AI Ethics & Bias — Explained Simply

By Leslie Amadi | Last updated: November 2025

📌 Introduction

AI is powerful, but it isn't always fair. Behind every system — whether it's used for hiring, insurance, or policing — are decisions made by humans. These systems can accidentally learn *bias* from the data they are trained on.

🤖 What Is AI Bias?

AI bias happens when an AI system gives unfair results that favour or harm a certain group of people.

Example

- A hiring AI rejects CVs with African-sounding names more often.
- A facial recognition system is less accurate for darker skin tones.

Why does this happen?

✅ The AI was trained on unbalanced data
✅ Most of the training examples came from one gender, race, or region
✅ No human expert reviewed or corrected it

⚖️ What Is AI Ethics?

AI ethics is about making sure AI is used **responsibly, fairly, and transparently**.

Important Principles

Principle	Explanation
Fairness	The system should treat everyone equally
Transparency	People should know how decisions are made
Accountability	Someone must be responsible for mistakes
Privacy	AI must protect user data

🛠️ How Writers Can Help

You don't need to be a developer to support ethical AI.

✔ As a Technical Writer, You Can:

- **Explain limitations clearly** in documentation
- **Add warnings or disclaimers** if AI can produce incorrect or biased results
- **Write ethical usage guidelines**
- **Document dataset sources** and how they may influence results
- **Include fairness testing steps** in API documentation

✅ Summary Checklist

✅ Ethical AI Writing Checklist
Mention algorithm limitations
Warn users about possible bias
Explain what data the AI uses
Add human review recommendations
Encourage responsible usage

📎 Downloadable Version