The Ethics of AI in Decision-Making

Supriya Dutta
M.Sc. in Big Data Analytics

18th July 2025

Introduction

- AI is increasingly making decisions in health, finance, law, and more.
- Some serious questions:
  - **Is AI always fair?**
  - **Can we really trust a system we don't understand?**
  - **And if something goes wrong — who is responsible?**
- Importance of aligning AI with human values.

Biased Behavior in AI models

AI gives unfair results, often favoring or harming certain groups.

Why it happens:

- Biased training data
- Lack of diversity in the data
- Design flaws that developers didn't notice

Effects:

- Unfair hiring or loans
- Unequal services

Real world incident :
Amazon's hiring AI was shut down after it **preferred male candidates**.

Transparency and Accountability in AI

- Transparency = understanding how AI makes decisions
- Accountability = someone is responsible for AI outcomes

Why it happens:

- Black-box models (e.g., deep learning)
- Lack of disclosure by developers/companies
- No clear laws or responsibility rules

Effects:

- No explanation for decisions (e.g., loan/job rejection)
- Hard to detect or correct unfairness

Real world incident :

- UK A-Level Grading Scandal (2020)
- **Robert Williams**, a Black man from Detroit, was **wrongfully arrested** at his home in front of his family.

Privacy Concerns in AI

Often collect, analyze, and store large amounts of personal or sensitive data.

Why it happens:

- Large-scale data needed for AI training
- Data collected without clear consent
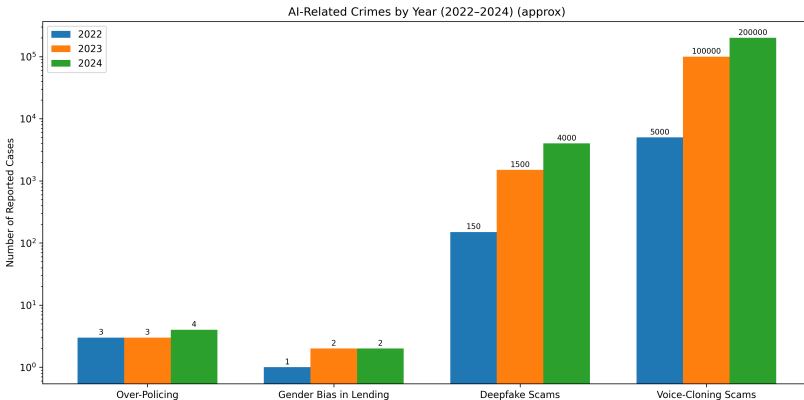- Data reused or shared across platforms

Effects:

- Privacy violations
- Data leaks, identity theft
- Unwanted tracking/surveillance
- Loss of trust in AI systems

Real world incident:
Facebook–Cambridge Analytica Scandal (2018)

Crime Trends Due to AI (Approx)

AI-Related Crimes by Year (2022–2024) (approx)

Regulatory Standards and Policies in AI

AI Acts/Laws Introduced Globally

- **EU AI Act (2024)** – Risk-based regulation of AI systems
- **USA AI Executive Order (2023)** – Safe AI development use
- **India's DPDP Bill (2023)** – Data protection and user consent

Focus Areas

- Data privacy and user consent
- Bias prevention and fairness
- Accountability and transparency

Goal: Ethical, safe, and trustworthy AI

Conclusion

Ethical AI must be:

- **Fair** – no discrimination
- **Transparent** – decisions must be explainable
- **Accountable** – someone must take responsibility
- **Privacy-respecting**

To achieve this, we need:

- **Strong regulations**
- **Global cooperation**
- **Human oversight**