

**Yu. Mielkov**

Institute of Higher Education of the NAES of Ukraine, Ukraine  
9, Bastionna Street, Kyiv, 01014  
uka7777@gmail.com  
<https://orcid.org/0000-0002-8186-0357>

## **AI CHALLENGES AND HUMANIST FOUNDATIONS FOR THE TRANSFORMATION OF HIGHER EDUCATION**

**Abstract.** Challenges related to the growing usage of technologies based on generative AI are considered from the perspective of philosophical analysis. It is argued that the AI is but a program that follows algorithms, and thus it is not capable of thinking or being a subject of any activity at all. The current AI issues are argued to lead to the transformation of higher education on the principles of humanism, so that technologies could augment human beings and their activities, while allowing humans to have control over AI and to evaluate them in an appropriate way.

**Keywords:** artificial intelligence, higher education, humanism, human thinking.

### **Introduction**

The rapid development of the technologies based on the special kind of artificial intelligence – generative AI, an advanced version of it having been presented to the general public at the end of 2022 – has led to many discussions that started to appear both in mass media and in academic publications. The very place of human person in the today's world of artificial intelligence and machines that are supposedly able to “think” is now being questioned.

If we review in brief the responses that appear in relation to the possible challenges and opportunities brought by AI to the human society, we can see quite a broad spectrum ranging from pure optimism to predictions of a threat to the very existence of humanity. On the one hand, the benefits of using AI in finance, national security, health care, and research and education as well are outlined, referring to its ability to combine information from a variety of sources, to analyze it instantly, and to act based on that analysis. As stated by Darrel West and John Allen, with massive improvements in storage systems, processing speeds and analytic techniques, AI technologies are capable of sophistication in analysis and decision making, thus providing the tremendous opportunities for economic development [1].

On the other hand, according to at least some of the leading experts in the sphere of machine learning, the development of AI-based technologies could be a real threat to humankind. In 2022 in a survey conducted

amongst leading researchers on the topic, the question “What probability do you put on future AI advances causing human extinction or similarly permanent and severe disempowerment of the human species?” had 5% positive replies – with as many as 10% of the experts answering also positively on the similar question about the possibility of human extinction caused by human inability to control future advanced AI systems [2].

Of course, machine learning researchers are not futurologists, and there was no elaboration on exact ways the AI development could lead the humankind to extinction; nevertheless such survey results demonstrate both the need to be cautious in evaluating the outcomes of the research (namely the existential risks coming out of it) and the need to investigate those outcomes further under the growing amount of debates on the said topic, where arguments, as properly noted by Jaron Lanier, aren't entirely rational [3].

Still, if we can't be quite sure of the ways and probability of the AI to cause human extinction, we can much more clearly envision less critical but still radical changes those technologies could lead to in industry and education. If machines are capable of thinking and generating knowledge, then they could definitely surpass their creators in many forms of activities that used to be traditionally associated with human beings. As it follows from the “Future of Jobs Report 2023” presented by World Economic Forum, “All but two technologies are expected to be net

job creators in the next five years: humanoid robots and non-humanoid robots” [4, p.6].

In other words, the threat of growing unemployment caused by AI technologies being introduced into new spheres of economy is quite evident – and has already led to numerous strikes in 2023. While the already quoted survey predicts the future unemployment for the majority of truck drivers, surgeons and salespersons (and unsurprisingly not for AI researchers!), one of the most vulnerable spheres is that of education, and especially higher education. As noted by scholars, with AI and its applications finding their way into educational practices, it could well mark the end of higher education as we know it [5]. Thus, it is even more evident that such perspectives requires no less thorough consideration than the abstract threat of a destruction, and *this paper is aimed* to provide some musings and discussion on the possible challenges AI technologies manifest for our society and for the system of higher education.

### **Results and Discussion**

First of all, I would like to argue that the issues related to AI technologies and their impact on human activities are issues of philosophy – amongst else, if not in the first place, – and not just of technology, including machine learning. In particular, it is philosophy that is engaged both in analyzing the terms we use and their correctness of that usage – and especially in considering those senses and meanings we attach to such terms.

And the first problem here is that the very term “artificial intelligence” is at least partly misleading and obscure – due to the lack of clarification of what *intelligence* actually is. Moreover, the Ukrainian translation of the term, which rather means “artificial intellect” literally, is thus doubly misleading as it hints that some human-made entity can possess the abilities similar or close to that of a human mind. And while nothing like mind or consciousness is meant here in the English version of the notion, it is still unclear what are those exact abilities that are being designated by the very idea of artificial intelligence?

On the other hand, if any philosophical considerations are usually being conducted post-factum (as noted already by Hegel), then we have to deal with notions that have been coined and gained widespread usage in social and academic discourse without any kind of profound analysis. Jaron Lanier, an American scientist and philosopher, who has coined one such popular term himself (that one is “virtual reality”), finds such a situation not only misleading, but even dangerous. In his essay with the eye-catching title “There is no A.I.” [3], Lanier argues that it is much more appropriate to think of A.I. as a tool, not a creature capable (or not capable) of possessing a kind of intelligence – the latter would be just mythologizing the technology.

And the technology in question is not some self-sufficient entity, but an innovative form of social collaboration between humans. In other words, the technology we call artificial intelligence can’t create anything in principle – it is not the subject of any activity at all, to say nothing of it being an intelligent or a cognizing subject. That technology can but combine something already created by humans basing on its programming coded again by humans themselves – as Lanier puts it, such programs “mash up work *done by human minds*”, and the one innovation here that enables us talking about a significant improvement is that the mashup process in question becomes even more guided and constrained now, providing more elaborate and sound results. “This is a significant achievement and worth celebrating – but it can be thought of as illuminating previously hidden concordances between human creations, rather than as the invention of a new mind” [3].

If we turn back to considering the origins of the term “Artificial Intelligence”, we would see that the very idea of it is not at all new – the notion was coined in about 1950s, amidst the then growing optimism about science and its ways to improve and reconfigure human society. Moreover, already in the middle of the 19<sup>th</sup> century, at the dawn of industrialization, it was proclaimed that technologies could both improve and augment humans in their activities (thus threatening them with unemployment in the long run).

Charles Babbage, who is credited with the original concept of a digital programmable computing machine, used to stress it as early as 1832: “One great advantage which we may derive from machinery is from the check which it affords against the inattention, the idleness, or the dishonesty of human agents” [6, p.54].

On the other hand, such harsh words, which could be explained by Babbage’s personal inclination towards misanthropy (that was noted by Charles Darwin), provide a sound judgment on the abilities of machines to do purely *mechanical work* and thus to replace human in precisely that type of labor. In particular, Babbage used to argue that a skilled mathematician should not spend his or her time and efforts performing simple calculations as a basis for more elaborate research – and because such simple tasks that could well be done by a less qualified person could be with even more success left to a computing machine. After all, the latter could do such mechanical and thus non-creative tasks much faster than humans, whose “inattention” could be explained already by the inhuman nature of that mechanical job.

Still, while nobody would argue that a machine is capable of doing simple calculations or processing data at a rate that exceeds any abilities of a human person, it is still not enough to call that capability *intelligence* – at least not in the philosophical meaning of the word. I would argue that the notion of “Artificial Intelligence” reflects the discussions about *whether a machine could think* that were quite popular in the middle of the 20<sup>th</sup> century. When Alan Turing put that question up in the 1940s, he was not referring to any philosophical or psychological concept – the idea of thinking has rather been reduced to a technological approach of whether a machine could be mistaken for a conscious human being while blind-checking the replies we get from it while questioning or communicating. That is what the Turing test is all about, after all, – and it is hardly necessary to prove that an ability to respond to some questions in a more or less logical way can’t be made equal to the ability to think.

For example, when in 1960s Joseph Weizenbaum has created ELIZA, one of the first computer programs that could communicate with humans using a kind of natural language and matching patterns of speech and thus imitating a psychotherapist or someone else of the same sort – it was quite evident that this early kind of AI can at least attempt to pass the Turing test, but can’t be credited with the actual ability of thinking or understanding. All that was quite clear to Joseph Weizenbaum himself, but not to some representatives of the general public or even science, it seems, as the creator of ELIZA had to describe the illusion of “a thinking machine” as the confusing and dangerous one. In particular, he noted that “an entirely too simplistic notion of intelligence”, which is peculiar to both popular and scientific thought, is partly responsible for the emergent “perverse grand fantasy” on the nature and ability of artificial intelligence: “Man is not a machine. I shall argue that, although man most certainly processes information, he does not necessarily process it in the way computers do. Computers and men are not species of the same genus” [7, p.203].

“Intelligence”, however broad the usage of the term would be, is not a synonym of “thinking”, and “generative” as an attribute of artificial intelligence does not actually mean “creative”. A machine being “intelligent” is an illusion that follows out of that machine being programmed by humans to look like a human (to calculate, to communicate etc.) – but any program, even the one based on the newest technologies of machine learning and the one realized in the form of digital neural networks still remains a program.

From a philosophical perspective, such “intelligence” is reduced then to but simple operations of formal logic [8]. Thinking in algorithms, according to Albert Rutherford, could even be useful to humans themselves [9], – however, I would argue, that it is totally not the whole thinking! Any schemes used in thinking appear to deny the reality of being with its plurality and diversity [10]. The simple idea of following the “if... then...” scheme while allowing AI with computational powers to follow rather well artificial situations with pre-set rules (like, say, to play

chess quite skillfully and better than most human players), limits its creative reasoning. It is virtually not possible for any program to “think out of the box” or to follow dialectical or any non-classical logic other than simple formal logic.

My favorite example of such kind of limited reasoning is the experiment conducted by Alexander Luria, one of the founders of the Ukrainian school of psychology, in the very beginning of the 1930s in Middle Asia. In order to investigate and to compare the logic of traditional common sense and that of modern school, Luria proposed residents of Uzbekistan a question representing a simple syllogism: it is known that there are no camels in Germany, and Berlin is the capital of Germany, – so can you please tell if there are any camels in Berlin?

It turned out that the representatives of the younger generation who have been attending modern schools, replied to the inquiry without any hesitation: no, there are no camels in Berlin. At the same time, elder people answered the question in an opposite way: yes, they said, there are camels in Berlin. When asked to explain their answer, they replied that Berlin is a big city, so it is quite probable that an Uzbek with a camel could come and visit a bazaar in Berlin [11].

What we can see here is the opposition of *formal reasoning* – and, so to say, *contentual thinking*. Formal logic and algorithms are examples of the former, and common sense with all its lack of academic discipline is a follower of the latter. It could be noted that while reasoning by syllogism is entirely correct formally, it does not reflect the actual state of events. That is, the procedure of making a conclusion out of two premises does not question the validity of those premises: if Germany is indeed in no way a natural habitat of camels, that does not necessarily mean that there are no camels at all in the country in general and in its capital city in particular. In other words, those elder people who used to follow the common sense and the traditional informal reasoning, could have actually been more correct in their conclusions than the younger generation that was following the rules and the algorithms and was not considering the question by its

real context. That is, a single camel may in fact be actually found in Berlin after all – if not in a bazaar, then at least in a zoo or in a circus.

That example not only shows us the profound distinction between the two types of reasoning in question – the mechanical deduction and the human thinking, – but also draws our attention to education and its role in understanding and using thinking. After all, it was the younger generation of people accustomed to the traditions of formal schooling who, while facing no difficulties in solving syllogisms or following algorithms, did not even try to, well, actually think about the question they were asked! Could it be the fault of that very schooling that impoverishes human thinking rather than enriches it? And could it be that the current situation in education is even more mechanically inclined than it was in the middle of the 20<sup>th</sup> century with all its standardization, massification, formalization and adherence to discipline in reasoning? French existentialist philosopher Bernard Stiegler was one of the first to note the danger of a user of digital technologies being transformed into data him- or herself, so that the artificial intelligence could actually make humans more stupid [12; 13]. Would not thinking in algorithms also lead humans to be less creative and similar to machines in their formal and thus limited reasoning?..

It is indeed clear that the development of technologies based on generative artificial intelligence (however philosophically unsound that term could be) leads us to the necessity to consider many new issues related to the sphere of education, and especially higher education. While many researchers refer to such obvious topics as digital literacy and the problems of academic plagiarism, with ChatGPT-4 and other similar tools being extensively and relatively effectively used for, say, writing term papers instead of students – while all these topics are still important, the AI challenges relate to the very understanding of the goal and the mission of higher education in the first place.

And that’s why the most important task of today’s higher education becomes positioning the AI-based technologies in the correct way and outlining the principles of its

usage. While researchers agree that generative AI holds immense educational potential due to its ability to handle unstructured and natural language data, they still argue that it is important to assess all the risks (like the possibilities of “AI hallucinations”, i.e. cases of AI-based resources providing false data) and to allow AI programs to enhance the learning experience by putting humans in total control over what they are using and for what purpose exactly [5].

Considering all that was said about the limitations of AI’s ‘thinking’, it could be argued that such technologies are not capable of replacing educators (who are thus less threatened by the possibility of unemployment than salesmen or truck drivers) – but the fact is that teachers do have less digital literacy than students, which hinders the process of adopting a set of norms and rules of using AI for study and research. Even with all the attempts to establish a kind of AI ethics [14], there is no clear distinction between situations when using AI should be allowed, and when it should not. AI skepticism and bans would not work anyway: as surveys show, the majority of students who already use ChatGPT and other similar tools indicate that they will continue to use it even if their professors or institutions would ban the technology [15].

However, I would argue that the main problem here is not the lack of digital literacy of teachers (that could be taken care of by organizing a series of instructive webinars and so on) nor the lack of normative regulations (which could be established either on national or institutional levels as well). I think that the main problem is the lack of critical evaluation of the AI-based technologies (and of the information provided by those tools and programs) on the part of students. In accordance with what was noted about the AI technologies replacing humans in non-human work, it becomes even more necessary to follow the classical ideal of university in educating creative and critical thinking and the ability for a person to follow one’s own mind, which is exactly what the philosophical conception of humanism is all about. And, as noted by Martin Bekker, critical thinking as a vital part of education is

at threat when there is a tendency among students to treat the AI-based technologies as something that is superior to humans – and thus to trust anything ChatGPT would tell [16].

The idea of trust and “horrors” related to some new emergent technologies is not exactly new: some people used to (and some still do) believe anything a television would say to them, as earlier they tended to trust every written word they would get. The new digital mass media is in no way an exception here. However, the difference is that the new type of media are much more accessible, so that each person can become a kind of creator – and the possibility of, say, a fake message distributed as a propaganda by central media controlled by an authoritarian or even a totalitarian state is being replaced by a whole ocean of information. Here it becomes quite difficult to distinguish a true message from a fake message already due to the enormous volume of messages available.

Of course, that’s not just the problem of AI-based technologies, but the problem of trust in general. As argued by contemporary researcher of digital media Christian Fuchs, many people now just distrust facts, truth, experts and research – they tend to believe only what seems to be emotionally comforting and ideologically acceptable to them. Moreover, the public sphere gets fragmented and polarized, with right-wing extremists steering hatred online against minorities and quality media. Which is especially visible under the current state of the war in Ukraine and elsewhere: “Digital technologies also shape warfare. Digital warfare has extended and intensified the destructive capacities of military technologies” [17, p.2].

I would argue it is the lack of education and general culture that leads many humans to distrust experts – while trusting bloggers who are just ‘simple’ people like themselves who talk in common everyday language. The phenomenon of Wikipedia, a free online encyclopedia where virtually anyone can write an article on any topic, serves as the best example of this new type of media. There are of course editors who are to maintain the credibility of the articles submitted, but it

would still be inappropriate to consider any Wikipedia article to provide but true, well-grounded and complete information.

That example is especially striking if we compare Wikipedia to Scholarpedia – the resource created in pursuit of credibility, so that a similar free online encyclopedia would feature articles written by invited or approved expert authors only, to say nothing of the peer review and other similarities to academic journals. Launched in 2006, Scholarpedia has only 1817 content pages by April 2024 and in English only [19], while Wikipedia has more than 6.8 million of articles, more than 60 million pages and more than 47 million registered users in its English version by the same date – with still more millions of articles, pages and active users in 348 other languages [20].

The reference to Wikipedia is quite appropriate here: the researchers like the already quoted Jaron Lanier are right in defining new AI-based neural networks like GPT-4 as “something like a version of Wikipedia” [3], as the problem of trust is practically the same in relation to Wikipedia in particular and digital mass media in general – and in relation to AI. For a layman, who does not have neither skills nor inclination to rational inquiry and academic research, it would be difficult to spend time checking the information from an article in Wikipedia – it would be a matter of trusting or distrusting the information provided, just like in the case of “communicating” with ChatGPT (especially with its free versions that, unlike the more recent 4.0, do not provide references).

That means that we have a kind of *circulus vitiosus* here: you have to be an expert in some field of knowledge to check and confirm the information provided by either an article in Wikipedia or a reply by ChatGPT, as well as to be in control of AI-based technologies. But if somebody is already an expert, he or she is in no need to consult Wikipedia or ChatGPT, – after all, those are first of all tools for people to get something they do not know yet!

However, the conclusion should not be that pessimistic, after all – if we do remember that Wikipedia or ChatGPT are *tools* capable

of fast processing the high volumes of information *provided by humans*, and they are *not experts nor even subjects of judgments*. And one more use those technologies could be of is actually the translation into many natural languages – something that was not available before to billions of people. “La traduction est la langue de l'Europe”, as was once said by Umberto Eco, who stressed that the future challenge for the united Europe is going toward multilingualism, toward a polyglot community [18]. I'd say that the noted trend is useful for the whole world, and not just for Europe: AI-based translation plays even more important role in education than processing the information, as it enables humans to communicate with each other and to access knowledge that used to be limited to only one language.

### **Conclusions**

We can thus make some conclusions on the topics considered. The AI-based technologies represent many challenges for humans, especially in the sphere of education, but many possibilities as well. In order to make those technologies useful and to be in control of them, it is important to position them in the correct way and to acknowledge both their limitations and their abilities.

IT is argued that the technology we call artificial intelligence can't actually be the subject of any activity at all, to say nothing of it being an intelligent or a creative subject. The machine cannot ‘think’ in principle – but it can imitate human reasoning by following algorithms and formal logic due to being programmed by humans. Thus, no machine could be capable of exactly “replacing” a human person in his or her human activities: AI can only take on the mechanical tasks humans have to spend their time and efforts on. It does not make much sense to fear any sound rise of unemployment due to the development of AI technologies: only humans that do not perform their job in a creative way may have such fears.

AI could well replace human workers in doing routine, mundane tasks that could be automated – thus helping humans and allowing them to dedicate themselves to more creative work. And in order to envision and to

use ChatGPT and other similar technologies as tools capable of fast processing the high volumes of information, some transformations in higher education are desirable. It is necessarily to not only improve the digital literacy of teachers and to form and adopt a set of normative and ethical regulations for using the AI-based technologies, but also to follow the classical ideal of university in educating creative and critical thinking and the ability for a person to follow one's own mind – in order to be able to evaluate the results provided by the AI appropriately. The ways to realize that could be the possible topics for the new research on the problem.

### References

1. West, D. M. & Allen, J. R. (2018). How artificial intelligence is transforming the world. *The Brookings Institution*, 24 April. [Online]. Available: <https://www.brookings.edu/articles/how-artificial-intelligence-is-transforming-the-world/>
2. *2022 Expert Survey on Progress in AI* (2022). AI Impacts, 3 August. [Online]. Available: <https://aiimpacts.org/wp-content/uploads/2022/08/2022ESPAIV.pdf>
3. Lanier, J. (2023). There is no A.I. *The New Yorker*, 20 April 2023. [Online]. Available: <https://www.newyorker.com/science/annals-of-artificial-intelligence/there-is-no-ai>
4. World Economic Forum (2023). Future of jobs report 2023. [Online]. Available: [https://www3.weforum.org/docs/WEF\\_Future\\_of\\_Jobs\\_2023.pdf](https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf)
5. Matthijs, N. (2024). What AI means for higher education teaching: hype vs reality. *University World News*, 27 January. [Online]. Available: <https://www.universityworldnews.com/post.php?story=20240123095419557>
6. Babbage, Ch. (1846). *On the economy of machinery and manufactures*, 4<sup>th</sup> ed. London: John Murray.
7. Weizenbaum, J. (1976). *Computer power and human reason: From judgment to calculation*. New York & San Francisco: W. H. Freeman & Co.
8. Zinchenko, V., Mielkov, Y., Nych, T., Abasov, M., Trynyak, M. (2023). Human thinking in the age of generative A.I.: Values of openness and higher education for the future. *2023 International conference on electrical, computer and energy technologies (ICECET)*, Cape Town, South Africa, p. 1-6. DOI: 10.1109/ICECET58911.2023.10389449
9. Rutherford, A. (2023). *Thinking in algorithms: How to combine computer analysis and human creativity for better problem-solving and decision-making*. APB Publications.
10. Myelkov, Yu. (2004). From explanation to understanding: the paradigm shift in visioning the factuality. *Science of self-organization and self-organization of science* (I. Dobronravova and W. Hofkirchner, eds.). Kiev: Abris, p. 48–58.
11. Luria, A. R. (1976). *The cognitive development: Its cultural and social foundations*. Cambridge, MA: Harvard University Press.
12. Stiegler, B. (2001). *La technique et le temps. Vol. 3, Le temps du cinéma*. Paris: Editions Galilée.
13. Stiegler, B. (2003). Our ailing educational institutions. *The Culture Machine*, 5. [Online]. Available: <https://web.archive.org/web/20040711060841/http://culturemachine.tees.ac.uk/Cmach/Backissues/j005/Articles/Stiegler.htm>
14. Mukherjee, A. (2023). *AI and Ethics: A computational perspective*. Bristol, UK: IOP Publishing Ltd.
15. Osman, R. (2023). Intelligent responses needed for artificial intelligence. *University World News*, 12 May. [Online]. Available: <https://www.universityworldnews.com/post.php?story=20230512052612694>
16. Bekker, M. (2023). Heroes and horror stories: Thinking about ChatGPT. *Wits University Webinar*, 27 March. [Online]. Available: <https://www.youtube.com/watch?v=vdpt6795elo>
17. Fuchs, Ch. (2022). *Digital humanism*. Bingley, UK: Emerald Publishing.
18. Eco, U. (1994). Traduction et l'ange parfait. *Dixièmes assises de la traduction littéraire*. Arles: Actes Sud, p. 191–207.
19. <http://www.scholarpedia.org/article/Special:Statistics>
20. [https://en.wikipedia.org/wiki/Wikipedia:List\\_of\\_Wikipedias](https://en.wikipedia.org/wiki/Wikipedia:List_of_Wikipedias)

The article has been sent to the editors 30.04.24.

After processing 02.12.24.

Submitted for printing 30.12.24.

Copyright under license CCBY-NC-ND