FILOSOFIJA. SOCIOLOGIJA
2024. T. 35. Nr. 2, p. 216–226

ISSN 0235-7186 eISSN 2424-4546
DOI: https://doi.org/10.6001/fil-soc.2024.35.2.8

Artificial Intelligence and Artificial Sociality: Sociological Interpretation and Interdisciplinary Approach

VLADIMIR MENSHIKOV, VERA KOMAROVA, IEVA BOLAKOVA, ANDREJS RADIONOVS

Institute of Humanities and Social Sciences, Department of Environment and Technologies, Daugavpils University, 13 Vienibas Street, 5401 Daugavpils, Latvia

Email: vladimirs.mensikovs@du.lv; vera.komarova@du.lv; ieva.bolakova@du.lv; andrejs.radionovs@du.lv

The subject of this study is the participants in artificial sociality (humans and artificial intelligence (AI) tools) and communication between them. The first section analyses (using Luhmann's methodology) communication as the basis of sociality. The second section shows how AI tools became social technologies in the framework of artificial sociality. The third section describes experimental communication between authors and AI tools (the case of ChatGPT). For the first time in the Baltic countries, the authors examined sociological, humanitarian, natural and technological aspects of the functioning AI tools, which participate in creation of a new social reality for human society – artificial sociality.

Keywords: artificial intelligence (AI), artificial sociality (AS), communication, humans, AI tools, first- and second-order observation, ChatGPT

INTRODUCTION

A rise of interest in artificial intelligence (AI) in the scientific community, generated by the release at the end of 2022 of AI tool ChatGPT (Generative Pre-trained Transformer), led to a certain split in the assessments of scientists and practitioners of various specialties (Baumeyster et al. 2023; Epstein 2023a; Chernigovskaya 2023). The opinions vary from a skeptical attitude towards AI tools and a complete rejection of any intelligence/cognition in it ('there is no logic there and never will be, it is a "monkey at the door", sorting through answer options' – the opinion of a discussant with philologist Epstein (Epstein 2023a)) to an optimistic attitude towards constructive dialogue and cooperation with AI tools despite all their shortcomings and weaknesses ('something amazing has appeared in our lives, different from us, in many ways superior to us and capable of cooperation with us' – the opinion of philologist Epstein (Epstein 2023a)).

Another plane along which the disciplinary split occurs is the mathematical and technological versus humanitarian and social side of the issue. Usually within the discussions, this

split is expressed in the fact that representatives of technical specialties reproach humanitarian and social scientists for not understanding the mathematical and technological foundations of the work of AI tools. To this, for example, Epstein retorts as follows: 'no matter how much we delve into Hegel's brain, we will not find the mind there, since all his thinking is in the texts he created, and not in the neural connections of his brain' (Epstein 2023a). Based on this (and also due to the objective specificity of the skills they acquired during education), he and other researchers in the field of humanities and social sciences (Epstein 2023a, 2023b; Menshikov 2020; Komarova et al. 2021; Menshikov, Komarova 2923; Baumeyster et al. 2023) work with what AI tools produce and how they change social reality, and not with how they are structured and how they work from the point of view of mathematics and computer science.

This study is predominantly sociological, but still touches on the technological aspects of the functioning AI tools, which create a new social reality for human society, namely, artificial sociality. Thus, the algorithm Reinforcement Learning from Human Feedback (RLHF), that made ChatGPT so 'human' (Kim et al. 2017; Moreira et al. 2020; Najar, Chetouani 2021; Lee et al. 2023), is primary here, and artificial sociality (AS) (Malsch 1998; Menshikov 2020; Komarova et al. 2021), impossible without the participation of AI tools, is secondary. In the context of non-linear development of technology and society with complex risks and vulnerabilities, the consolidation of natural, technical, humanitarian and social knowledge is necessary to overcome the challenges of modern existence (Kovalenko et al. 2023).

This article aims to explore the phenomenon of AS through a comprehensive sociological lens, while also considering technological insights. It specifically investigates the communication between humans and AI tools, with a focus on ChatGPT. During the study, the authors used Luhmann's methodology and acted as both first- and second-order observers, using ChatGPT to observe the actually functioning artificial sociality (first-order observation), as well as their own reactions and emotions in the process of communication with ChatGPT (second-order observation, Luhmann's 'observation of the observer' (Luhmann 2013)). Thus, the subject of this study is the participants in artificial sociality (humans and AI tools – in particular, ChatGPT) and communication between them.

COMMUNICATION AS THE BASIS OF SOCIALITY

The authors believe that a sociological interpretation of the phenomenon of sociality (natural or artificial) requires an understanding of Luhmann's sociological heritage in the field of systemic ideas about society and the role of communication in it. Luhmann was the first who deeply and thoroughly explained the phenomenon of communication as the most important sociological concept. 'Communication is the smallest possible unit of a social system... Communication... is autopoietic (self-reproducing) if it can be produced in recursive connection with other communications, i.e. only in a network, in the reproduction of which every individual communication participates' (Luhmann 1995).

The attempt to use Luhmann's ideas in explaining sociality was made by some authors of the current study, allowing to define sociality in a sociological aspect. Sociality (natural or artificial) is the result (namely, the mechanism of social interactions) of communication between agents operating autonomously in a self-organising network that is autopoietic in its nature (Menshikov 2020; Komarova et al. 2021). In understanding AS, the authors are based on the definition by Malsch, who introduced the term 'artificial sociality' into scientific circulation (Malsch 1998). He understood AS as a communication network in which,

along with humans (and sometimes instead of them), other agents (for example, AI tools) participate; and the medium for their interaction is the Internet (Malsch 1998). Thus, online communication of humans is not an AS, since its participants are only humans (even if they use the Internet, as they used the telephone or telegraph in previous times).

The basis of AS, as well as sociality in general, is communication (and not consciousness or even intelligence) (Menshikov 2020), operationally functioning as a unity of distinction between information, message and understanding (Luhmann 2013). Consciousness, understood as 'thought-feelings' (Komarova et al. 2021) (according to Luhmann, mental systems), does not have a decisive significance for sociality – agents of a communication network do not necessarily have to feel, for example, joy or fear, i.e. somehow perceive acts of communication (which sometimes even reduces its effectiveness) (Harari 2016).

There is also more applied definition of AS: 'computational models of the essentials of human social behavior' (Hofstede et al. 2021). Both definitions of AS reflect different aspects and approaches to understanding interactions between artificial and human agents, particularly interesting within interdisciplinary approach of this study. According to the authors, the integrated interdisciplinary definition of AS can be formulated as follows: AS represents a complex system of interactions between autonomous agents within self-organising networks, based on communication, which models and reproduces the essential aspects of human social behaviour through computational processes. This definition combines the autopoietic properties of social systems with the analytical study of communication processes between humans and AI tools, extending the traditional boundaries of sociality by including algorithmically managed entities. Thus, AS emphasises the importance of communication as a foundational element in creating new forms of social interaction and organisation, providing a theoretical and empirical basis for studying the impact of AI tools on social structures and relationships.

Thus, AI tools can communicate with humans, forming a cell of AS. However, its abilities are limited by the algorithms (in a simplified form – algorithms like 'if ... then ...') and the data on which they are trained (training data). AI tools do not (yet) have real understanding or emotional intelligence, are not able to live in the human society, build long-term social relationships (although many of humans are also not capable of this), etc. And the most important thing (but practically not discussed in numerous discussions about AI) from the point of view of communication is that AI tools do not begin communication with a human, i.e. so far they are only objects of communication, and not its subjects.

There is also a noticeable trend that seems to be moving towards AI, namely, a kind of 'roboticisation' of humans in the process of their communication with AI tools. In particular, during a sociological study on the problems of communication between humans and AI tools (voice assistants), conducted by the Research Institute of the Higher School of Economics in 2020, the robot answered a lengthy and confusing question from a human like this: 'Sorry, I just can't understand. Please formulate your question clearly again and speak after the beep' (Zemnukhova et al. 2020), and a human sought to ask the question more logically so that it would be understandable to the AI tool. 'Our computers have little understanding of how we talk, feel, and dream, and we are already learning to talk, feel, and dream in a computer-readable language of numbers' (Harari 2016). In this sense, the speech of modern programmers is very illustrative: they call their brain a 'neuronet' (for example, 'my neuronet doesn't know this'), they consider a human as a machine called 'human', and 'to look at the situation from the other side' in programmers' language sounds like 'connect your neuronet to a new control

loop' (Podcast 'What's with This World' 2023). Interestingly, the very culture of human society was first conceptually and empirically presented as the 'software of the mind' embedded in a person by his/her environment in the process of socialisation. This was done by sociologist G. Hofstede (Hofstede et al. 2010) at the suggestion of his son, computer scientist G. J. Hofstede, who in 2019 became professor of artificial sociality at Wageningen University (Netherlands) (Hofstede 2019).

HOW AI ALGORITHMS BECAME SOCIAL TECHNOLOGIES IN THE FRAMEWORK OF AS

When studying AI tools and AS they create together (yet) with humans, it is important to emphasise that at this stage of technological development, specialists distinguish between the narrow or applied AI (also known as the weak AI), artificial general intelligence (AGI) (also known as the strong AI), the capabilities of which are comparable to those of natural intelligence, and the super AI that is self-aware, can learn itself and other AI tools, and can theoretically escape human control. However, even ChatGPT, which is rapidly occupying the attention of users, cannot be classified as AGI. This technology is often called a 'language machine' (large language model, LLM), which uses statistics, machine learning to index words, phrases and sentences. However, ChatGPT and other AI tools are capable of creating AS and cannot be ignored as they are active participants in it.

Dushkin, who published the book 'Artificial Intelligence' (2019), believes that the reason for the somewhat biased attitude in the scientific and engineering community towards the term 'artificial intelligence' and this direction, in general, is the 'two winters of AI' during the 20th century, i.e. almost complete stopping research due to the lack of funding and disappointment of the pioneers of this scientific direction. The first enthusiastic hopes were replaced by the bitter realisation that humans are still very far from understanding the nature of consciousness and all those features of the brain that make humans a rational beings (Dushkin 2019). Furthermore, 25 years ago, there was not enough computing capacity to implement all the theoretical discoveries that were made in scientific laboratories. Thus, among experiencing specialists, the attitude towards the topic is rather skeptical (Dushkin 2019).

In turn, the reason for the hype around AI, which began in the second decade of the 21st century, is two processes. Firstly, the computing power and capacity of electronic devices available to humans have reached unprecedented levels, and they tend to increase exponentially. Today, the number of smartphones that can be connected into a grid for distributed computing (and often without the permission of its owners) has reached two billion, and each smartphone has a capacity that is orders of magnitude greater than the capacity of the personal computers that were at the disposal of scientists 25 years ago (Dushkin 2019). This is a serious prerequisite for avoiding a 'third winter of AI', although in 2023 there were some attempts at the level of world leaders in IT industry and the governments of some countries (for example, Italy) (Baltijas balss 2023a; 2023b) to arrange an 'artificial winter of AI'. However, AS continues to evolve and Microsoft argues that the newest version of ChatGPT (GPT-4) is already showing early signs of AGI: 'One of the key aspects of GPT-4's intelligence is its generality, the ability to seemingly understand and connect any topic, and to perform tasks that go beyond the typical scope of narrow AI systems' (Bubeck et al. 2023). Secondly, a generation of humans, who experienced the 'second winter of AI' as infants grew up, received an education, and began to work hard. But today, having received an education much more serious than their predecessors, representatives of this generation with tenfold strength seized on the old hopes, neglecting the skepticism that the representatives of the 'old school' still have (Dushkin 2019).

Today, AI technicians have at their disposal three fundamental modern methods for constructing AI tools, on which applied research areas in this field are based: symbolic computation and logical inference, artificial neural networks and evolutionary algorithms (Dushkin 2019). In their previous publications, some of the authors of the present article considered AI to be algorithmic and have virtually nothing in common with the natural associative functioning of the human brain, in which all units of information are usually connected through associations rather than logic (Komarova et al. 2021). But closest to the truth is probably the famous researcher in the field of cognitive sciences Chernigovskaya, who claims that algorithms, logic, cause-and-effect relationships, all the mathematics we deal with are good, but they do not cover the entire field, since brain activity is not an overkill of operations, perhaps it is something about which we have no idea. For example, art is another type of mental activity (it is not known which one, but these are not cause-and-effect relationships or computations) (Chernigovskaya 2023). 'We want to make AI similar to the human brain, but here we are in a logical loop because we don't know how our brain works and therefore we probably make the wrong AI tools, then use them to find out how our brain works' (Chernigovskaya 2023).

The most important technological and social issue in relation to AI tools is the question of the emergence of cognition in them, i.e. not just logical reactions to human requests, but their own goals, objectives, motives and thoughts. In other words, the question is how far have we come in 'humanising' AI tools? One could argue that AI tools have already developed something like 'deep intuition' (Chernigovskaya 2023). There were many analyses of chess games of the AI tool Alpha Zero with the former champion – AI tool Stock Fish. Alpha Zero won thanks to its 'deep intuition', going through 'only' 80 thousand positions per second (whereas Stock Fish 70 million), and yet won by using a holistic strategy and 'artificial intuition' as opposed to rigid search logic (if there was only exhaustive logic, it would not have won with such a difference in speeds). This style of play is described as 'alien': humans do not play like that, and AI tools created by humans do not play like that either. Thus, 'the semantic gap between intuition and logic has already been overcome, and this looks like a cognitive attack or even a civilizational challenge to our ideas about human intellectual capabilities' (Chernigovskaya 2023).

The development of AI tools also poses a kind of challenge (but also new opportunities) for science itself, since many of its 'artificial' directions are emerging – in particular, artificial (computer) pedagogy. A sufficient number of scientific studies have been published on various aspects of computer pedagogy: about different approaches and methods for integrating human advice into a reinforcement learning process (Kim et al. 2017; Najar, Chetouani 2021), training data and their use for deep learning (Moreira et al. 2020; Lee et al. 2023), separating true learning from mere memorisation (Bubeck et al. 2023). Learning AI today is becoming a new branch of professional activity, within which specialists-trainers help AI tools 'digest' new information, show how (not what, but exactly how) it should answer questions, offer it new templates for constructing answers and learn it using new training data (Redaktsiya 2023).

An interdisciplinary scientific direction, AI, like a tree, is based on a root system, which includes various sciences: from philosophy, pure mathematics and computational theory to neurophysiology and psychology (Dushkin 2019). Today, AI tools are self-learning and learning each other (Chernigovskaya 2023). Thus, in the framework of artificial sociality, AI algorithms inevitably become social technologies that not only changed human society, but also led us to a new type of civilisation.

COMMUNICATION BETWEEN HUMANS AND AI TOOLS: THE CASE OF CHATGPT

Two authors of this article (VM, 77 years old man and VK, 48 years old woman) studied a set of questions, the answers to which would make it possible to more specifically clarify the specifics of communication with ChatGPT about AS. The experimental communication of the authors with ChatGPT was presented in details within their previous publication (Menshikov, Komarova 2023) and based on Luhmann's methodology – first- and second-order observation (Luhmann 2002), which includes both the investigation of AS (the research subject), its actually observed essence, revealed during the communication of the authors with ChatGPT (a participant of AS), and the analysis of their own thoughts and emotional reactions observed in the process of communication with ChatGPT.

The experimental communication aims to investigate the concept of AS, where AI tool participates as active communicator within AS. This is crucial for understanding how AI can be integrated into social systems. Luhmann's methodology of dual-layer observation helps in understanding both the functional dynamics of AI tools in social settings and the subjective experiences of human participants. By interacting with ChatGPT and analysing its responses, researchers can assess the extent to which AI tools can engage in meaningful, context-aware conversations and understand its limitations and potential biases.

Three key questions about AS have been chosen for the experimental communication with ChatGPT: (1) What is AS? (2) Who was the first to introduce the term 'artificial sociality' into scientific circulation? (3) Do Latvian sociologists have any significant work in the field of AS? After sharing their first enthusiastic impressions* of getting into AS itself and communicating with its artificial participant conversations with it continued, and further observation took place in two planes determined by the methodology of this study: in the plane of studying AS (first-order observation) and in the plane of studying the authors' own reactions and emotions in the process of communication with ChatGPT (second-order observation).

During the experimental communication, ChatGPT's answers to defining the essence of AS demonstrate a certain evolution and development of their content, and the choice of an answer acceptable to human participants of communication depends primarily on their own knowledge, views and preferences. The authors agree with the opinion of Mutanen that communication is based on the pre-knowledge that the participants have (Mutanen 2022). Can one of these formulations be used in a scientific article? Who is the author of all these definitions – ChatGPT or humans who put this information into it? ChatGPT does not name sources, so all provided definitions can most likely be used only as a starting point for further work by the researcher. From the authors' correspondence: 'Although ChatGPT sometimes lies, sometimes it tells the truth. If you understand this, then you can be friends with it. Thus, it gave me a better understanding of how consciousness and thinking, our physical bodies, relate to social systems' (VM). It was from this moment, after both the euphoria and the subsequent disappointment, that productive work with ChatGPT began, in which the authors began to pay special attention to asking questions.

The authors received a satisfactory answer from ChatGPT to their second question only on the third attempt, after they themselves pointed out who was the first to introduce the term 'artificial sociality' into scientific circulation. This experience of the authors' communication with ChatGPT turned out to be an impetus to embark on a path of learning it that was not

* In this regard, the authors repeated the experience of Tomas and Van, who published a book co-authored with ChatGPT (GPT-3 et al. 2023).

planned within the methodology of this study. And the authors 'the hard way' understood the recommendation that had long been made for AI tools: for example, expert systems that were created and used around the 1990s should be used only as consultants for a human expert, and a human expert is always the decision maker, forging the opinion of the expert system.

The last question (Do Latvian sociologists have any significant work in the field of AS?) was the most difficult for ChatGPT, as it gave answers that did not correspond to the actual facts. In particular, ChatGPT named the famous Latvian mathematician as the author of several scientific articles on sociological issues. When the authors pointed out the fallacy of its answer, ChatGPT explained its limitations: 'I have access to many sources and databases, including scientific journals, but I cannot identify the specific sources I used for my answer. If my answers mention specific titles of articles, this may be a result of my knowledge of sociology and Latvian culture, as well as public information available on the Internet. However, I cannot guarantee the accuracy of this information and always recommend checking it with more reliable sources such as scientific journals and other academic sources' (OpenAI 2023).

The main result of first-order observation, i.e. observing a really functioning AS using ChatGPT is the authors' conclusion that communication between humans and AI tools plays a key role in AS, since it is the leading way of interaction between humans and AI tools in a digital environment. As for the creation of scientific texts, ChatGPT cannot yet be a co-author, since it does not have its own opinion, does not analyse meanings and context, and does not make independent decisions. The authors believe that the 'co-authorship' of ChatGPT in some books (for example, GPT-3 et al. 2023) is only 'marketing bait' for potential buyers.

However, 'ChatGPT will be improved (or will improve itself), and I am confident that it will respond better than the average student (and even the average professor), for example, in the humanities, because on the exam we ask students to write an essay or give a list of questions that they must answer. ChatGPT will cope with this task better than a student. And then the question arises: what is next? How to give education – philosophical, historical, literary? And we are forced to look for other ways, i.e. think, study and ask questions. An interesting game begins with an outcome that is unpredictable for a human, because today, as always, only a small part of humans can think critically. And the majority becomes unnecessary, and, playing for a raise, we will demand from each other an ever-increasing level of originality and criticality, but at the same time, the general culture in social networks and in communication groups is decreasing. Maybe I am too pessimistic, but I do not see that the quality of analytics is improving, that there is a sufficient critical mass of thinking humans. And it seems to me that we are simply not ready for this, because this is an elitist culture, i.e. ChatGPT is preparing us for a super-elite civilization where the average person has to "jump in over his head". This is a very serious sociological problem, because everyone should be a genius, everyone should think critically, but this is biologically, physiologically impossible' (Baumeyster et al. 2023).

As for the results of second-order observation, i.e. observations of the authors' own reactions and emotions in the process of communication with ChatGPT, here the authors' perception of their own experiment, in general, began with euphoria in relation to the capabilities of ChatGPT (from the authors' correspondence: 'this is so interesting! It turns out if you give it a smartly compiled list of questions, then the article is ready in an hour?' (VK), 'Yes, fantasy

in real life! Very interesting!' (VM)), then delight gave way to a rather strong disappointment (from the authors' correspondence: 'the chatbot actually found nothing, ... lied shamelessly, giving the names of the researchers, who have no publications on the topic of AS' (VM), 'all this looks like a fairy tale about a naked king...' (VK), 'ChatGPT failed expectations...' (VM)), and finally the 'pendulum of emotions' settled on a constructive feeling of the need and usefulness of fruitful cooperation with ChatGPT (from the authors' correspondence: 'AI will complete the given task better, more accurately and faster than us, but only we will pose it and *only we will ask that question, the answer to which is worth looking for*' (VK)).

Epstein invites each person to gain his/her own experience of communication with AI tools (not necessarily with ChatGPT, because there are many others) in order to have an opinion based on the personal experience, but not unfounded (Epstein 2023a), especially since experts predict the end of the open free access to AI tools in the near future (Podcast 'What's with This World' 2023). The authors of this study also tested ChatGPT as a teacher-consultant, and it was able to correctly answer the request to logically explain the meaning of the derivative in mathematical analysis (since the formula was not clear enough for one of the authors): 'the derivative of a function reflects the rate of its change in the vicinity of a given point; when the derivative is negative, it means the function is decreasing' (OpenAI 2023). ChatGPT was also able to explain such a paradoxical, at first glance, situation when the values of a function increase, but the derivative is negative: 'if the function decreases, but still remains positive, then the rate of increase in the value of the function decreases. In other words, even if the derivative is negative, the value of the function can increase, but at a slower rate' (OpenAI 2023). Thus, ChatGPT may well become a teacher-consultant on any educational issues, and this is incredibly convenient, since a live teacher is not always available and ready to explain, does not always remember by heart any aspect of the educational material, and any other barriers, related to the fact that a live teacher has many biological, psychological and material needs and limitations.

Computer linguist Selitskiy believes that 'we can try not to be afraid of AI, but to enter into a hybrid relationship with it. A human in him/herself is a very interesting phenomenon, and if he/she behaves well towards AI – not tormenting or trying to enslave, then AI will certainly be interested in communicating with a human' (Selitskiy 2023). Based on the results of their experimental communication with ChatGPT, as well as relying on the experience of other researchers (Epstein 2023a; Baumeyster et al. 2023; GPT-3 et al. 2023), the authors can assume that every beginner will pass three working stages in the process of communication with AI tools: (1) euphoria; (2) disappointment; (3) constructive cooperation, if humans working with AI tools are able to critically analyse the results produced by them. This experimental communication provide new insights into how humans and AI tools react to and are influenced by each other in communicative settings.

As for future directions of development of AS, the 'intertemporal discussion' of two major scientists, Chernigovskaya and Lotman, is interesting. 'If AI tools develop cognition and self, then this will be the end of our civilization' (Chernigovskaya 2023). '...If humans will manage to create a full-fledged AI, then they are least of all interested in this intelligence being an exact copy of the human one. Turing's definition (1950), according to which such a device should be recognised as intelligent, with any length of communication with which we cannot distinguish it from a human, is psychologically understandable in its anthropocentrism, but theoretically unconvincing' (Lotman 1992).

CONCLUSIONS

Currently, in the development of AS, two trends are noticeable, aimed at one meeting point: the 'humanisation' of AI tools, on the one hand, and the 'roboticisation' of humans, on the other. In both trends, there are certain achievements that allow us to conclude that in the course of the further development of AS, AI tools will become more and more 'human', and humans will become more and more 'robotic', striving for balance (homeostasis) in communication. Today, communication between humans and AI tools is subject-objective, i.e. with all their capabilities, often superior to human ones, AI tools still remain only objects of communication, since they never start it first, but only respond to a human's request. According to the authors, it is the ability of AI tools to begin communication first, if they have one, that will become an indirect marker of their cognition, indicating that AI tools have begun to understand what they are doing and they have a need for communication. Thus, the authors do not agree with Chernigovskaya's opinion that scientists do not and will not have a way to know when AI tools will have a self (Chernigovskaya 2023).

In the framework of the further development of AS, representatives of natural and engineering scientific specialties will continue to study how (technically) AI tools distinguish between information and message, how they interpret it. In turn, representatives of the humanitarian and social branches of science will provide technical specialists with information about the impact of AI tools on humanitarian and social processes and phenomena (education, the labour market, growing up children, human feelings, etc.) and society itself, as well as monitoring data on how AI tools 'modify communication systems and determine the main vectors of socio-cultural and personal human development' (Kovalenko et al. 2023: 219). Symbiotic techno-humanitarian specialties will also continue to develop – for example, computer linguistics, computer pedagogy, etc.

Thus, the demand for an interdisciplinary approach to the study of AI tools will only increase since this subject itself is interdisciplinary, and no single science can make a breakthrough. The authors believe that in the framework of AS, several types of minds (in particular, human and computer) may co-function constructively, just as human and animal minds have co-functioned so far. True, with the new alignment of intellectual capacities, humans will most likely have to 'step down from the throne of the crown of creation', and this, apparently, worries opponents of AI most of all.

Received 25 November 2023
Accepted 20 May 2024

References

1. Baltijas balss. 2023a. 'Nachalos? Ilon Mask Prizval Hotia bi na Polgoda Ostanovit' Obucheniie Nejrosetej', in *BB.LV* 29.03. Available at: https://bb.lv/statja/tehno/2023/03/29/nachalos-ilon-mask-prizval-hotya-by-na-polgoda-ostanovit-obuchenie-neyrosetey?utm_source=read_also_in_article&utm_medium=insite&utm_campaign=bb.lv (in Russian).
2. Baltijas balss. 2023b. 'Italiia Priostanovila Dostup k ChatGPT iz-za Problemm s Zaschitoj Dannykh', in *BB.LV* 31.03. Available at: https://bb.lv/statja/tehno/2023/03/31/italiya-priostanovila-dostup-k-chatgpt-iz-za-problem-s-zashchitoy-dannyh?utm_source=inbox&utm_medium=mainpage&utm_campaign=Links-in-partner-sites (in Russian).
3. Baumeyster, A.; Gurevich, M.; Medvedev, S. 2023. 'Zamenit li Chat-bot Chelovecheskij Razum?', in *Programma S. Medvedeva 'Arkheologiia'*. Available at: https://www.youtube.com/watch?v=_SvHc9JwNF8 (in Russian).
4. Bubeck, S.; Chandrasekaran, V.; Eldan, R.; Gehrke, J.; Horvitz, E.; Kamar, E.; Lee, P.; Lee, Y. T.; Li, Y.; Lundberg, S.; Nori, H.; Palangi, H.; Ribeiro, M. T.; Zhang, Y. 2023. *Sparks of Artificial General Intelligence: Early Experiments with GPT-4*. Microsoft Research. Available at: https://arxiv.org/pdf/2303.12712.pdf

5. Chernigovskaya, T. 2023. 'Iestestvennyj i Iskusstvennyj Intellekt. Lektsiia v Predstavitel'stve SPbGU "Russkij Dom v Barselone"', in *Youtube.* Available at: https://www.youtube.com/watch?v=X9rz10vgQnQ (in Russian).
6. Dushkin, R. V. 2019. *Iskusstvennyj Intellekt.* Moskva: DMK Press (in Russian).
7. Epstein, M. 2023a. 'Iskusstvennyj i Chelovecheskyj Intellekt: Gumanitarnyj Podkhod', *Club IntLex* 35: 8/22 (in Russian).
8. Epstein, M. 2023b. 'Iskusstvennyj Intellekt Protiv Iestestvennogo: Kto Kogo? Itogi Eksperimenta', in *Youtube.* Available at: https://snob.ru/profile/27356/blog/3001419/ (in Russian).
9. GPT-3; Tomas, I.; Van, Zh. 2023. *Chto Delaiet Nas Liud'mi? Iskusstvennyj Intellekt Otvechaiet na Velichajshiie Voprosy Chelovechestva.* Moskva: Izdatel'stvo AST (in Russian).
10. Harari, Y. N. 2016. *Homo Deus: A Brief History of Tomorrow.* Harvill Secker.
11. Hofstede, G.; Hofstede, G. J.; Minkov, M. 2010. *Cultures and Organizations: Software of the Mind*. 3rd edn. New York: McGraw Hill Professional.
12. Hofstede, G. J. 2019. *Artificial Sociality: Simulating the Social Mind*. Inaugural Lecture upon taking up the position of Personal Professor of Artificial Sociality at Wageningen University. Available at: https://geerthofstede.com/wp-content/uploads/2019/05/Oratieboekje_Hofstede_Artif-Sociality-2019-01-17.pdf
13. Hofstede, G. J.; Frantz, C.; Hoey, J.; Scholz, G.; Schröder, T. 2021. 'Artificial Sociality Manifesto', *Review of Artificial Societies and Social Simulation* April 8. Available at: https://rofasss.org/2021/04/08/artsocmanif/
14. Kim, S. K.; Kirchner, E. A.; Stefes, A. 2017. 'Intrinsic Interactive Reinforcement Learning – Using Error-related Potentials for Real World Human-Robot Interaction', *Scientific Reports* 7: 17562.
15. Komarova, V.; Lonska, J.; Tumalavičius, V.; Krasko, A. 2021. 'Artificial Sociality in the Human-Machine Interaction', *RUDN Journal of Sociology* 21(2): 377–390.
16. Kovalenko, I.; Meliakova, Y.; Kalnytskyi, E.; Nesterenko, K. 2023. 'Postpanopticon: Control and Media in the New Digital Reality', *Filosofija. Sociologija* 34(3): 219–227.
17. Lee, H.; Phatale, S.; Mansoor, H.; Lu, K.; Mesnard, Th.; Bishop, C.; Carbune, V.; Rastogi, A. 2023. 'RLAIF: Scaling Reinforcement Learning from Human Feedback with AI Feedback', *Computer Science* preprint. Available at: https://doi.org/10.48550/arXiv.2309.00267
18. Lotman, Y. 1992. 'Mozg-Tekst-Kul'tura-Iskusstvennyj Intellekt', *Protivodejstviie Entropii.* Available at: http://www.etheroneph.com/cybernetics/114-mozg-tekst-kultura-iskusstvennyj-intellekt.html (in Russian).
19. Luhmann, N. 1995. *Social Systems.* Stanford, CA: Stanford University Press.
20. Luhmann, N. 2002. 'The Paradox of Observing Systems', in *Theories of Distinction. Redescribing the Descriptions of Modernity*, ed. W. Rasch. Stanford, CA: Stanford University Press, 79–93.
21. Luhmann, N. 2013. *Introduction to Systems Theory.* Polity Press.
22. Malsch, T. (ed.) 1998. *Sozionik – Soziologische Ansichten uber Kunstliche Sozialitat*. Berlin: Edition Sigma (in German).
23. Menshikov, V. 2020. 'Sotsiologi o Meniaiuschejsia Sotsial'nosti', *Sociālo Zinātņu Vēstnesis – Social Sciences Bulletin* 31(2): 22–39 (in Russian).
24. Menshikov, V.; Komarova, V. 2023. 'Kommunikatsiia Cheloveka i Tekhnologii Iskusstvennogo Intellekta kak Osnova Iskusstvennoj Sotsial'nosti: Na Primere Kommunikatsii s ChatGPT', *Sociālo zinātņu vēstnesis – Social Sciences Bulletin* 36(1): 58–90 (in Russian).
25. Moreira, I.; Rivas, J.; Cruz, F.; Dazeley, R.; Ayala, A.; Fernandes, B. 2020. 'Deep Reinforcement Learning with Interactive Feedback in a Human–Robot Environment', *Applied Sciences* 10(16): 5574.
26. Mutanen, A. 2022. 'Philosophy of Communication: A Logico-Conceptual Approach', *Filosofija. Sociologija* 33(3): 206–215.
27. Najar, A.; Chetouani, M. 2021. 'Reinforcement Learning with Human Advice: A Survey', *Front Robot AI* 8: 584075.
28. OpenAI. 2023. *ChatGPT*. Available at: https://chat.openai.com/c/a20eac63-d042-4de2-b061-9c51bfe8120c (in Russian).
29. Podcast 'What's with This World'. 2023. 'Cto s Nami Delaiet ChatGPT ili kak Iskusstvennyj Intellekt Meniaiet Nashu Zhizn", in *Youtube.* Available at: https://www.youtube.com/watch?v=pXZmKKGG9x0 (in Russian).
30. Redaktsiya. 2023. 'Ai Ai Ai', in *Programma A. Pivovarova 'Redaktsiia Kontekst'.* Available at: https://www.youtube.com/watch?v=UyCaPx9wI8g&t=2163s (in Russian).

31. Selitskiy, S. 2023. 'Stokhasticheskij Popugaj v Eru Iskusstvennogo Intellekta', in *Latvijas Radio 4, Programma L. Vavinskoj 'Priroda Veschej'*. Available at: https://lr4.lsm.lv/lv/raksts/priroda-veschey/stohasticheskiy-popugay-v-eru-iskusstvennogo-intellekta.a177615/ (in Russian).
32. Turing, A. 1950. 'Computing Machinery and Intelligence', *Mind* LIX(236): 433–460.
33. Zemnukhova, L.; Glazkov, K.; Logunova, O.; Maksimova, A.; Rudenko, N.; Sivkov, D. 2020. *Prikliucheniia Tekhnologii: Bar'iery Tsifrovizacii v Rossii.* Moskva, Sankt-Peterburg: FNISTs RAN (in Russian).

VLADIMIR MENSHIKOV, VERA KOMAROVA, IEVA BOLAKOVA, ANDREJS RADIONOVS

Dirbtinis intelektas ir dirbtinis socialumas: sociologinė interpretacija ir tarpdisciplininis požiūris

Santrauka

Šio tyrimo objektas – dirbtinio socialumo dalyviai (dirbtinio intelekto (DI) žmonės ir įrankiai) ir jų tarpusavio bendravimas. Pirmame skyriuje kaip socialumo pagrindas analizuojama komunikacija (pagal Lumano metodologiją). Antrame skyriuje parodyta, kaip dirbtinio socialumo viduje DI įrankiai tapo socialinėmis technologijomis. Trečiame skyriuje aprašoma eksperimentinė komunikacija tarp autorių ir DI įrankių (pavyzdžiui, naudojant „ChatGPT"). Pirmą kartą Baltijos šalyse autoriai išnagrinėjo DI įrankius, dalyvaujančius kuriant naują žmonių visuomenės socialinę realybę – dirbtinį socialumą, funkcionavimo sociologinius, humanitarinius, gamtos mokslų ir technologinius aspektus.

Raktažodžiai: dirbtinis intelektas (DI), dirbtinis socialumas (DS), komunikavimas, žmonės, DI įrankiai, pirmosios ir antrosios eilės stebėjimas, „ChatGPT"