Minds and Machines (2024) 34:8
https://doi.org/10.1007/s11023-024-09667-z

We are Building Gods: AI as the Anthropomorphised Authority of the Past

Carl Öhman[1]

Received: 27 October 2023 / Accepted: 21 February 2024 / Published online: 25 April 2024

Abstract
This article argues that large language models (LLMs) should be interpreted as a form of gods. In a theological sense, a god is an immortal being that exists beyond time and space. This is clearly nothing like LLMs. In an anthropological sense, however, a god is rather defined as the personified authority of a group through time—a conceptual tool that molds a collective of ancestors into a unified agent or voice. This is exactly what LLMs are. They are products of vast volumes of data, literally traces of past human (speech) acts, synthesized into a single agency that is (falsely) experienced by users as extra-human. This reconceptualization, I argue, opens up new avenues of critique of LLMs by allowing the mobilization of theoretical resources from centuries of religious critique. For illustration, I draw on the Marxian religious philosophy of Martin Hägglund. From this perspective, the danger of LLMs emerge not only as bias or unpredictability, but as a temptation to abdicate our spiritual and ultimately democratic freedom in favor of what I call a *tyranny of the past*.

Keywords Artificial Intelligence · ChatGPT · Critique · God · Large language models · Theology

The criticism of religion is the prerequisite of all criticism.

Karl Marx.

1 Introduction

The launch of OpenAI's ChatGPT3 and its subsequent incarnations has intensified an already heated debate on the risks and challenges of artificial intelligence (AI) in general, and Large Language Models (LLMs) in particular. The most commonly

✉ Carl Öhman
carl.ohman@statsvet.uu.se

[1] Department of Government, Uppsala University, Gamla Torget 6, 751 20 Uppsala, Sweden

raised critiques have pertained to various forms of bias and discrimination inherent to such models (see Bender et al., 2021; Hacker et al., 2023; Hartmann et al., 2023; Kirk et al., 2021; McGee, 2023; Rozado, 2023; Sison et al., 2023). Others have rather placed their focus on privacy and copyright concerns (Karamolegkou et al., 2023; Lee et al., 2023; Winograd, 2023) or on the implications for mis/disinformation (De Angelis et al., 2023; Sebastian, 2023; Sison et al., 2023). Yet another group of scholars have called attention to the unpredictability of the models, arguing for various technological and regulatory means to contain or otherwise stop them from wreaking havoc on the internet (Future of Life Institute, 2023). And some even fantasize about the risk of AI systems beginning to program themselves, or form consciousness, leading to some form of general intelligence beyond human control (Bostrom, 2014). The list goes on.

How we perceive the plausibility and seriousness of these threats depends on the words and concepts we use to understand what the models in question are and what they *do* (Murray-Rust et al., 2022). Indeed, already the term *artificial intelligence* is itself an invitation to (mistakenly) compare the technology to human cognition, which has had major consequences for both its technical development and its social implications (see Watson, 2019). In a similar way, terms like *models*, *systems*, *tools* and, to use Bender et al.'s (2021) example, *stochastic parrots*, each carry with them protocols for how to interpret and deal with a given technology and its associated risks.

In this article, I submit that many forms of AI, but LLMs in particular, should be interpreted as a form of *gods*. In a theological sense, a god would be an immortal being beyond the materiality of space and time (see e.g., Augustine, 1998/397 AD). Clearly, this is nothing like any kind of AI. In an anthropological sense, however, the concept of gods, whether monotheistic or as a pantheon, is rather to be understood the synthesis of a myriad ancestral voices into one—the anthropomorphized authority of the human past (see e.g., Durkheim, 1995; Steadman et al., 2016). A god, that is, is essentially a conceptual tool which channels the aggregated knowledge and authority of a groups' collective past into a single agency or voice. And *this* is exactly what LLMs are designed to do.[1]

Unlike traditional approaches to AI, such as expert systems, logic-based models and artificial agents, LLMs (like many other contemporary approaches to AI) are trained on vast amounts of human-generated data, absorbing and assimilating information from a wide range of sources that sometimes encompass significant shares of the open web. The models emerge by "learning" from the patterns, linguistic structures, and knowledge contained within the input—not by semantically interpreting it, but by drawing statistical inferences about which sequences of words will produce a satisfying output. Albeit without any actual understanding of what it communicates, a LLM chatbot thus becomes an embodiment—indeed a *personification*—of

[1] As is elaborated below, this kinship holds irrespective of whether the anthropological definition encompasses all worship of divine creatures. The point is the similarity between AI and an established anthropological interpretation of a social phenomenon. Whether the theory can explain every instance of religious behavior is beside the point.

the collective knowledge and sentiments encoded in its training data. In the context of this article, these training data represent the accumulated wisdom, insights, knowledge (and prejudice) of a community's past that are manifested in the presence of God, or, in polytheistic religions in the pantheon of gods. Just like gods merge the voices of ancestors into one, LLMs merge a multitude of voices and perspectives into what appears to the user as a single (extra human) agency, commonly embodied as a chatbot.

I argue that reconceptualizing the nature of LLMs as gods opens up new avenues of critique beyond the usual suspects that dominate the current debate. Specifically it allows us to draw on the vast theoretical resources assembled by critics of religion from (at least) the enlightenment and onwards, and turn these against LLMs. To illustrate the fertility of this approach, I draw on the religious philosophy of Martin Hägglund (2008, 2019). Unlike many other atheist philosophers, Hägglund's critique of God is based not merely on the implausibility, but the *undesirability* of a perfect being or mode of existence. Following Marx, he argues that, when placing our faith in something *beyond* the finitude of the material realm, we can never realize our *spiritual freedom*—referring to the uniquely human capacity to *choose* the ends that we live, die and take risks for—since such choices are only intelligible in light of finitude.[2] Instead of a religious faith in the immaculate, Hägglund thus proposes a *secular faith*—an embrace of the vulnerability and risk of living for finite causes. To realize our ability to live by such a faith, however, requires a fundamental *democratization* of the economy, so as to allow each individual to participate in the collective evaluation of which ends are worth pursuing in our limited time.

When applied to LLMs, this perspective unveils that, in addition to the commonly cited risks of losing control and reifying biases in the training data, LLMs may create what I call a *tyranny of the past* over the present and the future. The meaning of this is not that the models are coercing people into obeying their demands. Rather, I posit that, through their perceived extra-human superiority, LLMs create a temptation for users to abandon their *secular faith* and submit to the predictive accuracy of the model (i.e., the past personified).[3] Just like religious communities submit to the superior wisdom of the ancestors, channeled through the extra-human voices of gods, a cultural and political reliance on LLMs is a submission to the digital past of the open web, channeled through the voices of chatbots. The challenge of LLMs, and potentially many other classes of AI, hence emerges not only as one of alleviating bias and hate speech, but to maintain our faith in the human ability to begin anew, i.e., our spiritual and democratic freedom.

[2] Note that Hägglund uses several terms that would normally invoke religious connotation (e.g., spiritual freedom, faith etcetera) in an explicitly secular sense. "Spirit" is here to be understood as connected to the German word *geist,* or Swedish *ande*, not as denoting something eternal beyond the material world. Moreover, part of the goal of Hägglund's project is to separate the religious on the one hand, and the spiritual on the other, where the former is understood as a corrupt version of the latter. In religion, by Hägglund's interpretation, the value of the material world is merely instrumental so as to reach eternal salvation. Spirituality, on the other hand, recognizes the intrinsic value of the material world as fundamentally finite and fragile.

[3] Note that the word *temptation* is here used in a strictly colloquial sense.

The argument is structured as follows. The following two sections outline the kinship between gods and LLMs by detailing their mutual nature as synthetizations of the human past. I then turn to illustrate the advantages of this reconceptualization by pointing to how the critique of religion, here exemplified by Hägglund, can be mobilized as a means to articulate new risks related to LLMs. I conclude by suggesting an expansion of the argument to other forms of AI, arguing that, when seen through the lens of this religious critique, the social challenge of AI is not merely to reduce bias or risk, but must be seen also as a larger quest to restore our faith in our ability to *begin anew*, independently of the (personified) past, however accurate its predictions and superior its productivity.[4]

2 The Anthropological Essence of God(s)

What is a god? In a theological sense, a god would be some form of absolute being beyond time and space, commonly attributed properties such as omnipotence and eternal grace. This is not what we are interested in here. Rather, our interest is gods in an *anthropological* sense. Anthropologists commonly hold that the concept of gods emerged out of the practice of ancestor worship. God, or the gods, are in fact a mere mirror of society itself, or in any case its past, that has been molded into a unified agency—the anthropomorphized version of the collective past.

It is difficult to identify a single individual responsible for this interpretation. A philosophical parallel is notably present already in the early days of nineteenth century German idealism, not least in form of the so-called pantheism controversy (*pantheismusstreit*) (see Bienenstock, 2023). But the first to call for an explicitly *anthropological* interpretation of God is Feuerbach in *The essence of Christianity* (1841). Feuerbach develops a distinction between, on the one hand, a theological (or false) understanding, where God is a supernatural spirit; and, on the other, an anthropological (or true) one, where God is understood as a projection of qualities inherent in all humans—a conceptual yet material embodiment of our collective ideals. Collectively, the human race possesses certain attributes and aspirations, such as love, power, wisdom, and (some kind of) immortality, and in religion—or, more specifically, in Christianity—worshippers tend to externalize these qualities and project them onto a supernatural being *beyond* humanity. Ironically, what Christians worship is therefore an idealized mirror image of themselves—a super*human*, which is explicated in the body of Christ. When God assumes human form, argues Feuerbach, this must be read, also, as the *deification of man*. Thus, when Christ dies on the cross, it is merely the supernatural God that perishes. The true God exists in humanity as a whole, because he *is* humanity as a whole, or at least our projected ideals.

Feuerbach's larger point is that, by adopting a theological interpretation, attributing their ideals to an external divine entity, people alienate themselves from their

[4] The phrase"begin anew" is here used in an Arendtian sense. For closer examination of what is meant by this concept, see chapter 3 in (Nässtrom, 2021).

own potential and inherent value as human beings. The best parts of humanity are seen as belonging to another world. Only when worshippers realize the true anthropological essence of Christianity as an *anthropotheism* can it be turned into a force for good in this world. We shall have reason to revisit this perspective later in this article, especially in relation to the Marxian critiques of it. But the main takeaway for now is that, when deconstructed, Christianity appears as the worship of society, or better *humanity* itself. The abilities of God are the abilities of our collective force, his grace the grace of the human race, and his voice nothing but the collective voice of our kinsmen, channeled through scripture, clergy and other religious institutions.

Feuerbach's argument about the essence of Christianity resonates remarkably well with subsequent empirical findings in the anthropological study of so-called primitive religions. The most famous example here is Durkheim's *The elementary forms of religious life* (1912)—a detailed account of the totemic religious practices of Australian Aboriginals. In these "primitive" religious forms, tribes worship plants or animals, so-called *totems,* that are unique to their community. In Durkheim's interpretation, such a totem is at the same time a form of deity, a portal to the spiritual dimension, *and* a representation of society itself. This leads him to conclude that "if [the totem animal] is at once the symbol of the god and of the society, is that not because the god and the society are only one?" (Durkheim, 1912, 201). In other words, the totem is merely a channel, a medium through which the group's collective consciousness is expressed. When everyone in a group places their trust, faith and loyalty in the same object, the object will reflect this energy back upon them as a single force. It will appear to assume an agency of its own, larger than the sum of its parts. This mechanism, argues Durkheim, is the origin of all religion. The difference is only that, in more recent religious forms (Durkheim tends to see them as higher stages of development), such as the Abrahamic faiths, the materiality of the totem has become redundant. The collective force of the group is projected, not towards a physical object, but towards a personified abstract entity—God.

According to multiple anthropologists—including Durkheim (1912) and Spencer (1870), but also more recent scholars (see Steadman et al., 2016)—gods are not only representations of the collective here and now, but, more specifically, of the group *through time,* including (and perhaps primarily) its past members. All religions, the argument goes, originally emerged as the worship of ancestors. As the departed members of a tribe gradually became more and more temporally detached, they increasingly came to be perceived as a single collective presence rather than individuals. Eventually, they would be so distant that they were essentially seen as part of a world *beyond time,* or at least beyond the time reckoning accessible to preliterate societies.

Evans-Pritchard's (1939) study of the Nuer people of North Africa is an illuminating example. In Nuer time reckoning, valid history extends only about a century—thereafter begins "tradition." Tradition, in turn, extends only as far back as lineages can be traced orally, meaning maximally about 10–12 generations.[5] Everything beyond 10–12 generations belongs to the domain of myth, and in myth time,

[5] Only very important ancestors remain in oral memory for that long.

chronology collapses. One event no longer precedes another. The things that happen in myth, as well as the ancestors who are part of it, are thus located *beyond time.* It is easy to imagine how the voices of these collective ancestors, speaking from a sphere beyond human time, would eventually merge into the voices of gods; how all mothers from 10 generations back would simply converge into a single "mother god," all tricksters into a "trickster god," and how the imagined original ancestor would eventually emerge as the god of gods. Thus when anthropologists say that gods are the collective force of the group anthropomorphized, "the group" should be interpreted as a temporally extended entity, including both present and past generations. The gods, that is, are not merely representations of society here and now, but personifications of the authority of the human past.

Before moving on, it is important to note that gods are not to be equated to mere tradition or custom, for in order to be a god, an entity needs to be perceived as such. That is, even though gods are in essence anthropological, and even if they are the key communicators of (indeed, embodiments of) customs, traditions and law, they remain gods only so long as their anthropological nature is concealed and as long as they remain *personified*. Simply put, gods are the *unconscious* anthropomorphizing of the human past.

It should also be noted that recent research on religious beliefs and practices among hunter-gatherer tribes do not unambiguously support the theory that ancestor veneration is the foundation of *all religion*, as held by Durkheim, Steadman and Spencer. As Peoples et al., (2016) note, animism, not ancestor worship, appears to be the universal denominator of indigenous religion (which, supposedly, is the basis of all religious life).[6] Moreover, there may be individual instances of gods that do not fit neatly with the anthropological interpretation as presented above. Insofar as this is the case, it is too simplistic to say that *all* gods share a common nature. While stimulating, such debates have little bearing in the context of the present paper. For the argument proposed here does not hinge upon the anthropological interpretation of gods being true in all cases. What matters is that, as we shall see in the following section, the nature of LLMs is strikingly similar to the way classical anthropologists interpret the emergence of gods. Whether or not they are empirically correct in their historical analysis, they unquestionably discovered one of the most important social mechanisms of religious life, and this is our interest in the present argument.

3 The Anthropological Essence of Large Language Models

Unlike many traditional approaches to AI, such as expert systems and logic-based modelling, many of today's dominant approaches are based on a training process whereby some input data are matched with a desired output. Unlike a human, who would learn to identify, say, a cat after having seen only one or a few specimens, an

[6] According to their dataset, however, ancestor worship does appear to co-develop with the notion of high-gods, suggesting that the notion of deities depends on, or is even an extension of, ancestor worship. (See also Swanson, 1960).

AI model would normally have to parse through thousands of cats before it "learns" to consistently identify new ones. As such, the training processes of the vast majority of AI models employed today require large volumes of data.

LLMs have a particularly big appetite for data. OpenAI's GPT-3 is an illustrative example. The model is trained on a total of nearly 500 billion tokens (i.e., text-based units/words), and 175 billion parameters (variables used in the model's predictions), from a variety of sources (Brown et al., 2020). The biggest source (60%) is the CommonCrawl dataset, which is composed of more than 3.2 billion webpages, including images and captions, and trillions of contextualized hyperlinks. The second biggest (22%) is the WebText2 dataset, which is composed of more than 45 million webpages from the open web. Then comes Books1 and Books2 (16%), which allegedly contain all books in the public domain, i.e., everything published in English before 1928 and a considerable volume of books after that (although this has been seriously questioned). The final part of the data (5%) is collected from Wikipedia. To put this in context, a person of average reading speed would have to spend 3800 years, 24 h a day, to finish reading, let alone interpreting, the training data for GPT-3. OpenAI has not disclosed the training data for its newer model GPT-4, but it is believed to be around 1.76 trillion parameters—ten times that of its predecessor (Albergotti, 2023). This is to say that LLMs like ChatGPT, or one of its increasing number of competitors, are essentially the products of a significant portion of the digital heritage of Western (anglophone) civilization, and (albeit to a lesser extent) of humanity in its totality. Anyone who has published a book before 1920, or indeed posted anything to the open web, have probably contributed to the models' training.

It should be noted that LLMs do not just arise from the training data by themselves—nor does any other AI model. The training of a model—especially in supervised learning, which is part of most LLMs' design—is an active process, involving thousands of decisions from those designing it, including which datasets to use and how to curate them, how tokens are to be weighted, and what type of responses are deemed appropriate from the model, e.g., what type of language is desirable for conversations with a chatbot, and which topics the bot should avoid (such as sexually explicit language). As such, it is a mistake to treat LLMs as if were they neutral representations of humanity that just "arise" from the data (see Bennet et al., 2021). They are one out of many possible representations. Nevertheless, a model is for all intents and purposes *a* product of its training data, the same way a dish *is* its ingredients, albeit in processed form. In the case of LLMs, those ingredients are pieces of the past, and the model is a channel, a medium through which they are given a voice, even agency. In other words, when you are chatting with a bot like ChatGPT, you are more or less chatting with a personification of significant portions of the web prior to 2021 plus a serious chunk of public domain books. But when it answers, it will not be in a million voices, but one, and the answer will not be the only possible one, but one that the designers of the model favor.

This is why the essence of LLMs (but also other forms of AI), is *anthropological*. They are personifying the human past, giving it voice, agency, and—more importantly—*authority,* in the present. God, anthropologically speaking, is the presence of the group (including its living and dead members) molded together into what now appears to the worshipper as a single voice. Conversely, an AI

chatbot is the presence of (large chunks of) our digital past molded together into what appears to the user as a single voice speaking with almost infinite authority. It is a reflection of ourselves in the past. Perhaps it is not a neutral one, and it is far from complete or fair, but it is, nevertheless, an extension of the voices of those whose data it has been trained on. It works with the combined force of the patterns engraved in the web over its first decades of existence. In other words, LLMs emerge through processes that are strikingly similar to the way that anthropologists argue that gods come into being. When we talk about the increasing influence of LLMS in society, we are therefore inevitably talking about the influence *of other humans*, namely the individuals inhabiting our past, those who provided the data that trained the model. And as we shall see in the following section, this makes a big difference in terms of how we think about its social influence.

As with any analogy, the similarities between gods and LLMs are not absolute. One difference, for example, is that, while the former remain purely mental (albeit social) entities, the latter are directly visible to all. Gods have communicated only through shamans, prophets and (sometimes) scripture, i.e., human interpreters. Their voices appear between or within worshippers. Chatbots, on the other hand, speak to us directly. Unlike previous gods, LLM-chatbots actually communicate in letters and words that anyone can see and read. As such, the kinship described here is more than a "mere" metaphor. Rather, gods and LLMs are members of the same species, they belong to the same type of social phenomenon. We may not be building gods as understood by Abraham, St. Augustine or the Pharaohs, but we are *literally* building the gods of Feuerbach and Durkheim.

Before we move on, there is reason to underline the distinction between the anthropological nature of LLMs and the common notion of the internet as some kind of collective nervous system, or "embryonic technological autonomy" (Kelly, 2010, 14). The experience of interacting with a single independently reasoning entity is an *illusion*, both when it comes to gods and to AI (see Floridi & Chiriatti, 2020). In fact, the distinction between gods/LLMs and other representations of the past is exactly their concealment of their true anthropological nature. Search engines, for example, are also able to answer most questions, but do not hide the fact that the answers they provide are the answers *of other people.* They always make the human origin of the communicated knowledge explicit by design. A LLM chatbot like ChatGPT, on the other hand, presents its answers as if where they of its own making. As Chiang (2023) puts it "OpenAI's chatbot offers paraphrases, whereas Google offers quotes," and in this process, it is easy to forget that what you are interacting with is essentially a *medium* and not a machine intelligence.

Finally, it should be noted that the comparison between LLMs and gods pertains to their *essence*, not their social consequences. The argument presented above states only that, insofar as gods are personifications of the human past, LLMs fall within this definition. Whether or not they assume the same status and have the same cultural implications depends, I argue, on our ability to recognize the processes by which they are made. Thus, let us now turn to the ways the proposed reconceptualization may help articulate the possible threats and challenges that arise from LLMs.

4 New Avenues of Critique.

What is the point of conceptualizing LLMs as gods? I submit that this move opens up new avenues of critique beyond the usual suspects—bias, disinformation, unpredictability etcetera—accounted for in the opening section of this article. More specifically, it allows us to mobilize the theoretical resources accumulated by critics of religion, or rather of *God,* and turn these against LLMs, but also AI more broadly.[7] In the following section, I illustrate one way in which this can be done, by redefining the overarching danger of LLMs as a form of *tyranny of the past* over the present. I then turn to the religious philosophy of Martin Hägglund (2008, 2019) to show why our past, even when rid of its imperfections, is an undesirable ruler of current affairs.

4.1 The Tyranny of the Machine as a Tyranny of the Past

Generally speaking, the most commonly identified threat of AI is the loss of human control over our machines. This threat takes various guises in the literature. In some cases, it appears as outright science-fiction scenarios where an artificial agent develops general intelligence, takes over and dominates humanity into subjecting to its goals (Bostrom, 2014). Most scholars, though, tend to focus on more urgent and realistic forms of control—areas where humans are already being replaced or assisted by AI systems (Diab, 2023). The criminal justice system, where sentencing and recidivism assessments are increasingly assisted, or taken over by AI, is one such example (Dressel & Farid, 2018). Matters of accountability/trustworthy AI (High-Level Expert Group on Artificial Intelligence, 2019), credit score assessments (O'Neil, 2017) and autonomous weapons of various kinds (Lin et al., 2008), also belong to some of the most urgent matters where the fear is that humans are increasingly being controlled by their machines rather than vice versa. Similar fears are also discussed in the broader AI ethics literature, where calls are commonly made to align AI with "human values" (Dignum, 2017; Kasirzadeh & Gabriel, 2023) or make it "human-centric" (European Commission, 2019). Thus, irrespective of focus areas, it seems that the principal concern of AI is that humans will somehow lose control over it and fall into a form of *tyranny of the machine.*

In view of the argument proposed in this article, the supposed tyranny of the machine over humanity emerges rather as a tyranny of our own past over the present and future. Insofar as we accept that LLMs, like gods, are personifications of the authority of the past, their social and political influence is to be seen, not as a force external to humanity, but as the continuous influence of the humans on whose data the models were trained. As we have seen, a LLM chatbot is really a synthesis of the voices on which it is trained, often encompassing large segments of the web. The same goes for a machine learning program that has been trained on, say, thousands of doctors' assessments of tumor X-rays, or the recidivism of millions of convicts.

[7] As observed by Noble, [1997), the ties bet the cultural kinship to religion is not limited to AI but can, in some regards be extended to all technology. In the present article, however, it is only personalization of the past-aspect that is of interest, which falls outside of the scope of Noble's argument.

When a machine is taking over a job that would previously be done by a person, it is not merely a machine replacing a human, but a technological means to extend past human actions (commonly speech acts) into the present. Like gods, AI systems are not agents acting upon humanity from the outside, but reflections of *ourselves* in the past that have assumed a new agency. They are a medium by which our collective data traces continue to exert their influence upon us, and insofar as they become the dominant source of decisions*, a tyranny of the past* over the present.[8]

Why is a continuous reliance on the past necessarily a bad thing? Are not LLMs, after all, incredibly efficient and (for the most part) accurate? Based on most current critiques of LLMs, the answer would be that the past is imbued with various ethical and political sentiments that we may not want to reproduce. In the context of religious critique, this would largely be analogous to Marxist approaches. Marxists stress that, while Feuerbach's "anthropotheistic" interpretation of the Christian God is largely correct, gods are by no means neutral or democratic representations of humanity. Gods, in any society, must be understood as *designs* by the ruling class to legitimize and cement their power, a means to portray the current order of things as inevitable. As Marx puts it in *Theses on Feuerbach* (1969) "Feuerbach, consequently, does not see that the 'religious sentiment' is itself a social product, and that the abstract individual whom he analyses belongs to a particular form of society." Gods are social products that eliminate some voices from the past, while elevating others to universal truth speakers. In other words, the past is an unfit ruler because it stands in the way of true social progress. Or as Marx (1969) has it "the traditions of past generations [personified by God] weigh like a nightmare on the brains of the living."

Today, it is not the traditions, but the linguistic patterns and sentiments of past generations that continue to haunt the living through the AI models that increasingly inhabit our life worlds. Like gods, these models are by no means fair nor neutral representations of our past. As Bender et al., (2021) note, "In accepting large amounts of web text as 'representative' of 'all' of humanity we risk perpetuating dominant viewpoints, increasing power imbalances, and further reifying inequality." Enterprises that are over-reliant on LLMs "run the risk of 'value-lock', where the LM-reliant technology reifies older, less-inclusive understandings" (Bender et al., 2021, 614). The voice of the LLM becomes an amplifier of past speech acts that uphold a broader system of oppression. In the case of ChatGPT, for instance, there is considerable evidence that the model reproduces the values of Western, educated

[8] A tyranny of the past does not necessarily mean a stagnation of development or innovation. All human innovation is, in a sense, mixing things that already exist. What it does mean, however, is that, as innovation and development become increasingly reliant on AI, they also become increasingly reliant on those whose data trained the model—the people populating the world wide web during its first few decades of existence. Today's LLMs are trained mainly on data from online sources generated during the past few decades, but given the ease with which they produce content, it is likely that a significant share of the contents of the future web will be produced with AI assisted means. As pointed out by Bennet et al., (2021) these contents will in turn be the training data for future models, leading to a feedback loop where LLMs are increasingly trained on data generated by other LLMs. All this must not necessarily mean that the future will look increasingly like the past, but it does mean that it will increasingly be *a product of the past.*

and affluent populations (Mohammad et al., 2022). The same goes for e.g., banks that use AI to assess credit scores, or colleges automating admissions processes (see O'Neil, 2017). If the model is trained on data that reflect an undesirable or skewed world, it is likely to continue reproducing, indeed *reifying* that world in perpetuity. As such, the "voices of the past" that speak through the AI chatbot, are not just any voices, but often the voices of hegemony. Whether embodied as gods or as AI models, this is why the past is an undesirable ruler of the present and future.

As we have seen, the current literature gives a convincing answer to why we should be worried about our increasing reliance on the past—it is imperfect and stalls progress. But the theoretical resources from religious critique provide us also with another answer, namely that even a past without imperfections would be an undesirable ruler of our lives. Even when reflecting a completely just and equal past, and even with an accuracy far superior to any individual human, we would be wise not to outsource our decisions to LLMs or any other form of AI. To understand why, I propose we turn to the radical atheism of Martin Hägglund (2008, 2019).

4.2 Hägglund and the Elimination of Risk

The standard critiques of God would either be the Marxist one accounted for above, or one grounded in the *implausibility* of the existence of a divine being beyond the material world (e.g., Dawkins, 2016). By contrast, the critique that Hägglund (2019) proposes is based on the *undesirability* of such a state of being (notably still drawing heavily on Marx but also Nietzsche, Hegel and Kirkegaard). For a divine being, or mode of existence (here defined as an existence beyond the material world), would be the end of the finitude upon which human freedom depends. Or, put differently, the problem in believing in an absolute being or state of being, is not that it is unattainable, but that it would divorce us from the very thing that makes us human—the active process of choosing the ends for which we suffer, take risks and die, referred to by Hägglund as *spiritual freedom* (see Hägglund, 2019, 145).

Spiritual freedom refers to the fact that, unlike any other life-form, human beings can question not only how we pursue our ends, but also which ends are worth pursuing in life. We can choose the causes we are living for. The ultimate condition for this freedom, however, is *finitude*. Without the prospect of failure and ultimately death, the activity of choosing which ends to pursue is not intelligible. Indeed, in a state of perfection, nothing can change, nothing is at risk. And if nothing is at risk, there are, by definition, no ends to pursue, no causes to care for, since everything is already complete. To live a free human life, therefore, is not to live without constraints, struggle or the risk of failure, but is about "being confronted with the question of what you ought to do and what you ought to believe" (Hägglund, 2019, 271).

Which causes are worth taking risks for is not the type of question that can be answered once and for all. And it cannot be answered *for* us. Rather, it is a matter of what Hägglund refers to as *secular faith*. Religious faith, in Hägglund's definition, is the belief in something beyond the material realm, whose existence is independent of our actions and fidelity, and whose invulnerable nature can be shared with us

by salvation (e.g., infinity).[9] Secular faith, by contrast, is a belief in causes such as friendship, love or the survival of cultural institutions—things that we believe (but cannot prove) have an intrinsic and irreplaceable value and that require our care to be sustained. Since there can be no final proof that e.g., a friendship or an institution is worthy of one's time and effort, the challenge of choosing the objects of our faith is always fundamentally open. We must begin every day by asking anew, knowing that the answer may change tomorrow. It is a matter of *fundamental uncertainty.* Choosing how, and what we live and struggle for, i.e., what we have *faith* in*,* is thus not an instrument towards a righteous end destination—it is the very process by which we become human.

The fact that spiritual freedom requires fundamental uncertainty and fragility, argues Hägglund, is what makes the existence of an absolute being (a god) so undesirable. If there were such a being, or mode of existence, it would be the end of freedom. If there were a final answer to how and for what our lives should be lived, such as an answer given by God, we would cease to live as humans, since we would no longer need to—nor be able to—ask what causes are worthy of our time and risk-taking. In such a world, nothing could be worthy of our care and commitment, since nothing could, by definition, be lost. It would be like cheering for a sports team for whom it is impossible to lose. Even if the divine answers to life were true, and would make our lives more righteous and happier, the result would no longer be *our* lives.

What does this have to do with AI? The answer is that Hägglund's critique of God allows us to verbalize why even a perfect past, cured from the imperfections identified by the likes of Bender et al. (2021), Dressel and Farid, (2018), Diab, (2023) and O'Neal (2017), is an undesirable ruler. By outsourcing our decisions to the combined wisdom of the past, we would indeed make better decisions, find jobs or partners more fitted to our desires and make better assessments of the likelihood of recidivism etc. It would even rid us of the bias, discrimination and other flaws inherent to human decision making. But by submitting to such an authority, these would no longer be *our* decisions, and so we would no longer be living *our* lives. We would, in Hägglund's terms, abandon our secular faith in favor of the predictive authority of the models.

For illustration, consider the following scenario. Two young lovers ask an AI model whether their marriage will continue to bring them joy in life. The model has access to all of their personal data, and data on the mental health of millions of couples like them in the past. After analyzing these, the AI concludes that, with 98% certainty, the couple would be happier if they separated right away and would not regret their decisions later. Now, any rational agent who values their own wellbeing would heed the advice of the model. Choosing to stay together involves a huge risk. Yet, someone who lives their life this way loses something fundamentally human, namely their willingness, and ultimately *ability* to take risks for the causes they

[9] Note that religious individuals are perfectly capable of caring for the intrinsic value of finite things too. But insofar as one acts for causes that remain in the material realm, this cannot be characterized as religion (defined as the belief that overcoming finitude is the ultimate goal of life).

believe in, i.e., their secular faith. Were the couple to heed the advice of the AI, it would mean that they have more faith in the machine (i.e., the personified past) than in themselves and in each other. They would be treating themselves and each other as *mere* subjects to their history, as opposed to being the subjects *of* their history (paraphrasing Hägglund, 2019, 24). The cost of outsourcing our decision making even to a perfect model, is thus our ability to make decisions that we *mean* and that are truly *ours.*

But wherein lies the threat to freedom? Surely, the couple in the above scenario are perfectly free to disregard the advice of the AI councilor? Indeed, they are. And this is an important point because, through the lens of Hägglund's critique, the threat of AI to human freedom lies not in some form of coercion (as proposed by e.g., Sparrow, 2023), but in the *temptation* to abdicate from our faith in ourselves so as to reduce the risk of choosing ends that we may come to regret or fail to realize. When standing before what appears to the user as an infinitely superior intelligence with virtually unlimited access to information, the courage and confidence it takes to defy it is insurmountable. In each instance, it will seem less risky to merely submit to the authority of the past.

When taken to its extreme, this temptation manifests itself as a strive to eliminate risk all together. This is arguably what is happening in the so-called Singularitarian community. Singularitarians typically believe that, once super-intelligent general AI is invented, it will bring about a society of overabundance, basically erasing human finitude, both figuratively and literally. The champion of this camp, Nick Bostrom, is a telling example. He believes that, with superintelligence, disease, death and unhappiness will be overcome once and for all. Human existence will be one of overabundance. This is all provided that the super-intelligent AI does not turn evil and somehow destroys humanity, of course. Such a catastrophic event must, according to Bostrom, be prevented at *all* costs. In fact, reducing the prospect of this scenario by even "a mere one billionth of one billionth of one percentage point," he finds, is worth more than "a hundred billion times as much as a billion human lives" (Bostrom, 2013, p. 19).

It is not hard to see how such a calculation could get very ugly very fast, but the point here is rather that, as detailed by e.g., Noble (1997), Geraci (2012), Samuel (2023) and Floridi (2022), the underlying logic of the argument is the same as in the worship of an absolute God. The prospective individuals living in overabundance with the super-intelligent machine are the singulatarian version of humanity living beside God in infinity. Rather than caring about the real, material individuals affected by AI here and now, we are thus encouraged to care about the abstract prospect of a world beyond finitude, where all our limitations are overcome. Rather than focusing on the intrinsic value of the humans of the finite world, thus exposing themselves to the risk of losing them, the Singulatarians place their hope in the prospect of a world beyond finitude and the risk of failure. And so, the present world is reduced to a mere instrument, as the overcoming of finitude becomes the principal goal of all human conduct.

The Singularitarian case is fraud and dangerous, but it is merely an extreme form of a tendency present wherever human decision-making is outsourced to AI models—the temptation of reducing uncertainty at the cost of acting authentically out

of (secular) faith. Plenty of real-life examples are at hand: A man buying a birthday present for his wife, but to reduce the risk of buying something she may not like he asks a chatbot "What is a romantic gift for a 30-year-old woman?"; A mother preparing a wedding speech for her daughter but rather than taking the risk of writing a dull one, has a LLM write it for her. The gift will probably be well received and the speech eloquent, but the point here is that it is no longer the man's gift, nor the mother's speech since neither relies on their exposure to risk.

Perhaps more concerning is the increasing influence of various AI tools in political communication. Here too, these is no lack of examples, both hypothetical and real: Struggling to formulate an articulate response to a political argument on social media, a citizen asks a LLM to write one for him. The opponent uses the same tool to counter, and thus, all of a sudden, political deliberation has been outsourced to the machine; A politician is to make a statement on a contentious topic, but before deciding upon a stance, runs a simulation based on past voting patterns to see which position is most likely to win the votes of her audience. In each of these cases, the course of action may have reduced the risk of failure, but only at the cost of one's ability *to mean what one does and says.* Rather than taking risks for the causes they believe in, the actors in the examples are letting the past (embodied by the AI/God) do the living for them. As opposed to having *faith* in what they believe, they are treating themselves and those around them as if they were already dead, the future as if it had already happened. They are, in other words, submitting to the tyranny of the past.

On an individual level, what is at stake here is spiritual freedom, or perhaps what someone like Jean-Paul Sartre would have called *authenticity.* On a collective level, however, what is at stake is the very *spirit of democracy* (Näsström, 2021), because in collective form, our spiritual freedom is what Arendt (1993) called *democratic freedom,* i.e., our collective ability to form and communicate judgements independently of tradition or divine imperative—to begin *anew*.[10] In monarchical political forms, the question of who should rule is already settled once and for all by God (see Näsström, 2021, 75). The future is (or can be) *certain*. The same goes for totalitarian regimes, where the question of who should rule is once and for all settled by a supreme ideology rather than God. But in democracy, the questions of who should rule and how are always open—the future *fundamentally uncertain*. While each of us are free to form judgements about society and communicate these to our peers, we can never know for certain how these will be received nor whether they will still stand tomorrow—we can only have *faith* in our ideals and in our peers. From this perspective, the danger of AI is not only bias and lack of control, but also the collective temptation to reinstall God, figuratively speaking; to reduce the uncertainty of the future by letting our collective past take the place of our democratic freedom. The absence of God is what made the openness of democratic rule possible. We would be wise not to reinstall him in the guise of AI.

[10] See in particular Chapter 4"What is freedom?" in (Arendt, 2023) for a closer look at the concepts of democratic freedom and beginning anew.

If it is not perfectly clear how LLMs pose a *new* kind of threat to our spiritual and democratic freedom—after all, have not religion, capitalism, technology and social conformity always posed a threat to authenticity and freedom?—it is because this is exactly the point. LLMs may provide awesome powers to its users. But insofar as we understand them as temporal technologies, as personifications of our past, it is evident that the challenges they raise are not entirely new. We have seen these social mechanisms before. But more importantly—we have developed successful analytical and institutional means to deal with, and unmask them.

5 Concluding Remarks

Whatever words or metaphors we use to conceptualize new technologies, they always carry with them social and moral protocols for how to deal with their implications. As such, it is important that we continue debating the pros and cons of each concept. And that we keep developing new ones.

The contribution of this article is that LLMs should be interpreted as *gods* in the sense that they are personifications of the human past. Notably, the analogy is not necessarily restricted to LLMs, but could also be applied to multiple other forms of AI. Indeed, any machine learning system could, in a sense, be characterized as a form of synthesis of data, and all data are essentially traces from the past (see Öhman, 2020). Yet, delineating exactly which approaches fall within the scope of the analogy is a task beyond this article. For now, the point is merely that the analogy allows us to formulate new types of critique against LLMs and AI based on the resources of religious critique.

Here, I have used Hägglund's radical atheism as an example, making the case that even a perfectly unbiased, copyright compliant, and fair LLM would be an undesirable ruler of our individual and collective lives—not because its decisions would be bad, or because we would somehow be coerced by them, but because they would not be *ours.* In addition to bias and lack of control, the danger of LLMs is thus the *temptation* to place one's faith in the past over one's faith in oneself and one's peers. This example is by no means exhaustive. It is merely an illustration of the fruitfulness of the proposed approach. Plenty of other critics of religion—Feuerbach and Marx are perhaps the most prominent candidates—may be mobilized as a means of critiquing the role of LLMs and AI in society.

Yet, the proposed analogy does not only provide a means to criticize. It also tells us something about the challenge before us, about what we must do to avoid the identified risks, and where we should be heading. This challenge, I argue, is to have faith in our ability to begin anew, and courage to step into the categorically unknown. For the very the moment our faith in the personified past outweighs our faith in ourselves and each other, we succumb to its tyranny—we become merely subjects *to* and not of our history; the slaves, not the successors, of our ancestors.

This research was generously funded by Wallenberg AI, Autonomous Systems and Software Program – Humanity and Society

Funding Open access funding provided by Uppsala University.

References

Albergotti, R. (2023). *The secret history of Elon Musk, Sam Altman, and OpenAI*. Semafor. https://www.semafor.com/article/03/24/2023/the-secret-history-of-elon-musk-sam-altman-and-openai.

Arendt, H. (1993). *Between past and future*. Penguin Books.

Augustine, St. (1998). *The confessions*. Paris, France: Bibliothèque de La Pléiade. http://www.ourladyswarriors.org/saints/augcon10.htm.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? FAccT 2021 – Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency 610 623

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss., A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., & Amodei, D. (2020) Language models are few-shot learners https://doi.org/10.48550/arXiv.2005.14165

Dawkins, R. (2016). *The god delusion (10th anniv)*. Black Swan London.

De Angelis, L., Baglivo, F., Arzilli, G., Privitera, G. P., Ferragina, P., Tozzi, A. E., & Rizzo, C. (2023). ChatGPT and the rise of large language models: The new AI-driven infodemic threat in public health. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4352931

Diab, R. (2023). Too dangerous to deploy? The challenge language models pose to regulating AI in Canada and the EU. *University of British Columbia Law Review*. https://doi.org/10.2139/ssrn.4680927

Dignum, V. (2017). Responsible artificial intelligence: Designing AI for human values. *ICT Discoveries, 1*, 1–8.

Dressel, J., & Farid, H. (2018). The accuracy, fairness, and limits of predicting recidivism. *Science Advances, 4*(1), 1–6. https://doi.org/10.1126/sciadv.aao5580

Durkheim, E. (1912). *The elementary forms of religious life*. London: George Allen & Unwin Ltd.

European Commission (2019) Building trust in human-centric artificial intelligence: Communication from the commission to the european parliament, the council, the european economic and social committee and the committee of the regions https://ec.europa.eu/digital-single-market/en/news/communication-building-trust-human-centric-artificial-intelligence

Evans-Pritchard, E. E. (1939). Nuer time reckoning. *Journal of the International African Institute, 12*(2), 189–216.

Floridi, L. (2022) Ultraintelligent machines, singularity, and other sci-fi distractions Lavoro, Diritti, Europa https://www.lavorodirittieuropa.it/ https://ssrn.com/abstract=4222347

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines, 30*(4), 681–694. https://doi.org/10.1007/s11023-020-09548-1

Future of Life Institute (2023) Policymaking in the pause: What can policymakers do now to combat risks from advanced AI systems? https://futureoflife.org/wp-content/uploads/2023/04/FLI_Policymaking_In_The_Pause.pdf

Geraci, R. M. (2012). *Apocalyptic AI: Visions of heaven in robotics, artificial intelligence, and virtual reality*. Oxford University Press.

Hacker, P., Engel, A., & Mauer, M. (2023) Regulating ChatGPT and other large generative AI models. Fairness, Accountability, and Transparency (FAccT '23) 1112 1123 https://doi.org/10.1145/3593013.3594067

Hägglund Martin (2008) Radical Atheism : Derrida and the Time of Life Stanford University Press

Hägglund Martin (2019) This life: Why mortality makes us free (ebook) Profile books Ltd

Hartmann, J., Schwenzow, J., & Witte, M. (2023) The political ideology of conversational AI: Converging evidence on ChatGPT's pro-environmental, left-libertarian orientation SSRN Electronic Journal https://doi.org/10.2139/ssrn.4316084

High-Level Expert Group on Artificial Intelligence (2019) Ethics guidelines for trustworthy AI. High-level expert group on artificial intelligence In European Commission https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

Karamolegkou, A., Li, J., Zhou, L., & Søgaard, A. (2023) Copyright Violations and Large Language Models 7403 7412 https://doi.org/10.18653/v1/2023.emnlp-main.458

Kasirzadeh, A., & Gabriel, I. (2023). In conversation with artificial intelligence: Aligning language models with human values. *Philosophy and Technology, 36*(2), 1–24. https://doi.org/10.1007/s13347-023-00606-x

Kelly, K. (2010). What technology wants. *Penguin Books*. https://doi.org/10.5840/traddisc2010/20113 7341

Kirk, H. R., Jun, Y., Iqbal, H., Benussi, E., Volpin, F., Dreyer, F. A., Shtedritski, A., & Asano, Y. M. (2021). Bias out-of-the-box: An empirical analysis of intersectional occupational biases in popular generative language models. *Advances in Neural Information Processing Systems, 4*(NeurIPS), 2611–2624.

Lee, J., Le, T., Chen, J., & Lee, D. (2023) Do language models plagiarize? ACM Web Conference 2023 – Proceedings of the World Wide Web Conference, WWW 2023 3637 3647 https://doi.org/10.1145/3543507.3583199

Lin, P., Bekey, G., & Abney, K. (2008) Autonomous military robotics: Risk, ethics, and design In California Polytechnic State University http://ethics.calpoly.edu/ONR_report.pdf

McGee, R. W. (2023) Is chat Gpt biased against conservatives? An empirical study SSRN Electronic Journal 1 19 https://doi.org/10.2139/ssrn.4359405

Mohammad, A., Xue, M. J., Park, P.S., Blasi, D. E. & Henrich, J. (2022) Which humans? PsyArXiv https://doi.org/10.31234/osf.io/5b26t

Murray-Rust, D., Nicenboima, I., & Lockton D. (2022) Metaphors for designers working with AI DRS2022: Bilbao 0 19 https://doi.org/10.21606/drs.2022.667

Näsström, S. (2021). *The spirit of democracy: Corruption, disintegration*. Oxford University Press.

Noble D. (1997) The Religion of Technology: The divinity of man and the spirit of invention Alfred A knopf

O'Neil Cathy (2017) Weapons of math destruction: how big data increases inequality and threatens democracy Broadway Books

Öhman, C. (2020). A theory of temporal telepresence: Reconsidering the digital time collapse. *Time and Society, 29*(4), 1061–1081. https://doi.org/10.1177/0961463X20940471

Peoples, H. C., Duda, P., & Marlowe, F. W. (2016). Hunter-gatherers and the origins of religion. *Human Nature, 27*(3), 261–282. https://doi.org/10.1007/s12110-016-9260-0

Rozado, D. (2023). The political biases of ChatGPT. *Social Sciences, 12*(3), 148. https://doi.org/10.3390/socsci12030148

Samuel, S. (2023) Silicon Valley's vision for AI? It's religion, repackaged Vox https://www.vox.com/the-highlight/23779413/silicon-valleys-ai-religion-transhumanism-longtermism-ea

Sebastian, G. (2023) Exploring Ethical Implications of ChatGPT and Other AI Chatbots and Regulation. May

Sison, A. J. G., Daza, M. T., Gozalo-Brizuela, R., & Garrido-Merchán, E. C. (2023) ChatGPT: More than a "weapon of mass deception" ethical challenges and responses from the human-centered artificial intelligence (HCAI) perspective International Journal of Human-Computer Interaction 1 31 https://doi.org/10.1080/10447318.2023.2225931

Sparrow, R. (2023) Friendly AI will still be our master. Or, why we should not want to be the pets of superintelligent computers AI and Society, 0123456789 https://doi.org/10.1007/s00146-023-01698-x

Spencer, H. (1870). On ancestor worship and other peculiar beliefs. *Fortnightly Review, 13*(7), 535–550.

Steadman, L. B., Palmer, C. T., Tilley, C. F., Steadman, L. B., Palmer, C. T., & Tilley, C. F. (2016). The universality of ancestor worship. *Ethnology, 35*(1), 63–76. https://www.jstor.org/stable/3774025.

Swanson, G. E. (1960). *The birth of the gods: The origin of primitive belief*. University of Michigan Press.

Watson, D. (2019). The rhetoric and reality of anthropomorphism in artifcial intelligence. *Minds and Machines, 29*(3), 417–440. https://doi.org/10.1007/s11023-019-09506-6

Winograd, A. (2023) Loose-lipped large language models spill your secrets: The privacy implications of large language models Harvard Journal of Law & Technology36 2