Quote number 4

Kate Crawford states that artificial inteligence (AI) is not a neutral tool that may be used by human beings. She believes that its output is dependent on interests of people and institutions that have power and that is caused by the fact that it is built in a specific way to benefit them. Her conclusion is that use of AI in social context like politics, healthcare etc. will result in amplifying inequality and unjust hierarchies. There are three main points of her view, which are represented by each of three sentences above. I agree with the first one that modern AI based on neural networks is something more than mere calculating tool. However I think that social inequalities aren't the main danger of AI and its impact on society may be harmful to rational flow of information and knowledge in general. In order to argue for this position I will firstly describe the difference between logical AI that was the first paradigm of creating artificial inteligence and concepts of neural networks which constitute modern AI. This context will be important in the later parts of the essay. Then I will give reasons why I agree with Kate Crawford that the AI we currently know is more than just calculation tool. Finally I will argue that AI is something that may be a danger for humanity and we need to be careful if we want to keep control of our knowledge and future.

Development of AI was at first connected with mathematical logic and automatisation of mathematical proofs. These machines, called problem solvers, that were built in the begining of the second part of XXth century, were purely logical, but their use was reduced to some formal computations. Later there were attempts to create programs with a specific knowledge included, that would serve in more specific and profesional branches of science and engineering. Those two ideas, even if they differ in many ways, have in common that they were based on belief that AI should be created top-down. To understand this it might be good to use terminology of Daniel Dennett. That AI was an intelligent design of intelligent designer. Something that was based on a rational idea and understanding, not a blind process.

However, modern AI is based on artificial neural networks. The main difference between them and ideas of AI described above is that they are not something designed with specific information inprinted. They are given some input and the information they keep is based on a result of this learning from it. Some links between information are strengthened and some are weakened according to how often certain thing in the input is occuring. However, the result of this process is statistical. Neural network based AI may also learn from the interactions with the enviroment it is working and adjust to it. It is something dynamic and also autonomous. The process of its design is not purely rational, due to the fact that it is not fully supervised by human being, but it gives rational results, because AI often works really well and can be used successfully in practical situations. If we once again want to use terminology of Daniel Dennett, we may say that the information that neural networks has is an intelligent design without intelligent designer. Something that is functional and in some way rational, but which was not fully designed by intelligent creature aware of its purpose. Obviously the technical aspects are effects of human design, but the flow of information and its evolution in time is not fully designed, its rather statistical and probabilistic.

The use of Dennett's biological terminology was purposeful, because now I would like to argue for the view that modern AI, based on neural networks, is something which is closer to biological evolution and living organism than just a computational tool. It is why I agree with Kate Crawford that AI is not objective, but although it may be controlled it also has some range of agency. Both AI and regular calculating computers are something that according to input give certain output. However the second ones are based on algorithms and system of algebra that is easy to follow. There are specific rules and it may be traced why the result is what it is. These machines are static and may be compared to the simple natural phenomenon like falling rock. There is rarely anything unusual about them. On the other hand information structure in AI is not easily predictable. It follows some general regularities but we don't know how specifically it works or what process took place before the result was visible. We also don't see this process when we use simple calculator, but we can know about this if we check the algorithms. However it's nearly impossible for us, if not fully, to know what happened with neural network, we only see the result.

Neural networks are complex systems and the nature of that systems is often upredictable. This makes contact with artificial inteligence more similar to the interaction with wild animal or at least swimming in the deep ocean with huge waves. AI is something that is living on its own and we don't have enough knowledge to fully understand the processes taking place in it. I conclude then that this autonomy of AI makes it an agent. Most probably it is not aware of it but it is autonomous, it is changing and it is also not completly random, but shows some regularity.

Kate Crawford also belives that although AI is not objective, it serves the interests of people of power. What is more it is not objective, because it was designed in that way by them. I don't know the intentions of rulling class, however it might be true that there is a possibility of making AI supporting some social ideologies. It may depend on the input that is given during learing of AI or there may be other mechanism. Answer to this question should include the actual technical details of specific AI programs. This issue is certainly important, but there arise other problems, which are more general. There are to main of them. First one is that AI as autonomous agent is something out of human control. The second issue is to what extent can we trust AI.

As was stated in the paragraphs above, AI is something that is in many ways independent from human control. It is based on probabilistic processes that we cannot be fully aware of. Because of this the information given by the AI is not rational in a human sense. It is not intentional, there is no thought in it. AI text generators, that are popular nowadays, are good example of this. They often offer missinformation or something that seems to be gibberish, because they don't understand like humans do. They only find the most probable connections between words and this probability is based on the input that was given during the learning.

There might be a counter argument that it is only a technical issue and it will be improved. It probably will. However the real problem is not the content of the information. It is the process that creates it. AI may be truthful, however it is not purposeful. It's random process that seems to be correct, but communication is something more than just exchanging words. Words are sensual representations of mental states and thoughts. Interaction with AI is not connected with exchanging thoughts, it's something only resembling real communication, by its sensual appearence.

The information given by AI has meaning for humans, but it is not meant by AI itself. If we consider AI that base their learning on input made by other AI, we could see that it is dangerous. The probabilistic process of acquiring information made by other probabilistic process may be unpredictable. If we go further with iteration it is obvious that knowledge from AI will be less and less genetically connected with actual human ideas and discoveries. That may result in complete chaos in our education and data.

The second issue is connected with the first one. If AI is not intentional or communicating in the same way as humans do, how can we trust it with our lives? Once again there is a need to use Dennett's terminology. Information structure of AI may be referred to as a inteligent design without intelligent designer, it is created in a random process that is evolutionary (especially if we include evolutionary algorithms or competition between different AIs), but it results in a system that is well adjusted to reality. On the other hand information structures of humans are designed in a certain way to serve its purpose. Both systems are rational, however there is a huge difference between them, which also may be described in terms of evolutionary biology.

Rationality of AI and its evolution is limit (in a mathematical sense) of some process and due to this its development is similar to the biological evolution. Natural selection creates organisms that are well-adjusted to their habitat, but it is a result of many failed attempts, something that is still not perfect, but good enough to prevail. Death of an organism is no harm for nature, it is a cost that may be paid. However it is not enough for humans. How could we use airplanes that are just good enough to fly some distance? Or even if we gained very good airplane from the process of evolutionary design, the costs of many failed attempts is too huge.

Humans are designing their inventions and laws in a specific and rational way, because the value of their live is too important to risk it. Obviously both are changing and adjusting to the conditions, but it is change in intelligent design. On the other hand AI is not a rational designer, so humans should be suspicious when it comes to considering solutions given by it. There might be an

argument that people also may be dangerous or want to harm someone else. However this behaviour is also in a rational human discourse. Attempts of these people most of the times are in some way also an intelligent design. They understand some mechanisms in order to use or change it so they could achieve something. AI doesn't have this ability. What is more, humans also have morality, emotions and are able to be persuaded, the interaction with them is then much different.

A debate about the role of AI in future technological and social development of the humankind will probably emerge in the near future, so it is important to note those differences between human intelligent design and AI's "thinking". I personally don't feel competent to answer how much can we trust AI, but we should definately be careful when it comes to applying it to the scientific and technological research and solving social issues. What is more, this debate should include not only goverments, international organisations and corporations, but most importantly regular people from all over the world, so the won't come out as a tool of dictatorship and discrimination.

In this essay I responded to the views of Kate Crawford concerning AI. I agreed with her that it is more than just computation tool. I argued that it is much more autonomous than previous computer designs and it has its own agency to some extent, because its information structure is not fully designed by humans, but it arises from the separate process. Then I suggested that AI may be a danger to rational flow of information, because its answers doesn't base on representing thoughts, but on probabilistic connections between words. This also applies to the practical use of AI, which solutions are not an intelligent design, but are created through the random process of its learning. This also applies to solving social issues mentioned by Kate Crawford. Contrary to her, I don't believe that it is dangerous, because of the influence that power structures have on the development of AI, but it emerges from the essence of artificial neural networks. Because all of this modern AI development should be treated with special care and controlled by law, non-goverment organisations and include interests of regular people.