

UDK 004.8

DOI: 10.25688/2078-9238.2022.39.3.05

Robert P. Lockard⁽¹⁾,
Olesya A. Zmazneva⁽²⁾,
Илья N. Volnov⁽³⁾

Artificial Intelligence: Are Humans Protected from the Systems They Created?¹

The article appeared as a result of the main theses made on the Open International public discussion with both scientists and industrial participants from the USA, France and Russia that took place in Moscow Polytechnical university in April 2021. The questions under review are devoted to strength, weaknesses, opportunities and threatens that Artificial Intelligence brings to our life. AI have already changed the way we live and work. And from one side, AI can rewrite the social rules and human cognitive bias. At the same time humans can stay on a control position to AI by having risen from the level of intelligence, capable of working with the new as a new combination on a given set of known elements, to the level of thinking, capable of generating something else that is absent in the previous human experience and to the level of meanings that are inalienable from its storage medium. The authors, cybersecurity specialist Robert P. Lockard (USA) and scientists Ilya N. Volnov (MosPolytech) and Olesya A. Zmazneva (IT-department, MosPolytech, Russia) tried to look at AI connected questions from both pragmatic and philosophical points of view.

Keywords: AI; Artificial Intelligence; new skills; social rules; cognitive abilities.

For citation: Lockard Robert, P., Zmazneva, O. A., & Volnov, I. N. (2021). Artificial Intelligence: Are humans protected from the systems they created? *MCU Journal of Philosophical Sciences*, 3 (39), 47–55. DOI: 10.25688/2078-9238.2022.39.3.05

⁽¹⁾ *Lockard Robert P.* — an Oracle ACE Director, recognized internationally for his work in securing databases systems and is the winner of the 2015 Oracle Developers Choice Award for database design. The author of the book «Oracle Database Application Security» (2018).

¹ Текст печатается в авторской редакции (The text is printed in the author's editon).o

Oracle ACE Director (USA).

E-mail: rob@oraclewizard.com

- (2) *Zmazneva Olesya Anatolievna* — PhD (Philology), Associate Professor, associate professor of infocognitive department Moscow Poly.

Moscow Polytechnical University, Bolshaya Semyonovskaya st., 38, Moscow, 107023, Russia.

E-mail: ozmazneva@gmail.com

- (3) *Volnov Ilya Nikolaevich* — PhD (Engineering).

Moscow Polytechnical University, Avtozavodskaya st., 16, Moscow, 115280, Russia.

E-mail: ilja-volnov@yandex.ru

УДК 004.8

DOI: 10.25688/2078-9238.2022.39.3.05

Роберт П. Локард⁽¹⁾,

О. А. Змазнева⁽²⁾,

И. Н. Вольнов⁽³⁾

Искусственный интеллект: защищен ли человек от созданной им системы?

Статья представляет собой ряд основных тезисов, которые были сделаны ее авторами на Международной публичной онлайн-дискуссии представителей академического и индустриального сообществ России, Америки и Европы, состоявшейся в Московском политехническом университете в рамках Международного телемоста в апреле 2021 года. Основными вопросами, которые ставят перед собой авторы, являются вопросы недостатков и преимуществ систем искусственного интеллекта, их возможности и потенциальные угрозы для человека, их создавшего. Уже сегодня искусственный интеллект трансформировал социальные практики. И, с одной стороны, AI может существенно изменить социальные основы и поведенческие паттерны. Но, с другой стороны, человек имеет возможность контролировать AI, поднявшись с уровня интеллекта, оперирующего отчуждаемой от своего носителя информацией, на уровень смыслов, не отчуждаемых от своего носителя; с уровня интеллекта, способного работать с новым как новой комбинацией на заданном множестве известных элементов, на уровень мышления, способного порождать иное — отсутствующее в предыдущем опыте человека. Авторы рассматривают эти проблемы как в философском, так и в прагматическом аспектах.

Ключевые слова: AI; искусственный интеллект; новые навыки; социальные правила; когнитивные возможности.

Для цитирования: Локард Роберт П., Змазнева О. А., Вольнов И. Н. Artificial Intelligence: Are humans protected from the systems they created? [Искусственный интеллект: защищен ли человек от созданной им системы?] // Вестник МГПУ. Серия «Философские науки». 2021. № 3 (39). С. 47–55. DOI: 10.25688/2078-9238.2022.39.3.05

- (1) *Роберт П. Локард* — специалист по кибербезопасности, автор книги «Oracle Database Application Security» (2018).
Oracle ACE Director (США).
E-mail: rob@oraclewizrd.com
- (2) *Змазнева Олеся Анатольевна* — кандидат филологических наук, доцент, доцент кафедры инфокогнитивных технологий Московского политеха.
Московский политехнический университет, Большая Семеновская ул., д. 38, Москва, 107023, Россия.
E-mail: ozmazneva@gmail.com
- (3) *Вольнов Илья Николаевич* — кандидат технических наук.
Московский политехнический университет, Автозаводская ул., д. 16, Москва, 115280, Россия.
E-mail: ozmazneva@gmail.com

Introduction: philosophical approach

We've had three industrial revolutions: the steam engine, mass production, and the digital era. Each of these changed the way we live and work. The fourth industrial revolution is AI (Schwab, 2017). And the opposition that is often discussed today is «Human vs AI».

Lets start with considering some features of human intellectual activity. To do this, we will distinguish at least 2 levels in the structure of consciousness — intelligence and thinking. Often in modern psychology, neurophysiology and in the developers' environment (Vedyahin, 2021, p. 30), these terms do not differ and are considered synonymous. The maximum that AI developers can distinguish between is artificial intelligence (AI) and the phenomenon of consciousness (Tegmark, 2017), but we will show their fundamental difference and in this difference we will indicate the possibility of compensatory mechanisms.

Intelligence can be understood as a personalized mechanism for working with information for processing it into other types of resources according to known rules and algorithms (Schwab, 2017). It ensures the survival of the individual and is characterized by utilitarianism, a focus on the material, the desire for efficiency and simple, effective, pre-understandable, well-defined results. Beyond the boundaries of the intellect is thinking, understood as its opposite. Thinking can be defined as the ability of a person to perform activities that go beyond biological instincts and expedient behavior, it is not utilitarian and is turned to the ideal. Thinking tends towards complexity and uncertainty, and prefers the interesting to the effective. Thinking, unlike intelligence, is able to distinguish good from good, good from evil, new from different, important from unimportant.

In this comparison, we understand that AI is able to take over only the intellectual part of a person's consciousness.

Alan Turning said “machines can be constructed which will simulate the behaviour of the human mind very closely” (Turing, 1996, p. 256–260). Because the human brain is about two pounds of matter, it was assumed that we could create

Artificial Intelligence that mimicked the human brain. It was not until 1997 when Deep Blue beat Garry Kasparov, the world chess champion that AI was approached with a different mind set. Researchers did not look at how humans played chess, the researchers at IBM focused on playing chess well. This was a completely different approach from early AI research that focused on how humans accomplish a job; this is also when researchers started to be practical.

AI doesn't need to accomplish the task as a human would do it, it just need to do the task well. This is an important point that will keep people from being blindsided by the abilities of AI.

Thinking for AI intelligence is apparently not available and will remain another cognitive limit to the possibility of its development. As we have already seen, the problem here is not that the machine begins to think like a person. The problem is that a person starts thinking like a machine. A person voluntarily renounces thinking from that part of his consciousness that is not formalized and, therefore, cannot be digitized.

It is in thinking that the mechanisms of overcoming AI by a person, its cognitive advantages and the very possibility of holding a controlling position to AI are found. Let's list them:

– Despite the fact that a person psycho-physiologically can not work with big data (at the same time we can not hold more than 7–9 units of information in the mind), a person is able to work with “deep” data, i. e. in big data to distinguish important, necessary and significant information from unimportant, unnecessary and irrelevant. This is how our thinking works: it reduces big data to deep data, discarding redundant information, and builds a model for making the right decision based on the latter. All classical science is built according to this scheme.

– Intelligence works with information, thinking works with flushes. The difference between information and meanings is that information can be alienated from a person, but meanings cannot. This means that the AI works only with meanings reduced to information, and therefore such information will always be incomplete. No matter what big data is used by AI, they are always incomplete, and in the decisions made by artificial intelligence, there is a small, but fundamentally irremediable error.

– One of the arguments of the supporters of strong AI is that a person is not able to make good decisions quickly. However, this argument is valid only at the level of intelligence, at which the loss of human AI became obvious quite a long time ago, starting with the defeat of G. Kasparov to the chess supercomputer “Deep Blue”. But this statement is unfair for the level of thinking, where a person can instantly make the right decisions and there are enough examples of this. It is enough to list a few famous names of mathematicians Sofia Kovalevskaya and Srinivas Ramanujan, physicist Nikola Tesla. Russian shipbuilder Peter Akindinovich Titov, chief engineer of the Franco — Russian Shipbuilding Plant in St. Petersburg and an employee of the famous academician-shipbuilder A. N. Krylov, may be mentioned among the less well-known personalities. P. A. Titov did not have a special education, was not properly educated, but was the author of a lot of inventions.

Which he made “by eye”, and when people who received a European education came to check it, it turned out that all the sizes assigned by Titov strictly coincided with the scientific formulas.

– Through thinking, a person can work with environments characterized as infinite. Thus, within the framework of the probabilistically oriented model of consciousness created by prof. V. V. Nalimov (Vedyahin, 2021), the fundamental possibility of connecting our consciousness with the “semantic vacuum” — an infinite environment containing all possible meanings of the world and extracting specific meanings from it in certain life situations is shown. The ultimate transition from big data to infinite data is also apparently inaccessible to AI and is another of its cognitive limitations.

AI: cases, experience, facts, pragmatical approach

The fourth industrial revolution we live inside brought, as we have already said, by AI. One of the most popular threats is that AI will eliminate jobs. But the fact is that in the past 100 years only one job has been completely eliminated: the elevator operator (Klein, 2019). If you also look at wages and quality of life over the past 100 years, these three industrial revolutions have increased wages and improved quality of life for everyone.

As AI gets better at specific tasks than humans, we can use the AI machine to augment our work, and improve the quality of life. We are already seeing AI based medical systems bringing advanced medical care to remote villages (Guo, Li, 2018), and AI based fraud detection systems improve financial services. When the Automated Teller Machine (ATM) first came about, there was discussion that bank tellers would be eliminated. This turned out to be false, the number of bank tellers increased over the past thirty years. ATM’s allowed bank tellers to spend more time with customers to provide better financial services. “AI could contribute up to \$ 15.7 trillion to the global economy in 2030, more than the current output of China and India combined. Of this, \$ 6.6 trillion is likely to come from increased productivity and \$9.1 trillion is likely to come from consumption-side effects.” Economic forces will continue to drive AI forward in both research and implantation. We’ll continue to find new ways to put AI to work; many of these will be incremental improvements to the way we work, such as AI Operations in the data center that detect something going wrong and can be corrected before it becomes a large issue. Or AI in the diagnosis of disease; catching a medical issue before it becomes life threatening.

AI can rewrite the rules and our cognitive bias. Because AI requires massive amounts of information to learn, it can and will find patterns in the data that we as humans don’t see. This is because of two reasons, our cognitive bias, and the sheer amount of information AI can process. Our cognitive bias is a filter that causes us to overlook information that may be relevant. Let’s consider AlphaGo; in it’s

thirty-seventh move of the second game against Lee Sedol, AlphaGo placed a stone on the fifth line from the edge. This threw Lee Sedol for a loop; for thousands of years of the game go, there has always been the rule of thumb; avoid placing a stone on the fifth line from the edge (Hassabis, 2017). Why did AlphaGo make that move? This exposes an issue with AI, it will do things that are unexpected, and we are at a loss to understand the logic. This has led to research into understanding the logic behind why AI produces some results. AI does not equal nor does it need to equal human intelligence. It is because of this, we're going to see AI making more inroads into the job market; and some of those inroads will come as a surprise.

As a society we define jobs from the top down. Someone is an Accountant, Lawyer, Doctor, Professor, or Truck Driver, etc. Where AI and Automation in general work is from the bottom up; accomplishing discrete tasks that can be automated, freeing people up to focus on the tasks that require a human to perform. If a task has a well defined goal, there is enough training data, and we can define when that goal is achieved, then there is a high probability that task can be automated (Brynjolfs-son, Mitchell, 2017). There will be a chipping away at the corners of the task we perform in our work; bringing us more time to work on interesting tasks that AI can not do. Now just because AI can not do it now, does not mean it won't be able to do it in the future. Kai Fu Lee, in "AI Superpowers: China, Silicon Valley, and the New World Order" said "50 % of all jobs will be automated by AI inside of 15 years" (Lee, 2018, c. 155). We are not so much of a pessimists and believe the issue is much more nuanced. If you spent time writing out what AI can not do, by the time you're done, you'll need to restart your list.

Just as the ATM became the coworker of the bank teller, we believe as AI advances, it will become our coworker, augmenting our abilities, helping us make better decisions, and take over work that is dangerous for humans to do. When Oracle developed the Autonomous Database, that freed up Database Administrators (DBA) to focus on more interesting work. The database administrator knows the data he/she is responsible for better than anybody else. So, database systems that migrated over to the autonomous database freed their database administrators to become the de facto data scientist for the organization. This involved additional education in data science, and those DBA's that undertook the transition to data science were rewarded with more interesting work, and higher salaries, a 10.4 % salary increase in the US National Average (*Glassdoor salaries survey*, 2020).

With AI, jobs will evolve, and some will become less prevalent. In 1950, there were 342,000 telephone operators working for Bell Telephone Systems, in 2020, there were 4,630 telephone operators in the United States (*US Bureau of Labor Statistics*). The decline in telephone operators was a direct impact from advancing technologies. In the future we can see truck drivers, and taxi drivers impacted by advances in autonomous vehicles the same way as the telephone operator. As these jobs disappear, other jobs will take their place. AI will be a slow encroachment in all of our careers, gradually taking over the tasks we need to do. Because of this,

we'll need to rethink parts of our education systems and social safety nets to prevent people from becoming unproductive members of society.

Society has used the model of going to school and getting an education for the first twenty – twenty five years of life to get prepared for a career. Till today the situation looked like this: once you have chosen your selected career, in many industries, there is little continuing education. The information technology industry demands continued education simply to stay relevant. In 1979, the computer had 8K of memory, you loaded the bootstrap by hand, programs, and data was stored on tape. That skill to operate that computer system served well in 1979; however, today it's not much more than trivia. The mindset of continued education should be adapted to other industries that will be impacted by AI and automation.

People will be displaced with the advances in AI, if they don't have a constant learning stream, and guidance on where to focus their learning, people will fall into unemployment or underemployment, putting a burden on society. There is another side of this we tend to ignore, that is the impact on unemployment/underemployment on the person. Many people get a large part of their identity from the work they do, and without that work, there is an increase in mental health issues (Linn et al., 1985). AI will make its biggest impact initially in jobs that have little social interaction and require little dexterity. These would be jobs like, product inspection, garment worker, and truck driver. Although these jobs may not get completely eliminated, AI will make a large impact, reducing the number of people employed in these fields.

What are the skills people need to start picking up to make the ride into the AI revolution as smooth as possible? Two of the most prized skills for the future will involve Creativity and Social based skills (Zmazneva, 2016). These will be careers such as Public Relations, Criminal Defence Attorney, and Social Worker (Kai-Fu, 2018). All of the skills are high in human interaction and creativity; places where AI is stubbornly not able to make much headway in. Even in these "safe" jobs, there will be tasks that AI can perform, improving the decisions made, deciding on the proper way to frame a message before giving it. Because of this, we need to be prepared to adapt as AI takes on tasks we are accustomed to doing ourselves.

In the end, AI will improve the human experience: freeing up people to do interesting things in their work, removing the dull tasks few people enjoy doing, improving healthcare outcomes, and the increase in global GDP by \$ 15.7 Trillion by 2030, will increase wages across the board. Humans need to be ready to adapt.

Conclusion

At the level of intelligence, no one is protected from AI, especially a developer. As a programmer, a developer is forced to think in an artificial programming language that is extremely linear and unambiguous. That is, the strengths of natural languages in their nonlinearity, ambiguity and a wide semantic range of words,

reflected in the flexibility of the consciousness of the speaker of these languages, are completely absent here.

A person (any person), especially a programmer, in principle, cannot work with big data and obviously loses to AI in this field.

Human cognitive abilities are deployed in a different direction — in the ability to work with deep data, i. e. with valuable information and meanings. It is in this field that protection from AI should be sought, but in this field social groups engaged in various types of intellectual activity such as science and philosophy feel confident.

It is possible to remain in a control position towards AI by:

- having risen from the level of intelligence, operating with information alienated from its storage medium, to the level of meanings that are inalienable from its storage medium.

- having risen from the level of intelligence, capable of working with the new as a new combination on a given set of known elements, to the level of thinking, capable of generating something else that is absent in the previous human experience.

- having risen from the level of intelligence, which is completely in metrological, physical (dead) time, to the level of thinking, living in biological, living time with the uncontrollable insights of another form of time interwoven into it — the Greek Kairos.

References

- Aczel, A. D., Ono, K., & Ramanujan, A. S. (2016). *My search for Ramanujan: how I learned to count*. Springer International Publishing. 238 p.
- Brynjolfsson, E., & Mitchell, T. (2017). What Can Machine Learning Do? Workforce Implications. *Science*, 358, 6370. URL: <https://science.sciencemag.org/content/358/6370/1530>
- Dreyfus, H. L. (1992). *What Computers Still Can't Do: A Critique of Artificial Reason*. MIT Press. 776 p.
- Ford, M. (2018). *Architects of Intelligence. The truth about AI from the people building it*. Packt Publishing. 554 p.
- Glassdoor salaries survey* (2020). Retrieved from <https://www.glassdoor.com/>
- Guo, J., & Li, Bin (2018). The Application of Medical Artificial Intelligence Technology in Rural Areas of Developing Countries. *Health Equity*, 2 (1), 174–181. DOI: 10.1089/heq.2018.0037
- Hassabis, D. (2017). *Artificial Intelligence: Chess Match of the Century*. *Nature*. P. 544; 413–14. Retrieved from <https://www.nature.com/articles/544413a>
- Kai-Fu, Lee. (2018). *AI Super Powers China, Silicon Valley and the New World Order*. Retrieved from <https://www.technocracy.news/ai-superpowers-china-silicon-valley-and-the-new-world-order/>
- Klein, A. (2019). *Not All Robots Take Your Job, Some Become Your Co-Worker. Real Clear Markets*. Retrieved from https://www.realclearmarkets.com/articles/2019/10/30/robots_dont_take_your_job_they_become_your_co-worker_103962.html
- Linn, M., Sandifer, R., & Stein, S. (1985). Effects of unemployment on mental and physical health. *Am J Public Health*, 75 (5), 502–506. DOI: 10.2105/ajph.75.5.502

- Nalimov, V. V. (2007). *Spontannost` soznaniya: Veroyatnostnaya teoriya smy`slav i smy`slavaya arxitektonika lichnosti [Spontaneity of Consciousness: Probabilistic theory of Meanings and semantic architectonics of personality]* (2-e izd.). M.: Vodolej Publishers. 374 p. [in Rus.].
- Nikitin, V. (2013). *Inzhenernaya ontologiya. Inzheneriya kak stranstvie [Engineering Ontology. Engineering as a Journey]: Study Guide / V. Nikitin, S. Pereslegin, A. Paribok et al.* Ekaterinburg: Azhur. 230 p. [in Rus.].
- Penrose, R. (1989). *The Emperor`s New Mind*. Oxford University Press. 480 p.
- PwC (2018). *Sizing the Prize, PwC`s Global Artificial Intelligence Study: Exploring the AI Revolution*. Retrieved from <https://www.pwc.com/gx/en/issues/data-and-analytics/publications/artificial-intelligence-study.html>
- Russell, S. (2019). *Human Compatible: Artificial Intelligence and the Problem of Control*. Viking. 352 p.
- Schwab, K. (2017). *The Fourth Industrial Revolution*. Penguin UK.
- Schwab, K., & Malleret Th. (2020). *COVID-19: The great reset*. Forum publishing.
- Tegmark, M. (2017). *Life 3.0: Being Human in the Age of Artificial Intelligence*. New York: Alfred Knopf Publ. 364 p.
- Turing, A. (1996). Intelligent Machinery, A Heretical Theory. *Philosophia Mathematica*, 3 (4), 256–260. Retrieved from <https://cyberleninka.ru/article/n/97-04-009-tyuring-a-m-myslyaschie-mashiny-ereticheskaya-teoriya-turing-a-m-intelligent-machinery-a-heretical-theory-philosophia-mathematica>.
- US Bureau of Labor Statistics. Retrieved from <https://www.bls.gov>
- Vedyahin, A. (2021). *Sil`ny`j iskusstvenny`j intellekt: Na podstupax k sverxrazumu [Strong Artificial Intelligence: On the approaches to Superintelligence]*. M.: Intellektual`naya literatura. 232 p. [in Rus.].
- Zmazneva, O. (2016) Govorit`, chtoby` pobedit`: o novy`x programmax po razvitiyu kommunikativny`x navy`kov studentov na IT-fakul`tete Moskovskogo Politekha [Talking to win: about new programs for the development of students' communication skills at the IT Faculty of the Moscow Polytechnic University]. *Kachestvo obrazovaniya [Quality of Education]*, 9, 26–29. [in Rus.].