Q1: Define *algorithmic bias* and provide two examples of how it manifests in AI systems.

Algorithmic bias refers to systematic and unfair discrimination embedded in the decisions or outputs of an AI system, often due to biased training data, flawed model design, or unequal access to technology.
Examples:
Hiring Algorithms: AI models trained on historical hiring data may learn to prefer male candidates if past data reflects gender bias in hiring practices.
Facial Recognition: Studies have shown that facial recognition systems perform significantly worse on individuals with darker skin tones, particularly Black women, due to underrepresentation in training datasets

Q2: Explain the difference between *transparency* and *explainability* in AI. Why are both important?

Transparency refers to how open and accessible information about an AI system is (its design, training data, decision-making process, and deployment context.)
Explainability focuses on a user's ability to understand *why* an AI system made a specific decision or prediction (the logic or features that led to a given output)
Why they matter:
Transparency builds trust and accountability among developers, regulators, and users. Explainability is critical for end-users (e.g., doctors, judges) to make informed, ethical decisions based on AI outputs.

Q3: How does GDPR (General Data Protection Regulation) impact AI development in the EU?

The **GDPR** significantly shapes AI development in the EU by imposing strict rules on how personal data can be collected, processed, and used:
Data Minimization & Consent: AI developers must limit data collection to what is necessary and obtain clear consent for its use.
Right to Explanation: Article 22 grants individuals the right not to be subject to automated decisions without meaningful human oversight and to request explanations of those decisions.
Privacy by Design: AI systems must incorporate data protection principles from the earliest design stages.
Data Subject Rights: Individuals can request access, correction, or deletion of their data, which affects how AI models handle data storage and retraining.

Ethical Principles Matching.

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