

An Artificial God:

The Takeover of Artificial Intelligence

Alexander Hernandez Romanowski

GLBL 507: Making Decisions Under Uncertainty

Rice University

December 16, 2020

In light of the upcoming revolution in humans' lives connected with Artificial Intelligence (AI), the topic of AI research is addressing deeply relevant and acute problems of mankind. AI significantly influences all aspects of our society, even more than the past industrial and digital revolutions did [11]. This type of technology is found in common applications such as the filters and smart replies in email communication, the recommendation algorithms of the most popular sites like YouTube, Google, and Facebook, in navigation and ride sharing applications like Uber, Doordash, Google Maps and others. The implementation of AI is undoubtedly widespread, but current--and future, no doubt--research is rapidly broadening the day-to-day applications of technology in ways previously unimaginable. The pace of technological inventions has increased over time and will likely only increase in the future. Raymond Kurzweil, the famous American inventor, named this pattern of accelerating technological progress “exponential increase” [9].

Millions of years ago, our ancestors lived in a primitive society, hunting wild animals, making fire and collecting roots. Now a man, according to an outstanding Russian scientist Vernadsky, is a “driving geological force”. According to the 2017 Club of Rome Anniversary Report, we live in the era "Anthropocene", it's defining characteristic being humans as having the most significant impact at the planetary level [16]. Within this period is what is known as the “Great Acceleration”. This point in time in the history of mankind began around the mid 20th century and is characterized by the rapid development of almost all aspects of life.

Origin and Development of AI

We can positively state that by the 1950s we had a generation of scientists, mathematicians, and philosophers with the concept of artificial intelligence (or AI) culturally

assimilated in their minds. In December of 1955, Herbert Simon and Allen Newell developed the “Logic Theorist” which is considered the first artificial intelligence program.

Back in 1928, German mathematician David Hilbert attracted attention of the world community with the problem of resolution (Entscheidungsproblem), which is this: Is there some procedure that can be used to determine if a particular statement is provable or whether there remains the possibility that some statements¹ are destined to remain unresolved? He put it simply: “There is no such thing as an unsolvable task.” You just need to find an algorithm that would take a description of any problem as input decidability (formal language and mathematical statement "S" in this language) - and, after a finite number of steps, it would stop and give one of two answers: "True" or "False"[20]. In 1936 Alan Turing, who is now considered to be the founder of the modern theory of AI, together with the American mathematician and logician Alonzo Church, independently proposed a theorem, known now as the "Church-Turing theorem", about the absence of an algorithm that solves the problem of permissions, in other words, that there is no pure mechanical procedure which would make it possible to decide whether a given statement is from a set of mathematical axioms. [21]

In the article "On computable numbers, with an application to the problem of solvability", Turing developed the concept of abstract digital computing machine, later called the Turing machine, capable of simulating (if there is an appropriate program) any machine, the action of which performs a transition from one discrete state to another [15]. This was the first attempt to automate logical and mathematical calculations that at that time could only be performed by humans. In his next article "Computing Machines and the Mind", Turing outlined a mental experiment, later called the Turing test, in which two people, a man and a woman, communicate to a third person, the questioner through a channel on opposite sides of the wall, eliminating the perception of a voice. The questioner's goal is to determine the

¹such as, for example, mathematical riddles: Fermat's last theorem, Goldbach hypothesis or Collatz hypothesis

gender of each of their interlocutors; while the man tries to confuse the questioner, the woman helps the questioner to find out the truth. The Turing test was then adapted to having one human and one machine respondent to answer the question of whether a machine can participate in this "imitation game" indistinguishably from a human and will the questioner be mistaken in his conclusions more often or not. Back in 1950, the question first arised: "Can a machine think?" Or, more precisely, "Can a machine adequately communicate with a person in a natural language?" [10]. 68 years later, modern "intellectual" machines are already able to simulate individual intellectual functions, individual mental processes (pattern recognition, decision logical tasks, playing chess, etc.).

The famous French economist Jean-Baptiste Say wrote in 1828 about the possibility of replacing horses with cars: "However, no machine will ever be able to do what even the worst horses can – service in the transportation of people and goods through the hustle and bustle of the great city." [28] Say may have never dreamed, even in the most distant aspect of imagination, about self-propelled cars, drones, calls through Skype, supercomputers, smartphones, or intelligent robots. Technology that seemed like pure science fiction less than 200, 50, even 20 years back is available today, and some fantastic upcoming inventions are likely to be widely used during the next twenty years. The problem is to be realistic when predicting upcoming artificial intelligence technologies and not to get into the same myopic trap of Say and others.

The article by Makridakis [11] lists some of the most impressive modern projects over the past 25 years, such as a robot Polly created in the artificial intelligence lab

Massachusetts Institute of Technology by Ian Horsville, 1st mobile robot that can move at the speed of animals (1m per second) using computer vision to navigate; victory of the Deep Blue Chess Computer over the World Chess Champion Harry Kasparov on May 11, 1997 in a six-game match with a standard time. Kasparov won the first game, lost the second

and the following three draws. When Deep Blue won the final game, Kasparov refused to believe it. Watson's supercomputer created by the company IBM in 2010 beat Ken Jennings and Brad Rutter, two record holders for Jeopardy, a TV quiz game with simple questions in the fields of history, literature, sports, geography, science and other branches of knowledge. AlphaGo program won 18-times world champion in the game "GO" Li Sedol with a score of 4-1. [11] However, in all of these examples, the AI systems lack intellectuality in the true sense of the word - they are not capable of self-study, cannot meaningfully understand human speech and participate in a meaningful dialogue with a person, are not able to creatively approach the problem solving, and do not have the same flexibility of behavior, which is characteristic of a person. In this regard, we are coming to the other questions: "Will the computer ever have consciousness?", "Can human consciousness be computed by a machine?", "Can we talk about a fundamental difference between artificial and natural intelligence?" And though it is difficult for scientists to answer these questions, they do not seek solutions in the development of real "artificial intelligence", rather, the main efforts focus on solving specific, practically significant problems.

Perhaps this is due to the fact that about 30 years ago, when AI research was just beginning, scientists were convinced that they could create a thinking computer by simply increasing the volume of its memory and resources for data processing, without deep knowledge of how the real human brain works [24]. However, the development of the artificial intelligence problem should be closely related to the results of research on natural intelligence, which is not limited to cognitive functions but presents conscious activity in general. That is why a full-scale study of the structure of the biological brain is important. Such features as the interconnections between neurons will create an accurate simulation of artificial intelligence and reveal the secret of the perception of the surrounding world. In his book "Artificial Intelligence. Stages. Threats. Strategy" (2014) Swedish philosopher Nick

Bostrom distinguishes three stages of complete emulation of a human brain: detailed scanning of the human brain with its vitrification or a glass transition, reconstruction from scan data three-dimensional neural network responsible for cognition in the biological brain, by the method of automatic processing, and finally, the creation of digital copies of the original intelligence with memory intact and an intact personality type based on a neural network computing structure, loaded into a super powerful computer [18]. Right now scientists have succeeded to carry out the first stage described by Nick Bostrom and are rather close to the second.

AI and the Holy Trinity

As for the future development of digital intelligence, according to recent research, in the nearest future we will be facing the new “Holy Trinity”. It is composed of blockchain, the Internet of Things and AI. With Blockchain controlling all the “monetary” transactions (deposits) of the future, the Internet of Things will be supporting it as a network of devices (gadgets, sensors, computers, etc.) which communicate between themselves without involvement of humans. This means that the next stage can be about these devices becoming uncontrollable by humans, even if in the future we would like to regain control on them, e.g., when facing catastrophic consequences. We will elaborate on possible dangers of “The Holy Trinity” in a subsequent section.

Singularity of AI

The Singularity is the hypothetical future creation of superintelligent machines. Superintelligence would be characterized by the ability to surpass human intelligence and complete processes that are beyond what humans can do. Since the ability to design such a machine is also one of these intelligent actions, the superintelligent machine can build even

more advanced machines. This will undoubtedly be followed by an "intellectual explosion", and the human mind will lag far behind artificial intelligence. Thus, it is possible to conclude that the first superintelligent machine will become the last invention of humanity, provided that the machine will be submissive enough and tell us how to keep it under control.

What will be the consequences of this event? When progress is guided by an intellect superior to that of humans, it gets much faster. Actually, there is no reason to believe that progress will not produce more and more reasonable entities at an increasingly accelerated pace. The best analogy that can be carried out here - in the evolutionary past. Animals can adapt and be creative, but not faster than the natural selection process. In the case of natural selection, the world itself appears in the role of your own simulator. We, the humans, have the ability to absorb the world around us and build a causal communication in our heads, so we solve many problems thousands of times faster than the mechanism of natural selection. When an opportunity will arise for the machines to solve these problems at higher speeds, we will find ourselves in a situation, which differs from our human past no less radically, than we humans differ from the lower animals. Such an event cancels the entire principle of human dominance, perhaps in a blink of an eye. An uncontrollable chain reaction will begin to develop exponentially with no hope of regaining control of the situation.

Changes that were thought to take "thousands of centuries" (if they were to happen at all) are likely to happen in the next hundred years. It would be quite reasonable to call this event a Singularity. It is the point at which our old models have to be discarded and we accept a new reality. This is a world, whose outlines will become clearer, taking advantage ahead of the modern humanity, until this new reality will overshadow the surrounding reality, becoming a new normal.

Makridakis asks a vital question: "What will be the role of humans at a time when computers and robots will be able to perform all the tasks that people are currently doing, but better and much cheaper [11]?"

Danger of AI

Some researchers are concerned that intelligent machines can become our "ultimate invention" and put an end to human domination, which is why it is so important to predict the impact of the AI revolution. There is no doubt that AI has a huge potential, since computers and robots are likely to approach or even reach human intelligence within the next twenty years, becoming a serious competitor for all workplaces that are currently filled with people. For the first time, there may even be a question regarding the continued human domination.

Supporters of the "strong AI theory" believe in the fundamental possibility of creating computers, which would be able not only to think, but to understand and feel; with their creation the halo of mystery will be removed from consciousness. Their firm opponents ("weak AI") deny this possibility, pointing out fundamental limitations of computers: adherence to strict rules, lack of creative and flexible behavior inherent in a person, etc. However, supporters of "weak AI" take a flexible position. They admit that the computer can simulate the mentality, but it cannot duplicate it, i.e. have a full-blooded experience of consciousness, inner life and genuine understanding.

The international community may argue about the consequences of the emergence of smart machines but it admits one thing: the development of technology will have an unconditional impact on society, economy and relations between people in the future. So what awaits us in these two or three decades? Will Singularity become established in the human worldview? And what will be the onset of the Singularity itself? What can be said about the true nature of this event? It is likely to be the fastest technical revolution of all we

have known before. It will fall, most likely, like a snow on one's head - even for the scientists involved in the process. And what will then happen in a month or two after that? There is only one analogy, which I can use - the emergence of a new sort of humanity. We will find ourselves in a posthuman era.

The Club of Rome in its anniversary report "Come On! Capitalism, myopia, population and destruction of the planet" (2018) claims that there is a real danger of uncontrolled development and unethical use of technologies, and it is not yet clear how to avoid this. The authors agree with the experts who believe that Moore's Law will cease to operate in 2020 - 2025, so no singularity is expected. Also, promises of techno-utopians are demotivating to people: if a certain technology solves all problems, there is no need to search for complex solutions that require changes in the lifestyle [16]. All this can lead us to supercapitalism with superinequality, when the share of earned income approaches zero, and the share of income from capital will rise to the maximum, forcing the majority of people to live on benefits.

In the article by Makridakis [11] it is said at all that companies in developed countries will not transfer their production to developing countries to use their cheap labor as they will be able to achieve the same at a lower cost using technology (AI), thereby increasing the tendency for the process of returning the production lines back to their countries [4]. And the main consequence of these events will be a massive dropout from the middle class to the poor, where people will have to adapt by settling in unproductive sectors of the economy with less income.

Furthermore, 3.5 million truck drivers carrying cargo could be at risk in the United States alone, according to the latest report of the international energy agency [26]. There will also be no need for such traditional, creative and currently irreplaceable professions such as lawyers, doctors, economists, journalists, engineers, programmers, nurses, teachers,

designers, musicians. Already now, according to an experiment at the University of Tübingen, a neural network can paint a picture, plausibly forging styles of famous artists such as Picasso, Van Gogh and Munch, compose music and master the complete basics of small motor skills.

The Multi-Arm Unmanned Ground Vehicle is capable of using simple locksmith tools, tying knots, and even picking locks on doors and inspecting hard-to-reach areas. Most importantly, there may be a threat for the teachers. At the Moscow Technological Institute robot Alantim is already giving lectures to students. Thus, a depreciation of education can occur in many areas of our activity. In his book "Superintelligence", Nick Bostrom illustrates this with the following statement: "Suppose that the smartest person on Earth is conventionally twice as smart as the stupidest, and artificial intelligence evolving exponentially in a few years surpasses a man thousands, and then millions of times. At the level of this height both today's genius and today's dumbass will be equally negligible".

Finally, AI-powered robots can increase risk of discrimination: racism, sexism and other forms of it. A widely known scandal was related to the granting of Saudi Arabian citizenship to the humanoid robot Sophia on October 25, 2017 at the summit "Future Investment Initiative" in Riyadh. Sofia was designed by the Hong Kong company Hanson Robotics in a way to learn and adapt to human behavior and to work with people. Sophia has already had many meetings around the world. The event related to the issuance of a citizenship to her provoked controversy, as some commentators started asking questions, whether this implied that Sofia could vote or marry, and whether deliberate disabling of the robot can be considered murder. Social media users used this event to criticize the human rights situation in Saudi Arabia. The event caused a scandal in Saudi Arabia, as during conferences and announcements of citizenship Sofia was with an uncovered head and face, with no hijab and no male guardian. It is crucial to note that Saudi Arabia is a country where

women remain disenfranchised, while the robot, who clearly identifies itself as a woman, receives citizenship. One response to this is that Sofia's lawyers managed to find "loopholes" in the legislation of the country, and this allowed the female robot to obtain full civil rights, which, in turn, are denied for Saudi women. Additionally, critics noted that citizenship was given to the robot easily, while millions of labor migrants work in the country with limited rights. In particular, under the laws of Saudi Arabia, foreign workers are not allowed to leave the country without a permission of their employer [25, 23].

Can we mitigate the danger of AI?

With all the rising concerns regarding the progress of technologies, the most important question right now is whether we are actually able to mitigate or shoulder the consequences of an explosive AI development. Our answer is: yes, and we further propose some suggestions with examples to prove it.

From ancient times, philosophers developed the discipline of deontology, or the science or doctrine regarding ethics. However, according to Prof. Mellichamp: "It seems to me that a lot of needless debate could be avoided if AI researchers would admit that there are fundamental differences between machine intelligence and human intelligence. Differences that cannot be overcome by any amount of research. In other words, the "artificial" in artificial intelligence is real." AI cannot reach the level of human intelligence, consequently, no ethics is possible within it - that is why we, the humans, need to develop instruments to control it.

Let's start by saying that the statement of Prof. Mellichamp seems reasonable to us, just because no ethics or ethical norms can be traced in any computer or technological development process. Furthermore, we suggest that humanity should not expect from a computer, for instance, to control itself, but must take advantage of the current situation and

take the necessary precautions. Most importantly, such examples as the industry of autonomous military robots, which are used to kill and destroy based on technology without involving any human, prompt us to the idea that the AI may become an existential threat for our species (questioning the existence of humanity).

The most important suggestion we can propose is to implement the Man-machine rules - a set of principles approved, for example, by the Engineering and Physical Sciences Research Council of the UK, - into the current laws and judiciary of our countries. We claim that the existing rules in the robotics industry can be easily adjusted for AI technologies because the essence of the two is the same: an autonomous machine. Most important rule is called "Personal Data is a Thing" - in other words, there is data belonging to each specified individual, which cannot be used anywhere without their consent. The laws targeting data protection are already implemented by the European Union and the State of California - this proves that with the energy of the individuals all the necessary rules restricting AI can be put as laws with the appropriate enforcement - e.g., prosecution of the scientists consciously creating or inventing dangerous systems and AI mechanisms.

No less important, is which of "our" countries such restrictions should be processed first. By any means, this is most urgent for the developed countries, such as the US and China, as the problem with the AI control there is perhaps closest. Nevertheless, it would be relieving to see such norms being implemented into all the judiciary systems of the world's countries.

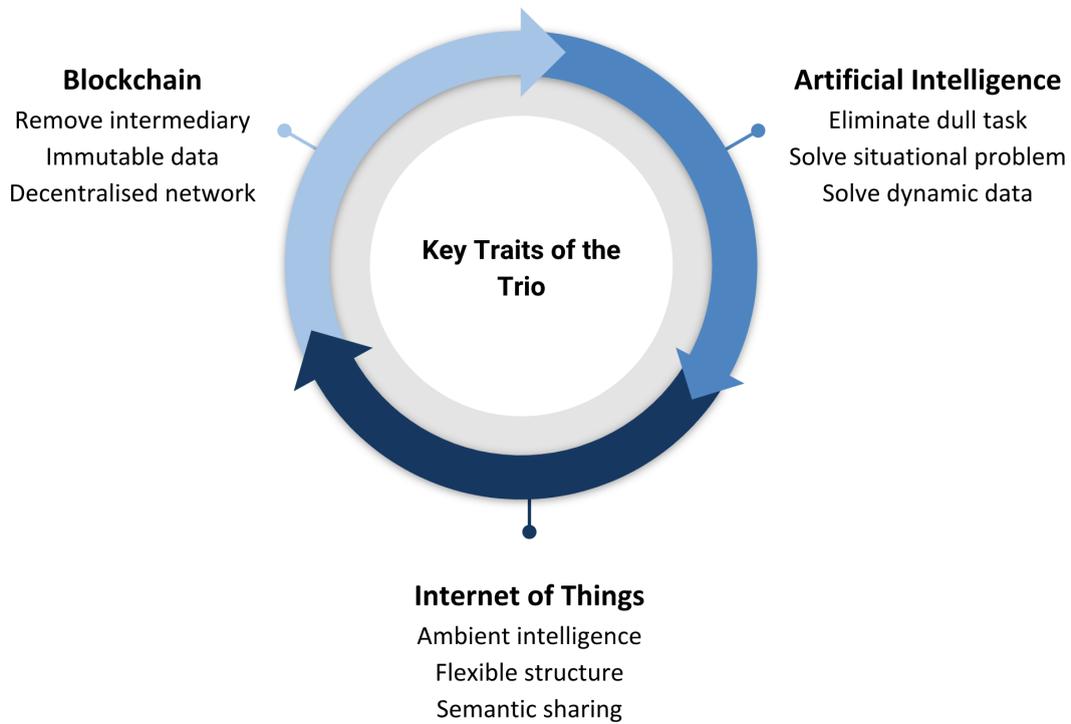
It is my strong belief that we can achieve the balance within our civilization, of which AI became an inextricable part, if we are guided with a culture of ethical values - as a result of our consciousness, not any computer program. We should not let the engineers, claiming the AI "safe" and the singularity problem "solved by itself", rule the opinion of our society - because they do it only for their product and business profit. They are trying to disable the

alert system of the society. In turn, science can play the leading role in shaping the people's perspective. Therefore, we can say that science and education here take on the most important mission to save the world.

In this time when we are constantly becoming more global, we need to strengthen our universal conscience to see our common values. The most important thing is to see that the biggest thing that we have, before race, ethnicity, tribe, religions, etc., is that we are the human species. Beginning this very instant, we need to strengthen our unity as the human race to preserve our human essence. The most powerful tools we have to do this is increase our love, our global leadership, our faith, our family values. With that we can increase our hope. To achieve peace, we need to fight for it. I am optimistic in spite of all of the negative realities. I remind myself constantly that the most sold book in the world is the bible, and the last book of the bible, Revelation, argues that good always triumphs. Darkness doesn't exist, it is merely the absence of light. We will keep shining our lights, however, from now and into the future.

Appendix:

FIGURE 1: Key Traits of the New Holy Trinity



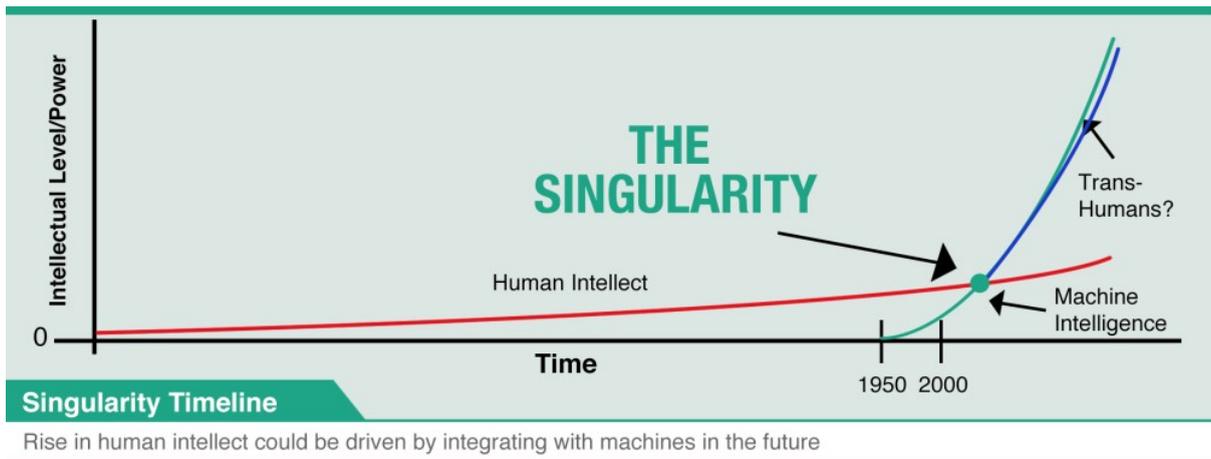
Source: Medium

FIGURE 2: Likelihood v. Consequence Matrix

	Consequence				
Likelihood	Insignificant	Minor	Moderate	Major	Severe
Almost Certain	Medium	High	High	Extreme	Extreme
Likely	Medium	Medium	High	Extreme	Extreme
Possible	Medium	Medium	High	High	Extreme
Unlikely	Low	Medium	Medium	High	High
Rare	Low	Low	Medium	High	High

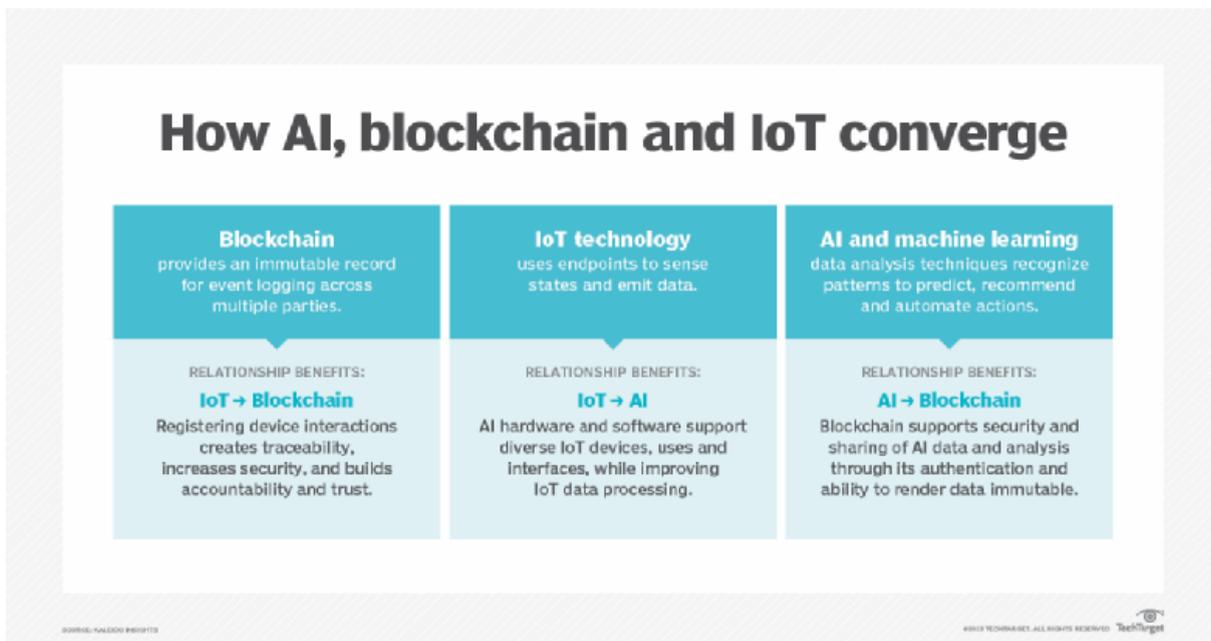
Source: Paladin Risk Management Services

FIGURE 3: The Singularity Timeline



Source: Colvin Consulting Group

FIGURE 4



Source: TechTarget

References:

1. Amunts K. et al. *BigBrain: an ultrahigh-resolution 3D human brain model* //Science. – 2013. – Vol. 340. – №. 6139. – P. 1472-1475.
2. Autor D. et al. *Polanyi's paradox and the shape of employment growth*. – Cambridge, MA: National Bureau of Economic Research, 2014. – P. 20485.
3. Ding S. L. et al. *Comprehensive cellular-resolution atlas of the adult human brain* //Journal of Comparative Neurology. – 2016. – Vol. 524. – №. 16. – P. 3127-3481.
4. Ford M. *Rise of the Robots: Technology and the Threat of a Jobless Future*. – Basic Books, 2015.
5. Glasser M. F. et al. *A multi-modal parcellation of human cerebral cortex* //Nature. – 2016. – Vol. 536. – №. 7615. – P. 171-178.
6. Harari Y. N. *Homo Deus: A brief history of tomorrow*. – Random House, 2016.
7. Huth A. G. et al. *Natural speech reveals the semantic maps that tile human cerebral cortex* //Nature. – 2016. – Vol. 532. – №. 7600. – P. 453-458.
8. Karabarbounis L., Neiman B. *The global decline of the labor share* //The Quarterly Journal of Economics. – 2013. – Vol. 129. – №. 1. – P. 61-103.
9. Kurzweil R. *The singularity is near* //Ethics and Emerging Technologies. – Palgrave Macmillan, London, 2014. – P. 393-406.
10. Machinery C. *Computing machinery and intelligence-AM Turing* //Mind. – 1950. – Vol. 59. – №. 236. – P. 433.
11. Makridakis S. *The forthcoming artificial intelligence (AI) revolution: Its impact on society and firms* //Futures. – 2017. – Vol. 90. – P. 46-60.
12. Penrose, R. *Shadows of the Mind*. L., 1993.
13. Rosenblatt F. *The perceptron: A probabilistic model for information storage and organization in the brain* // Psychological Review. — 1958. — V. 65. — P. 386–408.

14. Searle J. *Minds, brains and science*. The 1984 Reith lectures. L., 1984
15. Turing A. M. *On computable numbers, with an application to the Entscheidungsproblem* // Proceedings of the London mathematical society. – 1937. – Vol. 2. – №. 1. – P. 230-265.
16. Von Weizsäcker E. U., Wijkman A. *Come On!: Capitalism, Short-termism, Population and the Destruction of the Planet*. – Springer, 2017.
17. User. (2018, December 17). *Man-Machine Rules*. Retrieved December 17, 2020, from <https://madsiblog.tradoc.army.mil/106-man-machine-rules/>
18. Moroz, O. (n.d.). *Cyber ethics*. Retrieved December 17, 2020, from <https://etika.nplus1.ru/etichka/bibliograph>
19. *Deontology*. (n.d.). Retrieved December 17, 2020, from <https://iphlib.ru/library/collection/newphilenc/document/HASH015cc25c36368caea0a5d95c>
20. Written by Artur Kluz. (n.d.). How to be a leader in the digital age. Retrieved December 17, 2020, from <https://www.weforum.org/agenda/2016/05/how-to-be-a-leader-in-the-digital-age/>
21. Foffano, F., Scantamburlo, T., Cortes, A., & Bissolo, C. (n.d.). European Strategy on AI: Are we truly fostering social good?
22. The Future of Implications for Energy and the Environment. (n.d.). Retrieved December 17, 2020, from <https://iea.org/publications/freepublications/publication/Oil&GasSecurityNL2012.pdf>
23. Semantic Maps. (n.d.). Retrieved December 17, 2020, from <https://gallantlab.org/huth2016/>
24. GITTLESON, B. (n.d.). Saudi Arabia criticized for giving female robot citizenship, while it restricts women's rights. Retrieved December 17, 2020, from

<https://abcnews.go.com/International/saudi-arabia-criticized-giving-female-robot-citizenship-restricts/story?id=50741109>

25. Фролов, А. (2017, July 26). Поставщик Apple пообещал построить завод в США за \$10 млрд после требования Трампа перенести производство iPhone в Америку - Офтоп на vc.ru. Retrieved December 17, 2020, from <https://vc.ru/flood/25450-foxconn-usa-plant>
26. Barsanti, Sam. Saudi Arabia takes terrifying step to the future by granting a robot citizenship (англ.), The A.V. Club.
27. Sophia. (2020, September 01). Retrieved December 17, 2020, from <https://www.hansonrobotics.com/sophia/>
28. Say, J. B. (1828). Cours complet d'économie politique, Chez Rapilly, Paris.