



Ethics in Artificial Intelligence

Ajay Radhakrishnan Nair

Keraleeya Samajam Model College, Dombivli (W)

ABSTRACT:

People now live a life supported by intelligent systems designed and built by artificial intelligence. AI-based technologies have already had many successes, such as facial recognition, medical diagnostics, and self-driving cars. AI promises to benefit economic growth, social development, and improved human health and safety. However, the low level of translation, information bias, data security, data privacy and ethical issues based on AI technology pose a great risk to users, developers, people and society. As AI advances, the key question is how to address the ethical and moral issues associated with AI. Although the concept of "machine ethics" was proposed in 2006, artificial intelligence ethics is still in the conceptual stage. Artificial Intelligence Ethics is a field related to the study of artificial intelligence ethics. To address AI ethics, it is necessary to think about AI ethics and how to create ethical AI. This article will discuss AI ethics by examining AI ethics and AI ethics. What is the moral and ethical problem of intelligence?

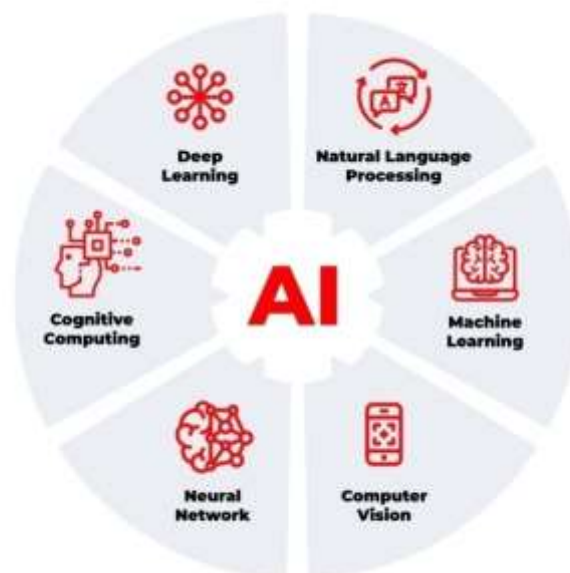
What are the ethical standards, rules, guidelines, rules and regulations that can resolve or at least reduce concerns about the ethics and integrity of AI? What are the necessary characteristics and features of ethical artificial intelligence? How is intelligence integrity monitored and how intelligence integrity is created?

Keywords: AI Ethics, Artificial Intelligence, Ethical AI, Ethics, Ethics of AI, Machine Ethics

1. INTRODUCTION:

Artificial Intelligence:

Artificial Intelligence (AI) is a broad branch of computer science that deals with the creation of intelligent machines that can perform tasks that normally require human intelligence. AI is an interdisciplinary field, but advances in machine learning are making changes in almost every aspect of the tech industry



Ethics:

Morality is a complex, complex and puzzling concept. Morality can be defined as the moral principles that govern the behavior or actions of a person or a group of people (Nalini, 2019). In other words, morality is a principle or rule guideline that helps determine what is good or right. In a broad sense, ethics can be defined as the discipline of right and wrong and the ethics and responsibilities of organizations such as: humans, intelligent robots, etc.

Ethics has been studied by many academics from various disciplines. Most people are familiar with ethics from childhood because it is a guide for parents and teachers to teach behavior that helps children behave well. Aristotle (Yu, 1998) believes that when a person acts virtuously, he feels good and successful. Ethics is the part of ethics that studies what makes behavior right and wrong.

It can be seen as one of the most ethical principles that helps people make moral decisions.

Normative Ethics	Reference
Ethics is the capacity to think critically about moral values and direct our actions in terms of such values.	Churchill, 1999
Ethics is a set of concepts and principles that guide us in determining what behavior helps or harms sentient creatures.	Paul & Elder, 2006
Ethics is the norm for conduct that distinguishes between acceptable and unacceptable behavior. Ethics is the discipline that studies standards of conduct, such as philosophy, theology, law, psychology, or sociology. Ethics is a method, procedure, or perspective for deciding how to act and for analyzing complex problems and issues.	Resnik, 2011
Applied Ethics	
Computer ethics is the analysis of the nature and social impact of computer technology and the corresponding formulation and justification of policies for the ethical use of such technology.	Moor, 1985, p. 266
Machine ethics is concerned with giving machines ethical principles or a procedure for discovering a way to resolve the ethical dilemmas they might encounter, enabling them to function in an ethically responsible manner through their own ethical decision making.	Anderson and Anderson, 2011, p. 1

Ethics of AI and Ethical AI:

AI Ethics is part of Advanced Technologies Ethics, which focuses on robots and other AI agents. It can be divided into robotics (robot ethics) and machine ethics. Robotics is about the human ethics of designing, building, using and interacting with AI agents and the impact of robots on humans and society. In this article, we refer to it as AI ethics, which addresses ethical issues related to AI, including those that may arise during the design and development of AI (e.g. Human bias in information, information confidentiality and transparency) and ethical issues raised by intelligence (for example, unemployment and wealth distribution). We should also consider the law of the robot, the idea that humans have a moral responsibility to intelligent machines as machines become more intelligent and may one day gain consciousness. It is similar to human rights and animal rights. For example, is it ethical to send clever soldiers into dangerous wars or to plant robots in dirty places? The right to freedom, the right to freedom of expression, the right to equality, and the right to think and feel fall into this category. Machine Ethics deals with the moral behavior of Artificial Moral Agents (AMA), a field of research that deals with the creation of moral agents. As technology and robotics get smarter, robots or AI workers must behave more ethically and efficiently. We think of the ethics of artificial intelligence representatives as ethical artificial intelligence.

Currently, the most famous rules for controlling AI agents are Isaac Asimov's Three Laws of Robots in 1950 (Asimov, 1950). First law, A robot must not harm or wrongly allow humans to be harmed. Second rule, the robot must follow the command given by the human unless the command is inconsistent with the first rule. Third rule, the robot must protect its own life unless the guard interferes with the first or second rule.

The table below describes two dimensions of AI ethics

For example, AI Ethics and AI Ethics) and how these two interact with AI, people, and society. The relationship between AIs is new at this writing. This is especially important for cognitive skills. Creating wisdom should not harm people and self-preservation, nor should it harm other people's wisdom. Therefore, the Three Principles of Robotics should be extended to include the impact of intelligent artificial intelligence

	AI	Human	Society
Ethics of AI	Principles of developing AI to interact with other AIs ethically	Principles of developing AI to interact with human ethically	Principles of developing AI to function ethically in society

Ethical AI	How AI should interact with other AIs ethically?	How AI should interact with humans ethically?	How AI should operate ethically in society?
------------	--	---	---

2. LITERATURE REVIEW:

2.1 The articles in this special issue offer different perspectives on the relationship between morality and intelligence. While the first two papers, written by Rahwan and Bryson, respectively, can be mainly divided into Ethics in Design and for parts in the area of Ethics for Designers. The last three articles were written by Vamplew et al., Bonnemains et al., Arnold and Scheutz made a different recommendation. ethical development approach

2.2 Iyad Rahwan's article "Society in the Loop: Programming the Algorithmic Social Contract". Focuses on control and management processes for autonomous systems. The vision of the article is that the algorithms that govern our lives must prove to be transparent, fair and accountable based on the values shared by those involved. This article explains the principles of running, debugging, and managing algorithmic contracts, which are contracts of many people on machines; here it establishes a community power in a loop

(SITL) approach to identify and discuss various cost Stakeholders.

AI system as a basis for monitoring whether the system complies with the social contract.

2.3 In "Patience Is Not a Virtue: The Design of Intelligent Systems and Systems of Ethics", Joanna Bryson contends that the place of AI in society is a matter of normative, rather than descriptive ethics. In the view exposed in this paper, the question of whether AI or robots can, or should, be afforded moral agency or patience is not one amenable either to discovery or simple reasoning, because we as societies constantly reconstruct our artifacts, including our ethical systems.

2.4 Peter Vamplew et al. focus on the need to ensure that the behaviour of AI systems is beneficial to humanity. In their paper "Human-Aligned Artificial Intelligence is a Multiobjective Problem", they discuss the requirement for ethical, legal and safety-based frameworks to consider multiple potentially conflicting factors. They demonstrate that these alignment frameworks can be represented as utility functions, but that the widely used Maximum Expected Utility (MEU) paradigm provides insufficient support for such multiobjective decision-making. They then propose a Multiobjective Maximum Expected Utility paradigm based on the combination of vector utilities and non-linear action-selection that can overcome many of the issues which limit MEU's effectiveness in implementing values-aligned artificial intelligence.

2.5 In "Embedded Ethics: Some technical and ethical challenges", Vincent Bonnemains, Claire Saurel and Catherine Tessier focus on a formal approach to what can be considered as artificial ethical reasoning by an observer. The approach includes formal tools to describe a situation and models of ethical principles that are designed to automatically compute a judgment, and to explain why a given decision is ethically, or not, acceptable.

2.6 The paper "The Big Red Button Is Too Late: An Alternative Model for the Ethical Evaluation of AI Systems", by Thomas Arnold and Matthias Scheutz presents existing proposals for an emergency button in AI systems, and discuss the viability of emergency stop mechanisms that enable human operators to interrupt or divert a system while preventing the system from learning that such an intervention is a threat to its own existence.

3. OBJECTIVE OF RESEARCH:

- Understanding and addressing ethical and moral issues related to AI
- To see whether people can access it or not.
- Is it safe for the mankind or not.
- To understand the general and common ethical principles, rules, guidelines and regulations.
- To understand the potential and moral issues of ethical AI.
- To study the benefits for economic growth, social development as well as human well-being and safety improvement.

4. METHOD:

- **Primary Data Collection:** Questionnaire Method
- **Secondary Data Collection:** Research Paper

Website Links

Observations

Documents and Records

Reports from Publication

Articles

5. DISCUSSION:

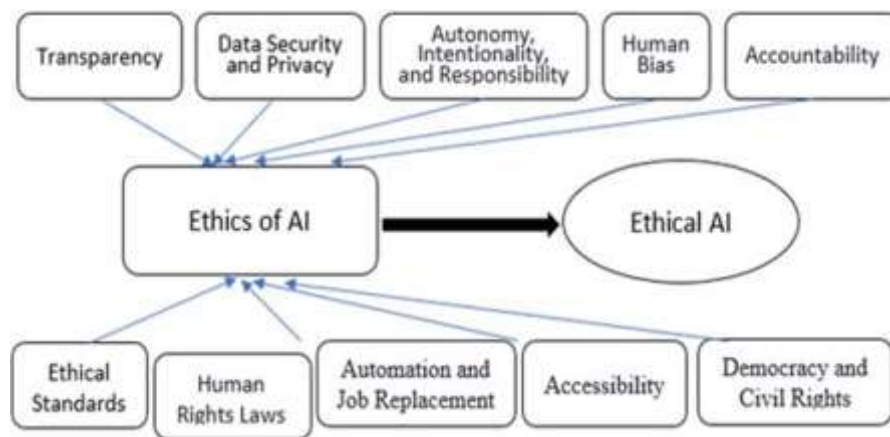
5.1 From Framework to Practice:

This diagram creates an AI ethics framework that lists the factors to consider when determining AI ethics to create an ethical AI. While defining AI ethics is multifaceted and complex, putting AI ethics into practice to create ethical AI is no small feat. How is intelligence? In its simplest form, we can define that fair AI should not harm humans. But what is the damage?

What constitutes human rights? Before designing and building an honest AI, there are many questions that need to be answered. Ethical education is necessary to make ethical decisions.

In theory, AI should be able to identify ethical issues. If AI can make decisions, how can we design and build ethical AI?

Unfortunately, it is not easy to use and in practice. It requires longterm effort and support. Still, understanding and understanding the importance of developing a fair AI and starting to work on it is a big step forward.



Many organizations, such as Google, IBM, Accenture, Microsoft, and Atomium-EISMD, began working to create an ethical principle to guide the development of AI. In November 2018 the Monetary Authority of Singapore (MAS), together with Microsoft and Amazon Web Services, launched the FEAT principles (i.e., fairness, ethics, accountability, and transparency) for the use of AI. Academics, practitioners, and policymakers should work together to widen the engagement to establish ethical principles for AI design, development, and use.

5.2 Ways to Educate AI to Be Ethical

Moor (2006) indicates three potential ways to transfer AI to ethical agents: to train AI into “implicit ethical agents”, “explicit ethical agents”, and “full ethical agents”. Implicit ethical agents mean constraining the machine’s actions to avoid unethical outcomes. Explicit ethical agents mean stating explicitly what action is allowed and what is forbidden. Full ethical agents mean machines, as humans, have consciousness, intentionality, and free will. An implicit ethical agent can restrict the development of AI. An explicit ethical agent is currently getting the most attention and is considered to be more practical (Anderson and Anderson, 2007). A full ethical agent is still an R&D initiative and one is not sure when a full ethical agent will be a reality.

When a full ethical agent is realized, how to treat an AI agent that has consciousness, moral sense, emotion, and feelings will be another critical consideration. For instance, is it ethical to “kill” (shut down) an AI agent if it replaces human jobs or even endangers human lives? Is it ethical to deploy robots into a dangerous environment? These questions are intertwined with human ethics and moral values.

6. CONCLUSION:

Understanding and addressing the moral and ethical issues associated with AI is still in its infancy. AI morality is not about "right or wrong", "good or bad" and "virtue and vice". It is not a problem that can be solved by a small group of people. But the ethical and moral issues related to intelligence are important and should be discussed now. This study aims to demonstrate the urgent need to focus on the ethics of AI representatives from various stakeholders. As we strive to build AI ethics to support the evolution of AI ethics, we will also better understand human ethics, improve existing ethics, and strengthen our interactions with AI content in the age of AI. AI ethics should be a priority in creating an AI agent, not an afterthought. The future of humanity will depend on the development of artificial intelligence!

7. REFERENCES:

Journal Papers:

-
- Allen, C., Wallach, W., & Smit, I. (2006). Why machine ethics? *IEEE Intelligent Systems*, 21(4), 12–17. doi:10.1109/MIS.2006.83
 - Anderson, M., & Anderson, S. L. (2007). Machine ethics: Creating an ethical intelligent agent. *AI Magazine*, 28(4), 15–26.
 - Anderson, M., & Anderson, S. L. (2011). *Machine ethics*. Cambridge University Press. doi:10.1017/CBO9780511978036
 - Bossmann, J. (2016). Top 9 ethical issues in artificial intelligence. World Economic Forum. Retrieved from <https://www.weforum.org/ethical-issues-in-AI>
 - Bostrom, N., & Yudkowsky, E. (2014). The ethics of artificial intelligence. In *The Cambridge handbook of artificial intelligence* (pp. 316-334). Cambridge Press. doi:10.1017/CBO9781139046855.020
 - Nalini, B. (2019). The Hitchhiker's Guide to AI Ethics. Medium. Retrieved from <https://towardsdatascience.com/ethics-of-ai-a-comprehensive-primer-1bfd039124b0>
 - Moor, J. H. (1985). What is computer ethics? *Metaphilosophy*, 16(4), 266–275. doi:10.1111/j.1467-9973.1985.tb00173.x

Books:

-
- AI Ethics by Mark Coeckelbergh
 - The Oxford Handbook of Ethics of AI edited by Markus D. Dubber and Frank Pasquale
 - Ethics of Artificial Intelligence edited by S. Matthew Liao
 - Responsible Artificial Intelligence: How To Develop And Use AI In A Responsible Way by Virginia Dignum

Web References:

-
- <https://blog.growthbot.org/the-ethics-of-artificial-intelligence>
 - <https://blog.einstein.ai/ethics-in-ai-research-papers-and-articles>
 - <https://ethicsblog.crb.uu.se/tag/artificial-intelligence>