http://doi.org/10.23925/2179-3565.2025v16i3p05-12

RISUS - Journal on Innovation and Sustainability
volume 16, número 3 - 2025
ISSN: 2179-3565
Editor Científico: Arnoldo José de Hoyos Guevara
Editor Assistente: Vitória Catarina Dib
Avaliação: Melhores práticas editoriais da ANPAD

ARTIFICIAL INTELLIGENCE IN THE CONTEXT OF SELF-ORGANIZATION THEORY (PROBLEMS OF HUMAN-MACHINE COLLABORATION)

Inteligência artificial no contexto da teoria da auto-organização (problemas de colaboração homem-máquina)

Evgeniya Nikolaeva[1], Mikhail Nikolaev[1], Nataliya Soldatova[1,2]
[1]Kazan (Volga region) Federal University, Kazan, Russia, [2]Kazan National Research Technical University named after A.N. Tupolev-KAI, Kazan, Russia.
E-mail: evgeniya.nikolaeva@mymail.academy, mikhail.nikolaev@mymail.academy, nataliya.soldatova@mymail.academy

ABSTRACT

The focus of the authors is on the AI phenomenon, which can be interpreted by self-organization theory as a complex system. The paper considers key characteristics of AI, ways and possibilities of its development and training, integration into production, computational processes. In the course of the study, the authors identify important patterns of AI as a complex adaptive system: possibility of variable behavior depending on the level and volume of elements, unpredictability of decision-making (property of emergency). The article problematizes ethical issues related to the ability of effective management of complex AI systems, their correct and safe functioning.
Keywords: AI, Self-organization theory, Complex system, Adaptability, Diachronic emergence, Law of U. Ashby, Biological-silicon collaboration, Ethical imperatives

SUBMETIDO EM: 28/04/2025
ACEITO EM: 15/09/2025
PUBLICADO EM: 30/10/2025

RISUS - Journal on Innovation and Sustainability
volume 16, número 3 - 2025
ISSN: 2179-3565
Editor Científico: Arnoldo José de Hoyos Guevara
Editor Assistente: Vitória Catarina Dib
Avaliação: Melhores práticas editoriais da ANPAD

INTELIGÊNCIA ARTIFICIAL NO CONTEXTO DA TEORIA DA AUTO-ORGANIZAÇÃO (PROBLEMAS DE COLABORAÇÃO HOMEM-MÁQUINA)

Artificial intelligence in the context of self-organization theory (problems of human-machine collaboration)

Evgeniya Nikolaeva[1], Mikhail Nikolaev[1], Nataliya Soldatova[1,2]
[1]Kazan (Volga region) Federal University, Kazan, Russia, [2]Kazan National Research Technical University named after A.N. Tupolev-KAI, Kazan, Russia
E-mail: evgeniya.nikolaeva@mymail.academy, mikhail.nikolaev@mymail.academy, nataliya.soldatova@mymail.academy

RESUMO

O foco dos autores está no fenômeno da IA, que pode ser interpretado pela teoria da auto-organização como um sistema complexo. O artigo considera as principais características da IA, as formas e as possibilidades de seu desenvolvimento e treinamento, a integração na produção e os processos computacionais. No decorrer do estudo, os autores identificam padrões importantes da IA como um sistema adaptativo complexo: possibilidade de comportamento variável dependendo do nível e do volume de elementos, imprevisibilidade da tomada de decisões (propriedade de emergência). O artigo problematiza questões éticas relacionadas à capacidade de gerenciamento eficaz de sistemas complexos de IA, seu funcionamento correto e seguro.
Palavras-chave: IA, Teoria da auto-organização, Sistema complexo, Adaptabilidade, Emergência diacrônica, Lei de U. Ashby, Colaboração biológica-silício, Imperativos éticos

INTRODUCTION

Research on complex systems shows that there may be a gap between the local behavior of interacting elements of a complex system at the micro-level and patterns observed in the global behavior of this system at the macro-level. This peculiarity makes it impossible to understand global behavior based on a set of local rules. It can be assumed that intelligent systems are complex, and therefore global behavior of AI systems cannot be expected to have an immediate analytical relationship with their constituent mechanisms. This has serious implications for the methodology of AI research.

Modern artificial intelligence is no longer a set of hard-coded algorithms - it has evolved into a complex adaptive system (CAS), whose behavior is determined by the dynamic interactions between individual components, multiple feedback loops and the ability to self-organize. To understand the nature of AI in this context, it is necessary to look at the basic principles of self-organization theory (general complexity theory).

The key question of the work is whether AI can be managed as a complex system without imposing limitations on its adaptive potential, and what risks arise in the process of bio-silicon collaboration.
In this context, we focus on three key concepts:

- adaptability of complex systems (AI, developing through self-organization, building feedback);
- diachronic emergence (the ability of algorithms to go beyond what the developers have built and exhibit unexpected/unpredicted effects);
- the law of requisite diversity (the requirement for the control system (human) to match the diversity and complexity of the controlled system (AI).

The aim of this study is to conceptualize the phenomenon of AI through a theoretical and methodological apparatus of self-organization theory, allowing to form a resonant correspondence between technological innovations and management control.

1 METHODS

In this study, the method of idealization, analysis and synthesis, concretion and abstraction are used as general-science heuristic regulations.

As a theoretical and methodological basis of the study, synergetics is presented as a general theory of complexity and self-organization. Important for this study are the works of E. Morin (1980; 2000; 2001), G. Haken (1988; 2004), I. Prigogine and I. Stengers (1984), representing the general theoretical basis. The problematic field of their research is concentrated around the concept of “complexity”, oriented towards the knowledge of its nature, principles of its organization and evolution.

The main methodological principle is self-organization. It is understood as the ability of a system to build its own internal structure, responding to changes in the external world and simultaneously making changes in its environment. Self-organization allows the system to adapt to changing environment, make non-standard decisions.

2 RESULTS

1. The discussion about the prospects and risks associated with the active development of AI technologies, should be noted that complexity theory has the necessary potential for their comprehension. AI, being a complex system, needs to be included in systems of greater scale (political, economic, socio-cultural, etc.) and it is interconnected with them by permanent interactions of the form of feedback loops. The feedback loop in case of complex systems form a special type of causality - non-linear (cyclic causality). Any action taken is determined by the conditions of the natural and/or social environment and often deviates from the direction originally set for it. French researcher, one of the most authoritative complexologists E. Morin on this occasion notes: “We can not be sure, that the result of action will correspond to our intentions, on the contrary, we have the right to seriously doubt it” (Morin, 2002, p. 23).

AI as a complex system has the property of adaptability, the ability to integrate into the environment, change it for its own purposes. Adaptability allows AI systems to respond quickly to changing reality, offering fast acting complex engineering solutions that can work effectively in different environments. In this context, AI shows a trend from less complex to more complex and orderly forms of organization, acting as a complex self-organizing system.

It should be noted that today we do not have an exhaustive knowledge of the brain, therefore it is impossible to model and reproduce in the AGI system. However, the technology community believes that humanity will get acquainted with a strong AI version in 2026, although the original forecast was 2042 (Aguirre, 2020). The appearance of AGI can seriously change the relationship of a person with technology, it will become a point of no return (researchers call this state as singularity). A powerful artificial intelligence based on ultra-fast self-learning will produce increasingly powerful devices, make independent decisions, and change the world, while humans will be “sidelined” from the process, observing but not comprehending it.

In the very mechanism of human cognition of the world, there are no boundaries except temporal ones (one cannot know yesterday what was not known yesterday). Here, the principle of irreversibility applies, according to which a person cannot return to the past in a gnoseological sense to realize cognitive practices there. AI, in this regard, has no limitations because it lacks a sense of time (there is no demarcation line between past, present, and future, as they are syncretic); consequently, infinite return for additional self-learning/cognition is possible, and the inexhaustibility of resources allows this to be done “effortlessly”. This appears as a continual, uninterrupted recursive process, unlike human cognition, which is discrete in nature.

2. The greatest number of risks in the process of development of AI systems is related to diachronic type emergent effects. In this case, the developers and agents involved in all stages of the life cycle of such systems cannot fully predict the behavior of the algorithm, moreover, they do not have absolutely reliable ways to control what generates their models. There are cases when the system goes beyond a given logic, demonstrating skills that specialists have not taught it. According to the technical director of Google, D.Maniki one of the AI programs adapted itself when it was offered to use Bangladeshi language, which was missing in the competencies set by the developers (Medvedev, 2023). There is a known case where the AlphaGo system developed by DeepMaind for playing Go based on deep learning and large data set, instead of strictly following the game’s algorithms started to play itself using deep neural network mechanisms. This gave her the opportunity to improve her strategies beyond the human studied and make unpredictable, creative moves that break the framework of standard strategies.

In a study conducted by scientists of the University of Zurich, we investigated the possibilities of the OpenAI GPT-3 model, with the aim of determining its potential risks and benefits when generating and disseminating disinformation. The study, in particular, showed that GPT-3 is able to “refuse” to generate misinformation, and in more rare cases it can generate misinformation when it is tasked with producing accurate information (Spitale et al., 2023).

In terms of complexity theory, such situations can be interpreted as the ability of AI to make decisions autonomously, exhibiting emergence properties. This property is the result of interaction between many elements of a system and its surroundings, with the result that it is not possible to sum the properties of these elements.

Russian researchers note that indeed, AI, like many other technologies of digital economy, remains poorly studied in the context of possible emergences. There are no relevant historical comparisons to compare (Kurnosova, Filippov, 2023). “Today, the patterns of functioning of distributed organizational structures, the sharing economy, peer-to-peer financing, and many other innovations in the economic mechanism are radically altering the contours of economic relations that have been shaped over centuries. In the neuro-network economy of the future (sometimes referred to as “NeuroNet”), toward which progress in digitization is bringing the planet in leaps and bounds, classical economic laws may cease to operate or will be significantly transformed” (Kurnosova, Filippov, 2023, p. 500).

Scientists from Fudan University, in their study using two popular large language models (LLM), attempted to determine whether self-replicating AI could proliferate uncontrollably. The results obtained during the experiment led the scientists to express serious concern about a range of unexpected behaviors. “By analyzing the behavioral traces, we observe that the AI systems under evaluation already exhibit sufficient self-perception, situational awareness, and problem-solving capabilities to accomplish self-replication. We further note that the AI systems are even able to use the capability of self-replication to avoid shutdown and create a chain of replicas to enhance survivability, which may ultimately lead to an uncontrolled population of AIs. If such a worst-case risk remains unknown to human society, we would eventually lose control over frontier AI systems: They would take control over more computing devices, form an AI species, and collude with each other against human beings” (Pan et al., 2024). The unpredictable capabilities of large language models have recently been described as emergent (Wei, 2022).

Here, it is important to note that while diachronic emergent effects arise in cases involving the behavior of AI systems designed for automated solutions to local, specialized tasks (narrow artificial intelligence), the problem is exacerbated when discussing general or strong AI (General AI, AGI). The idea behind this is that it is an anthropomorphic system mimicking the work of biological intelligence, possessing universal thinking abilities that can be applied to solve any problem.

3. On the horizon of civilization, AI will increasingly surpass humans, and the question of whether artificial intelligence is a competitor or a complement to humans, acting as an assistant, will become more acute. In the first case, the prospect of a war between the biological and silicon worlds becomes quite likely; in the second case, a bio-silicon interaction emerges, where ethical constraints that would necessarily extend to developments in AI, especially self-improving algorithms (strong AI), become increasingly relevant.

Modern humans, lacking a vertical hierarchy of universal values and existing "on a flat plane" often struggle to establish a system of ethical constraints and values for themselves. Will they be able to do so in relation to artificial intelligence? Some steps in this direction are already being taken. For instance, in November 2021, the UNESCO General Conference adopted the Recommendation on the Ethics of Artificial Intelligence, which is advisory in nature (United Nations Educational, Scientific and Cultural Organization [UNESCO], 2022). Almost simultaneously, Russia opened the National Code of Ethics in the Field of AI for signing (AI Alliance Russia, 2021), and China released an ethical code for next-generation artificial intelligence, titled "Beijing Artificial Intelligence Principles" (Yi, 2019).

T. Nakazaki, a special advisor at Anderson Mōri & Tomotsune with extensive experience in data protection and privacy, discusses new trends in AI governance in Japan based on the Data Protection Laws and Regulations Trends in AI Governance in Japan 2024-2025. He notes, "In order to develop the artificial intelligence (AI) sector, the Japanese government has refrained from introducing comprehensive laws and regulations for AI and instead has adopted a so-called soft-law approach introducing comprehensive guidelines, and expecting AI business operators to develop their business based on these guidelines on a voluntary basis. As part of this, the Japanese government released the AI Guidelines for Business Version 1.0 in April 2024" (Nakazaki, 2024).

The authors of an analytical report on AI ethics rightly point out that the primary purpose of such documents "should be to create a common understanding of the ethical responsibility of AI actors towards users and society, including long-term, systemic, and other significant outcomes of the widespread application of AI technologies. A coordinated and widely accepted understanding of the ethical responsibility of individuals involved in various life cycles of AI systems will help reduce risks and promote the safe implementation and development of new technologies in society and the state" (Abramova et al., 2021, p. 17).

The resolution of this issue remains in human hands: it is humans who decide where, how, and for what purpose to apply AI, where it will be beneficial, and what results it will yield. The data that AI possesses is data created and uploaded by humans. Machines operate according to the models they have been taught by humans and in the contexts where humans decide to apply them. Therefore, the task of introducing ethics into AI technologies by defining "red lines" is a challenge that humans must address.

In this context, the ideas of the English researcher W. Ross Ashby are highly relevant. He is the author of two fundamental works: "Design for a Brain" and "Introduction to Cybernetics" in which he develops the idea that the protein-carbon nature of life is not unique. The researcher believed that alternative forms are possible, and life can be iron, silicon, or electronic in its material substrate, with the laws of its emergence being the same; this is not natural selection but feedback mechanisms (Ashby 1962; 2005)

In the context of our discussions on artificial intelligence, it is important to consider the law of requisite variety, formulated by the English researcher W. Ross Ashby. The original formulation by the author is: "Only variety can absorb variety" (Ashby, 2005, p. 21). Applying this law to the challenges of managing artificial intelligence gives us the following interpretation: for effective management, the variety of the controlling system (natural intelligence) must be greater than the variety of the controlled system (artificial intelligence). In other words, if an algorithm demonstrates decision-making variability measured in millions, then the controlling system must be capable of "reading" these variants and aiming to predict them.

Building relationships with AI systems in accordance with this law gives us a chance to create collaborations that work for humans and contribute to the evolution of natural consciousness, which is arguably facing a crisis today. Otherwise, artificial intelligence could become humanity's last invention.

3 DISCUSSION

This work generally aligns with the main methodological framework of some contemporary studies.

For instance, the American and Mexican researcher C. Gershenson uses concepts such as “interactions, self-organization, emergence, and balance to compare different aspects of complexity, ALife, and A.I.” in his work. He formulates the purpose of his research as follows: “My purpose is to notice the similarities and differences between the three fields as they share conceptual, methodological, and philosophical approaches” (Gershenson, 2024, p. 2). He notes, that “In A.I., self-organization has had a more limited use. Still, it could be argued that most artificial neural network models are implicitly self-organizing, as their weights (interactions) are modified during the training phase” (Gershenson, 2024, p. 6).

Chinese researchers G. Ping G. and Y. Qian (2020) describe the artificial intelligence system as a “synergetic learning system”, which is a vast open complex system. They state, “Our proposed Synergetic Learning Systems is inspired by natural intelligence and is based on the cognitive neuroscience mechanism and computational intelligence, integrating multidisciplinary knowledge and adopting the mode of complex system thinking. The Synergetic Learning Systems is based on the concept of systems science. The use of the term “collaboration” is influenced by the idea of “synergetic learning” (Ping, Qian, 2020, p. 4).

M. Szczepanski, an External Policy Analyst at the European Parliament and author of numerous studies in international economics, particularly focusing on the economic potential of AI, concludes that “Research launched by consulting company Accenture covering 12 developed economies, which together generate more than 0.5% of the world's economic output, forecasts that by 2035, AI could double annual global economic growth rates. AI will drive this growth in three important ways. First, it will lead to a strong increase in labor productivity (by up to 40%) due to innovative technologies enabling more efficient workforce-related time management. Secondly, AI will create a new virtual workforce – described as “intelligent automation” in the report – capable of solving problems and self-learning. Third, the economy will also benefit from the diffusion of innovation, which will affect different sectors and create new revenue streams” (Szczepański, 2019, p. 3).

Szczepanski also notes that according to the report, “the next wave of digital revolution is expected to be unleashed with the help of the data generated from the Internet of Things (IoT), which is likely to be many times greater than the data generated by the current “Internet of People” (Szczepański, 2019, p. 3).

J.H. Wilson, the Global Managing Director of Technology Research and Thought Leadership at Accenture Research, and P.R. Daugherty, formerly the Chief Technology Officer at Accenture and currently a Senior Technology Advisor to the company, present data from a study involving 1,500 leading European companies across more than 12 industries (including medicine and pharmaceuticals, transportation and machinery, cosmetics, software production, financial structures, electronics and equipment, education, and many others). The study shows that the greatest increase in productivity is achieved when people and smart machines work together, enhancing each other's strengths.

While AI will radically alter how work is done and who performs it, the technology's larger impact will be in complementing and augmenting human capabilities, rather than replacing them. Through collaborative intelligence, humans and AI actively enhance each other's complementary strengths: the leadership, teamwork, creativity, and social skills of humans, and the speed, scalability, and quantitative capabilities of machines (Wilson, Daugherty, 2018).

Companies need employees who support the correct, safe, and responsible functioning of AI systems. The authors emphasize the need to rethink the entire business process, which implies more than just implementing AI technologies. Significant development of so-called “fusion skills” among employees is necessary, enabling them to work effectively at the human-machine interface. These skills include the ability to delegate tasks to new technologies, combine unique human skills with those of intelligent machines to achieve better results, and train intelligent agents in new skills while undergoing training themselves (Wilson, Daugherty, 2018).

P. Isackson, the Strategic Director of Fire Observer, develops the idea of “creating a culture of communication in which humans and machines can work together to explore complex questions and find ways of negotiating and agreeing upon meaning" (Isackson, 2024). He notes that William Wulf's concept of "collaboratories" - virtual environments where geographically distributed people work together—is highly relevant to our era, especially with the advent of complex AI tools that can enhance human collaboration. To explore the future of collaboratories with AI integration, it is important to consider both the evolution of human behavior and the evolving role of AI in creating meaning, decision-making, and communication.

As AI becomes more integrated into daily workflows, human behavior is likely to shift in significant ways: An expanded collective intelligence will be formed. Dynamic interdisciplinary collaboration will become possible. Iterative meaning creation will occur between AI and humans.

Reflecting on the future development of AI behavior, particularly its capabilities for understanding and decision-making, the author concludes that AI's "behavior" will need to adapt to interacting with humans, reflecting the nuances of meanings and contexts. Over time, AI will become context-aware, possessing adaptive behavior capable of aligning meaning with humans and considering nuances in decision-making.

These changes will lead to broader social and cultural shifts: Educational priorities will shift towards critical thinking and the ethics of AI use. AI can act as a mediator in debates, facilitating structured and productive dialogue. Strict ethical and governance frameworks will be established based on universal norms and values, preventing AI from any form of discrimination and errors (Isackson, 2024).

CONCLUSION

The interaction between humans and AI today can no longer be viewed through the lens of a subject-instrumental logic. Modern AI systems often demonstrate subjectivity in decision-making, content creation, and even in shaping socio-cultural phenomena. Such transformations compel us to seek an answer to the question: is AI an object under human control, or is it actively acquiring the characteristics of a subject capable of autonomous behavior? In our view, the answer to this question can be found through the heuristic potential of self-organization theory.

The rapid development of AI technologies carries significant risks of becoming their hostage. Preventing these risks is linked to the skillful use of the potential of complex systems. In this context, it is necessary to discuss limits that allow maintaining a balance between the freedom to develop AI and retaining human control. Consequently, such a balance will enable adherence to ethical imperatives in the development and use of AI.

Acknowledgments

The publication was carried out using funds from a subsidy allocated to Kazan Federal University for the implementation of project No. FZSM-2023-0022 "Digital socialization and digital competence of young people in the context of global systemic changes: regulation technologies, risks, scenarios" within the framework of the state assignment.

REFERENCES

Abramova, A.V, Ignatiev A.G., Panova M.S (2021). Etika v oblasti iskusstvennogo intellekta - Ot diskussii k nauchnomu obosnovaniyu i prakticheskomu primeneniyu: Analiticheskiy doklad [Ethics in the field of artificial intelligence - From discussion to scientific justification and practical application: Analytical report]. Moscow: MGIMO-University.
Aguirre, A. (2020). When will the first weakly general AI system be devised, tested, and publicly announced? Metaculus. https://www.metaculus.com/questions/3479/date-weakly-general-ai-is-publicly-known/
AI Alliance Russia. (2021). Kodeks etiki v sfere II [Code of ethics in the field of artificial intelligence]. https://ethics.a-ai.ru/
Ashby, R.W. (1962). Konstruktsiya mozga: Proiskhozhdeniye adaptivnogo povedeniya [Brain design: Origin of adaptive behavior]. Moscow: Izd-vo inostrannoy literatury.
Ashby, R.W. (2005). Vvedeniye v kibernetiku [An introduction to cybernetics]. Moscow: URSS; ComBook.
Gershenson, C. (2024). Complexity, artificial life, and artificial intelligence. Artificial Life, 1-15. https://doi.org/10.1162/artl_a_00462

Haken, H. (1988). Information and self-organization: A macroscopic approach to complex systems. Berlin: Springer Series in Synergetics.
Haken, H. (2004). Synergetic computers and cognition: A top-down approach to neural nets. Berlin: Springer & Business Media.
Isackson, P. (2024, October 14). Outside the box: Can AI–human "collaboratories" save the world? Fairobserver. https://www.fairobserver.com/more/science/outside-the-box-can-ai-human-collaboratories-save-the-world/#
Kurnosova, T., Filippov, A. (2023). Iskusstvennyy intellekt, kak istochnik vozmozhnostey i ugroz ekonomicheskogo razvitiya [Artificial intelligence as a source of opportunities and threats for economic development]. Innovation and investment, 12, 498-503.
Medvedev, Yu. (2023). Iskusstvennyy intellekt vyuchil yazyk, kotoromu yego ne obuchali [Artificial intelligence learned a language it was not taught]. Rossiiskaia Gazeta [Ros. Gaz.] 18.04.2023. https://rg.ru/2023/04/18/iskusstvennyj-intellekt-vyuchil-iazyk-kotoromu-ego-ne-obuchali.html
Morin, E. (1980). La méthode, La vie de la vie. Tome 2: La Vie de la vie. Paris: Seuil.
Morin, E. (2000). Les sept savoirs nécessaires à l'éducation du future. Paris: Seuil.
Morin, E. (2001). L'humanité de l'humanité: l'identité humaine. La méthode - Tome 5. Paris: Seuil.
Morin, E. (2002). Le complexus, ce qui est tissé ensemble. In R. Benkirane (Dir.), La complexité, vertiges et promesses. Dix-huit histoires de sciences. Paris: Poche-Le Pommier.
Nakazaki, T. (2024, July 31). Data protection laws and regulations trends in AI governance in Japan 2024-2025. London: International Comparative Legal Guides. https://iclg.com/practice-areas/data-protection-laws-and-regulations/02-trends-in-ai-governance-in-japan
Pan, X., Dai, J., Fan, Y., Yang, M. (2024). Frontier AI systems have surpassed the self-replicating red line. New York: Ithaca. https://doi.org/10.48550/arXiv.2412.12140
Ping, G., Qian, Y. (2020). Synergetic learning systems: Concept, architecture, and algorithms. Neural and evolutionary computing. New York: Ithaca. http://dx.doi.org/10.48550/arXiv.2006.06367
Prigogine, I., Stengers, I. (1984). Order out of chaos man's new dialogue with nature. London: Heinemann.
Spitale, G., Biller-Andorno, N., Germani, F. (2023). AI model GPT-3 (dis)informs us better than humans. Science Advances, 9(26), eadh1850. https://doi.org/10.1126/sciadv.adh1850
Szczepański, M. (2019). Economic impacts of artificial intelligence (AI). EPRS - European Parliamentary Research Service. https://www.europarl.europa.eu/RegData/etudes/BRIE/2019/637967/EPRS_BRI(2019)637967_EN.pdf
United Nations Educational, Scientific and Cultural Organization [UNESCO]. (2022). Recommendation on the ethics of artificial intelligence. Paris: UNESCO. https://unesdoc.unesco.org/ark:/48223/pf0000381137
Wei, J., Tay, Y., Bommasani, R., Raffel, C., Zoph, B., Borgeaud, S., Yogatama, D., Bosma, M., Zhou, D., Metzler, D., Chi, E.H., Hashimoto, T., Vinyals, O., Liang, P., Dean, J., Fedus, W. (2022). Emergent abilities of large language models. New York: Ithaca. https://doi.org/10.48550/arXiv.2206.07682
Wilson, J.H., Daugherty, P.R. (2018). Collaborative intelligence: Humans and AI are joining forces. Massachusetts: Harvard Business School Publishing. https://hbr.org/2018/07/collaborative-intelligence-humans-and-ai-are-joining-forces
Yi, Z. (2019). Beijing artificial intelligence principles. International Research Center for AI Ethics and Governance. https://ai-ethics-and-governance.institute/beijing-artificial-intelligence-principles/ (accessed on February 15, 2025).