

Part1

Q1: Algorithmic bias refers to systematic and unfair outcomes produced by artificial intelligence or machine learning systems due to flawed assumptions, incomplete or skewed training data, or biased design decisions.

Examples of how algorithmic bias can show up in AI systems:

Facial recognition systems misidentifying people of color

Facial recognition technologies have been shown to be significantly less accurate for individuals with darker skin tones.

2. Hiring algorithms favouring certain demographics

Some companies have used AI to screen job applicants based on past hiring data. If that historical data reflects gender or racial bias—say, favouring male candidates—then the algorithm can learn to replicate those preferences.

Q2: Difference between transparency and explainability in AI.

Transparency

This is about how open and accessible the design, data sources, and decision-making processes of an AI system are. It includes:

- Disclosure of the AI's training data and algorithms
- Information about who built the model and how it was tested.

Explainability

This focuses on whether the system's outputs and decisions can be understood by humans especially non-experts. It includes:

- Breaking down why the AI made a particular decision.
- Offering plain-language reasoning for its outputs
- Helping users grasp cause-effect relationships inside the model.

Q3: The **GDPR has a profound influence** on how AI is developed and deployed in the EU, especially when personal data is involved.

2.Ethical Principles Matching

Principles

Definition

A) Justice	Fair distribution of AI benefits and risks
B) Non-maleficence	Ensuring AI does not harm individuals or society
C) Autonomy decisions	Respecting users' right to control their data and
D) Sustainability	Designing AI to be environmentally friendly

