



## Mains Practice Question

**Q.** Do you think Artificial Intelligence (AI) compromises ethical principles? How can AI be transformed into responsible decision-makers? (250 words)

22 Feb, 2024 GS Paper 4 Theoretical Questions

### Approach

- Give a brief Introduction to Artificial Intelligence (AI).
- Discuss the ethical challenges of AI.
- Explain the methods through which AI can be transformed into responsible decision-makers.
- Conclude suitably.

### Introduction

Artificial Intelligence (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. These tasks include understanding natural language, recognizing patterns, learning from experience, and making decisions. AI systems aim to mimic human cognitive functions such as reasoning, problem-solving, perception, and learning.

### Body

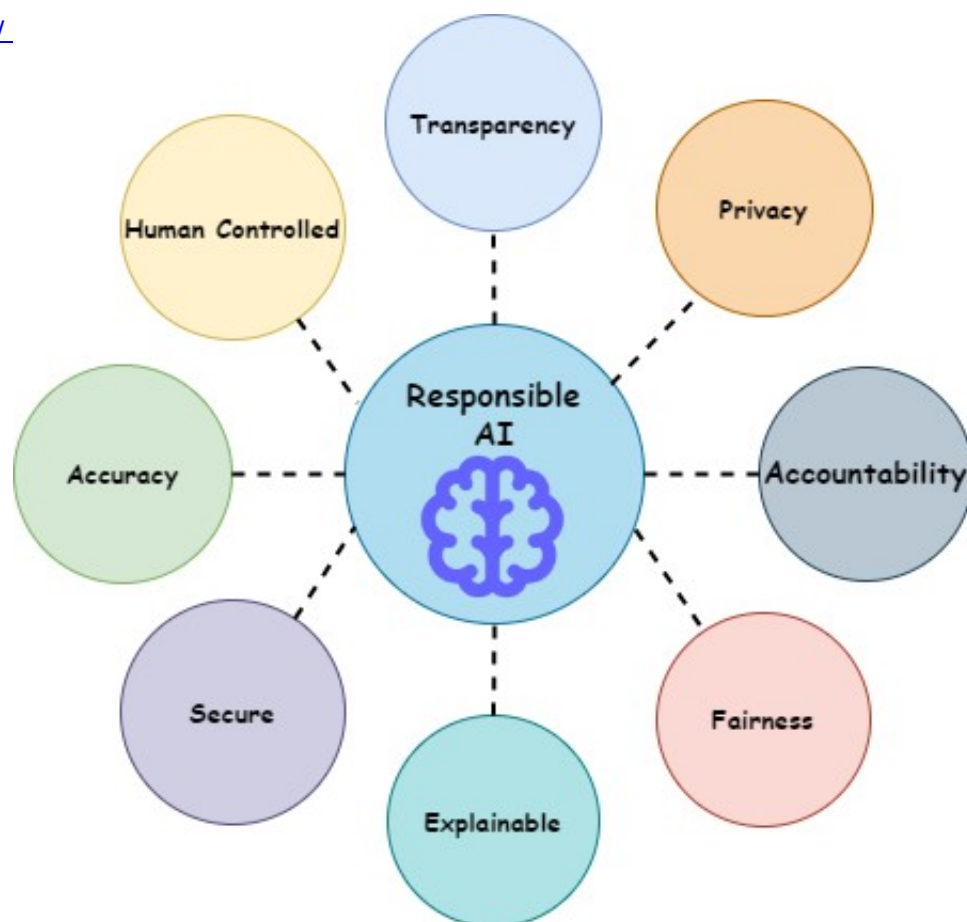
**Some of the prominent ethical challenges of AI include:**

- **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 organizations in addressing these consequences.
- **Violation of Right to Privacy:** AI systems can be used to collect and analyze personal data without the knowledge or consent of the individuals involved. This raises concerns about informed consent and the right to privacy.
- **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.
- **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.
- **Safety and Security:** AI systems can pose risks to safety and security if they malfunction, are hacked, or are manipulated for malicious purposes. The development and deployment of autonomous weapons systems raise concerns regarding their potential for indiscriminate harm and the erosion of human accountability in warfare.
- **Codifying Ethics For Robots:** 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.

Some of the methods through which AI can be transformed into responsible decision-makers:

//



- **Ethical Design Principles:** Incorporating ethical considerations into the design phase of AI systems is essential. **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.
- **Fairness and Bias Mitigation:** Employing techniques to detect and mitigate bias in AI algorithms is crucial for ensuring fairness in decision-making. This includes techniques such as data preprocessing to remove bias, algorithmic fairness constraints, and fairness-aware learning algorithms.
- **Explainability and Transparency:** Enhancing the transparency of AI systems by providing explanations for their decisions helps users understand how AI algorithms work and build trust.
- **Accountability Mechanisms:** Establishing mechanisms for holding AI systems accountable for their decisions and actions is essential. This includes traceability measures to track the decision-making process, auditability features for assessing AI system performance, and mechanisms for recourse or redress in case of errors.
- **Regulatory and Governance Frameworks:** Developing robust regulatory frameworks and governance mechanisms to oversee the development, deployment, and use of AI systems is

crucial for ensuring responsible AI. This includes legal standards, industry guidelines, certification programs, and oversight bodies tasked with monitoring AI applications.

- **Ethics Training and Education:** Providing education and training programs on AI ethics for developers, data scientists, and other stakeholders is essential for promoting ethical awareness and responsible AI practices. This includes ethical AI curriculum development, professional certification programs, and workshops on ethical decision-making in AI development.

## Conclusion

By prioritizing ethical considerations in AI development and deployment, we can harness the potential of AI technologies to create a more inclusive, equitable, and beneficial future for all.

PDF Reference URL: <https://www.drishtias.com/mains-practice-question/question-8173/pnt>

