

Mariusz Wojewoda, prof. UŚ dr hab.

Uniwersytet Śląski w Katowicach

Wydział Humanistyczny, Instytut Filozofii

mariusz.wojewoda@us.edu.pl; tel. 604-590-850

Sztuczna inteligencja – problemy etyczne (Artificial Intelligence - Ethical Problems)

Słowa kluczowe: system techniki, technicyzacja zachowań ludzki, nowe kompetencje, „syntetyczny” podmiot moralny

W moim wystąpieniu chciałbym się skupić na problemie AI rozumianej jako system informatyczny - wariant systemu techniki. Tworzenie AI jako przyjaznej dla ludzkiego środowiska oznacza, że potrafimy, lub w najbliższym czasie będziemy potrafili tworzyć systemy, które będą kształtować ludzkie zachowania. Przede wszystkim chodzi o poprawę efektywności działania organizacji i form komunikacji społecznej. Od strony technologicznej prace nad stworzeniem coraz wydajniejszej AI są zawansowane. Jednak poziom obaw związanych z jej wprowadzeniem w obszar życia społecznego jest bardzo duży. Społecznie oczekiwania dotyczą stworzenia AI przyjaznej i życzliwej dla człowieka. To stawia przed informatykami i filozofami techniki zadanie stworzenia etycznej maszyny, wrażliwej na ludzkie wartości. Autor wystąpienia stawia tezę, że to jest możliwe. Jednak należy pamiętać o całym szeregu zagrożeń, które związane są z funkcjonowaniem AI w przestrzeni życia społecznego.

In my speech, I would like to focus on the problem of AI being understood as an information system - a variant of a system of technology. Creating AI as friendly to the human environment means that we can, or soon, will be able to develop systems that will shape human behavior. First, it is about improving the efficiency of organizations and forms of social communication. Technically, work to create increasingly efficient AI is well advanced. However, the concern about its introduction into social life is grave. Socially, the expectation is to create AI that is friendly and kind to humans. This sets computer scientists and philosophers of technology to create an ethical machine sensitive to human values. The author of the speech posits that this is possible. However, it is crucial to keep in mind a range of risks associated with the functioning of AI in the space of social life.

Dusek, V. (2006). *Philosophy of Technology. An Introduction*. Blackwell Publishing Ltd. Oxford.

Larson, E. (2021). *The Myth of Artificial Intelligence. Why Computers Can't Think the Way We Do*. The Belknap Press of Harvard University Press. Cambridge, Massachusetts, London.

Maas L.J., Lukas Snoek, Claire E. Stevenson (2021). How much intelligence is there in artificial intelligence. A 2020 Update? *Intelligence*. Vol. 87, issue C. DOI: [10.1016/j.intell.2021.101548](https://doi.org/10.1016/j.intell.2021.101548).

Mitcham, C. (2022). What is Living and What is Dead in Classic European Philosophy of Technology. In: *Oxford Handbook of Philosophy of Technology*. Ed. Shannon Vallor, The Oxford University Press, New York 2022, 19–34. DOI: [org/10.1093/oxfordhb/9780190851187.013.3](https://doi.org/10.1093/oxfordhb/9780190851187.013.3).

Neubauer, A.C. (2001). The future of intelligence research in age of artificial intelligence – With a special consideration of the philosophical movements of trans- and posthumanism. *Intelligence* 87. DOI: [10.1016/j.intell.2021.101563](https://doi.org/10.1016/j.intell.2021.101563).

Sautoy, du M. (2020). *The Creativity Cod. How AI is Learning to Write, Paint and Think*, Harper Collins Publishers, New York.