AI for Software Engineering Assignment

WEEK 7 (AI Ethics PART 4)

DATE : November 2025

INSTITUTION : PLP Academy

Group members

- **Obuye emmanuel chukwuemeke**
- **Eunice fagbemide**
- **antonia othetheaso**
- **mark ireri daizy**
- **jepchumba kiplagat**

Reflection: Ensuring My AI Project Adheres to Ethical Principles

When working on any AI project—whether it is a smart monitoring system, predictive model, or recommendation tool—I believe ethical considerations must be integrated throughout the entire development process, not added at the end. To begin with, I will ensure **fairness and justice** by carefully examining the datasets I use. This includes checking for unbalanced representation, historical prejudice, or missing values that may lead to biased predictions. I will apply techniques such as data balancing, fairness metrics, and cross-group testing to make sure that all user groups receive equitable treatment and that the system does not reinforce existing social inequalities.

To promote **transparency and explainability**, I will document every step of the development process—from data collection to model selection and evaluation. This documentation will be shared with users or stakeholders in a way they can easily understand. Additionally, I will use interpretable models where possible or add explanation tools (like feature importance analysis) to complex models. This allows users to see why a decision was made and gives them confidence in the system's reliability.

Respecting **user autonomy** is also essential. I will design the project to collect only the data necessary for its function and give users clear choices: consent forms, opt-in and opt-out options, and the ability to review or delete their data. The system will comply

with regulations like GDPR, which emphasize user rights and responsible data handling. This ensures individuals maintain control over their personal information.

The principle of **non-maleficence**, or "do no harm," will guide my risk assessment throughout development. I will conduct ethical impact assessments to identify potential negative consequences—such as misuse, misinterpretation, or unintended discrimination—and create safeguards to prevent harm. This may include human oversight, bias mitigation strategies, and strong security measures to protect the system from cyberattacks.

Finally, I will integrate **sustainability** into the project by considering the environmental impact of the AI model. This means choosing energy-efficient model architectures, optimizing code to reduce computational load, and using cloud or hardware resources responsibly. Sustainable design ensures that my AI project contributes positively without causing unnecessary environmental strain.

By embedding these ethical principles—fairness, transparency, autonomy, non-harm, and sustainability—at every stage, I will build AI systems that are not only functional but also trustworthy, responsible, and beneficial to society.