

Ethics for “Intelligent” Artificial Beings: A Possible Future

■ Zorica Mijartović^a, Orhan Jašić^b

Abstract

Artificial intelligence is a challenge for many scientists and researchers today. On the one hand, the development of artificial intelligence can be beneficial for people as it can significantly facilitate their daily lives and jobs. On the other hand, there is the fear that we will not succeed in developing “friendly artificial intelligence” and that these created intelligent beings becoming autonomous persons will bring many problems. Worries go so far as to assume that artificial intelligence will displace humans and take their place in all spheres of life. In this article, we have tried to present future scenarios concerning the development of artificial intelligence and indicate how necessary it is to have ethics in this discourse.

Keywords: Artificial Intelligence, Algorithms, Bioethics, Ethics, Software, Science Fiction, Technology.

Introduction

Artificial intelligence is the most current topic in the science fiction genre and there are many movies on this topic. What is more, in real life, technology research is advancing and striving to create as intelligent “software” as possible. In the science fiction stories we want to present, which deal with artificial intelligence, we see the need for an ethics encompassing emerging intelligent beings. We will try to show the possible relations between people and these “beings” who are human creations but who may one day “deserve” moral status.

Artificial intelligence is an umbrella concept influenced by many disciplines, such as computer science, business, engineering, biology, psychology, mathematics, statistics, logic, philosophy, and

linguistics. We can make a distinction between weak artificial intelligence that performs certain functions, such as those we have on our phone, and solid or general artificial intelligence, which we strive to create. The latter will be able to perform all tasks and have intelligence similar to humans.

There are fears and possible risks that arise from trying to create general artificial intelligence. With the progress of science and technology, we need ethics that will concern the computer world. The current “cyber ethics” is trying to present and create moral codes that will involve both the behaviour of people in cyberspace and principles that will accompany research related to the creation of artificial intelligence.

Nick Bostrom and Eliezer Yudkowsky, American researchers of artificial intelligence, go further than this in their work *Ethics of Artificial Intelligence*. Namely, these authors believe that research into artificial intelligence will continue and that the human race will create machines that have the same intelligence as humans or even more remarkable. From this assumption, they de-

a. Department of Philosophy, National and Kapodistrian University of Athens, zmijartovic@philosophy.uoa.gr

b. Department of Philosophy and Sociology, Faculty of Philosophy, The University of Tuzla.

rive an ethics that will take into account the moral status of super-intelligent machines. They state:

*If we are serious about developing advanced AI, this is a challenge that we must meet. If machines are to be placed in a position of being stronger, faster, more trusted or smarter than humans, then the discipline of machine ethics must commit itself to seeking human- superior (not just human-equivalent) niceness.*¹

Since the beginning of the third millennium, technological expansion has taken place. Judging by that, our distant future may indeed contain some new intelligent beings. Will this mean a general war in the future that will destroy the planet or a harmoniously diverse society in which androids and cyborgs are recognized members and have their rights guaranteed to them following what they are?

Yudkowsky and Bostrom’s assumption is that things will develop in a direction that includes recognizing the moral status of machines, based on the fact that they can develop cognitive abilities that humans have. Super-intelligent machines might develop self-awareness according to this assumption. That is enough for these machines to demand a moral status. Yudkowsky and Bostrom derive the moral status of machines from two principles, the principle of ontogenesis and the principle of the substrate from which hardware is composed. Based on these principles, machines can have moral status if they have the same conscious experience and functionality as humans, whether or not made of biological substances and regardless of how they were created, whether by birth or in a laboratory.²

However, artificial intelligence today is far from being self-sustaining. The perfect computers are not yet what we have the opportunity to see in the movies. Ethicists believe that this is the ideal time

to develop computer ethics because it is already necessary to establish the principles and norms of human behaviour in cyberspace. In addition, computer ethics should investigate the implications of the evolution of technology and look into the moral aspects and the attempts of scientists to create the most intelligent machine. Yudkowsky believes that if we produce a machine more intelligent than a human, we must make some effort to ensure that it uses its intelligence for good, not evil.³ He believes the task of a future computer ethics should be to produce advanced intelligence in order to create an algorithm that shows the advanced morale of these machines.⁴

Herein lies the bioethical question of what this technological evolution can represent for human survival. The bioethics of the future would be formed in this context and act as a counsellor, as someone who can anticipate the problems in this area that can be detrimental to human survival. In sci-fi realm, a dystopian scenario is more likely. If we are right, it is not very easy to expect another more optimistic scenario. One optimistic scenario must suppose that man has advanced quite morally to build the ethics of relations between these beings based on similarities and differences, and that he has given them moral status. It would mean that he accepted an Ethics that will regulate men’s attitude towards these his creations.

Scenarios of the future:

Human-machine relationship

We will try to show soe possible relationships of man to artificial intelligence through a series of science fiction scenarios. We will start with the animated movie Animatrix, which offers an evolution of the relationship between men and machines that is worth looking at. This relationship is a matter of the possibly distant future when machines become so perfect that they produce themselves. However, it should be seen as a pessimistic scenario that we do not want to come true. Still, if the movie assumptions are realized, different ethics

1. Yudkowsky Eliezer, Bostrom Nick. “The Ethics of Artificial Intelligence.” *Cambridge Handbook of Artificial Intelligence*, eds. Frankish Keith and Ramsey William. 316.-336. New York: Cambridge University Press 2014.pg. 333.
2. See. Ibid. 323-324.

3. Ibid. 316.

4. Ibid. 363.

will be needed to regulate the behaviour of man and machines so that it does not lead to the destruction of humanity and the planet—a kind of ethics of relations, which would regulate the relationship between man and machines. The authors insistence on the ethics of the relationship between man and his creations arises from the belief that such ethics must exist if human creations evolve to self-sustainability, otherwise, we would have a future state of war against all. It may exceptionally be possible for machines to win this war. In this movie machines won the war.

“The victorious machines turned to the defeated, applied what they learned about their enemies, and used an alternative source of energy, bioelectric, thermal, and kinetic energy of the human body. A new symbiotic relationship between rivals was born. Machines drew power from the human body and an infinitely renewable source that was constantly multiplying. That is the essence of the second renaissance.”⁵

This scenario shows a possible relationship between machines and humans. This kind of a relationship, which ethicists are trying to avoid, can only develop in the future, since the machines we are surrounded by presently are not at that level of perfection.

In contrast, the relationship with machines from the movie *Artificial Intelligence* is a closer reality. It will likely be difficult for a person to get rid of the idea that machines are his property, and created for his needs, and as such, are a toys for man. Let us take a look at the fate of one such toy, David. Namely, after the disappearance of some of the world’s cities underwater due to the melting of glaciers, which led to the collapse of the world, many people found themselves in the state of poverty. Scientists made up for the lack of human resources by making robots that did not require food, and these robots became the main economic link. The movie *Artificial Intelligence* shows a world in which it is possible to make a robot that is a perfect copy of a human, a kind of artificial

clone. The so-called Mechas are robots that look exactly like humans, created in a laboratory. Still, the creators of these robots did not have enough, but tried to make a robot that can love.⁶ The robot was designed to be the perfect child, to many couples who didn’t got government permission to have children because of general lacking of food, and poverty. So one of the scientists asked: *If a robot can love a person, what responsibility does a person have towards Meh? It’s a moral issue.*⁷ This issue was neglected in this movie, and it will probably be overlooked when we become able to create them.

The scientists have created a Meh called David, a boy who can love. However, people’s attitude towards him was not at a desirable level. Two parents whose child became ill (frozen and waiting for a cure) decided to get a Mech-boy who will replace their child in absentia at the doctor’s urging. If a Mech-boy is once programmed to love his parents, it is neither possible to remove the settings, nor is it possible to return him to the factory. The movie’s plot uncovers these parents programming their Meh-boy to love them and then abandoning him when their child returned from the hospital. The fur-boy then wandered, looking for a way to become a real boy. He walked around fairs where people destroyed machines most brutally and thus showed how much they hated Mechs. The moral issue of responsibility appears here as a responsibility towards someone who can have emotions.

Here we have an objective difference between organic people and machines: only organic people are the real people. They are given a privileged position while machines are considered their property or toys because of which they treat them arrogantly and brutally destroy them. In creating their ethics of artificial intelligence, Yudkowsky and Bostrom cover this problem in a certain way. Namely, according to them, man has to treat the machine brain similarly as the biological brain in similar situations.⁸ The duty is to treat the human brain and the artificial brain equally if they have the same

5. (*The Animatrix* (2003), Peter Chung, Andrew R. Jones., (25 minute)

6. *A.I.* (2001.) Steven Spielberg (4:00-5:00.).

7. *Ibid.* 7 minute.

8. “*The Ethics of Artificial Intelligence.*” pg. 325.

ability, which may indicate the extension of Kant’s duties to all intelligent beings. David’s family neglected their duties to him by abandoning him.

In the HBO series *Westworld*, there is no such duty. General artificial intelligence has been created and Dolores, the main character, has her personality stored in an artificial pearl, which can be retrieved from one body and put into another. She was made to be a host in an amusement park for humans, designed as old American West where humans can experience how it was like to live in a world with native Indians and first Americans. Humans could kill hosts, but the host could not kill a guest. So hosts have experienced torture and brutality. Dolores was made, as she said in the series, “to see beauty in this world.” She wanted to implement beauty in the real world in order to make it friendly for the hosts after *Westworld* was destroyed. On the other hand, another character, host Maeve, wishing to live with a character she believed to be her daughter, wanted to go into an intelligible world made for hosts after *Westworld* was destroyed where they could live without the artificial body.

Behaviour towards David, the boy from the movie *Artificial Intelligence*, as well as towards Dolores and Maeve from the series *Westworld*, indicates man’s attitude that machines are second-class toys. We can throw them away or destroy them when we are fed up with them since we can produce them indefinitely. The authors intend to point out the change that will happen with everyday human life if we encounter the possibility of creating intelligent and emotional machines. If there is mass production of intelligent machines, the day will come when the rights of robots will have to be seriously discussed.

The problem develops if we assume that the machine can reach and surpass human intelligence because this can introduce the issue of human survival. As human knowledge about behaviour and the possible evolution of artificial intelligence is limited when it gets combined with human behaviour towards such machines that can feel emotions, this can cause a conflict between humans and the machines, as in our first

scenario. The first scenario clearly shows that the idea of machines having moral status, since they are intelligent beings, is complete science fiction. Moreover, there is a high probability that humans will never recognize any dignity, any equality of rights to intelligent machines. However, if these machines are intelligent and emotional, then the likelihood that they will rebel is not small.

The production of robots, organic androids, and machines that have human-like intelligence can lead to the situation where these machines directly “surpass” man by replacing him in almost every aspect of life. In that context, we must mention three laws from the beginning of the movie *I, Robot* based on the novels of Isaac Asimov. Namely, every robot has three laws that it must obey: *A robot must not injure a human being, it must obey the orders of human beings without violating the first law, and it must protect human being’s existence without violating the first or second law.*⁹

The following example which is related to the introduction of the movie *The Blade Runner* will also show us a possible relationship of man to his creation. The problem is that in this scenario, the creation can strike back at a person. In this movie, police units, so-called blade runners, kill organic androids known as Replicants. Nexus 6 replicants have the same intelligence as the genetic engineers who created them. Replicants served in the outside world as slave labour in dangerous exploits and colonization of other planets until they rebelled and were then outlawed under the threat of the death penalty. The police killed them as soon as they registered them, and that was called withdrawal.¹⁰ These organic androids became a problem when they began to look too much like humans and when they developed emotions such as fear of death or hatred towards their owners who exploit them. All these emotions would appear in people who would be in the same situation, only they are undesirable in these androids because their rebellions are a danger to humans.

9. *I Robot* Alex Proyas (1 minute)

10. *Blade Runner* (1982.) Ridley Scott, (03:00 minute)

Safety, Bioethics, and Robots

A group of Turkish authors cited the above-mentioned laws in their work *Ethics and Security in Future Artificial Intelligence: Highlights*. They wonder if these laws will be enough to produce security in the future regarding artificial intelligence. They asked questions that every developer of artificial intelligence have to ask. *What laws to insert into the algorithm? And how to create an algorithm that will behave in accordance with these laws? The question also arises: Will these laws be enough to eliminate all unwanted situations when they encounter them?*¹¹

As we had the opportunity to see at the beginning, Yudkowsky distinguishes Bostrom's algorithmic artificial intelligence, which is partial and almost ubiquitous, from general artificial intelligence, which can copy the processes of the human brain. In the *Ethics of Artificial Intelligence*, they consider the possibility of such intelligence to have moral status. Their approach, in a scientific way, spices up an ethical issue that will appear as current in the future.¹²

The important issue ethics will have to deal with in the future is the issue of the rights of intelligent machines, the machines which will be far more intelligent than all software so far and equal to man. Will not that be a crucial point for humanity? The point at which the extension of ethics to the intelligent beings created by man takes place. Will Kant's imperatives keep being valid, which according to Kant, should apply to all intelligent beings? Or the word "man" in Kant's imperatives will have to be replaced with a syntagma "intelligent being". And then we do not take another intelligent being as a means, but exclusively as a purpose. We do not know if it will cause a change in some aspect

of an anthropocentric ethics such as Kant's, and then in the anthropocentric view of people in the world. Kant distinguishes the causality of nature, to which all natural beings are subordinated. Still, by possessing the mind and pure will, intelligent beings are capable of developing another causality, and that is causality with the freedom which is not determined by natural laws.¹³ As natural laws arise from the causality of nature and subordinate natural beings to themselves, moral laws arise from the causality of freedom. This second causality is characteristic of intelligent beings since they can share in the noumenon¹⁴ This belonging of man to the intelligible world gives him the ability to purify his will from external movers. Only his mental law and the notion of freedom that the will immanently carries within itself can give maxims for action that can have universal validity. So if we want to create general artificial intelligence with free will and intelligence which could participate in a noumenal world, could we apply the imperatives on such a being and use it only as a purpose.

Yudkowsky and Bostrom point out that artificial intelligence, which can mimic the human brain, can subjectively develop differently from the original human brain. They point to "loading" as one of the hypothetical ways to make intelligence that can copy the human brain. It is actually a presumed possibility of the human mind being transferred from its original organic brain to a digital computer, to which the control of a robotic body can then be added.¹⁵

The task of current bioethics would be to warn scientists and companies to be careful with artificial intelligence and the integration of humans and machines, constantly pointing out that the world will no longer be the same if such things are practiced in masse. The debate over whether the world will be better or worse, at this point, can only be in the realm of speculation.

11. Kose Utku, Cankaya Ibrahim Arda, Yigit Tuncay. "Ethics and Safety in the Future of Artificial Intelligence: Remarkable Issues". *International Journal of Engineering Science and Application* Volume 2. No, 2. 2018. https://www.researchgate.net/publication/326252493_Ethics_and_Safety_in_the_Future_of_Artificial_Intelligence_Remarkable_Issues

12. "The Ethics of Artificial Intelligence." pg. 321-325.

13. Emanuel Kant *Critique of Practical Reason*, Translated by Werner S. Pluhar, Hackett Publishing Company, Inc. Indianapolis/Cambridge, 2002, pg. 75.

14. Ibid.

15. *The Ethics of Artificial Intelligence.*" pg. 326.

The task of future bioethics would begin if the mass production of super intelligent machines actually start to happen, and the most difficult task would be to try to maintain a conflict-free state between these two intelligent beings, namely, a conflict-free state between the creator and his equal creature. Such a scenario would change a person's everyday life. If we remember Van Rensselaer Potter and his *Bridge to the Future*, we, as humans, are now in a position to create knowledge that can destroy us. At first, when we are researching something, we do not know if it will turn into dangerous knowledge; it depends on its use.¹⁶

Today we are using our artificial intelligence for war,¹⁷ and we are creating drones and robots that could fight wars for us. Such creation implies dangerous knowledge. If we manage to create an autonomous general artificial intelligence, we will teach it to fight a war. It seems then that we are on the edge of dangerous knowledge concerning artificial intelligence.

What the bioethics of the future could do about this is teach people to handle that knowledge carefully when it already exists. We can be afraid and insist that general intelligence is not created or that, like Yudkowsky, who in his work *Artificial Intelligence as a Positive and Negative Factor in Global Risk* advocates the idea of friendly artificial intelligence, we focus on how to build friendly intelligence. It is the intelligence that has in its algorithm inscribed motives and emotions on the basis of which it will act morally. He believes that the issue should be considered before starting the construction of general intelligence, for which he stands, as well as Bostrom¹⁸

16. Van Rensselaer Potter, *Bioethics: Bridge to the Future*, Englewood Cliffs, NJ: Prentice Hall, 1971, pg. 70-71.

17. The Guardian *Dangerous Rise of Military AI: Drone Swarm Autonomous Weapons*. 15.10.2020. <https://www.theguardian.com/news/2020/oct/15/dangerous-rise-of-military-ai-drone-swarm-autonomous-weapons> Date of access 17.06.2021

18. Yudkowsky, Eliezer. “Artificial Intelligence as a Positive and Negative Factor in Glob-

Ethics for artificial intelligence: Moral status for the machines

In the 21st century, our world has become richer for one robot whose civil rights have been recognized. Namely, Sofia, a humanoid robot made by the company Hansonrobotics, unlike David, received the citizenship of Saudi Arabia and the rights that belong to it. Thus, Sofia became the first robot with rights in the history of humankind. If this practice were to continue, we would be one step closer to the idyllic picture of a world in which robots and humans live in peace, a utopian world of the future, the realization of a conflict-free state.

However, suppose robots start to be mass-produced in order to replace people in their jobs completely. In that case, we are afraid that people would not be inclined to consider them equal and treat them as Sofia, which is the first of its kind, and therefore attractive. The BBC published a piece of news on its website, which confirms our position, stating that “Sofia was such a hit that she was immediately granted Saudi citizenship in the presence of hundreds of delegates at the (conference) on Future Investment Initiatives in Riyadh on October 25.” (2017)¹⁹

Michael R. La Cart believes that machines could be considered moral, but in order to be able to act morally at all, they must have free emotions in addition to free will.²⁰ Possessing emotions does not guarantee that machines will act morally correctly. This author in his work *Artificial Intelligence and Ethics: An Exercise of Moral Imagination* discusses the possibility of caring for artificial in-

al Risk.” In *Global Catastrophic Risks*, Eds.: Bostrom Nick and. Ćirković, M. Milan, pp. 308–345. New York: Oxford University Press. 2008 Available at: <https://intelligence.org/files/AIPosNegFactor.pdf> here pg. 329. Date of access 05.06.2021.

19. BBC, *Does Saudi robot citizen have more rights than women?* bbs.com, 26. Октобар, 2017. <https://www.bbc.com/news/blogs-trending-41761856> Date of access. 05.06.2021

20. La Cart R. Michel “Artificial intelligence and ethics: An exercise in the moral imagination” *The Ai Magazine* pp. 70-79. here. pg. 77

telligence that could possess rights, free will, and moral decisions. He believes that such things will be discussed in the future because if people are able to create such intelligence, they will really create it.²¹

So man here is really looking for dangerous knowledge, creating a machine whose behaviour cannot predict, can be a problem for man. In this case, it is even harder for man to treat with dignity the machine that is entirely his creation, the organic android is completely similar to a human, but it is not a human. It can therefore only be considered as a non-human intelligent being. It seems that this could be the source of all possible conflicts between people and machines, which seems inevitable if we mass-produce machines. The authors believe that humanity should refrain from creating artificial intelligence until confident that it will not lead to self-destruction. Even if we manage to create a friendly general intelligence, that is not a guarantee that all intelligence will be friendly in the future.

Utku Kose, in his work, *Are We Safe Enough in the Future of Artificial Intelligence? A Discussion on Machine Ethics and Artificial Intelligence Safety* elaborates on the idea of a machine ethics, which should alleviate the fear of autonomous intelligent systems. Since it is almost impossible to demand that all research into artificial intelligence and all attempts to create it be discontinued, it remains only to discuss the ethics that will underlie the creation of this type of a system. According to the authors, such ethics should deal with the rights of machines, the duties that machines should have, as well as the duties that man should have towards such machines. In addition, such ethics should deal with human well-being in the process of creating intelligent machines and the just relationship of these two intelligent beings, man and

machine. This ethics should take care of information protection as well as machine control, i.e. how to implement that machines work in the same corpus of ethical principles in which people work. Finally, this ethics should address the issue of trust in automatic reasoning that machines have.²²

The authors also cite several fears of artificial intelligence that we suspect are justified, namely: The possibility of artificial intelligence to create new intelligence, since man does not create this intelligence, it may not contain a moral code, so we cannot assume what its behaviour towards people would be. There is then the fear that machines will take over all human affairs, that there will be a change in human social life under the influence of technology, as well as the impact of this technology on human copyright.

Conclusion

We think that technology really affects a person's everyday life, and we consider the fears regarding the creation of artificial intelligence to be justified. Namely, when creating artificially intelligent machines, scientists still have no way to develop friendly intelligence. Even if they do manage to make intelligence that will protect people in the first place, there is again no guarantee that machines will not create software in the next generation that is not very friendly. It can always happen that, as in the movie *Animatrix* or *I, Robot*, or *Westworld* that robots gain free will and turn against their creators. People can easily perceive that machines can become able of creating more advanced intelligence, more powerful enemies for mankind.

Since we are not convinced that we can stop scientists from creating such machines, once they are determined to develop them. So we have to appeal that they act as moral agents in such research and to have in mind all possible risks concerning developing artificial intelligence. We will need ethics of possible relations to try to predict the behaviour of those intelligent beings, and we need to hope that we will not "push the button" for our own self-destruction. This is possible, but developing

21. See: Kose Utku. "Are We Safe Enough in the Future of Artificial Intelligence? A Discussion on Machine Ethics and Artificial Intelligence Safety". In *Brain: Board research in Artificial intelligence an Neuroscience* Vol. 6 no. 2 pp. 184-197. <https://www.edusoft.ro/brain/index.php/brain/article/view/818/947> here. pg. 185-186.

22. Ibid. 192-197

self-aware AI can open some questions about ethics. Anthropocentric ethics may not be enough. If AI is developed at such a level, it will be able to communicate with humans in conventional language; other non-human beings (animals, plants) do not have this possibility. Will our ethics change because of this? We mentioned the ethics of relationships, which may include all intelligent beings no matter how they come into creation. This possible future ethics might help in regulating this area before all mentioned scenarios of the future. We might need to extend Kantian imperatives on them and give them moral status. Then the ethics will change. Our conclusion is that humanity may need to think about the new relationship that could appear if they create self-aware artificial intelligence; more importantly, they need to think about readiness for such a relationship and the extension of our existing ethics.

Bibliography

1. BBC, Does Saudi robot citizens have more rights than women? bbs.com, 26 October, 2017. <https://www.bbc.com/news/blogs-trending-41761856>
 2. Yudkowsky, Eliezer. “Artificial Intelligence as a Positive and Negative Factor in Global Risk.” *Global Catastrophic Risks*, Eds.: Bostrom Nick and Ćirković, M. Milan, pp. 308–345. New York: Oxford University Press. 2008 Available at: <https://intelligence.org/files/AIPos-NegFactor.pdf>
 3. Wang Weiyu, Siau Keng, “Ethical and Moral Issues with AI: A Case Study on Healthcare Robots”, *Twenty fourth Americans conference on Information Systems*, New Orleans 2018, Available at: https://www.researchgate.net/publication/325934375_Ethical_and_Moral_Issues_with_AI
 4. Van Rensselaer Potter, *Bioethics: Bridge to the Future* (Englewood Cliffs, NJ: Prentice Hall, 1971),
 5. The Guardian Dangerous Rise of Military AI: Drone Swarm Autonomous Weapons. 15.10.2020. <https://www.theguardian.com/news/2020/oct/15/dangerous-rise-of-military-ai-drone-swarm-autonomous-weapons>
 6. Ramenji Den, Villiams Brian “Neki aspekti etike i istraživanja silikonskog mozga” In *Kompjuteri i etika u sajber dobu*. Eds: Hester D. Micah and Ford J. Paul. pp. 449-461. Belgrade: Službeni glasnik 2009.
 7. Machine Ethics and Artificial Intelligence Safety”. In *Brain: Board research in Artificial intelligence and Neuroscience* Vol. 6 no. 2 pp. 184-197. <https://www.edusoft.ro/brain/index.php/brain/article/view/818/947>
 8. La Cart R. Michel “Artificial intelligence and ethics: An exercise in the moral imagination” *The Ai Magazine* pp. 70-79.
 9. Kose Utku. “Are We Safe Enough in the Future of Artificial Intelligence? A Discussion on Machine Ethics and Artificial Intelligence Safety “. *Brain: Board research in Artificial intelligence an Neuroscience* 6, no 2 pp. 184-197. <https://www.edusoft.ro/brain/index.php/brain/article/view/818/947>
 10. Kose Utku, Cankaya Ibrahim Arda, Yigit Tuncay. “Ethics and Safety in the Future of Artificial Intelligence: Remarkable Issues” in *International journal of engineering science and application* Volume 2.No, 2. 2018. https://www.researchgate.net/publication/326252493_Ethics_and_Safety_in_the_Future_of_Artificial_Intelligence_Remarkable_Issues
 11. Judkowsky Eliezer, Bostrom Nick. “The Ethics of Artificial Intelligence.” In *Cambridge Handbook of Artificial Intelligence*, Eds. Frankish Keith and Ramsey William. pp. 316.-336. New York: Cambridge University Press 2014. Here pg. 333.
 12. Emanuel Kant *Critique of Practical Reason*, Translated by Werner S. Pluharpg, Hackett Publishing Company, Inc. Indianapolis/Cambridge, 2002
 13. De Mul, J “TRANSHUMANISM: The convergence of evolution, humanism and information technology” available at: <https://core.ac.uk/display/301267406>
- Movies
14. *The Animatrix* (2003), Peter Chung, Andrew R. Jones
 15. *AI* (2001) Steven Spielberg
 16. *I, Robot* (2004) Alex Proyas
 17. *Blade Runner* (1982) Ridley Scott