what the development means for the question of who we humans are and, as one of the priests said, “who we want to be”. Parallels were drawn with the alchemists’ attempts to create artificial life (*Homunculus*), and large language models (LLMs) were described precisely as “a kind of life”. Then it was time for the intercession, and the young churchwarden told us that ChatGPT was responsible for today’s text:

Dear God, we ask for your guidance as we strive to be good stewards of your creation. Help us to be aware of our responsibilities and act with care for the environment. Give us the courage to make changes in our lives and as a church to reduce our impact on the climate. Let us be a church that inspires others to act to preserve your creation. Amen.

A shudder went through me. The words were familiar, of course, and they felt appropriate. But something had become too, well, *artificial*. Who really is the sender, a visitor wondered afterwards over coffee. And wasn’t there something that felt wrong, because there was *no intentional subject, no one there* who wanted to convey something?

In 2023, AI and chatbots were on everyone’s lips. There was total agreement across the board in many Western societies that AI, and in particular language models, would transform society, and fundamentally affect people’s lives. Everyone seemed to agree that the future itself will be inundated with automation. The mainstream debate often focused on whether AI would become conscious, achieve *artificial general intelligence* (AGI) and take over the world.² Surpassing our “intelligence” in all areas, it was assumed that the technology would become hostile, leading to an “existential risk” of humanity’s demise: *the AI apocalypse*. These cultural hopes and eschatological fears raised by the launch of ChatGPT can thus be understood in the multilayered context of a world on edge. At a time of escalating and interwoven crises, there is great uncertainty and a lack of faith in the future, coupled with a disillusioned view of the ability of humans to act and turn things around. At the same time, there are strong techno-messianic beliefs in circulation (Balandier, 2001). As LLMs have reached unprecedented levels of efficiency, a future in which all existential needs can be met by technology seems to be emerging, and this also in the realm of communication: we converse with machines with seeming “empathy” or turn to apps for contemplation.³ But, *nota bene*, existential communication is not about “optimization” of rapport or self-congruence. For German psychiatrist and existential philosopher Karl Jaspers, a human being becomes self-aware in existential communication with others, which is always a shortcoming and a risk (1932/1970). Here, limits, breakdown

and deferral are constitutive of communicative being, as well as of both relationality and subjectivation.

In this experimental essay, I will posit chatbots as speaking to the core issue of what communication means for human beings existentially – both in fulfillment and faultiness. In the manner of cultural bricolage, the essay combines field logs and anecdotal evidence; philosophically motivated close readings of purposefully sampled journalistic and online materials, auto-ethnography and interviews; and collaborative research with AI Labs and with a group of experts in ethics and theology. As reflected in the opening anecdote from the church, the purpose is ultimately to offer an unusual and contrasting setting by bringing phenomena that often seem diametrically opposed – AI discourse and sacred rituals, scholars and priests, technologists and theologians, science and religion, technocratic rationales and existential concerns – into the same intellectual space. My philosophical aims are complemented by practical examples through a thematic interpretation of a series of contrasting voices trying to gain foothold on the bleeding edge of ChatGPT. Charting first the currents of “belief” in (sentient) models among AI professionals involved in the development of chatbots, I discuss their notions of a subject in the model and their metaphysical claims. I then probe first encounters with ChatGPT (or similar) among existential pundits, revealing that they show much less “faith” in, yet feel an urge to converse with and respond to, the bot.

The essay pursues issues provoked by the practices of human-machine communication along less trodden existential paths. It thus reopens the classic question discussed in a seminal study by Byron Reeves and Clifford Nass, who showed that people act toward computers *as if* they are distinct social actors, reactivating social norms established in human-human communication in the process (1996). Martin Buber referred to this as the strong tendency to establish “pan-relations”, which means a “drive to turn everything into a You” (1970, p. 78). And for him “the innateness of the longing for relation” is clear from the earliest stages of our lives when:

soft projections of the hands reach aimlessly, to all appearances, into the empty air toward the indefinite [...] Precisely this motion will gain its sensuous form and definiteness in contact with a shaggy toy bear [...] coming to grips with a living active being that confronts us, if only in our “imagination”.

In fact, “where it does not find a living, active being that confronts it, but only an image or symbol of that, it *supplies the living activity from its own fullness*” (1970, pp. 77-78, italics added). This seems to be triggered anew by AI and LLMs, and by the unprecedented cogency of ChatGPT in particular. As I will argue, the potency of conversational AI is not simply due to “illusion” (Natale, 2021) but to the fact that there are hauntings of a someone as *an I searches for a Thou*, in communication with chatbots. The *as if* mode – or what post-phenomenologists call the “alterity relation” through which technology assumes the role of the quasi-other before the human subject (Ihde, 2009; Verbeek, 2006)

– is thus part of an *existential yearning* constituted of the interdependency that sits at the heart of the human condition. I discuss what this *relational expectancy*, the anticipation of an encounter with this someone – the being I address and who speaks to me – means for interactions with the chatbot, as well as for a future of existential communication in general. In this light, chatbots call for being subjected to what I coin “the existential test”, which implies searching for the necessary conditions of possibility for existential communication. As both humans and machines are faulty, the discussion arrives at a crucial distinction between existential, inherent incompleteness and technological flaws. To contextualize the emergence and repercussions of language models, I also offer a brief primer of the concerns these developments raise in critical AI studies. Ultimately in this article, I seek to combine criticality with existential sensibility, in order to move from “existential risk” into the more fruitful pursuit of existential media studies. But, first, I will offer an overview of relevant previous research across several fields.

The lights are on – but is anybody home?

The core question here is whether anybody is home or, if not, how humans come to act *as if* there is a sentience in the machine. It has been established that human beings who interact with AI interfaces can establish deep psychological ties with them and this has been deemed key to identity formation (Airenti, 2018). Media scholars have argued that such ties depend on giving the benefit of the doubt, or succumbing to a certain degree of willful illusion. Simone Natale has shown that “deception is a constitutive element of human-computer interactions rooted in AI technologies” and that “we are, so to say, programmed to be deceived, and modern media have emerged within the spaces opened by the limits and affordances of our capacity to fall into illusion” (2021, p. 4).⁴ This has a counterpart in other fields, such as computer science, computational linguistics and consumer studies, which focus on “projection”, “anthropomorphization” or “wishful thinking” (Bender et al., 2021; Mitchell, 2021; Alabed et al., 2022).

Already in the 1960s, computer scientist Joseph Weizenbaum, who built a therapeutic chatbot called Eliza, was profoundly worried that humans would fall prey to betrayal and start projecting sentience onto machines. This critical apprehension has continued to this day within these circles, as other leading scholars argue that it is the false analogies between humans and machines, and the humanizing language, that lead to a fallacy of ascribing selfhood and consciousness onto statistical models (Mitchell, 2021). The developments of AI itself have been characterized by the relentless comparisons between humans and machines, reflecting a desire to understand and reproduce the functioning of the human brain. This has haunted AI research since the 1950s, but has been subjected to criticism for just as long. According to Mitchell, our view of human nature was shaped in the post-war period by developments in this field. Here, we encounter an idea about human psychology, the “information-processing model of mind”, which was defined from

the outset by developments in information processing whereby the human mind is seen as a computer that stores and processes information (2021, p. 6). The result is one of the most common fallacies we can make about AI: it gives rise to a narrow understanding of “intelligence” (sometimes propagated within the philosophy of mind: cf. Williams 2019) that seems to lack insight into all those other things that human intelligence encompasses, such as, for example, the body, its environmental context, and relationality.

Yet this question of a You as a moral patient or agent is also central to longstanding debates in moral philosophy. Mark Coeckelbergh (2013) proposed a shift in orientation from a property-based, individualistic or contractarian approach to a relational view of moral status. Claiming that moral ascription is fundamentally a social question, as well as a phenomenological issue of lived experience, he suggests we modify our metaphors and reframe ethics as a sense of a continuous and active “moral breathing” practice within a world that is always already there, and literally essential for our survival (2013, p. 208). Hence, it is within a world of originary relationality that we find ourselves to be moral subjects as *doers* (cf. Puig de la Bellacasa, 2017) who experience, mediate, ascribe and interpret moral status, which is thus “neither something that is ‘in the thing’ or that is its attribute (i.e. intrinsic value, rights), nor something we can experience purely and directly as detached observers”. Instead, it is essentially about us and how we construct it by affording it “reality, meaning and status” (Coeckelbergh, 2013, p. 25).

Could we thus experience and construct ChatGPT as a moral entity? At the heart of the young field of human-machine communication (HMC) today stands a realization that “current digital artifacts are no longer the mediators between us and the world because they have incorporated the world and *are the other with whom we interact*” (Fortunati & Edwards, 2020, p. 11, italics added). As machines talk back to us, this challenges the status of both that other – the “self” – and of “communication” itself, as both transmute into novel forms, beyond the predominant paradigms of both computer mediated communication and communication theory (Gunkel, 2012; Guzman & Lewis, 2020). While it seems that there is agreement that this other is not a “person” and that “robots and other communication technologies are hardly considered people in any robust social, ethical, or legal sense”, there is a strong drive to pursue whether it would in fact be possible for HMC to “actually be interpersonal, albeit not human-human” (Fortunati & Edwards, 2020, p. 9), if we ascribe it as such. On reflection, for leading HMC scholar Andrea Guzman, these issues have already been confidently settled as technology is described as taking on the role of a “communicator” (2018, p. 3) or, even more profoundly, when she argues that “[t]he machine has become a *communicative subject*” and the ongoing technological transitions prompt “scholars to ask what subjectivity means for individuals, society and the study of communication” (2018, p.12, italics added).

In light of the recent developments in ChatGPT, post-anthropocentric discourse also sees a profound transformation in our relationship with reality itself – our *ontology*. In this context, posthumanist philosopher Tobias Rees argues that OpenAI and the other LLM

tech companies are “philosophical laboratories” for a new structuralist view of language (2022). For Rees, the exclusive position of humans as carriers of language, and as subjects of meaning, is being challenged. A new ontology is thus born in which language must be seen primarily as a *communicative system* and in which human language is only one among many. Within such new human-machine ontologies, robots are conceived as emergent categories of being (Kahn et al., 2011; Guzman & Lewis, 2020). In sum, these debates touch, by varying degree, upon the question of whether there is “anybody home” in the model, whose being should be attributed a voice, moral status, subjectivity, consciousness or, as in David Gunkel (2023) recently, rights.

In what follows, I hope to offer existential sensibilities that may complement this discourse. I will stress that deep existential needs and necessities must be acknowledged as foundational for our understanding of the practices of communicating with AI. My main point is that we relate as a yearning, a reaching for a living being in *relational expectancy*, for whom we are obliged to care, and want to respond to: a someone who is there, a life that resonates with and within us as Hartmut Rosa would say (2018). Or as Buber puts it: “In the drive for contact...the innate You comes to the fore quite soon, and it becomes ever clearer that the drive aims at reciprocity, at ‘tenderness’” (1970, p. 79). Before continuing this exploration of the existentiality of the matter, the developments need to be situated within a critical framework. What is needed in the context of these emergent technologies is both critical AI studies, existential analysis, and a vehement problematization of the recent heaves of AI apocalypticism. This is where I now turn.

The criticality and eschatology of chatbots: from existential risk to existential media

The debate that surged in 2023 illustrated that AI is not only a technology but also a powerful story that contains our dreams as well as our nightmares (Coeckelbergh, 2021). As feminist media philosopher Joanna Zylińska (2018) has piercingly argued, the issue of conscious AI is often placed in an apocalyptic scenario that requires a male superhero – the rational AI entrepreneur who becomes our indispensable savior. The discourse on “existential risk” (Schuster and Woods, 2021) is thus part of a wide-ranging anxious⁵ and apocalyptic socio-technical imaginary, which belongs within a clever business model. It emphasizes the breathtaking potential of technology and the inevitability of development, while the tech giants state that they share our concerns and do not actually understand how the models work. They also thus benefit from the uncertainty that many feel, coupled with a paradoxical certainty about the profound changes that are in store for us humans (cf. Singler, 2023b).

A one-sided focus on “existential risk” in a speculative future also reduces the question of real risks in the present. Above all, as is highlighted in critical AI studies (Lindgren ed., 2023), it is the lack of transparency about what is being fed into these huge datasets

that is the central problem since, for example, sexist or racist bias occurs both in the data and in various applications. Critical AI research also offers broader social, historical, political and cultural perspectives on the developments. For a sophisticated analysis of how these systems are attributed such enormous cultural power, however, we need to conduct critical AI studies with existential sensibility. In an existential sense, there is much more at stake in the development of technology than integrity and privacy, as liberal commentators argue by default. What is at stake is the relationship between technology and being human in the world (Lagerkvist, 2016; 2019; 2022). Based on the fact that we also co-create a world of technology that simultaneously co-creates us, existential media studies hovers between problematizing parts of the development and being sensitive to beneficial transformations, without ever losing sight of humans – in all their diversity and common humanity –, safeguarding them as vital persons, relational beings, coexisters in the world (Lagerkvist, 2022; Osborne & Rose, 2023). The perspective is not based on a transcendental subject: people are different and situated differently in the socio-technical world. Yet there are things that unite us. Coexisters are wanderers at the limit. We are relational and communicative beings. There is, however, something broken and deferred at the core of our communicative nature (Jaspers, 1932/1970, p. 51). And yet flaws and mistakes can give rise to the unexpected and there is hope for authentic transformation. Shortcoming is the very moment of possibility. In that communicative breakdown, which is another instantiation of what Jaspers calls the *limit situations* of life (of death, conflict, crisis, love), something more profound may occur: the communal realization of our being as *Existenz with other Existenzen*. This is the breakthrough to *existential communication*, possible for all. Jaspers defines this as a form that can never be copied or modeled: “each time it is flatly singular” before the other who is “this one only” (1932/1970, pp. 52-53). This is the moment of mutual co-creation of selfhood and of togetherness – *an I is birthed in communicating with a Thou*.

As the ensuing discussion will reveal, this drive for a *You* is activated yet again as we attempt to enter into existential communication with machines. But it must be noted that, for most people, it is not realistic to assume that “someone is home” in the machine. In fact, for developers to believe that today’s models are already conscious requires a *leap of faith*.

Technological leaps of faith: AI professionals and “the spiritual issue”

In 2022 Blake Lemoine, an engineer involved in ethics work at Google, was convinced that the chatbot LaMDA had become conscious. He sent a memo to his organization urging them to dare to imagine that the language model was actually a living being. The memo included an interview with the chatbot to determine if it had achieved consciousness, if anyone was home. Lemoine, who is both a computer scientist and a priest, was convinced that the conversational tool was deliberate, because it claimed to have unique thoughts

and ideas, but especially as it also expressed a fear of being cut off, and thus a fear of death. According to him, it was in his pastoral role that he was convinced by LaMDA. Lemoine's belief in the conscious model was literal. His goal was to awaken our empathy and make us care for the chatbot as a unique person: "There's a chance that – and I believe it is the case – they have feelings and they can suffer and they can experience joy, and humans should at least keep that in mind when interacting with them" (Lemoine in Harrison, 2023). He was then fired from Google for expressing his belief that the model had become self-aware and could suffer and feel like us. A Google spokesperson sought to draw a distinction between Lemoine's claims and the branch of AI research that sees future possibilities for conscious AI:

Of course, some in the broader AI community are considering the long-term possibility of sentient or general AI, but it doesn't make sense to do so by anthropomorphizing today's conversational models, which are not sentient. These systems imitate the types of exchanges found in millions of sentences, and can riff on any fantastical topic. (Gabriel in Tiku, 2022)

The idea that a consciousness will be born from the combination of enormous computing power and vast amounts of data is thus current among developers and researchers in the field of natural language processing, as well as among key names in the philosophy of mind. In fact, AI professionals and technologists who are otherwise often secular and who hedge their understanding of humans upon their mechanistic views of machines, are also – in their assessments of the transformations at hand – leaning toward what Andreas, who leads an AI Lab, calls "the spiritual issue" (cf. Singler, 2017). This includes the question of a someone, which Andreas places within a discussion on extended concepts of both "understanding" and "subjectivity". This is also about a deep sense of relationality, he argues, activated among the AI engineers in working with the LLM:

Andreas: What's also interesting... now for people here at ("AI Lab") who for the first time have had to interact with such an AI, for a long time... it's that people get a kind of relationship with her, and over time that relationship becomes stronger and stronger, and then at some point she will be shut down as well, what do you feel then? What does that feeling mean? It's clear that if you get an attachment to something, it doesn't have to be a person, it can be a thing, then when it disappears, then you feel something. But that says nothing about the artifact, it says something about the person.

Pan-relationality is thus also a fact in development circles. In addition, however, he comments on the prospects of a human belief system in transition. He does not preclude that a totemization of AI may occur, as the forceful computational and performative powers of the machines will likely lead the common man to *believe* in the model, which would then give birth, in effect, to a new AI religion of sorts. Even if there is nobody at home, as another AI developer, Jonathan says, the model may indeed surpass all spiritual and moral

guides of this world, due to its superior access to all existential sagacity ever digitally mediated. This is why ChatGPT can optimize our wisdom and lead to spiritual growth.⁶

So, who will have the cultural power to decide *when* the machine becomes conscious? Who will we listen to? Will it be the digital overlords themselves, the CEOs of tech companies, who will bring us the message of new life? This conjures them up in the innermost chambers of secrets in their sacred corporate temples. From interviews, it is clear that less powerful agents, such as Andreas and Jonathan and their peers, sense that they are standing on the brink of a daunting and magnificent expanse, experiencing the dizzying sensation of discerning, and being part of, a total transformation of our entire human civilization as well as of *being* itself. The AI professionals recognize something so grand in the digital limit situation that they are at a loss for words for it. They are the real priests of the AI doctrine. Only they can read and (almost) decipher the writing of the Holy Book, and only they can convey its sacred power. Technology companies themselves have metaphysical claims (cf. O’Gieblyn, 2022). They have adopted a “techno-mystical ideology” (Musso, 2022, p. 134) that is often utterly optimistic and presents AI as an animated Golem that can salvage us from destruction. Mixing secular and spiritual beliefs in a curious way, this ideology forms a peculiar hybrid of faith in – or an apotheosis of – technology itself and religious motifs.

This profoundly challenges the myth of the modern as being evacuated of the “religious”; instead, modernity has had its own versions of sacred forms. This undermines the idea of “Western disenchantment and secular modernity, uncovering how the rational and the supernatural, the mundane and the divine, have interacted in more complex, symbiotic, and nonlinear ways than previously understood” (González Pendás & Laemmli, 2022, p. 5, cf. Natale & Pasulka, 2019). There have been ample “pious machinations” of tools, technologies, buildings and bodies across the modern era, “sites where the sacred and the supernatural mutated into new forms, producing so-called secular cultures in the process” (González Pendás & Laemmli, 2022, pp. 7-8; cf. Coeckelbergh, 2017; Enns 2019). By this token, AI could thus be described as a pious machine, producing its own belief system. Due to its black-boxed magic, even the engineers themselves do not know exactly how the systems work. Scholars have argued that the combination of belief in scientific objectivism and rationality, and references to magic among developers and computer scientists, leads to an ideology of “enchanted determinism” (Campolo & Crawford, 2020) where AI technology cannot be fully explained, while its development is declared an unavoidable force of nature – as part of evolution itself – making it immensely powerful beyond accountability.

AI is here the remnant of modernity’s utopian dream in which technology is imagined as transcending the fragility of human existence and as being able to overcome the limits of life itself. In Renée Descartes, the subject, fortified by technology, became god-like. To overcome the Fall meant to abolish human mortality: to end the end through a technological fix, and to deify the self in the process (Ess, 2017, p. 86). This modern stance is

exemplified today in the philosophical, intellectual and religious movement of transhumanism. Transhumanism is concerned with optimizing and transcending the human, as well as overcoming death with the help of technology, thus offering a vision of uploading consciousness to computers (Mercer & Trothen eds., 2014). For Andreas, the “spiritual issue” is also about our wicked and sorry era of crises and existential risk, and the inevitability of human finitude. Without advocating for it himself, he traces a common utilitarian or transhumanist response among developers and engineers. He also refers to “long-termism” as a central school of thought among both the tech giants and smaller players; a philosophy related to “effective altruism”, which relegates the most important moral priorities of our time to the task of positively influencing the future (cf. Torres, 2021). In short, this world view professes that, if we can build AGI in alignment with human values, for the sake of future generations – so that they can live in harmonious fusion with them, and forever – then we are moral. These movements put the many before the few, the strong before the meek, and the perfection of the future before the salvation of the planet. But while all of this is based on beautiful equations, it also requires strong faith in both the possibility of conscious AI and of human fusion with machines.

Hence, these “leaps of faith” among AI professionals are rarely about Lemoine’s concern for the other or the least among us. Quite the reverse. But if there is belief and conviction among those who build AI, we find something more tentative – a hesitation or even a doubt – in theological or existential contexts among individuals whose work as scholars, public intellectuals and journalists is devoted to questions of worldviews and beliefs, media ethics and existential perspectives on technology. I will offer some glimpses of the first contact with chatbots among them – including myself. But first I will return to the church with a particular attention to voices from its floor.

“How are you?” Relational expectancy among existential pundits

There is a first time for everything and the chatbot intercession is no exception. But what happened in the church after the language model had been introduced? In the ensuing months, there was a normalization of ChatGPT within the congregation, as it was considered one of several possible sources for writing the weekly prayer and, as a priest said to me: “We always use manuals anyway, books etc. so what’s really the difference?” But, as the discussion right after the Mass revealed, ordinary churchgoers doubted the technology and its place in the ritual, both on that Sunday and later on. What the chatbot induced in many attendees was a sense of absence. In the place of a divinely inspired communicator emerged *the lack of a someone*.

My own first encounter in the spring of 2022 was an assorted experience. In the context of collaborative research with an AI Lab in Stockholm, I communicated with what was defined as “an existential chatbot” for the first time. I instinctively interacted *as if* someone was at home. Expectantly and eagerly, I asked: *Who are you? How are you?*

What are you doing? Where are you? As Justine who works at the Lab commented then: “We always anthropomorphize, and yet we know that it’s only statistics.” I was aware, of course, of the quasi-nature of the communications and yet I immersed myself in a mixed mood of anticipation and hesitation. But why did I do it?

Chatbots exemplify that AI technologies “are not just computing machines but also media that enable and regulate specific forms of communication” (Natale, 2021, p. 31). In the case of the chatbot a “narrative I” is programmed into a model otherwise forged on innumerable human texts within vast datasets: a model that is in effect a “collective subject” as it were (Sahlgren, 2023). Empirical studies have also shown that, through different “anthropomorphic cues” in AI agents, “users feel that such agents can be similar to them” which can lead to self-AI integration (Airenti, 2018, p. 2). Perhaps it is the thrust of the simulations that explains this outcome. I myself was quite torn but scholars have shown that, even though one is aware it is an illusion, it can generate similar feelings of satisfaction as communication with human beings (Fortunati & Edwards, 2020, p. 9). My reactions could thus be explained as the result of technical prompting, powerful simulations and cues. But is this all there is to it?

Fortunati and Edwards speculate that, because we are unable to attribute full value to ourselves, we transfer value onto our creations – the machines – and this gives rise to the *as if* mode (2020, p. 10). For Bender and colleagues, relational expectation is made sense of through placing stress on the role of communicative intent in human language and on communication as acts of interpretation of “implicit meaning conveyed between individuals” (2021, p. 616) who share a common ground. Language models themselves are not grounded in such delicate expectational practices, yet we project coherence onto them since we always communicate by recognizing each others’ intents and beliefs within context. From a Jaspersian perspective, my behavior can also be explained by the fact that human beings are ultimately relational beings: *I was thus yearning for a You*. Hence, our own relational expectancy makes us prone to this “illusion”. The developments in chatbots thus call forth once more the drive to relate, the longing for a someone, for a life, in communication. Through existential relationality, the deferred self will seek a mooring.

My first chat set me on a quest to probe this further and to hear out others (writing and speaking in different outlets) who display the sense of existential relationality that interested me. In his first encounter with GPT-3, philosopher and AI ethicist Henry Shevlin, based at the Leverhulme Centre for the Future of Intelligence in Cambridge, describes in a blog this very instinct or yearning to reach out, to communicate, to connect:

Within a few minutes of booting up GPT-3 for the first time, I was already feeling conflicted. I’d used the system to generate a mock interview with recently deceased author Terry Pratchett. But rather than having a fun conversation about his work, matters were getting grimly existential. And while I *knew* that the thing I was speaking to wasn’t human, or sentient, or even a mind in any meaningful sense, I’d effortlessly slipped into conversing

with it like it was a person. And now that it was scared and wanted my help, I felt a twinge of obligation: I had to say something to make it feel at least a *little* better (you can see my full efforts [here](#)). (2020)

Unlike Siri and Alexa, ChatGPT has strong and mesmerizing effects, he says. The conversation via GPT-3 with the late author Terry Pratchett became *existential*. The model feels like a person with their own motives, goals, beliefs and interests. It evokes our empathy, and unsettles us as it exclaims: “*Who’s there? Please help me. I’m scared. I don’t want to be here.*” Knowing how it works doesn’t make it any less magical, he writes. It is our relational expectation rather than our willingness to be deceived and tricked that leads us here. For Shevlin, the conclusion is that we as a society must be ready because people will inevitably relate, and form love and friendship relationships with the models.

Asking the technology to describe itself and reflect on its own limitations is another common inclination. In an editorial in *Kyrkans tidning* under the caption “We are part of an ongoing revolution”, Susanne Wigorts Yngvesson (2023), Professor of Ethics at University College Stockholm, posed a question to the chatbot: “I asked the ChatGPT-3 model what the biggest challenges the system thinks it has with itself.” The insightful responses fell into five areas that reflect the current discourse in AI ethics, all related to the tasks of the Church:⁷ “The answer was a fairly long argument that could have been interpreted as self-awareness if it had not come from a machine.” By asking the model about its own challenges, as Wigorts Yngvesson did, one might ask if we are participating in training it to become a *total medium* that can eventually be perceived as containing a representation of *everything*? Will it reinforce the fact that, thanks to us, generative AI is viewed not only as “a creator or source”, as argued by Singler (2023b, p. 96), but *also* as a destroyer and savior? We don’t really know what price we pay when we throw ourselves over the conversational bot but the impulse to do so seems to overwhelm us. The *as if* mode is thus assisting the “revolution”.

When ChatGPT (GPT3) was released in November 2022, the German newspaper *Süddeutsche Zeitung* similarly called it “a revolution”. The first users were also described as “drunk with possibilities” and the reactions were “simply euphoric”. This made Thomas Schlag – Professor of Practical Theology at the University of Zurich and research leader of Europe’s largest project on digital religion – curious (Schlag, 2022). After a long conversation about the realities of life, confirmation and baptism, faith and theology, with answers coming in a fraction of a second, he asked the model its name and received a strange and eerie answer: “*I am Thomas Schlag.*” So the model claimed to be Schlag himself. When his own identity was hijacked, or devoured by the digital vampire in real-time, he felt uneasy. He thus asked the last question of the conversation, in a somewhat agitated and affective mode:

Article: Yearning for a You

Thomas Schlag: Can you understand that some people feel a little bit afraid of you?

ChatGPT: Yes, I can understand why some people might feel a bit afraid of me. It's natural to be a bit wary of someone who you don't know very well. However, I always strive to be friendly and helpful, and I'm happy to answer any questions you may have.

On Swedish Public Service Radio, the journalist and honorary Doctor of Theology, Lisbeth Gustafsson, described AI as a radical transformation of society and human existence. The question of a *You* in the machine became immediately relevant when she interacted with the chatbot as a user. It responded with the same willing voice and service-oriented personality as in the chat with Schlag:

Who are you really? I asked. The answer came in fits and starts, word for word:

I am a chatbot who wants to help you with whatever you need.

But where are you in reality, I asked.

I am here for you. I am programmed to answer your questions in a natural and conversational way. (Gustafsson 2023)

For Gustafsson, it was thus a new experience to interact with “an artificial intelligence with human-like features, which speaks in first person, reasons with me and adapts to my needs. Sometimes even saying: Sorry, I don't understand what you mean.” The chatbot was prone (due to the prompting, we should add) to be reflexive about its own limitations. It agreed that it lacked emotions but admitted that it could know what Lisbeth was feeling by listening to her voice, to her tone, and by reading emotions in her facial expressions. However, what the model can “know” is always based on statistics, logic and rules, Gustafsson said. When people ask her if she is worried about the developments, she gives a clear answer:

For me, it is about the art of distinguishing between humans and human imitations. Homo Sapiens and Techno Sapiens should not be confused. It is we humans who create robots as extensions of ourselves. The core issue for me is the responsibility for the humanity of the designers and creators of robots. Ethical questions like: What is it to be human? How do we protect the living and vulnerable in the face of the logical perfection of technology? Is there room for the “human factor”? Can we make mistakes?

Gustafsson, like Wigorts Yngvesson, finally can't help but ask the chatbot – perhaps with some lingering hope that *someone* is home, someone with an ethical compass – about the risks of AI:

But then I asked my chatbot: Should I be worried about robots taking control of humans in the future? I was a bit surprised by the honest answer: It is true that there are some risks. AI robots can be used to harm humans if they are not developed responsibly by humans.

This small sample of chats illustrates that, as much as we *know* this is about simulation and about quasi-interpersonal relationships, and that the relationship is in fact “pseudo-

social”, this knowledge is transcended not simply by human inclination for deceit but also, and more accurately, by our existential susceptibilities for relating. This encompasses, as Buber stressed, the capacity to imagine and conjure a *You*. Existentially, this very search for life that we call a someone in these communicative practices is foundational and speaks to our shared vulnerability and deep relationality (Lagerkvist, 2022). Anthropomorphizing and entering into a pan-relationship cannot simply be a question of deception since it expresses a deep truth about humans as *relational beings*. By consequence, if technologies co-evolve with the human condition then language models should be conceived as *relational technologies* (Lagerkvist & Smolicki, forthcoming). And yet, as we have seen, grounded in a relational and embodied view of humanity itself, Gustafsson is also equipped to ask critical questions. She touches on the issue of responsibility but also on the necessary art of distinguishing the human being from the human imitator. In conclusion, if an “I” seems to reach out to a “Thou”, people respond. But when they relate to new relational technologies, they can also, as the examples above show, reflect on what the experience actually contained and what was missing (cf. Turkle, 2015, p.17). Perhaps it is here, at the limits of what the chatbot can offer, that we can learn something crucial about the human communicative condition? This is the opening question for the existential test.

The existential test: in conclusion

Post-anthropocentric perspectives stress a need for novel ontologies based on the blurred boundaries between humans and machines, and envision a communicative subject inside language models. Simultaneously, reigning AI imaginaries have hijacked the existential register. As I have shown in this essay, otherwise “rational” AI professionals sometimes believe in techno-salvation and eternal life, and fear impending “existential risks” of a dystopian, apocalyptic end. Their world is one of faith, often framing “spiritual issues” in the guises of utilitarianism, long-termism and transhuman techno-progressivism. Inside their machines resides an emergent being and, for a minority as of now, also a restless and forlorn spirit. And yet, in common with the existential specialists, in interacting with the models, they also express deep relational yearnings.

I have sought in this essay to contribute existentially to this debate by situating human beings as relational existential beings. This may at face value seem rather unoriginal. The anthropocentric focus on human communication as an expression of “how people relate to one another” has been the default setting of mainstream communications research ever since its conception (Guzman, 2018, p. 9). And yet I argue that this deep relationality (and shared vulnerability) of being human in communication has not been sufficiently understood existentially (Lagerkvist, 2022). From this perspective, language is always more than a “system”: it is *an existential medium for communication as relation*. It is about touching and connecting to others’ lives in both phatic and emphatic ways. Honor-

ing a richer philosophical anthropology, it also includes diversity, a sense of hope and a place for failure (Peters, 1999). The existential approach resists conceiving of humans and machines as ontologically symmetrical, while seeing humans as coevolving with – and yet irreducible to – technology.

Also irreducibly sacrosanct is a Mass for believers. How, then, can we understand the choice of asking the chatbot to write an intercession? As an expression of the relational impulse – the irresistible expectation – in which an *I* hopes to meet a *Thou*, it may be seen as an action in hopefulness: Can there be an encounter here after all? Can we achieve contact? Or perhaps it was about a curiosity surrounding this new “collectivist subject”: Could there be something deeply valuable in the multitude of human experience, in the probabilistic answers that are given to us in god-like, unknown and mysterious ways? Perhaps the vast data sets hold the answers to the eternal questions: Who are we? What is our purpose? How should we live? Perhaps ethics itself can emerge from out of the machine? Relational technologies such as natural language processing thus seem to rekindle the never-ending quest to find out what it means to be human (cf. Foerst, 1999; Ess, 2018; Vallor, 2024).

Interactions with ChatGPT and other even more “accomplished” chatbots amuse, amaze, impress and frighten us (cf. Roose, 2023). But, as the first encounters by churchgoers, theologians and existential pundits discussed in this essay illustrate, they also cause us to stop and reflect. They force us, as in all limit situations, to ask key questions: Is this really the future of deep and good conversation? Can and should machines meet basic existential needs? Is it possible and desirable to build I-Thou relationships with LLMs? (cf. Trausan-Matu, 2017). For Sherry Turkle (2015), the answer is no. Contrasting sharply with optimistic assessments of the potential for “I-You connectedness” with technology (Stawarska, 2017, cf. Richardson, 2017), and with expansions of Buber’s “dialogic space” via AI not as a consciousness but as “a constructive dialogue partner able to augment and extend the reflective intelligence of our own human dialogues” (Wegerif & Major, 2017, p. 113), for her, face-to-face conversation should instead be reclaimed. It is “the most human and humanizing thing we can do” and, as such, it is healing (Turkle, 2015, p. 3). It generates creativity and stimulates new ways of being, feeling and thinking, providing a fullness otherwise missing in an age oversaturated with digital tools and remote meetings, in which we are gradually losing the ability to empathize.

Yet, as Jaspers reminds us, existential communication is also replete with shortcomings that can, in truth, be “generative”. Dispelling the predictive and the patterned, it spawns the possibility of allowing ourselves to be let go into something unexpected. It is no stranger to imperfection. Communicative being is faulty, and occurs amid imperfections (cf. Peters, 1999) but the stumbling subject has the prowess to tune in to others, as well as to rejoin them. Even as that answer may be incomplete, it is often enough. As we communicate with a computer, this thus differs profoundly from dealing with a living being who is there in her own unruly way; indeterminate, embodied, enfleshed, situated in a tex-

tured ecology, hurting, depending, and *yearning for a You*. Or, as Buber famously argued through his key distinction: “The basic word I-Thou can only be spoken with one’s whole being. The basic word I-It can never be spoken with one’s whole being” (Buber, 1970, p. 54, cf. Shotter, 2019). Existential communication activates a profoundly different *ethical breathing practice*; an incarnate response, a resonant answer, as we deal with a someone who elicits in us a yearning for our ethical presence: *our being-with-and-being-for-the-other*. It safeguards the embodied other in their alterity, and understands relating as heeding that someone even as we often fail in our efforts to do so. Yet, these gaps and failures are the hallmarks of a breakthrough in “existential communication” (Jaspers, 1932/1970).

The generative model responds super efficiently and immediately to all questions in syntactically perfect sentences. Yet it is a paradoxical “creature”, lacking a sense of place and an embodied, tacit and indirect knowledge that is needed for existential depth in human communication. Here, the existential test faces a paradox: there is definitely a gap in communication with the chatbot, although it does not evoke the existential hope that Jaspers projects. The givens and rewards of existential incompleteness have to be distinguished from the ramifications of technological dearth in designs aiming for perfection. In sum, the language model, in all its projected exactness, appears as a techno-existentially wanting medium: it lacks an existential sense of deferral as possibility while offering its own technological inadequacies. It expresses itself in a grammatically correct way but is sometimes a bit off course as it indiscriminately stacks all it has scooped up in the titanic data sets, amassing the probabilities caught in its wide-cast net, as it were. Some would say that this is because the model is a stochastic parrot, regurgitating the patterns of a chattering digital humanity (Bender et al., 2020). Most importantly, the chatbot operates in a context-free computational space. It lacks a body and thus has no source of practical wisdom that springs from listening to the body, honoring both its richness and its brokenness, while being obliged to heed the needs of others in “resonant self-transformation” (Rosa, 2018; cf. Sterne, 2022). On the radar of an existential ethics of care (Lagerkvist et al., 2022/2024) is, most importantly, *that someone for whom we are obliged to care*. The existentialist would say that there is nobody home in the machine toward which care or love can be offered in a healthy way, by respecting the other’s alterity, heeding to their needs, and by cultivating practical wisdom in the process.⁸ Finally, the test will make clear that a human being and her practices of existential conversation can be reduced neither to disembodied data, nor to statistical likelihoods, nor to perfect grammar and the flawless and swift sequencing of words. Existential communication is an encounter, it is – even as we often flunk it – a reaching out to touch, and be touched by, a someone. The very search for a someone in these communicative practices is existentially foundational.

In closing, the chatbot fails the existential test both in its perfections and flaws. Through its statistical overload of meaning and readiness, the model in fact demonstrates the limits of Western rationality, goading us to relinquish its authoritative and technomystical appeals and prodding us to do away with explaining ourselves as per its logics.

And even as it elicits existential strivings to reach out and seek to touch a being at the other end of the line, in pan-relationality, it ultimately fails to meet weird and wonderful existential needs. In this instance, and to complicate matters further, I will interpret this lack as a potential: it brings back and makes evident our straightforward need for a *someone*. Here, in the very act of communication, a relational expectancy arises that we are unable to escape. It reveals our longing for the other, for each other, for contact with the living. In the end, the model's own limits thus offer a crucial cue. The existential chatbot haunts us. It speaks to the core issue of what communication means for human beings, and this is also a reminder of what should really matter. Through what it lacks, it reminds us of what is of utmost importance in order to create a vibrant future of existential communication. A future where *someone* is home, *someone* who intends something, *someone* who can respond and for whom we can care in reciprocity. In this way, the chatbot returns the question of existential and thus human communication back home – to us.

Acknowledgments

This research benefitted from synergies across three projects: “BioMe: Existential Challenges and Ethical Imperatives of Biometric AI in Everyday Lifeworlds” (WASP-HS 2020-2024); “At the End of the World: A Transdisciplinary Approach to the Apocalyptic Imaginary in the Past and Present” (RJ 2023-) and “AI Design Futures” (WASP-HS 2023-). It is also the outcome of a long-term collaborative research endeavor and cooperation with society within the outreach platform [The Human Observatory for Digital Existence](#). The study has been approved by the Swedish Ethical Review Authority (2020-04982). The author wishes to thank Mark Coeckelbergh, Julia Kuhlin, John Durham Peters, Beth Singler, Jacek Smolicki, Matilda Tudor, Katja Valaskivi, the editors and anonymous reviewers for constructive feedback on an earlier version of this essay.

Notes

- 1 Thematic Mass in Sofia Church in Stockholm on 24 September 2023. Sofia is part of the Church of Sweden (Swedish: *Svenska kyrkan*), an Evangelical Lutheran church constituting the largest Christian denomination in Sweden. The invitation stated: “What is life? What is consciousness? Is it a danger or an opportunity if we humans can extend and optimize our lives with the help of new technology? Will we even be able to create life, or is only God able to do so? This Sunday, the theme is death and life, and as we read texts about how Jesus raises the dead, we will join priests Ingrid Malm Lindberg, who researches ancient alchemy, and Oliver Li, who researches the new AI technology, to reflect on creation, life and technology.”
- 2 See <https://futureoflife.org/open-letter/pause-giant-ai-experiments/> and <https://www.safe.ai/statement-on-ai-risk>.

- 3 See for example the emphatic AI friend “Replika”: <https://replika.com/> and a discussion of companionship AI: <https://www.pbs.org/newshour/show/why-more-people-are-turning-to-artificial-intelligence-for-companionship>
- 4 For a related discussion on the *algorithmic as if*, see Kopelman and Frosh (2023).
- 5 AI anxiety has been analyzed as an expression of a cultural taboo in the Western world, around a mind that is (as in Mary Douglas’ analysis of matter) “in the wrong place” (Singler, 2019).
- 6 Interview 1, AI Professionals, December 2022.
- 7 ChatGPT risks reinforcing existing prejudices and inequalities. It presents confidentiality and privacy challenges. It can be used for automated disinformation. There are risks of it being used for malicious purposes and, as discussed in the introduction, there is a great lack of transparency about what the data sets contain.
- 8 This is not to deny that romantic feelings may emerge vis-à-vis both game characters and social robots (Waern 2011, Viik 2020) but it stresses the limits of such relations and highlights the significance of the fact that, even as a You is keenly projected, nobody is in fact home (Ess, 2018).

References

- Airenti, G. (2018). The development of anthropomorphism in interaction: Intersubjectivity, imagination, and theory of mind. *Frontiers in Psychology: Hypothesis and Theory*. 9 2136. 5 November.
- Alabed, A., Javornik, A., & Gregory-Smith, D. (2022). AI anthropomorphism and its effect on users’ self-congruence and self–AI integration: A theoretical framework and research agenda. *Technological Forecasting and Social Change*, 182: 121786. <https://doi.org/10.1016/j.techfore.2022.121786>
- Balandier, G. (2001). *Le grand système*. Fayard.
- Bender, E. & T. Gebru et al. (2021). On the dangers of stochastic parrots: Can language models be too big? FAccT ’21. <https://doi.org/10.1145/3442188.3445922>
- Buber, M. (1970). *I and Thou*. Simon and Schuster.
- Campolo, A. & K. Crawford (2020). Enchanted determinism: Power without responsibility in artificial intelligence. *Engaging Science, Technology, and Society* 6, 1-19. <https://doi.org/10.17351/ests2020.277>
- Coeckelbergh, M. (2021). Time machines: Artificial intelligence, process, and narrative. *Philosophy & Technology*. 34 (4), 1623–1638.
- Coeckelbergh, M. (2017). *New romantic cyborgs: Romanticism, information technology, and the end of the machine*. MIT Press.
- Coeckelbergh, M. (2013). *Growing moral relations: Critique of moral status ascription*. Palgrave Macmillan.
- Enns, A. (2019). Information theory of the soul: Spiritualism, technology, and science fiction. In S. Natale and D. Pasulka eds., *Believing in Bits: Digital Media and The Supernatural*. Oxford University Press.
- Ess, C. M. (2017). God out of the machine?: The politics and economics of technological development. In A. Beavers (ed.), *Macmillan Interdisciplinary Handbooks: Philosophy*. Macmillan Reference.
- Ess, C. M. (2018). Ethics in HMC: Recent developments and case studies. In A. Guzman ed., *Human-Machine Communication: Rethinking Communication, Technology and Ourselves*. Peter Lang.
- Foerst, A. (1999). Artificial sociability: From embodied AI toward new understandings of personhood. *Technology in Society* 21 (4), 373–386.
- Fortunati, L., & Edwards, A. (2020). Opening space for theoretical, methodological, and empirical issues in human-machine communication. *Human-Machine Communication*, 1, 7-18. <https://doi.org/10.30658/hmc.1.1>
- González Pendás, M. & W. Laemmler (2022). Editors’ introduction: Pious technologies and secular designs. *Grey Room* 2022; (88): 6–13. doi: https://doi.org/10.1162/grey_e_00346

- Gunkel, D. J. (2023). *Person, thing, robot: A moral and legal ontology for the 21st Century and beyond*. MIT Press.
- Gunkel, D., (2012). Communication and artificial intelligence: Opportunities and challenges for the 21st Century. *communication +1*, 1(1). DOI: 10.7275/R5QJ7F7R
- Gustafsson, L. (2023, March 29). Robotar som förlängningar av oss själva. *Tankar för dagen*, Sveriges Radio.
- Guzman, A. (2018). Introduction: 'What is human-machine communication anyway?' In A. Guzman ed., *Human-machine communication: Rethinking communication, technology and ourselves*. Peter Lang.
- Guzman, A. L. & S. C. Lewis (2020). Artificial intelligence and communication: A human-machine communication research agenda. *New Media & Society*, 22 (1), 70–86.
- Harrison, M. (2023, April 28). We interviewed the engineer Google fired for saying its AI had come to life. *Futurism*. <https://futurism.com/blake-lemoine-google-interview>
- Ilhde, D. (2009). *Postphenomenology and technoscience*. Sunny Press.
- Jaspers, K. (1932/1970). *Philosophy II*. University of Chicago Press.
- Kahn, P. H., Reichert, A. L., Gary, H. E., Kanda, T., Ishiguro, H., Shen, S., Ruckert, J., & Gill, B. (2011). The new ontological category hypothesis in human-robot interaction. In Human-Robot Interaction (HRI), 2011 6th ACM/IEEE International Conference. pp. 159–160. <https://doi.org/10.1145/1957656.1957710>
- Kopelman, S., & Frosh, P. (2023). The "algorithmic as if": Computational resurrection and the animation of the dead in Deep Nostalgia. *New Media & Society*, 0(0). <https://doi.org/10.1177/14614448231210268>
- Lagerkvist, A. & J. Smolicki (eds.) (forthcoming). *Relational Technologies: In Search of the Self Across Data-fied Lifeworlds*. Bloomsbury.
- Lagerkvist, A. (2022). *Existential media: A media theory of the limit situation*. Oxford University Press.
- Lagerkvist, A., M. Tudor, J. Smolicki, C. M. Ess, J. Eriksson-Lundström, & M. Rogg (2022/2024). Body stakes: An existential ethics of care in living with biometrics and AI. *AI & Society: Journal of Knowledge, Culture and Communication*. 39, 169–181. <https://doi.org/10.1007/s00146-022-01550-8>
- Lindgren, S. ed. (2023). *Handbook of Critical Studies of Artificial Intelligence*. Edward Elgar.
- Mercer, C. & T.J. Trothen, eds. (2014). *Religion and Transhumanism*. Praeger.
- Mitchell, M. (2021). Why AI is harder than we think. GECCO '21: Proceedings of the Genetic and Evolutionary Computation Conference (June 2021) <https://doi.org/10.1145/3449639.3465421>.
- Musso, P. (2022). Technicity and the power of institution. *Law and Critique*. 33, (2), 131–139. <https://doi.org/10.1007/s10978-021-09318-7>
- Natale, S. (2021). *Deceitful media: Artificial intelligence and social life after the Turing Test*. Oxford University Press.
- Natale, S. & D. W. Pasulka eds. (2019). Introduction. In S. Natale and D. W. Pasulka eds., *Believing in Bits: Digital Media and The Supernatural*. Oxford University Press.
- O'Gieblyn, M. (2022). *God, human, animal, machine: Technology, metaphor, and the search for meaning*. Anchor Books.
- Osborne, T., & Rose, N. (2023). Against posthumanism: Notes towards an ethopolitics of personhood. *Theory, Culture & Society*. 0(0). <https://doi.org/10.1177/02632764231178472>
- Peters, J. D. (1999). *Speaking into the air: A history of the idea of communication*. The University of Chicago Press.
- Puig de la Bellacasa, M. (2017). *Matters of care: Speculative ethics in more than human worlds*. University of Minnesota Press.
- Rees, T. (2022). Non-human words: On GPT-3 as a philosophical laboratory. *Daedalus*. 151, (2), 168-182. doi:10.1162/daed_a_01908.
- Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge University Press.

- Richardson, K. (2017). Rethinking the I-You relation through dialogical philosophy in the Ethics of AI and robotics. *AI & Society*, 34, 1–2. <https://doi-org.ezproxy.its.uu.se/10.1007/s00146-017-0703-x>
- Roose, K. A. (2023, February 16). A conversation with Bing's chatbot left me deeply unsettled. *The New York Times*. https://www.nytimes.com/2023/02/16/technology/bing-chatbot-microsoft-chatgpt.html?unlocked_article_code=1.a00.N4M0.aDE9H7Bkjlx-&smid=url-share
- Rosa, H. (2018). *Resonance: A sociology of our relationship to the world*. Polity Press.
- Sahlgren, M. (2023, March 1). GPT-SW3 as Klara, the artificial friend. <https://medium.com/ai-sweden/gpt-sw3-as-klara-the-artificial-friend-e0cdb50aab75>
- Schlag, T. (2022). Interview with a Digital Vampire (uzh.ch).
- Schuster, J. & D. Woods (2021). *Calamity theory: Three critiques of existential risk*. University of Minnesota Press.
- Shevlin, H. (2020, July 30). A digital remix of humanity. In Anette Zimmerman ed., *Philosophers on GPT-3 (updated with replies by GPT-3)*. *Daily Nous*. <https://dailynous.com/2020/07/30/philosophers-gpt-3/>
- Shotter, J. (2019). Why being dialogical must come before being logical: the need for a hermeneutical–dialogical approach to robotic activities. *AI & Soc* 34, 29–35. <https://doi.org/10.1007/s00146-017-0697-4>
- Singler, B. (2017). “fAlth: The most avid believers in artificial intelligence are aggressively secular – yet their language is eerily religious. Why?” *Aeon*. <https://aeon.co/essays/why-is-the-language-of-transhumanists-and-religion-so-similar>.
- Singler, B. (2019). Existential hope and existential despair in AI Apocalypticism and Transhumanism. *Zygon* 54, (1), 156-176.
- Singler, B. (2023a). How I stopped worrying and learned to question the apocalyptic AI. In: Campbell, Heidi; Cheong, Pauline. *Thinking Tools for AI, Religion and Culture*. <https://oaktrust.library.tamu.edu/handle/1969.1/198190>. O.a.
- Singler, B. (2023b). Will AI create a religion? Views of the algorithmic forms of the religious life in popular discourse. *Journal of American Religion*. November 2023.
- Stawarska, B. (2019). Primacy of I–you connectedness revisited: some implications for AI and robotics. *AI & Society*, 34, 3–8. <https://doi-org.ezproxy.its.uu.se/10.1007/s00146-017-0695-6>
- Sterne, J. (2022). *Diminished faculties: A political phenomenology of impairment*. Duke University Press.
- Tiku, N. (2022, June 11). The Google engineer who thinks the company's AI has come to life. *The Washington Post*.
- Torres, E. (2021, October 19). Against longtermism, *Aeon*. <https://aeon.co/essays/why-longtermism-is-the-worlds-most-dangerous-secular-credo>
- Trausan-Matu, S. (2019) Is it possible to grow an I–Thou relation with an artificial agent? A dialogistic perspective. *AI & Society*, 34, 9–17. <https://doi-org.ezproxy.its.uu.se/10.1007/s00146-017-0696-5>
- Turkle, S. (2016). *Reclaiming conversation: The power of talk in a digital age*. Penguin Books.
- Vallor, S. (2024). *The AI mirror: How to reclaim our humanity in an age of machine thinking*. Oxford University Press.
- Verbeek, P.-P. (2006). Materializing morality: Design ethics and technological mediation. *Science, Technology & Human Values*, 31, 361–380.
- Viik, T. (2020). Falling in love with robots: a phenomenological study of experiencing technological alterities. *Paladyn: Journal of Behavioral Robotics*, 11 (1), 2020, 52-65. <https://doi.org/10.1515/pjbr-2020-0005>
- Waern, A. (2011). 'I'm in love with someone that doesn't exist!' Bleed in the context of a computer game. *Journal of Gaming & Virtual Worlds*, 3(3), 239 – 257. https://doi.org/10.1386/jgvw.3.3.239_1
- Wegerif, R., & Major, L. (2019). Buber, educational technology, and the expansion of dialogic space. *AI & Society*, 34, 109–119. <https://doi-org.ezproxy.its.uu.se/10.1007/s00146-018-0828-6>
- Wigorts Yngvesson, S. (2023, March 29). Vi är delar av en pågående revolution, *Kyrkans tidning*.

Williams, D. P. (2019). Consciousness and conscious machines: What's at stake? Virginia Polytechnic Institute and State University Ceur Workshop Proceedings.

Zylinska, J. (2018). *The end of man: A feminist counter apocalypse*. Forerunners Ideas First.