

UNIVERSITY PARTNER



UNIVERSITY OF
WOLVERHAMPTON



HERALD
COLLEGE
KATHMANDU

Concepts and Technologies of AI (5CS037)

Project Report Ethics of AI

Student Id	: 2227486
Student Name	: Nayan Raj Khanal
Group	: L5CG4
Module Leader	: Mr. Siman Giri
Asst. Module Leader	: Mr. Anmol Adhikari
Instructor	: Mr. Rohit Parajuli
Cohort	: 8
Submitted on	: 21-01-2023

Abstract:

Ethics in AI is an increasingly important topic as technology continues to advance and become more integrated into society. AI's potential to outpace human knowledge and ability raises questions about control and accountability. The use of AI in areas such as surveillance and weapon systems brings up ethical concerns about privacy and responsibility. Additionally, the possibility of job displacement and the blurring of the distinction between humans and machines add to the moral maze surrounding AI. As AI continues to evolve, it is crucial to establish and adhere to ethical principles to ensure that the technology is used for the benefit of humanity. The development of ethical guidelines for AI is vital to ensure the responsible and fair use of this powerful technology.

Contents

1. AI Ethics Blueprint: Building Ethical AI:	1
2. AI's Moral Maze:	2
3. AI Ethics in Action: National and International Initiatives:	3
4. AI Ethics Framework: Building Fair, Explainable and Ethical AI:.....	4
5. Building an Ethical AI: An Analysis and Opinion	5
6. Bibliography	6

1. AI Ethics Blueprint: Building Ethical AI:

Ethics is an area of philosophy concerned with figuring out what is morally right or wrong. The Ancient Greek word êthikos, which implies "related to one's character," is where the English word ethics derives from. (Annabelle, 2017)

Artificial Intelligence is a field of computer science focused on creating machines that can perform tasks that would typically require human intelligence. This technology enables robots to simulate or duplicate human intelligence and cognitive processes including learning, problem-solving, and decision-making. (Schroer, 2022)

As a result of the rapid advancement of AI in the past decade, experts have developed safeguards to protect against the potential dangers that AI poses to humans. The main goal of these safeguards is to ensure that the use of AI benefits humanity and aligns with the values and well-being of all stakeholders. (Lawton, 2008)

On the other hand, "Ethical AI" refers to the use of artificial intelligence that adheres to established ethical principles such as protecting individual rights, maintaining privacy, avoiding discrimination, and preventing manipulation. (C3, 2020)

There are several reasons why building ethical AI is important, amongst them:

Human rights protection: Ethical AI makes sure that the rights and welfare of people aren't jeopardized by the application of AI technology. (Jagreet, 2022)

Promoting equality and non-discrimination: Ethical AI works to eliminate bias and discrimination from decision-making systems, ensuring that everyone is treated equally. (Jagreet, 2022)

Maintaining trust: By upholding moral standards, AI systems can maintain user confidence, which is crucial for AI's long-term success and adoption. (Jagreet, 2022)

2. AI's Moral Maze:

AI has been growing since its initial development but for the last decade or so it has seen a massive uprise. As humans learn from their mistakes, AI technology also allows them to learn but unlike the human, it is non-tiring and ever-learning. Thus, it can be brought to light that AI are capable of competing with humans under fair conditions. Hence, this similarity brings several major ethical and moral debates associated with AI and its impact on society.

Ever-learning: AI being restless and tireless can turn itself into an ever-learning being and might one day surpass humans in knowledge, which can make us question who's in control. (World Commission on the Ethics of Scientific Knowledge and Technology, 2019)

Invasion of Privacy: AI poses questions about data security and the possibility of the government's abuse of power through the surveillance of individuals. (World Commission on the Ethics of Scientific Knowledge and Technology, 2019)

Job displacement: As AI develops and becomes more capable, it has the potential to replicate a variety of human-performed tasks and the result of this might lead to the collapse of the economy and loss of jobs. (World Commission on the Ethics of Scientific Knowledge and Technology, 2019)

Weapon's Abuse: The development of weapons like homing missiles that can choose and engage targets without the need for human intervention, raises questions about who's accountable and responsible for the consequences. (World Commission on the Ethics of Scientific Knowledge and Technology, 2019)

Cybernetics: The ever-developing AI is integrated into many aspects of life; it is becoming easier to enhance human abilities. This raises questions about what makes us human and blurs the distinction between humans and machines. (World Commission on the Ethics of Scientific Knowledge and Technology, 2019)

These are some of the most important ethical and moral discussions surrounding AI and the cause for many ongoing discussions and debate about how to use AI responsibly and ethically.

3. AI Ethics in Action: National and International Initiatives:

Today's world revolves around AI. Countries like the United States, China, the United Kingdom, and India are miles above the rest when it comes to technological advancement in AI.

In comparison, Nepal is still in the juvenile phase. There is no major technological advancement but the concept has started to seep into courses at many universities. It's quite clear that the technology is fairly new to many natives and it will take some time before it catches up with the rest of the world.

As discussed earlier, AI's advancement brings forth many ethical and moral dilemmas. Hence, to counter this various institutes have taken initiatives to build fair, explainable and ethical AI. Provided below are some international institutions' works on AI ethics and their objectives.

Future of Life Institute: The key objective of this paper is "do not harm". The development of AI should promote trust, benefit society and avoid competitive racing. (Siau & Wang, 2020)

International Association of Privacy Professionals: It analyzes how big data and artificial intelligence may increase hazards to privacy, equity, transparency, equality, and many other problems. (Siau & Wang, 2020)

Institute of Electrical and Electronics Engineers: The suggested design outlines procedures for establishing an AI governance framework, including a practical approach to data management, affective computing, economics, legal issues, and other areas. (Siau & Wang, 2020)

4. AI Ethics Framework: Building Fair, Explainable and Ethical AI:

There are currently no agreed-upon solutions to the ethical issues relating to AI. However, a wide range of AI specialists, decision-makers, and organizations have been discussing and propagating their own sets of ethical frameworks and principles. (Rai, 2022)

The above discussed the objectives of various institutes that have taken initiatives to build fair, explainable and ethical AI. Now let's look into the framework that they suggested for building AI.

Future of Life Institute: In the year 2017, they published a paper which revolves around the key phrase “do not harm”, meaning for it to be ethical the use of AI should not harm the safety, values, privacy and control of an individual. Failure to be transparent be it in a workspace or governance, i.e., the relevant data should be disclosed, both the disadvantage and benefits. (Siau & Wang, 2020)

Ethical Framework: Safety, Failure Transparency, Judicial Transparency, Value Human Values, Personal Privacy, Liberty and Privacy, and Human Control. (Siau & Wang, 2020)

International Association of Privacy Professionals: They put forth a framework in 2018 that analyzes privacy hazards as well as other problems including fairness, openness, and equality. They gave a general introduction to data ethics and how to incorporate ethical issues into decision-making in organizations. (Siau & Wang, 2020)

Ethical Framework: Data ethics, Privacy, Bias, Accountability, Transparency, Human Rights. (Siau & Wang, 2020)

Institute of Electrical and Electronics Engineers: In the year 2019, IEEE proposed a design which lays out practices for setting up an AI governance structure. One key priority it proposed was to increase human well-being as a metric for AI progress. (Siau & Wang, 2020)

Ethical Framework: Human Rights, Well-being, Data Agency, Effectiveness, Transparency, Accountability, Awareness of Misuse, Competence. (Siau & Wang, 2020)

5. Building an Ethical AI: An Analysis and Opinion

AI as the name suggests is a copy of human intelligence. For it to thrive it needs human input or learning material. The more it feeds, the more it can copy and hence the more intelligent it can be. The performance of AI solely depends on the training data and programming. This working procedure of AI very much resembles that of humans, which eerily leads us to morally question our use of AI. Is the relationship between humans and machines that of master and slave or does AI must have the same rights as humans? That is a discussion of its own and it differs from individual to individual.

Maybe we can never fully comprehend the moral and ethical implications of AI but it is crucial to monitor the use and abuse of AI for the so-called copy of humans is ultimately in the end human. We're but two sides of a coin, hence when creating AI, AI ethics should not be a secondary issue because, in the end, the ethical development of AI technology may determine humanity's fate in the future!

6. Bibliography

Annabelle, L., 2017. *Medium*. [Online]

Available at: <https://medium.com/the-ethical-world/ethics-defined-33a1a6cc3064>

[Accessed 19 January 2023].

C3, 2020. *c3*. [Online]

Available at: <https://c3.ai/glossary/artificial-intelligence/ethical-ai/>

[Accessed 19 January 2023].

Jagreet, K., 2022. *XenonStack*. [Online]

Available at: [https://www.xenonstack.com/blog/ethics-artificial-](https://www.xenonstack.com/blog/ethics-artificial-intelligence#:~:text=It%20is%20a%20must%20to,%2C%20and%20human%2Dcentered%20AI.)

[intelligence#:~:text=It%20is%20a%20must%20to,%2C%20and%20human%2Dcentered%20AI.](https://www.xenonstack.com/blog/ethics-artificial-intelligence#:~:text=It%20is%20a%20must%20to,%2C%20and%20human%2Dcentered%20AI.)

[Accessed 19 January 2023].

Lawton, G., 2008. *TechTarget*. [Online]

Available at: <https://www.techtarget.com/whatis/definition/AI-code-of-ethics>

[Accessed 19 January 2023].

Rai, P., 2022. *Naamii*. [Online]

Available at: <https://www.naamii.org.np/ethical-ai-principles-brief-summary-of-frameworks/>

[Accessed 19 January 2023].

Schroer, A., 2022. *BuiltIn*. [Online]

Available at: <https://builtin.com/artificial-intelligence>

[Accessed 19 January 2023].

Siau, K. & Wang, W., 2020. *Artificial Intelligence (AI) Ethics: Ethics of AI and Ethical AI*, Chicago: IGI Global.

World Commission on the Ethics of Scientific Knowledge and Technology, 2019. *UNESCO*.

[Online]

Available at: <https://unesdoc.unesco.org/ark:/48223/pf0000367823>

[Accessed 19 January 2023].