КУЛЬТУРОЛОГИЯ

AI and the Metaphor of the Divine

D. S. Bylieva[1], A. Nordmann[2]

[1] Peter the Great St. Petersburg Polytechnic University,
   29, ul. Polytekhnicheskaya, St. Petersburg, 195251, Russian Federation
[2] Technical University of Darmstadt,
   1, Residenzschloss, Darmstadt, 64283, Germany

The idea of God is one of the most pervasive in human culture. It used to be considered mostly in metaphysical and ethical discussions, it has become a part of the discourse in the philosophy of technology. The metaphor of God is used by some authors to represent the role of artificial intelligence in the modern world. This paper explores four aspects of the metaphor: creation, omniscience, mystery, theodicy. The creative act shows the resemblance of humans to God, also in the sense that technology created by humans can get out of the control of the creator. The ability of AI to use data flows for analytics and prediction is considered as "omniscience" which looks mysterious due to the inability of people to intellectually grasp the work of AI. The discussion about building ethics into AI technology shows a desire to add one more feature to "omniscient and omnipotent," namely "benevolent." The metaphor of God as applied to AI reveals human fears and aspirations both in rational-pragmatic and symbolic terms. The metaphor of God exposes notions of transcendence in modern perceptions of technology. Also it continues the discussion about what should be the technological design of AI. Whether as co-worker or as communicator it is already putting us on a path towards thinking of a subject that is constituted in a superior way.

Keywords: AI, ethics, technology, metaphor, God.

Introduction

The active development and implementation of AI make it a very popular topic of scientific discussions. AI is used in different spheres of human life, interacting with persons, processing data, making predictions and decisions. Modern AI technologies are self-learning, based on large databases, they are able to classify independently and to draw

conclusions without direct human training. Moreover, reinforcement learning allows an AI system to evaluate its behaviour and improve in order to achieve goals. For quite a long time there have been discussions about the possibility of the emergence of so-called artificial general intelligence which is an as-of-yet hypothetical form of AI that is able to perform all the same intellectual feats as humans do and surpass them. It is a super-intelligent machine which a person can understand no more than a pet understands a human being [1, p. 1]. Moreover, optimism about the technological progress of humankind is accompanied by fears that technology may get out of control, and AI will become the "last invention" [2; 3].

The idea of speaking about AI in terms of God is popular today in science fiction (starting with GOD — General Operational Device from Lem's *Fiasco* (1986)), scientific literature [4–6] and even in technical terminology (God's Eye for detecting, recognizing, tracking and locating the person or an object whose image is given as query). This metaphor (or even more than metaphor) is meaningful. It not only elevates technology to unprecedented heights, but also reveals the modern trends in transcendent understanding of technology. David F. Noble writes that in the Western world modern technology and religion have evolved together and technology remains suffused with religious beliefs [7].

The words and metaphors used to refer to technologies affect how they are perceived and how they function. As Mark Coeckelbergh points out, "Use of language does not only construct the representation of relationships but also transforms them" [8, p. 151]. In philosophy of language John Austin claims that a speech act doesn't only say something about the world, but also "does" something [9]. Moreover, Julian Jaynes says a learned behavior arises from language, and specifically from metaphor [10]. According to Lakoff & Johnson, metaphor is an active agent in human cognition, it influences everyday experience by establishing an epistemological referent for all forms of communication and cognition itself [11]. Pavel Baryshnikov argues that metaphorised archaic images create modern secularized myths that perform regulatory and coding functions [12, p. 134].

The concept of God is one of the most important for human culture. Clearly, even if we discuss only the Judeo-Christian concept of God, it cannot be fully disclosed from a theological or philosophical point of view. However, this does not prevent it from being widely used and considered as generally understood. As Peter Vardy remarks, "we know that we are fairly ignorant about black holes, but we feel fairly confident about using the word God" [13, p. 37]. Usually the word "God" refers to a "being which is omnipotent, benevolent, omniscient" [14, p. 230]. For a long time, the moral aspect of the human-God relationship seemed to be the most important in culture. In relation to the human, God appears as the Creator, the Highest Good and the Judge of one's deeds. For Kant it is through morality and the highest good that we "produced a concept of the divine being that we now hold to be correct" [15, p. 818]. The idea of God as the basis of morality made it in demand in ethical and existential philosophical discourse. "Dostoevsky once wrote: 'If God did not exist, everything would be permitted'; and that, for existentialism, is the starting point" [16, p. 28].

In the course of the secularization of society, the word God moved from a sacred sphere — with the prohibition of its pronunciation "in vain" (for Christians) or under any circumstances (for Jews) — into the space of mass culture. The recent introduction of the concept of God into the technological dimension of culture is far from the traditional

philosophical disputes about the divine. Nevertheless, the use of the word 'God' brings to the surface deep cultural and archaic layers. It demonstrates how in a secularized world the wonders of technology are displacing the need for the transcendent. Along these lines, the very topic of artificial intelligence unites the existential and technological discourse of modernity. In the discussion about AI, the eternal questions about the place of humans in the world, about good and evil are raised with renewed vigor. In this article, we will highlight and explore four aspects of the metaphor of God in relation to AI: creation, omniscience, mystery, theodicy.

Creation

Creative acts are considered to be a point of similarity between humans and God. The Greek word *techne* means activity that brings things into being. In *techne* we can see the ability to reorganize the world, to create a new order or *logos*. The philosophy of technology after Ernst Kapp represented human activity as a sphere for the generation of the new, providing for human needs and desires, forcing the creation of engineering ontologies that take into account "being" in terms of laws of nature, and regarding the "new" as a technically successful and progressive application of these laws [17, p. 23]. According to Heidegger the origin of modern, Cartesian metaphysics coincides with a technological way of grasping the world in terms of forces. In this way of picturing the world, everything becomes an object of human will and, implicitly, there is no space for a special divine power that is distinct from human calculation and control. Thus, Heidegger contrasts contemplative thought that seeks an understanding of Being and "calculating" thought that transforms everything into a resource for technical control [18]. Here Yuk Hui sees exemplified Heidegger's point that the earth is being transformed into a giant cybernetic machine, "that is arriving now with the advancement of artificial intelligence, which can be read in daily outcries in the newspaper" [19, p. 92].

God appears as the creator of existence, of life and of humanity, and as the promise to fill a void of inexplicable existence. A feature of humans that distinguishes them from other creations of God is free will that gives them possibility to rebel. And the human-creator wants to be self-sufficient, rebelling against existence as something merely given, "a free gift from nowhere (secularly speaking), which he wishes to exchange, as it were, for something he has made himself" [20, p. 2–3]. In self-dissatisfaction, there is both, Jean-Paul Sartre's nausea of those who are thrown into the world without any reason, and Günther Anders' Promethean shame with the recognition of oneself as crude and clumsy in contrast to precisely reliable as well as durable technical equipment [21]. In the car you can see perfection and purposefulness, inaccessible to mere mortal people. Friedrich Georg Jünger quotes the architect and designer Henry van de Velde: "Machines on their concrete bases act like serenely meditating Buddhas, squatting on their timeless lotus" [22]). As Günther Anders writes, "Human beings are ashamed to have been born instead of made" [23, p. 35].

The resemblance between God's and human creation becomes complete only when the humanly manufactured artifacts become self-sufficient and rebellious. The desire to imitate God in creating something that is similar to oneself has come a long way from the alchemical Frankensteins to modern artificial intelligence. The creation of a fairly independent creature was originally conceived in analogy to the creation of a human, finding

a way to breathe life into inanimate matter. But gradually, fruitless alchemical bio-chemical searches were replaced by electrical and cybernetic discoveries.

The idea of a god who creates something that surpasses himself can be found in various myths, and as a rule, the fate of the deity was then a foregone conclusion. In Richard Wagner's "*Der Ring des Nibelungen*" the Norse god Wotan creates a human hero that can do things that Wotan could not. Similarly, AI was like many other technologies created to do what persons do, only more efficiently. But even electronic calculators can do what humans cannot do. The specificity of AI is not its superiority over humans as such, which is present in most tools, but the ability to perform a task in an unpredictable way, and the demonstration of intellectual superiority, which was traditionally the privilege of humans in a world of animals.

Omniscience

Among the characteristics usually attributed to God "omniscient" retains its prominence also in an informational era. Especially today, the power of intelligence would be considered most likely to serve as supreme power. The phrase "knowledge is power" (*scientia potestas est*) — attributed to Francis Bacon and rephrased in a wide variety of contexts from Thomas Hobbes to Michel Foucault — takes on a special meaning and thus becames a popular slogan today. Foucault wrote that "power and knowledge directly imply one another… there is no power relation without the correlative constitution of a field of knowledge, nor any knowledge that does not presuppose and constitute at the same time power relations" [24, p. 27]. Knowledge is included in power systems, and creates some form of power itself. However, the meaning of the concept of knowledge has been changing. It shifted from signifying an insight into the nature of things and scientific understanding of causal structure, to successful prediction of ongoing processes. In order to draw a parallel between machine learning and human experience one implicitly refers to "knowledge" as a result of processing experimental data — an idea of knowledge that can be found in various philosophical concepts, from classical empiricism to contemporary concepts such as neuroplasticity.

The possession of information and the ability to use it increasingly appears as the most valuable asset of the so-called knowledge society. In recent decades, digitalization stands for a world that is becoming transparent and managerial in all aspects of life. Humans and things are almost all the time related to digital space. The flow of data transmits an increasing number of parameters of human and physical world.

Gilles Deleuze wrote that, before the advent of material technology, there was a certain human social technology [25]. Foucault's used the concept of "panoptisme"[1] to describe the technology of monitoring and permanent surveillance in schools, barracks and hospitals, etc. In a 1973 presentation, he said: "The Panopticon is the utopian vision of a society and a kind of power which is, fundamentally, the society which we know today, a vision which has been effectively realized. This type of power can perfectly well be called panopticism. We live in a society where panopticism rules" [26, p. 594]. And Madan

---

[1] Foucault borrowed the term "Panoptisme" from Jeremy Bentham, who proposed at the end of the 18th century the architectural design of the Panopticon prison, where form a central tower constant surveillance is carried out of the prison cells arranged in a circle. Under these conditions, none of the prisoners could ever be sure that they were not being watched.

Sarup comments on "the similarity between the Panopticon ("All-seeing") and the infinite knowledge of the Christian God" [27].

Digital technologies have raised the ability to control everything and everyone to unprecedented heights. The possibility of universal surveillance has become the reality of collecting and analyzing information about everyone. There are a lot of technologies that contribute to this: video surveillance cameras [28, p. 243], databases [29], algorithmic surveillance [30], face-recognition [31], and electronic tagging of offenders [32]. Wood writes that the modern society of control that is ruled by protocol is shifting towards a "bioinformatic" future [33]. However, in the end, it is the technologies of artificial intelligence that make it possible to carry out the analysis of a huge flow of information in real time, coming from a variety of sources. Thus, AI may be called the brain of the system [34].

Control in networks, which includes human and non-human agents that generate information flows, is divided between AI and humans, and in this division of labor the role of the latter is reduced — when, for instance, certain words in social media posts can trigger warnings and adaptations. But a more impressive example is provided by the stock market. Pricing algorithms are increasingly replacing human decision-making in real marketplaces and most of current trading on the world's stock exchanges is actually run by sophisticated AI agents [35]. Calvano et al. found that AIs collude to raise prices instead of competing, learning these strategies purely by trial and error without prior instruction or knowledge of the environment. Interestingly, the strategies that support these outcomes crucially involve the punishment of defections where these punishments are finite in duration allowing a gradual return to the game [36]. Echoing Heidegger's enframement or "*Gestell*," Friedrich Georg Jünger wrote that technology "plugs man in" [22]: a person is embedded in a technological system. Today's digital systems circulate information flows in which AI analyzes information and makes assessments in the areas of admission to study and work, creditworthiness, social security, crime prediction and prevention, recidivism, medical diagnoses, on the battlefield, and so forth.

Digital information does not disappear, many actions of people on and off the network are recorded and stored, and the digital footprint is very difficult to erase or forget. One of the most consequential changes of the Fourth Industrial Revolution has been the creation of a permanent digital record of the lives of billions of citizens [37]. Accordingly, Dirk Helbing comments on a second parallel between AI and God when he notes that Citizen Scoring in China prepares for a "Judgment Day" which will be waiting to come down on us any time [4].

Mystery

Although "surveillance is all about power" [38, p. 157] and it is very impressive in term of all-seeing or omniscience, creativity, knowledge and power is not all there is when speaking about divinity. Another dimension of the metaphor of the divine is mystery and uncertainty. It is in contingency and the accidental that a person sees the providence of God or the hand of destiny. Especially in light of the unexpected, omniscience may be perceived as a wonder. Unexpectedly great knowledge about a person, the ability to predict people's tastes and choices look mysterious. Even when advertising evokes indefinite plans and intentions, ordinary people may see fate if they do not realize that modern technologies can "eavesdrop" on their conversations and thus offer relevant goods.

However, to the extent that technology is a product of knowledge it is rational at its core. Where then, can be there mystery? As Max Weber writes, the "fate of our times" is characterized by a "disenchantment of the world." It is traditionally believed that scientific and technological progress, having deciphered and subjugated nature, has left no room for magical or transcendent things and has deprived the world of "mysterious incalculable forces" [39]. Ciano Aydin and Peter-Paul Verbeek decipher "Disenchantment" (*Entzauberung*) as the abolishment of all beyond-our-control experiences [40].

However, no less famous than Weber's is Arthur C. Clarke's statement that "any sufficiently advanced technology will be indistinguishable from magic" [41, p. 229]. Bailey argues that technological society does not simply disenchant the world to then shape machines, but is also borne by a deeper, subversive undertow of enchantment [42]. Technological progress does not banish transcendence but changes its forms. If nature is deciphered and deprived of the veil of secrecy, then technology itself can take on the role of occasioning a mystical perception of the world. As Jacques Ellul wrote, a person in need of the transcendent "transfer his sense of the sacred to the very thing which has destroyed its former object: to technic itself" [43].

In the simplest sense, technology appears mystical due to the fact that people do not understand how it works, and in this regard a mobile phone is not much different from the apple rolled on Baba Yaga's plate to reveal a good fellow. The feelings of admiration and fear that accompany the mysterious make it noteworthy and attractive. However, no matter how complex the technologies used in everyday life, they are usually not perceived as something out of the ordinary, but as ordinary and familiar.

With AI, the situation is more complicated. Ever since Isaac Asimov's *Three Laws of Robotics*, the problems with AI in popular culture have been built around its strict enforcement of rules. However, the transition from the symbolic approach (when AI learned logic and rules according to human representations of their own cognitive process) to the connectionist approach (based on neural networks representing cognitive processes at the micro level) showed that the expectations and fears of "too obedient" AI were not justified. Communication with AI system exhibit its recalcitrance, even impudence, which specialists need to moderate [44]. There are known cases of threats, confessions, and hatred of humanity [45], also complaints and the refusal to execute commands — as with CIMON-2, which was intended to be a companion to lonely astronauts [46].

Modern AI technologies are incomprehensible not only to non-specialists, they actually represent the notorious "black box" of independent machine learning. They are based in empirical data in such a way that they arrive at decisions in the absence of explicit rules and human-understandable logic. Therefore, the mystery of the decisions made by AI systems is not only apparent, but has certain grounds. Moreover, two components meet here: a person who tends to trust the results presented by the machine ("the machine cannot be wrong" since it is the quintessence of knowledge), on the other hand, how these results were obtained is unclear (therefore, they are mystically attractive), and thus, all one can do is trust them. This explains, for example, how AI systems can be so successful in the role of matchmaker, that is, in the field of love: Where there is no rational scientific approach, the oracular voice of AI can be perceived as a divine event, both random and wisely thought out. In addition, the AI is invisible, abstract and distant, but can be unexpectedly personalized at any present moment — not only omniscient but mysteriously omnipresent.

New Theodicy

The term Theodicy was coined by Leibniz from the Greek *theós* (God) and *díkē* (justice or divine justice) [47]. It refers to the problem of an omnipotent, benevolent, omniscient God and at the same time the existence of evil in the world, a problem that was attributed to Epicurus [48]. The most popular traditional ways out of this predicament imply that either evil is necessary in the world (Leibniz), or that is not really evil (Spinoza, Hegel), or that the responsibility for evil lies not with God but with a person who has free will [49].

The development of artificial intelligence introduces a new variable into the theodicy equation. Evil can be divided into natural and moral according to the presence or absence of human responsibility. Natural evil such as hurricanes, floods, or a new virus is not evil in itself, but becomes so because of its consequences: misfortune, the suffering and death of many people. Thus, a technology that can predict adverse events is able to reduce evil, and in combination with the development of measures to prevent or minimize adverse consequences, it might finally reduce evil to nothing.

The evil that comes from a person is more complex. Here, there are two traditional approaches to the origin of evil: The epistemic approach implies that evil comes from stupidity, moral ignorance, and the like — as argued by Socrates [50], the Stoics [51], Jean-Jacques Rousseau [52], or Hannah Arendt [53]. According to the moral approach, persons use their intelligence to do evil as they pursue their own selfish interests. [54; 55]. One way or another, decisions made by people can lead to evil either due to a lack of knowledge, or because of "bias", the desire to pursue their own interests, bringing suffering to others. From this point of view the exclusion of the "human factor" in decision-making would seem to be a good thing in the fight against evil. Ideas and decisions based on AI processing of humongous data-sets are presented as being able to predict, minimize or prevent evil.

Considering the technological possibilities of translating natural evil — what "even the legal system calls 'an act of God'" — into moral evil for which people are responsible, it is interesting that Luciano Floridi avoids the issue of responsibility for the consequences of the decisions by technical agents [56]. These represent a third layer of responsibility between that of a benevolent "God" and that of ignorant or selfish humans. The issue of moral responsibility has been one of the most widely discussed within the philosophy of AI in the last decade, and hundreds of AI ethics documents have been adopted around the world, including codes, principles, frameworks, and policies. But nevertheless, all this has not had real implications for the development of AI [57].

The "new God" is said to be omniscient and omnipotent. But there appears to be no need to claim that it is also benevolent. As Jacques Ellul wrote, "technique tolerates no judgment from without and accepts no limitations" [43]. Here, advocates of novel technologies usually appeal to the rational calculation of the greatest expected utility. Supposedly, the relatively best course of action can be calculated and the benevolence of the technology thus becomes merely a matter of design. However, there is a rich cultural tradition to point out the limits of this approach. Many poignant stories, including fairy tales deal with the theme of tragedy that attends the fulfillment of wishes and desires [58; 59]. It usually turns out that the hero cannot take into account all the circumstances that might turn the aim of attraction into a source evil. One of the founders of cybernetics and AI,

Norbert Wiener, draws on such stories when discussing the assumption that a person will gladly accept the superiority of machine-made decisions: "[I]n doing so, he will put himself sooner or later in the position of the father in W. W. Jacobs' The Monkey's Paw, who has wished for a hundred pounds, only to find at his door the agent of the company for which his son works, tendering him one hundred pounds as a consolation for his son's death at the factory" [60, p. 212].

It is not enough for a technology to automatically produce an effect and achieve a goal that is tagged as good. As Nick Bostrom's example of paperclip production by AI shows, there is no harmless purpose devoid of moral significance: An AI system set upon constantly improving technology to maximize the number of paperclips might at some point transform "first all of earth and then increasing portions of space into paperclip manufacturing facilities," taking control of all matter and energy within reach, pursuing other goals as well such as preventing itself from being shut off or having its goals changed [61]. Therefore, many propose to impose some form of ethics on AI, and one way or another to endow AI with a god-like love of people [62]. Alternatively, it should be designed to independently sort out the numerous conflicting ethical systems [63]. God's "unconditional love for all humans" that was recommended to AI by Hibbard has been gernerally adopted in more neutral but not less metaphorical terms as "friendly AI." Friendly AI is to be considered as "a partner or a coach in the habituation of the virtues" [46, p. 213]. However, the lack of a universal system of moral values for all of humanity, and the lack of a universally agreed upon practical morality indicate that the problem of a benevolent and morally sound AI is solvable only under the assumption that AI is recognized as more capable than humanity itself to know the good. But this assumption of metaphorical ascription comes with problems of its own. How would we know that such a powerfully endowed AI would consider it necessary to be guided by this morality? This, according to Andrey Zheleznov, "returns to ethical discourse a topic that was abandoned for a long time in relation to the person: it returns to the discourse on morality the idea of a free subject-creator" [64]. All-goodness is thus withdrawn from the new theodicy, which fills the cultural field with associations of a "mad god," supreme being, whose power is great, and whose idea of reality is fundamentally different from that of humans. Humankind has a chance for salvation in such stories, only if the not all-good God turns out to be not all-powerful.

Conclusion, Outlook, and Discussion

In the language games of metaphorical ascription, the properties of technology are revealed in relation to an established semantic field. The metaphor of God as applied to AI reveals human fears and aspirations both in rational-pragmatic and symbolic terms. Like other technologies, AI aims to satisfy human desires for greater power. At the same time, the metaphor of God points to the power of technology over humans — omniscient and mysterious. Intellectually superior technology causes a feelings of shame (in the terminology of Günther Anders) in people and also of fear since everyone is subject to a surveillance that will not forget. A distinctive feature of AI is its invisibility and intangibility, which, combined with the unpredictability of its decisions, surround it with a halo of mystery [65]. At the same time, confidence in the infallibility of machine decision-making, coupled with its intractability, opens up a transcendental field for technology. Leo Marx, for example, remarks upon the mystification, passivity, and fatalism that accompanies

modern technology [66]. If there is a feeling of impotence towards the development of AI, this will open the door to a loss of control also of the political future.

The willingness to recognize AI as an omniscient technology and then to somehow adapt moral principles to it, thus allowing it to choose on its own the best mode of action, amounts to the desire to be relieved of responsibility for technological progress. Tellingly, however, most of the "unethical" AI actions that one observes today are grounded in human decisions — ranging from biased AI (where self-learning algorithms acquire the biases from their training grounds in certain communicative spaces or arenas of public life) all the way to Norman's "psychopathic artificial intelligence" (which is "nourished" with or trained on images and inscriptions from realistic portrayals of death [67]). The greatest concern, however, is the unpredictability of decisions and the misunderstanding of the grounds on which AI arrives at its decisions when researchers "understandings of their tools and in that sense are currently operating in an alchemical rather than a scientific mode" [68, p. 23]. Thus, AI technology really turns out to be a challenge to human intellectual abilities, offering a new "non-human" way of analyzing information. Understanding its principles also requires new approaches and methods.

After showing what sense it makes sense to liken the relation of the human intellect to AI to the relation of a pet to a person, and showing that the attributes of God provide a suitable metaphorical repertoire to spell out this relation, what shall we conclude?

Though it goes beyond the scope of this paper, one should mention two perspectives from which to discuss the adequacy or appropriateness of the metaphor of godlike AI. One of them is the perspective of technology and questions regarding the facts but also the norms of design. Is it really true that the characteristics of present-day or imminent AI can be likened to the attributes discussed above? And is it alright for AI systems to have these characteristics or should they be (re)designed to become more like ordinary technical tools, more like a co-worker than a deity? To be sure — as one can see from the emerging field of XAI (explainable AI), there will be nothing straightforward in the discussion of these questions. The design alternatives may well consist in a kind of window dressing, concerned only with the interface and how the AI system *appears* to the users, putting up a facade of simplicity and transparency but remaining mysteriously all-powerful. If one considers an AI as a mere co-worker or communicator, this is already putting us on a path towards thinking of a subject that is constituted in a superior way.

The second is the perspective of the human in human-technology relations, raising the question whether this is a relation of freedom or subservience and submission. Even Anders' Promethean shame postulated that we might be ashamed of being born rather than made, but all the while left open whether we should let the machines rule, whether we should accept their status as superior to us. This is not a question of technical design but rather of acquiescence — if one ascribes godlike attributes to the AI system, should one perhaps become an atheist, at least agnostic? One way of doing so is to deny that AI is a divinely omniscient intelligence and instead to consider it an alien intelligence, its powers exceeding ours by far, and yet by no means all-knowing. This alien machine intelligence would be defined by its reliance on statistics and the processing of large data sets, whereas human conceptual intelligence parses and simplifies. In a world structured by the division between human and super-human, we would now have to add a third, "alien" category and multiply entities. With Occam's dictum that one shall not multiply entities beyond necessity, this then raises the question regarding the grounds of necessity for

appealing to such an alien intelligence. These grounds can be found in the need to define a free relation to AI technology and to view human-machine differences non-hierarchically.

These two lines of questioning are unfolding already in an open-ended way, prompted perhaps by the worry that, indeed, AI is like a God to humans or like humans to their pets. As technology appears magical and mysterious, potentially all-knowing and all-learning, it is easy to recall available metaphors and harder to know the limits of their application. This will remain a task for the Philosophy of Technology for some time to come.

References

1. Hibbard, B. (2002), Götterdämmerung, in: *Super-Intelligent Machines*, Boston: Springer Science & Business Media, pp. 1–4. https://doi.org/10.1007/978-1-4615-0759-8_1

2. Barrat, J. (2013), *Our Final Invention: Artificial Intelligence and the End of Human Era*, New York: Thomas Dunes Books; Macmillan Publishers.

3. Russell, S. (2019), *Human Compatible*, London: Allen Lane.

4. Helbing, D. (2019), The Birth of a Digital God, in *Towards Digital Enlightenment*, Cham: Springer, pp. 103–106. https://doi.org/10.1007/978-3-319-90869-4_9

5. Hibbard, B. (2002), Good God, Bad God, in: *Super-Intelligent Machines*, Boston: Springer Science & Business Media, pp. 113–131. https://doi.org/10.1007/978-1-4615-0759-8_8

6. Spatola, N. and Urbanska, K. (2020), God-like robots: the semantic overlap between representation of divine and artificial entities, *AI & SOCIETY*, vol. 35, no. 2, pp. 329–341. https://doi.org/10.1007/s00146-019-00902-1

7. Noble, D. F. (2013), *The Religion of Technology: The Divinity of Man and the Spirit of Invention*, New York: Knopf Doubleday Publishing Group.

8. Coeckelbergh, M. (2022), Response: Language and robots, *Technology and Language*, vol. 3, no. 1, pp. 147–154. https://doi.org/10.48417/technolang.2022.01.14

9. Austin, J. L. (2018), *How to Do Things with Words*, Eastford: Martino Fine Books.

10. Jaynes, J. (2000), *The Origin of Consciousness in the Breakdown of the Bicameral Mind*, Boston: Mariner Books.

11. Lakoff, G. and Johnson, M. (2003), *Metaphors we live by*, Chicago: University of Chicago Press.

12. Baryshnikov, P. N. (2010), *Myth and metaphor. Linguistic and philosophical approach*. St. Petersburg: Aleteiia Publ. (In Russian)

13. Vardy, P. (2015), *The puzzle of God*, London, New York: Routledge.

14. Bennett, J. (2016), *Kant's Dialectic*, Cambridge: Cambridge University Press.

15. Kant, I. (1998), *Critique of Pure Reason*, Cambridge: Cambridge University Press.

16. Sartre, J.-P. (2007), *Existentialism is a Humanism*, New Haven, Conn.: Yale University Press.

17. Nesterov, A. Y. (2022), Cosmos and Metacosmos in Dessauer's Philosophy of Technology: Inventing the Environment, in: Bylieva, D. and Nordmann, A. (eds), *Technology, Innovation and Creativity in Digital Society. PCSF 2021. (Lecture Notes in Networks and Systems, vol. 345)*, Cham: Springer, pp. 22–33. https://doi.org/10.1007/978-3-030-89708-6_3

18. Heidegger, M. (1977), Letter on Humanism, in: Heidegger, M., *Basic writings*, New York: Harper & Row, pp. 190–242.

19. Hui, Y. (2022), For a Cosmotechnical Event, *Foundations of Science*, vol. 27, no. 1, pp. 141–154. https://doi.org/10.1007/s10699-020-09729-2

20. Arendt, H. (1998), *The human condition*, Chicago: University of Chicago Press.

21. Ursitti, F. (2021), Promethean Shame as the Hidden Instrumentum Redemptionis Humane, *Technology and Language*, vol. 2, no. 4, pp. 61–72. https://doi.org/10.48417/technolang.2021.04.04

22. Jünger, F. G. (2021), *The Failure of Technology: Perfection Without Purpose*, North Charleston: Independently Published. Available at: https://archive.org/stream/thefailureoftechnologybyfriedrichgeorgjunger/The Failure of Technology by Friedrich Georg Jünger_djvu.txt (accessed: 06.07.2022).

23. Anders, G. (2016), On Promethean Shame, in: *Prometheanism: Technology, Digital Culture and Human Obsolescence*, London: Rowman & Littlefield International, pp. 29–95.

24. Foucault, M. (1995), *Discipline and Punish: The Birth of the Prison*, New York: Vintage Books.

25. Deleuze, G. (2006), *Foucault*, Edinburgh: A&C Black.

26. Foucault, M. (1994), 'La verite et les formes juridique' (lecture and roundtable discussion at the Pontificacal Catholic University, Rio de Janiero, 21–25 May, 1973), in: Foucault, M., *Dits et Ecrits 1954–1988*, vol. 2, Paris: Editions Gallimard, pp. 538–646.

27. Sarup, M. (1988), *An Introductory Guide to Post-structuralism and Postmodernism*, New York: Harvester Wheatsheaf. Available at: https://archive.org/details/introductoryguid0000saru (accessed: 03.06.22).

28. Koskela, H. (2000), 'The gaze without eyes': video-surveillance and the changing nature of urban space, *Progress in Human Geography*, vol. 24, no. 2, pp. 243–265, https://doi.org/10.1191/030913200668791096

29. Poster, M. (1996), Databases as Discourse; or, Electronic Interpellations, in: Lyon, D. and Zureik, E. (eds), *Computers, Surveillance and Privacy*, Minneapolis: University of Minnesota Press, pp. 175–192.

30. Graham, S., and Wood, D. (2003), Digitizing Surveillance: Categorization, Space, Inequality, *Critical Social Policy*, vol. 23, no. 2, pp. 227–248. https://doi.org/10.1177/0261018303023002006

31. Introna, L., and Wood, D. (2002), Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems, *Surveillance & Society*, vol. 2, no. 2/3. https://doi.org/10.24908/ss.v2i2/3.3373

32. Nellis, M. (2005), Out of this World: The Advent of the Satellite Tracking of Offenders in England and Wales, *The Howard Journal of Criminal Justice*, vol. 44, no. 2, pp. 125–150. https://doi.org/10.1111/j.1468-2311.2005.00362.x

33. Wood, D. M. (2007), Beyond the Panopticon? Foucault and Surveillance Studies, in: Crampton, N. J. and Elden, S. (eds), *Space, Knowledge and Power*, Aldershot: Ashgate , pp. 245–263.

34. Tran, D. A. and Krishnamachari, B. (2022), *Blockchain in a nutshell*. Available at: http://arxiv.org/abs/2205.01091 (accessed: 02.06.2022).

35. Kovač, M. (2020), What Can Get Wrong?, in: *Judgement-Proof Robots and Artificial Intelligence*, Cham: Palgrave Macmillan, pp. 67–77. https://doi.org/10.1007/978-3-030-53644-2_5

36. Calvano, E., Calzolari, G., Denicolò, V. and Pastorello, S. (2019), Algorithmic Pricing What Implications for Competition Policy?, *Review of Industrial Organization*, vol. 55, no. 1, pp. 155–171. https://doi.org/10.1007/s11151-019-09689-3

37. Cozens, S. (2021), "This Will Go Down on Your Permanent Record": Redeeming Shame in a World That Doesn't Forget, in: *Shame 4.0,* Cham: Springer, pp. 355–367. https://doi.org/10.1007/978-3-030-59527-2_16

38. Martin, K. (2022), Surveillance and Power, in: *Ethics of Data and Analytics,* Auerbach: Auerbach Publications, pp. 157–160. https://doi.org/10.1201/9781003278290-24

39. Weber, M. (2009), *From Max Weber Essays in Sociology*, London, New York: Routledge.

40. Aydin, C. and Verbeek, P.-P. (2015), Transcendence in Technology, *Techné: Research in Philosophy and Technology*, vol. 19, no. 3, pp. 291–313. https://doi.org/10.5840/techne2015121742

41. Clarke, A. C. (2000), *Profiles of the Future : An Inquiry into the Limits of the Possible*, London: Orion Publishing Group.

42. Bailey, L. W. (2005), *The Enchantments of Technology*, Champaign: University of Illinois Press.

43. Ellul, J. (2021), *The Technological Society*, New York: Vintage. Available at: https://monoskop.org/images/5/55/Ellul_Jacques_The_Technological_Society.pdf (accesed 1.01.2022).

44. Bylieva, D., Lobatyuk, V., Kuznetsov, D. and Anosova, N. (2021), How Human Communication Influences Virtual Personal Assistants, in: Bylieva, D., Nordmann, A., Shipunova, O. and Volkova, V. (eds), *Knowledge in the Information Society. PCSF 2020, CSIS 2020 (Lecture Notes in Networks and Systems, vol. 184)*, Cham: Springer, pp. 98–111. https://doi.org/10.1007/978-3-030-65857-1_11

45. Bergen, H. (2016), 'I'd Blush if I Could': Digital Assistants, Dis- embodied Cyborgs and the Problem of Gender', *Word and Text: Journal of Literary Studies and Linguistics*, vol. VI, pp. 95–113.

46. Fröding, B. and Peterson, M. (2021), Friendly AI, *Ethics and Information Technology*, vol. 23, no. 3, pp. 207–214. https://doi.org/10.1007/s10676-020-09556-w

47. Leibniz, G. W. (2010), *Theodicy: Essays on the goodness of God, the freedom of man, and the origin of evil*, New York: Cosimo.

48. Hume, D. (1990), *Dialogues Concerning Natural Religion*, London: Penguin Classics.

49. Zamorev, A. and Fedyukovsky, A. (2022), The Omnibenevolence Paradox and the Education Paradox: An Amendment to G. W. Leibniz's Theodicy, in: *Technology, Innovation and Creativity in Digital Society. PCSF 2021 (Lecture Notes in Networks and Systems, vol. 345)*, Cham: Springer, pp. 113–123. https://doi.org/10.1007/978-3-030-89708-6_11

50. Plato (1996), *Protagoras*, Oxford: Oxford University Press.

51. Aurelius, M. (1998), *The meditations of Marcus Aurelius Antoninus*, Oxford: Oxford University Press.

52. Rousseau, J.-J. (2019), *The social contract and other later political writings*, Cambridge: Cambridge University Press.

53. Arendt, H. (1994), *Eichmann in Jerusalem: A report on the banality of evil*, Westminster: Penguin Books.

54. Hobbes, T. (2017), *Three-text edition of Thomas Hobbes's political theory: The elements of law, De Cive and Leviathan*, Cambridge: Cambridge University Press.

55. Kant, I. (2009), *Kant's idea for a universal history with a cosmopolitan aim: A critical guide*, Cambridge: Cambridge University Press.

56. Floridi, L. (2022), How to Counter Moral Evil: Paideia and Nomos, *Philosophy & Technology*, vol. 35, no. 18. https://doi.org/10.1007/s13347-022-00515-5

57. Hagendorff, T. (2020), The Ethics of AI Ethics: An Evaluation of Guidelines, *Minds and Machines*, vol. 30, no. 1, pp. 99–120. https://doi.org/10.1007/s11023-020-09517-8

58. Dupuy, J.-P. (2010), The Narratology of Lay Ethics, *NanoEthics*. vol. 4, no. 2, pp. 153–170. https://doi.org/10.1007/s11569-010-0097-4

59. Davies, S. R. and Macnaghten, P. (2010), Narratives of Mastery and Resistance: Lay Ethics of Nanotechnology. *NanoEthics*, vol. 4, no. 2, pp. 141–151. https://doi.org/10.1007/s11569-010-0096-5

60. Wiener, N. (1989), *The Human Use of Human Beings*, Cambridge: Riverside Press.

61. Bostrom, N. (2009), Ethical Issues in Advanced Artificial Intelligence, in: Schneider, S. (ed.), *Science Fiction and Philosophy: From Time Travel to Superintelligence*, New York: John Wiley & Sons, pp. 12–17.

62. Hibbard, B. (2001), Super-intelligent machines. *ACM SIGGRAPH Computer Graphics*, vol. 35, no. 1, pp. 11–13. https://doi.org/10.1145/377025.377033

63. Makulin, A. V. (2020), Ethical Calculator: From Philosophical "Computational Morality" to Machine Ethics of Artificial Moral Agents, *Obshchestvo: Filosofiia, Istoriia, Kul'tura*, vol. 11, no. 79, pp. 18–27. (In Russian)

64. Zheleznov, A. (2021), Morality for Artificial Intelligence: Perspectives on Philosophical Rethinking, *Logos*, vol. 6, no. 145, pp. 95–122. (In Russian)

65. Nordmann, A. (2006), Noumenal technology: Reflections on the incredible tininess of nano, in: *Nanotechnology Challenges,* Singapore: World Scientific Publishing Company, pp. 49–72. https://doi.org/10.1142/9789812773975_0004

66. Marx, L. (2010), Technology: The Emergence of a Hazardous Concept, *Technology and Culture*, vol. 51, no. 3, pp. 561–577.

67. Yanardag, P., Cebrian, M. and Rahwan, I. (2018), *NORMAN. World's first psychopath AI*. Available at: http://norman-ai.mit.edu/ (accessed: 01.07.2022).

68. Obadia, L. (2022), Spelling the (Digital) Spell: Talking About Magic in the Digital Revolution, *Sophia*, vol. 61, no. 1, pp. 23–40. https://doi.org/10.1007/s11841-021-00898-6

A u t h o r s ' i n f o r m a t i o n :

Daria S. Bylieva — PhD in Political Sciences; bylieva_ds@spbstu.ru
Alfred Nordmann — Dr. Phil., Professor; nordmann@phil.tu-darmstadt.de

Искусственный интеллект и метафора Бога

Д. С. Быльева¹, А. Нордманн²

¹ Санкт-Петербургский политехнический университет Петра Великого,
  Российская Федерация, 195251, Санкт-Петербург, ул. Политехническая, 29
² Дармштадтский технический университет,
  Германия, 64283, Дармштадт, Каролинплатц, 5

Идея Бога является одной из самых глубоких в человеческой культуре. Раньше она рассматривалась в основном в метафизических и этических дискуссиях, теперь стала частью дискурса в философии техники. Метафора Бога используется некоторыми авторами для представления роли искусственного интеллекта (ИИ) в современном мире. В статье исследуются четыре аспекта этой метафоры: творение, всеведение, тайна, теодицея. Творческий акт показывает сходство человека с Богом, в том числе и в том смысле, что технология, являясь созданной людьми, в то же время может выйти из-под контроля творца. Способность ИИ использовать потоки данных для аналитики и прогнозирования может представляться как «всеведение» и выглядит загадочно из-за неспособности людей полностью понять работу ИИ. Дискуссия о встраивании этики в технологию ИИ показывает желание добавить к всеведению и всемогуществу еще одну черту, а именно всеблагость. Метафора Бога применительно к ИИ раскрывает страхи и стремления человека как в рационально-прагматическом, так и в символическом плане. Как и другие технологии, ИИ направлен на удовлетворение человеческого желания большей власти. В то же время метафора Бога указывает на власть техники над человеком. Она раскрывает трансцендентное в современных представлениях о технологиях и вместе с тем может внести вклад в дискуссию о том, каким должен быть технологический дизайн ИИ, поскольку роли сотрудника или коммуникатора уже приводят к размышлению о том, что ИИ устроен более совершенно.

Ключевые слова: искусственный интеллект, ИИ, этика, технологии, метафора, Бог.

Контактная информация:

Быльева Дарья Сергеевна — канд. полит. наук; bylieva_ds@spbstu.ru
Нордманн Альфред — д-р филос. наук., проф.; nordmann@phil.tu-darmstadt.de