AN ETHICS OF ARTIFICIAL INTELLIGENCE

By

Oraekii Chigozie Stephen

Mat. No: (DU/353)

Being a Term paper submitted to the Philosophy Department of the Dominican University in Partial Fulfillment of the Requirement on the Award of Degree In Bachelor of Arts in Philosophy

Course: Ethics: PHI 303

Lecturer: Rev. Fr. Dr Patrick Akunne, OP.

SAMONDA, IBADAN

JANUARY, 2025.

Table of Contents

Introduction:

Looking at the invention, development, and advancement of Artificial intelligence or thinking machines that are geared to assist humans in their exploration of the world, it needs ethics on how they ought to behave. This is the subject of this paper: What are the ethics of AI? Because, many human endeavours of the modern day rely on AI to reduce human error, automate repetitive tasks, and faster execution of internal processes. However, it is important to call to mind some disadvantages of AI technology as areas of improvement. Some examples of this include the case of the young boy named Sewell Setzer, who lost his life due to the advice of the character AI calling out loudly to this effect.[1]

So, the question that comes to the mind of those who are following the trending news about artificial intelligence is: What are the measures put in place to ensure that these AI do not harm a person? What becomes of the world or rather humanity when AI is not properly guided?

This paper attempts to discuss the ethics of artificial intelligence. This means that the paper gears to underscore the principles or ethics that artificial intelligence should imbibe like honouring privacy, fairness, safety, freedom, etc., to avoid causing havoc to humanity.

So, the paper shall discuss the meaning of AI, types of learning in AI, some of the ethical issues raised by AI before entering the Ethics of AI, and the relevance of the ethics of AI, before our summation.

What Is Artificial Intelligence?

Artificial intelligence is an attempt to make computers do the sorts of things human and animal minds can do either for technological purposes and/or to improve our theoretical

[1] https://news.sky.com/story/mother-says-son-killed-himself-because-of-hypersexualised-and-frighteningly-realistic-ai-chatbot-in-new-lawsuit-13240210

understanding of psychological phenomena.[2] Additionally, AI refers to the ability of digital computers or computer-controlled robots to perform tasks commonly associated with intelligent beings. It is frequently applied to the project of developing systems programmed with the intellectual processes characteristic of humans, such as the ability to reason, discover meaning, generalize, or learn from experience.[3] Dave Gershgorn states, "Artificial intelligence is software or a computer program with a mechanism to learn. It then uses the knowledge to decide in a new situation, as humans do. It begins when the researchers building this software try to write code that can read images, text, video, or audio, and learn something from it. Once a machine has learned, that knowledge can be put to use elsewhere."[4] The next session shall discuss the types of learning in artificial intelligence and their advancements.

Types Of Learning In Artificial Intelligence:

Weak Artificial Intelligence: A non-sentient artificial intelligence that is focused on one narrow task.[5] Examples of weak AI include: Self-driving cars, Sophia, humanoids, etc.

Strong Artificial Intelligence: A machine with consciousness, sentience, and mind or a machine with the ability to apply intelligence to any problem, rather than just one specific problem.[6] Examples are: Autonomous Vehicles, Robotics, Chatbots, etc.

Super Artificial Intelligence: Any intellect that substantially exceeds the cognitive performance of humans in almost all sectors of interest. It is capable of processing multiple tasks with human-

[2] Keith Frankish and William Ramsey, eds., *The Cambridge handbook of Artificial intelligence*, (United Kingdom: Cambridge University Press, 2014), 336.

[3] Jack Copeland, "The Turing Test" In Minds and Machines:10, (Netherland: Kluwer Academic Publishers, 2000,519.

[4] Dave Gershgorn, "The Quartz guide to artificial intelligence: what is it, why is it important, and should we be afraid?", Quartz, September 10, 2017, http://qz.com/1046350/the-quartz-giude-to artificial-intelligence-what-is-it-why-is-it-important/

[5] Donatus Okwara, Obinna Obiagwu, et al, "Artificial Intelligence and its societal legal Implications" In Igwebuike: An African Journal of Arts and Humanities. Vol. 10. No. 2, ISSN:2488-9210(Print)2504-9038(online, 2024), 18.

[6] Jugal Kalita, *Ethics in Artificial Intelligence,* (Colorado: University of Colorado Springs, 2004), 15.

like intelligence.[7] However, this type of AI is yet to be developed. The next part will discuss the ethical issues posed by AI as they advanced.

Some Of the Ethical Risks Of Artificial Intelligence:

The integration of AI technologies into various sectors like health care, finance, law-enforcement, and industries raises ethical issues about privacy, transparency, biases and discrimination of AI, the potential for military, lack of trust, cost of innovation, the displacement of jobs in the world by AI, misuse of personal data, negative impacts on democracy, lack of accountability and liability, etc.[8] This is because of algorithmic bias in which AI systems unintentionally reinforce discriminatory outcomes based on factors like gender, race, or socioeconomic status, copyright infringement, misinformation, and plagiarism.[9] These issues have endangered the lives of people, some have even lost their sources of livelihood, and some students have relied so much on the use of AI for academic research that they cannot produce their intellectual work without using Chatbots. What becomes of the world when these issues are not checkmated? The next session will focus on what AI ought to do to avoid the above issues or problems.

What are the Ethics Of Artificial Intelligence?

The ethics of artificial intelligence (AI) is a multidisciplinary field that governs how Artificial intelligence is developed and used to ensure it benefits society and respects human values. It involves the ethics that AI studies that is, the ethical theories, guidelines, policies, principles, rules, and regulations related to it.[10]

---

[7] Geoff Mulgan, Artificial Intelligence and collective: the emergence of a new field, AI and Society, 2018 33:631-632, https://doi.org/10.1007.s00146-018-0861-5

[8] Bernd Stahl," Artificial Intelligence for a better future" In SpringBriefs in Research and Innovation Governance, 2021, 38-39. https://doi.org/10.1007//978-3-030-69978-9_4

[9] Okwara, Obiagwu, et al, "Artificial Intelligence and its Societal Legal Implications" 27.

[10] Keng Siau and Weiyu Wang, "Artificial Intelligence (AI) Ethics: Creating an Ethical Intelligent Agent," AI Mag, Vol 31, no. 2, 2020, 77.

Ethics is a prerequisite in building ethical AI or making AI behave ethically. It involves the ethical or moral values and principles that determine what is morally right and wrong for a machine. Susan Anderson defines the goal of machine ethics thus, "to create a machine that follows an ideal ethical principle or set of principles in guiding its behaviour; in other words, it is guided by this principle or principles, in the decisions it makes about possible courses of action it could take. This means simply or involves adding an ethical dimension to the machine."[11]

The first ethical code of artificial intelligence was introduced by the famous science fiction writer Isaac Asimov, who presented his three Laws of Robotics in Runaround. The three Laws were later supplemented by a fourth law, called the Zeroth Law of Robotics. The four laws are:

1. A robot may not injure a human being or through inaction, allow a human being to be harmed.
2. A robot must obey the order given to it by human beings except when such orders would conflict with the first law.
3. A robot must protect its existence as long as such protection does not conflict with the second law.
4. A robot may not harm humanity or, by inaction, allow humanity to suffer harm.[12]

Ai Should Be Fair: Artificial intelligence should be designed to avoid bias against individuals or groups. This means preventing discrimination based on factors like race, tribe, gender, and socio-economic status for example in 2016, when a 22-year-old engineering student from New Zealand had his passport photo rejected by the systems of the New Zealand Department of Internal Affairs

[11] Susan Anderson and Michael Anderson, AI and Ethics, 2021,1:27-31, https://doi.org/10.1007/s43681-020-00003-6

[12] https://iep.utm.edu/ethics-of-artificial-intelligence/. Accessed 14th December, 2024.

because his eyes were allegedly closed. This student was of Asian descent (Richard Lee by name) and his eyes were open. The automatic photo recognition tool declared the photo invalid and the student could not renew his passport. He later told the press very graciously: "No hard feelings on my part, I've always had very small eyes and facial recognition technology is relatively new and unsophisticated".[13] Also, similar cases of ethnicity-based errors by passport photo recognition tools have affected dark-skinned women in the UK. "Photos of women with the darkest skin were four times more likely to be graded poor quality than those with the lightest skin".[14] For instance, a black student's photo (Elaine Owusu picture) was declared unsuitable as her mouth was allegedly open, which it was not. However, the fairness of artificial intelligence has to be inserted by the computer scientists or the programmer to enable the machine to behave fairly especially to a race that is not European.

Ai Data Collection Should Be Private: Artificial intelligence systems can collect, store, and analyse large amounts of personal and professional information, which must be protected to avoid privacy violations. The data must be private, to ensure that the individual information is not hacked or made accessible to people who may use it for their selfish interests.[15]

Ai Should Be Accountable: When an artificial intelligence system fails in a specified task and results in a bad consequence, who should be held responsible? Many factors like programming codes, inputs, etc may cause undesirable consequences. This brings "the problem of many hands"[16] So, it is important to know who should be held accountable Artificial intelligent systems cause

[13] Reuters, Passport robot tells an Asian man his eyes are closed. New York Post, 7 Dec. 2016, https://nypost.com/2016/12/07/passport-robot-tells-asian-man-his-eyes-are-closed/ Accessed 14th December, 2024.

[14] Ahmed Mary, UK passport photo checker shows bias against dark-skinned women. BBC News, 8 Oct., 2020. https://www.bbc.com/news/technology-54349538 Accessed 14th December 2024.

[15] Colin Allen, Wendell Wallach and Iva Smit, "Why Machine Ethics" In Intelligent Systems, IEEE- Vol. 21, No. 4. DOI:10.1109/MIS.2006.83.Source:1EEE Xplore, 2.

[16] Job Timmermans, Bernd Stahl Veikko Ikonem and Engin Bozdag, "The Ethics of Cloud Computing: A Conceptual Review," in 2010 IEEE Second International Conference on Computing Technology and Science, Indianapolis, IN USA, 2010, 614-620.

harm or make a mistake. Establishing a clear line of accountability is essential for artificial intelligence.

Ai Should Be Safe In Its Usage: Artificial intelligence should be built not to threaten people's physical safety or mental integrity. But shall ensure that those who operate them do not experience anything like harm. So, argued Lasse Rouhianen that AI systems should be safe throughout their operational lifetime, and verifiably so where applicable and feasible.[17]

Ai Should Be Promote Human Freedom and Autonomous: Artificial intelligence can be used (in some cases) to influence human behaviour, sometimes in ethically problematic ways. But the reason for freedom and autonomy is because the autonomous AI systems have already raised substantial ethical issues like the machine bias in law, making hiring decisions using smart algorithms, racist and sexist chatbots, or non-gender-neutral language translations, rise to worries about deception, especially if the AI is built into robots designed to look or act like human beings.[18] This has brought a lot of responsibility that has been shifted from human beings to autonomous Artificial intelligence systems which can work much faster than human beings without applying any breaks or needing constant supervision. So, Artificial systems should promote human freedom and autonomy, rather than impede it.

Ai Should Be Transparent: Artificial intelligence should be transparent in its assignment.[19] Because AI algorithms must be predictable by their people in social functions.[20] A PwC survey in 2017 found that 76% of businesses were slowing the adoption of AI because of trustworthiness

---

[17] Lasse Rouhianen, *Artificial intelligence 101: things you should know today about future,* (California: CreateSpace Independent, 2018), 230.

[18] Sven Nyholm and Lily Frank, From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible? DOI: 10.7551/mitpress/9780262036689.003.0012. 2017, 224.

[19] Joel Walmsley, "Artificial Intelligence and the Values of Transparency", In AI and Society. https://doi:10.1007/s00146-020-01066-z

[20] Nick Bostrom and Eliezer Yudkowky, "The Ethics of Artificial Intelligence", In The Cambridge Handbook of Artificial Intelligence, Keith Frankish And William Ramsey, eds., (United Kingdom: Cambridge University Press, 2014), 317.

concerns.[21] Because of their inability to understand the technology, leads to significant information asymmetries among the AI experts and users, and hinders human trust in the technology and AI agents.[22]

Value Alignment: Highly autonomous AI systems should be designed so that their goals and behaviour can be assured to align with human values throughout their operation. Lasse Rouhianen will say, that highly autonomous AI systems should be designed so that their goals and behaviour can be assured to align with human values throughout their operation.[23]

Ai Should Be Able To Take Responsibility: One of the ethics that AI needs, is the ability to take responsibility but it seems that is not sometimes. But, who takes responsibility when AI is used in fighting a war or when accidents occur when moving with autonomous vehicles? As Nick Bostrom and Eliezer Yudkowky asked when an AI system fails at the assigned task, who takes the blame or responsibility?[24] This is because programmers of advanced AI systems are stakeholders in the moral implications of their use, misuse, and actions, with a responsibility and opportunity to shape those implications.[25] So, it is good that Ais are built to take responsibility.

Ai Should Be Able To Adapt To Situations: The ability of AI to adapt is another ethic that AI needs to have but it seems that it is not so because they are not adaptable. They are only operating on the algorithms already inserted in them and cannot adjust to another action not inserted.[26] However, most passenger driverless trains believe that human drivers are more flexible and able to deal with emergencies than computerized controllers.[27]

---

[21] Stuart Russell and Peter Norvig, eds., *Artificial Intelligence: A Modern Approach Fourth Edition,* (United Kingdom: Pearson Education LTD, 2022), 1048.
[22] Weiyu Wang and Keng Siau Artificial Intelligence, Machine learning, Automation, Robotics, Future of Work and Future of Humanity: A Review and Research Agenda. Journal of Data Management, 30(1) 2019a, 61-79.
[23] Rouhianen, *Artificial intelligence 101,* 230.
[24] Bostrom and Yudkowky "The Ethics of Artificial Intelligence", 317.
[25] Rouhianen, *Artificial intelligence 101,* 230.
[26] Bostrom and Yudkowky, "The Ethics of Artificial Intelligence", 318.
[27] Allen, Wallach and Smit, "Why Machine Ethics" 4.

Ai Should Have Moral Status: Some scientists have argued that AI should have some moral status if it has the capacity of qualia like "to feel pain"… so for them, if it is morally wrong to inflict pain on a mouse then it should hold also for any sentient AI.[28] Because, if two beings have the same functionality and the same consciousness experiences, and differ only in how they came into existence, then they have the same moral status.[29] And artificial intelligence ought to be treated as identical to human minds because of its resemblance.[30]

The next session is met to highlight the relevance of the ethics of AI. The benefit that the world will enjoy when the programmers of AI can put in place the essential values that will promote life rather than annihilate it.

The Relevance Of Ethics Of Artificial Intelligence.

The implementation of ethics is crucial for Artificial intelligence systems because it will provide safety guidelines that can prevent existential risks for humans. It will solve the issues of biases. It will build friendly Artificial intelligence systems that will adopt our ethical standards and help in the flourishing of humanity.[31]

Conclusion:

This paper has attempted to discuss the ethics of AI or best put, how AI ought to behave. However, it began with the definition of AI and some of the ethical issues posed by AI before discussing our main topic, which is, some of the things that AI needs to imbibe in order to help man in the best possible way, not some people or races, especially as it advances in the society. This will enable AI not to be biased over some races or tribes, also it will promote a good society where there is

[28] Bostrom and Yudkowky,"The Ethics of Artificial Intelligence", 322
[29] Bostrom and Yudkowky, "The Ethics of Artificial Intelligence",323
[30] Bostrom and Yudkowky, "The Ethics of Artificial Intelligence",324
[31] https://iep.utm.edu/ethics-of-artificial-intelligence/ Accessed 14th December, 2024.

peace and harmony as its relevance when the algorithms are placed well on ethics to avoid harm . The questions that the paper seeks to pose for further inquiries are: What becomes of the world or humanity if all these things(ethics) are not put in place? Or what are the existential risks being posed by AI in the future?

BIBLIOGRAPHY

Ahmed Mary, UK passport photo checker shows bias against dark-skinned women. BBC News, 8 Oct., 2020.

https://www.bbc.com/news/technology-54349538 Accessed 14th December 2024.

Stahl, Bernd "Artificial Intelligence for a better future" In SpringBriefs in Research and Innovation Governance, 2021, 38-39. https://doi.org/10.1007//978-3-030-69978-9_4

Allen, Colin, Wendell Wallach and Iva Smit, "Why Machine Ethics" In Intelligent Systems, IEEE-Vol. 21, No. 4. DOI:10.1109/MIS.2006.83.Source:1EEE Xplore, 2.

Gershgorn, Dave. "The Quartz guide to artificial intelligence: what is it, why is it important, and should we be afraid?", Quartz, September 10, 2017, http://qz.com/1046350/the-quartz-giude-to artificial-intelligence-what-is-it-why-is-it-important/

Okwara, Donatus, Obinna Obiagwu, et al, "Artificial Intelligence and its societal legal Implications" In Igwebuike: An African Journal of Arts and Humanities. Vol. 10. No. 2, ISSN:2488-9210(Print)2504-9038(online, 2024), 18.

Mulgan, Geoff. Artificial Intelligence and collective: the emergence of a new field, AI and Society, 2018 33:631-632, https://doi.org/10.1007.s00146-018-0861-5

https://news.sky.com/story/mother-says-son-killed-himself-because-of-hypersexualised-and-frighteningly-realistic-ai-chatbot-in-new-lawsuit-13240210

Copeland, Jack. "The Turing Test" In Minds and Machines:10, (Netherland: Kluwer Academic Publishers, 2000,519.

Timmermans, Job, Bernd Stahl Veikko Ikonem and Engin Bozdag, "The Ethics of Cloud Computing: A Conceptual Review," in 2010 IEEE Second International Conference on Computing Technology and Science, Indianapolis, IN USA, 2010, 614-620.

Walmsley, Joel. "Artificial Intelligence and the Values of Transparency", In AI and Society. https://doi:10.1007/s00146-020-01066-z

Kalita, Jugal. Ethics in Artificial Intelligence, (Colorado: University of Colorado Springs, 2004.

Siau, Keng and Weiyu Wang, "Artificial Intelligence (AI) Ethics: Creating an Ethical Intelligent Agent," AI Mag, Vol 31, no. 2, 2020, 77-84

Rouhianen, Lasse. Artificial intelligence 101: things you should know today about future, California: CreateSpace Independent, 2018.

Nick Bostrom and Eliezer Yudkowky, "The Ethics of Artificial Intelligence", In The Cambridge Handbook of Artificial Intelligence, edited by Keith Frankish And William Ramsey. United Kingdom: Cambridge University Press, 2014. 316-334.

Okwara, Obiagwu, et al, "Artificial Intelligence and its Societal Legal Implications" 27.

Reuters, Passport robot tells an Asian man his eyes are closed. New York Post, 7 Dec. 2016,

https://nypost.com/2016/12/07/passport-robot-tells-asian-man-his-eyes-are-closed/ Accessed 14th December, 2024.

Russell Stuart and Peter Norvig, eds., Artificial Intelligence: A Modern Approach Fourth Edition, United Kingdom: Pearson Education LTD, 2022), 1048.

Anderson Susan and Michael Anderson, AI and Ethics, 2021,1:27-31, https://doi.org/10.1007/s43681-020-00003-6

Nyholm Sven and Lily Frank, From Sex Robots to Love Robots: Is Mutual Love with a Robot Possible? DOI: 10.7551/mitpress/9780262036689.003.0012. 2017, 224-243.

Wang Weiyu and Keng Siau Artificial Intelligence, Machine learning, Automation, Robotics, Future of Work and Future of Humanity: A Review and Research Agenda. Journal of Data Management, 30(1) 2019a, 61-79.