

The challenges of modernity: Artificial intelligence. Ethical aspect

© M.V. Moiseenko

Peoples' Friendship University, Moscow, 117198, Russia

The article considers artificial intelligence as an integral part of modern society life, and analyzes further prospects for its development and ways to prevent the adverse impact of artificial intelligence on humanity. The cost of the error, if it occurs in the system, increases every day due to the penetration of artificial intelligence in most areas of human life. The article describes ways to prevent errors in the work of artificial intelligence, using modern technologies. The article also considers possible scenarios of technology development, their analysis results in the conclusion that without observance of ethical principles in the development of artificial intelligence harmonious interaction between people and machines is not possible. Following the results of the Asilomar conference held in January 2017, a number of universal ethical principles were adopted; their implementation can reduce existential risks while preserving the prospect of making the biggest leap forward in the development of mankind. Although the prospect of creating superintelligence seems to be far away, right now the developers all over the world must follow the accepted ethical rules and be responsible for the technologies being created, and training talented programmers — future socially responsible ethical leaders must become a priority in education.

Keywords: artificial intelligence, ethical principles, ethics, Asilomar conference, existential risks, superintelligence

REFERENCES

- [1] Levy. S. How Elon Musk and Y Combinator Plan to Stop Computers From Taking Over. *Medium*. Available at: <https://medium.com/backchannel/how-elon-musk-and-y-combinator-plan-to-stop-computers-from-taking-over-17e0e27dd02a> (accessed July 6, 2018).
- [2] Bostrom N., Yudkowsky E. *The Ethics of Artificial Intelligence*. Cambridge, Cambridge University Press Publ., 2011.
- [3] Rahwan I., Cebrian M. Machine Behavior Needs to Be an Academic Discipline. *Nautilus*. Available at: <http://nautil.us/issue/58/self/machine-behavior-needs-to-be-an-academic-discipline> (accessed June 4, 2018).
- [4] Hills T. Does my algorithm have a mental-health problem? *Aeon*. Available at: <https://aeon.co/ideas/made-in-our-own-image-why-algorithms-have-mental-health-problems> (accessed May 28, 2018).
- [5] Cussin J. Developing Ethical Priorities for Neurotechnologies and AI. *Future of life institute*. Available at: <https://futureoflife.org/2017/11/09/developing-ethical-priorities-neurotechnologies-ai/> (accessed June 2, 2018).
- [6] Leonov V.V. Dvadtsat tri printsipa Asilomara [Twenty-three Asilomar Principles]. *Sovremennoe mashinostroenie — Sovmash.com*. Available at: <https://www.sovmash.com/node/348> (accessed May 22, 2018).
- [7] Asilomar AI Principles. *Future of life institute*. Available at: <https://futureoflife.org/ai-principles/> (accessed May 22, 2018).

- [8] ‘The Business of War’: Google Employees Protest Work for the Pentagon. *New York Times*. Available at: <https://www.nytimes.com/2018/04/04/technology/google-letter-ceo-pentagon-project.html> (accessed June 1, 2018).
- [9] Artificial Intelligence at Google. Our Principles. *Google AL*. Available at: <https://ai.google/principles> (accessed June 10, 2018).
- [10] Muehlhauser L., Helm L. *Intelligence Explosion and Machine Ethics*. Berlin, Machine Intelligence Research Institute Publ., 2012.
- [11] Benefits and Risks of Artificial Intelligence. *Future of life institute*. Available at: <https://futureoflife.org/background/benefits-risks-of-artificial-intelligence/> (accessed May 25, 2018).
- [12] Bostrom N. *The Superintelligent Will: Motivation and Instrumental Rationality in Advanced Artificial Agents*. Oxford, Oxford University Press Publ., 2012.
- [13] What is SPARC? *SPARC*. Available at: <https://sparc-camp.org/> (accessed June 1, 2018).
- [14] Tsvyk I.V. *Vestnik RUDN, seriya Filisofiya — Peoples’ Friendship University Journal of Philosophy*, 2017, vol. 21, no. 3, pp. 379–388.

Moiseenko M.V., Cand. Sc. (Philosophy), Associate Professor, Department of Ethics, Faculty of Humanities and Social Sciences, Peoples’ Friendship University of Russia. Research interests: spiritual and moral heritage of the Russian philosophers abroad, the ethical aspects of the art, ethical aspects of religion, the balance between politics and morality, problems of professional ethics, ethical aspects of the development of artificial intelligence. e-mail: moiseenko_mv@rudn.university