



## Regulating Artificial Intelligence

**For Prelims:** Regulating [Artificial Intelligence](#), [OpenAI's ChatGPT](#), Black Box, Eu's Artificial Intelligence Act, [Responsible AI for All report](#).

**For Mains:** Regulating [Artificial Intelligence](#) and Way Forward.

### Why in News?

The European Parliament has reached a preliminary agreement on a new draft of the [Artificial Intelligence Act](#), which aims to regulate systems like [OpenAI's ChatGPT](#).

- The legislation was drafted in 2021 with the aim of bringing transparency, trust, and accountability to AI and creating a framework to **mitigate risks to the safety, health, [Fundamental Rights](#)**, and democratic values of the EU.

### What is the EU's Artificial Intelligence Act?

- **About:**
  - It defines AI as **software that generates outputs such as content**, predictions, recommendations, or decisions.
  - It prohibits the use of AI technologies in the **highest risk category**, including real-time facial and **biometric identification systems in public spaces**, social scoring of citizens, subliminal techniques to influence behavior, and technologies that exploit vulnerable people.
- **Focus:**
  - It focuses on AI systems that have the **potential to harm people's health**, safety, or fundamental rights.
    - These include AI in healthcare, education, employment, law enforcement, and access to essential services.
  - Before high-risk AI systems can be sold, they will **undergo strict reviews to ensure they are transparent**, explainable, and allow human oversight.
  - AI systems with lower risks, like spam filters or video games, have **fewer requirements**.
- **Objective:**
  - It aims to **address ethical questions and implementation challenges** in various sectors ranging from healthcare and education to finance and energy.
  - The legislation seeks to strike a **balance between promoting “the uptake of AI while mitigating or preventing harms** associated with certain uses of the technology”.
    - Similar to how the EU's 2018 [General Data Protection Regulation \(GDPR\)](#) made it an industry leader in the global data protection regime, the AI law aims to “strengthen Europe's position as a global hub of excellence in AI from the lab to the market” and ensure that AI in Europe respects the 27-country bloc's values and rules.

### What is the Need for Regulating Artificial Intelligence?

- **Uncertainty in Risks Involved:**
  - The use of artificial intelligence is increasing, and as technology becomes more advanced and capable of various tasks such as recommending music, driving cars, detecting cancer etc., there are also **increased risks and uncertainties associated with it.**
- **Black Box:**
  - Some AI tools are so complicated that **they are like a "black box."** This means that even the people who create them **can't fully understand how they work** and how they come up with certain answers or decisions.
  - It's like a secret box that generates an output, but nobody **knows exactly how it does it.**
- **Inaccuracy and Biases:**
  - AI tools have already caused problems such as **mistaken arrests due to [Facial Recognition Software](#)**, unfair treatment due to biases built into AI systems, and more recently, with **[Chatbots](#)** based on large language models like GPT-3 and 4 creating content that **may be inaccurate or use copyrighted material without permission.**
  - These chatbots are capable of producing high-quality content that is difficult to **distinguish from content written by humans but** may not always be accurate or legally permissible.
- **Unsure of Future Behavior:**
  - AI poses a unique challenge because, unlike in traditional engineering systems, designers cannot be sure how AI systems will behave. When a traditional automobile was shipped out of the factory, engineers knew exactly how it would function. But with self-driving cars, the **engineers can never be sure how it will perform** in novel situations.

## How is Global AI currently Governed?

- **India:**
  - **[NITI Aayog](#)**, has issued some guiding documents on AI Issues such as the National Strategy for Artificial Intelligence and the **[Responsible AI for All report](#)**.
  - Emphasises social and economic inclusion, innovation, and trustworthiness.
- **United Kingdom:**
  - Outlined a light-touch approach, asking regulators in different sectors to apply existing regulations to AI.
  - Published a white paper outlining five principles companies should follow: safety, security and robustness; transparency and explainability; fairness; accountability and governance; and contestability and redress.
- **US:**
  - The US released a Blueprint for an **AI Bill of Rights (AIBoR)**, outlining the harms of AI to economic and civil rights and lays down **five principles for mitigating these harms.**
  - The Blueprint, instead of a horizontal approach like the EU, endorses a **sectorally specific approach to AI governance**, with policy interventions for individual sectors such as health, labour, and education, leaving it to sectoral federal agencies to come out with their plans.
- **China:**
  - In 2022, China came out with some of the world's first nationally binding regulations targeting specific types of algorithms and AI.
  - It enacted a law to **regulate recommendation algorithms** with a focus on how they disseminate information.

## Way Forward

- Regulating artificial intelligence **involves the creation of a simple regulatory framework** that defines the capabilities of AI and identifies those more susceptible to misuse.
- The government should prioritize data privacy, integrity, and security **while ensuring businesses have access** to data.
- Mandatory explainability should be enforced to **eliminate the black-box approaches**, which will bring **transparency and help businesses understand** the rationale behind every decision

made.

- To formulate effective regulations, policymakers must try to **strike a balance between the scope of the regulation and the vocabulary used**, and they should seek input from a variety of stakeholders, including industry experts and businesses.

### UPSC Civil Services Examination, Previous Year Question (PYQ)

**Q1. 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)**

**Q2. Consider the following pairs: (2018)**

	Terms sometimes seen in news	Context/Topic
1.	Belle II experiment	Artificial Intelligence
2.	Blockchain technology	Digital/Cryptocurrency
3.	CRISPR-Cas9	Particle Physics

**Which of the pairs given above is/are correctly matched?**

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

**Ans: (b)**

**[Source: TH](#)**

**[Buy Now](#)**