



## Fostering Ethical AI

**For Prelims:** Fostering Ethical AI, [Ethical Artificial Intelligence \(AI\)](#), [Machine Learning \(ML\)](#), AI facial recognition, Generative AI.

**For Mains:** Fostering Ethical AI, Need for establishing Ethical AI.

[Source: TH](#)

### Why in News?

Recently, some business leaders emphasized the Imperative of collaboration among governments, industry, and ecosystem players to develop [Ethical Artificial Intelligence \(AI\)](#).

### 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 rationalize and take actions that have the best chance of achieving a specific goal. A subset of **AI is [Machine Learning \(ML\)](#)**.
  - **ML** is a method of teaching computers to learn from data, without being **explicitly programmed**. It involves using algorithms to analyze and draw insights from data, and then using those insights to make predictions or decisions.
  - Deep Learning (DL) techniques enable this automatic learning through the absorption of huge amounts of unstructured data such as text, images, or video.

### What is Ethical AI?

#### ▪ About:

- Ethical AI, also known as Moral or Responsible AI, refers to the **development and deployment of AI systems** in a manner that aligns with **ethical principles, societal values**, and human rights.
- It emphasizes the **responsible use of AI technology** to ensure that it benefits individuals, communities, and society as a whole, while minimizing potential harms and biases.

#### ▪ Key Aspects of Ethical AI:

- **Transparency and Explainability:** AI systems should be designed and implemented in a way that **their operations and decision-making processes are understandable and**

- **explainable** to users and stakeholders. This promotes trust and accountability.
- **Fairness and Bias Mitigation:** Ethical AI aims to **mitigate biases and ensure fairness in AI algorithms** and models to prevent discrimination against certain individuals or groups based on factors like race, gender, ethnicity, or socioeconomic status.
- **Privacy and Data Protection:** Ethical AI **upholds individuals' right to privacy and advocates for the secure and responsible** handling of personal data, ensuring consent and compliance with relevant privacy laws and regulations.
- **Accountability and Responsibility:** Developers and organizations deploying AI systems **should be accountable** for the outcomes of their AI technologies. Clear lines of responsibility and mechanisms for addressing and **rectifying errors or harmful impacts are essential.**
- **Robustness and Reliability:** AI systems should be **robust, reliable, and perform consistently** across different situations and conditions. Measures should be in place to handle adversarial attempts to manipulate or subvert the AI system.
- **Benefit to Humanity:** AI should be **developed and used to enhance human well-being**, solve societal challenges, and contribute positively to society, economies, and the environment.

## What are the Ethical Concerns related to Artificial Intelligence?

- **Risk of Unemployment:**
  - The hierarchy of labor is concerned **primarily with automation.** Robotics and AI companies are building **intelligent machines** that perform tasks typically carried out by **low-income workers: self-service kiosks** to replace cashiers, fruit-picking robots to replace field workers, etc.
  - Moreover, the **day is not far when many desk jobs will also be edged** out by AI, such as accountants, financial traders, and middle managers.
- **Exacerbating Inequalities:**
  - Using artificial intelligence, a company can drastically cut down on relying on the human workforce, and this means that revenues will go to fewer people.
  - Consequently, individuals who have ownership in AI-driven companies will make all the money. Also, AI could compound digital exclusion.
  - Further, **investment is likely to shift to countries where AI-related work is already established,** widening gaps among and within countries.
- **Tech Addiction:**
  - Technological addiction is the **new frontier of human dependency.** AI has already become effective at directing human attention and triggering certain actions.
  - When used right, this **could evolve into an opportunity to nudge society** towards more beneficial behavior.
  - However, in the wrong hands, it could prove detrimental.
- **Discriminating Robots:**
  - We shouldn't **forget that AI systems are created by humans,** who can be biased and judgemental.
  - It can lead to **AI facial recognition** and surveillance technology to discriminate against people of color and minorities.
- **AI Turning against Humans:**
  - What if artificial intelligence itself turned against humans, imagine an AI system that is asked to eradicate cancer in the world.
  - After a lot of computing, **it spits out a formula that does,** in fact, bring about the end of cancer – by killing everyone on the planet.

## What are the Global Standards for Artificial Intelligence Ethics?

- **In 2021, the Recommendation on the Ethics of Artificial Intelligence** was adopted by UNESCO.
  - It aims to fundamentally shift the balance of power between people, and the businesses and governments developing AI.
- **UNESCO** members have agreed to use affirmative action to make sure that women and minority groups are fairly represented on AI design teams.

- The recommendation also underscores the importance of the proper management of data, privacy and access to information.
- It calls on **member states to ensure that appropriate safeguards** are devised for the processing of sensitive data and effective accountability and redress mechanisms are provided.
- The **Recommendation takes a strong stance that**
  - AI systems should not be used for social scoring or mass surveillance purposes
  - Attention must be paid to the psychological and cognitive impact that these systems can have on children.
  - Member states should invest and promote not only digital, media and information literacy skills, but also socio-emotional and AI ethics skills.
- UNESCO is **also in the process of developing tools to help assess** the readiness in the implementation of the recommendations.

## Way Forward

- AI models should be designed in a way that allows for a clear understanding of their functioning and decision-making processes.
- AI models should be developed with a strong focus on data privacy and ensuring that individuals' sensitive information is handled appropriately and securely.
- There is a need for **governance norms developed in collaboration with industries** and stakeholders rather than arbitrary legislation, pointing to advanced thinking and ongoing discussions at the government level.
- There is a need for clarity regarding foundational models and data usage in AI systems.
- Ethical AI can be a transformative force, capable of empowering over a billion dreams and bridging the digital divide, not only in India but globally.
- AI and Generative AI must reach diverse populations, being accessible in various languages and regions.

## 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 Refernece URL: <https://www.drishtias.com/printpdf/fostering-ethical-ai>

