Part 1:

1. Short Answer Questions

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

Algorithmic bias refers to systematic and repeatable errors in AI outputs that unfairly favor or disadvantage certain individuals or groups, often arising from skewed training data, flawed assumptions, or incomplete problem framing.

- **Example 1 – Facial Recognition:** Commercial facial-recognition systems have been shown to misidentify individuals with darker skin tones at significantly higher rates than lighter-skinned subjects, reflecting underrepresentation in training datasets and resulting in unequal error rates.

- **Example 2 – Loan Approval:** Credit-scoring models trained on historical lending data may encode discriminatory lending practices, leading to higher loan-rejection rates for minority applicants despite equivalent creditworthiness.

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

- **Transparency** denotes the degree to which an AI system's inner workings, data sources, model architecture, and decision-making processes are openly documented and accessible for inspection.

- **Explainability** refers to the extent to which an AI system can provide human-interpretable rationales for individual decisions or predictions, often through feature attributions or rule-based summaries.

  Both are crucial because transparency builds stakeholder trust by allowing independent auditing of data and code, while explainability empowers end-users and regulators to understand, challenge, and correct specific outcomes, thereby reducing the risk of unintended harm and enhancing accountability.

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

GDPR imposes strict requirements on the collection, storage, and processing of personal data, mandating lawful bases for data use (e.g., consent, legitimate interest), data minimization, and purpose limitation. It grants individuals rights such as access, rectification, and erasure ("right to be forgotten"), and introduces the "right to explanation" for automated decisions. Consequently, AI developers in the EU must integrate privacy-by-design principles, conduct Data Protection Impact Assessments (DPIAs) for high-risk processing, and implement robust technical and organizational safeguards to ensure compliance and avoid substantial fines.

2. Ethical Principles Matching

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