



Exploring the Ethical Implications of AI

This editorial is based on [Can AI be Ethical and Moral?](#) which was published in The Hindu on 24/08/2023. It talks about how programming ethics into machines is complex, and the world must proceed cautiously while using AI.

For Prelims: Artificial Intelligence (AI), Weak AI, Strong AI, Machine Learning (ML), Deep Learning (DL).

For Mains: Ethical Challenges of AI, Ethical Considerations of AI, Artificial Moral Agents (AMAs)

Increasingly, **machines and [Artificial Intelligence \(AI\)](#) are assisting humans in decision-making, particularly in governance.** Consequently, several [countries are introducing AI regulations.](#) Government agencies and **policymakers are leveraging AI-powered tools to analyse complex patterns,** forecast future scenarios, and provide more informed recommendations.

However, **the use of AI in decision-making comes with challenges.** AI can have built-in biases from the data it learns from or the viewpoints of its creators. This **can result in unfair outcomes, posing a significant obstacle to effectively incorporating AI into governance.**

What is Artificial Intelligence (AI)?

▪ About:

- AI is the **ability of a computer, or a robot controlled by a computer to do tasks that are usually done by humans** because they require human intelligence and discernment.
 - Although **there is no AI that can perform the wide variety of tasks an ordinary human can do,** some AI can match humans in specific tasks.

▪ Characteristics & Components:

- The ideal characteristic of AI is its **ability to rationalise and take actions that have the best chance of achieving a specific goal.** A subset of AI is [Machine Learning \(ML\).](#)
 - Deep Learning (DL) techniques **enable this automatic learning through the absorption of huge amounts of unstructured data** such as text, images, or video.

▪ Different Categories:

- Weak AI/ Narrow AI
- Strong AI

How Does AI Relate to Certain Philosophical Ideas?

▪ Kantian Ethical Philosophy:

- Immanuel Kant's ethical philosophy underscores three key principles:
 - **autonomy** (the ability to make one's own decisions),
 - **rationality** (using logic and reason to make choices), and
 - **moral duty** (following ethical obligations).
- **Application to AI in Governance:** The act of delegating decision-making processes to AI systems carries the risk of eroding the capacity for nuanced moral reasoning. Letting machines decide instead of humans might weaken the important ideas of Kantian ethics.

▪ Bounded Ethicality:

- In 2022, two researchers researched bounded ethicality using Delphi, a prototype for modeling human moral judgments. They discovered that machines, like Delphi, **might act immorally if the scenario is framed in a way that separates ethics from the action itself.**
 - This shows that the machine version of bounded ethicality is similar to **how humans sometimes act against their own morals without feeling guilty, often using justifications.**

Note: Bounded ethicality is people's ability to make ethical choices is often limited or restricted because of internal and external pressures.

• Parallel with Asimov's 'Three Laws of Robotics':

- Asimov's laws were created to guide robots to behave ethically. However, in Asimov's fictional scenarios, these laws often **resulted in unexpected and paradoxical outcomes, demonstrating the complexity of ethical decision-making** even in machines designed to act ethically.

• The Intersection of Kant and Asimov:

- Kant's emphasis on rational moral agency and Asimov's fictional exploration of ethical guidelines for robots are interconnected. This combination serves to illustrate the **ethical difficulties and complexities that arise when human responsibilities and functions are delegated to artificial entities.**

The Asimov's Laws:

- A robot **may not injure a human being** or, through inaction, allow a human being to come to harm;
- A robot **must obey the orders given to it** by human beings except where such orders would conflict with the First Law;
- A robot **must protect its own existence as long as such protection does not conflict with the First or Second Law."**
- Asimov later added another rule, known as the **fourth or zeroth law**, that superseded the others. It stated that **"a robot may not harm humanity, or, by inaction, allow humanity to come to harm."**

What are the Ethical Challenges of AI?

- **Job Displacement and Socioeconomic Impact:** Automation powered by AI can lead to job displacement in certain industries. The **resulting socioeconomic impact**, including **unemployment** and **income inequality**, poses ethical questions about the responsibilities of governments and organisations in addressing these consequences.
- **Threat to Moral Reasoning:** When decisions that were traditionally made by humans are handed over to algorithms and AI, **there's a risk that the capacity for moral reasoning could be compromised.** This implies that relying solely on AI might **diminish the human ability to**

engage in thoughtful ethical thinking.

- **Challenges of Codifying Ethics:** Attempting to translate ethics into explicit rules for robots or AI-driven governmental decisions is highlighted as a challenging task. Human morals are very complex, and **it's tough to make these complicated ideas fit into computer instructions.**
- **Lack of Accountability & Transparency:** It can be difficult to assign responsibility when something goes wrong with an AI system, **especially when it involves complex algorithms and decision-making processes.**
 - The inner workings of many AI systems are often opaque, **making it difficult to understand how decisions are being made.** This lack of transparency can lead to mistrust and skepticism among users.
- **Informed Consent:** AI systems can be used to collect and analyse [personal data](#) without the knowledge or consent of the individuals involved. This **raises concerns about informed consent and the [right to privacy](#).**

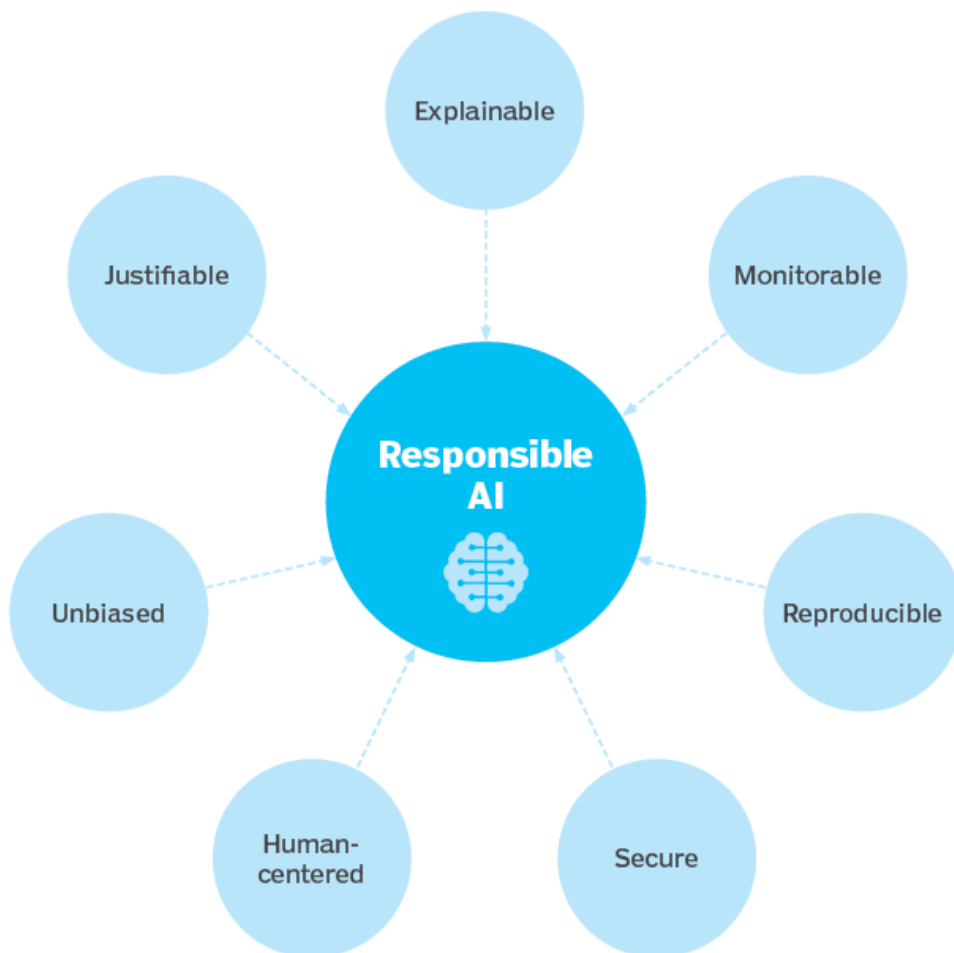
Can Machines or AI be like Moral Decision-Makers/ Artificial Moral Agents (AMAs)?

- Some research says that **machines can, in a way, be responsible for their actions in ethical terms.** **James Moore** (Professor at Dartmouth College) **classified machine agents related to ethics into four groups:**
 - **Ethical Impact Agents:** These machines, like robot jockeys, don't make ethical choices themselves, but **their actions have ethical effects.** For example, they could change how a sport works.
 - **Implicit Ethical Agents:** These machines have built-in safety or ethical rules, like the autopilot in planes. They **follow set rules without actively deciding what's ethical.**
 - **Explicit Ethical Agents:** These **go beyond fixed rules.** They use specific methods to figure out the ethical value of choices. For instance, systems that balance money investments with social responsibility.
 - **Full Ethical Agents:** These **machines can make and explain ethical judgments.** Adults and advanced AI with good ethical understanding fall into this category.

What are the Ethical Considerations of Responsible AI?

//





Conclusion

Currently, many machine predictions help with decisions, but humans still make the final call. In the future, governments might let machines make simple decisions. But what if a decision made by a machine is wrong or unethical? Who's responsible? Is it the AI system, the one who made the AI, or the person who used its data?

These are some of the tough questions that the world is going to face. Putting ethics into machines is tough, and everyone needs to be careful moving forward.

Drishti Mains Question:

“Programming a computer to be ethical is much more difficult than programming a computer to play world-champion chess”. Discuss.

UPSC Civil Services Examination, Previous Year Questions (PYQs)

Prelims

Q. With the present state of development, Artificial Intelligence can effectively do which of the following? (2020)

1. Bring down electricity consumption in industrial units
2. Create meaningful short stories and songs
3. Disease diagnosis
4. Text-to-Speech Conversion
5. Wireless transmission of electrical energy

Select the correct answer using the code given below:

- (a) 1, 2, 3 and 5 only
(b) 1, 3 and 4 only
(c) 2, 4 and 5 only
(d) 1, 2, 3, 4 and 5

Ans: (b)

Mains

Q. "The emergence of the Fourth Industrial Revolution (Digital Revolution) has initiated e-Governance as an integral part of government". Discuss. **(2020)**

PDF Reference URL: <https://www.drishtiias.com/printpdf/exploring-the-ethical-implications-of-ai>

