

# Part 4: Ethical Reflection

## Personal Project: AI-Powered Resume Screening Tool

### 1. Ethical Risks Identified:

- **Bias in Training Data:** If historical hiring data favors certain demographics (e.g., male candidates from elite universities), the model may replicate these biases.
- **Lack of Transparency:** Applicants might not understand why their resumes were rejected.
- **Privacy Concerns:** Storing sensitive applicant data without proper consent.

### 2. Mitigation Strategies:

- **Diverse & Representative Data:**
  - Collect resumes from a wide range of genders, ethnicities, and educational backgrounds.
  - Use **synthetic data augmentation** to balance underrepresented groups.
- **Bias Detection & Fairness Metrics:**
  - Implement **disparate impact analysis** to ensure equal selection rates across groups.
  - Use **IBM's AI Fairness 360** or **Google's What-If Tool** for bias audits.
- **Explainability & User Control:**
  - Provide **clear feedback** to rejected candidates (e.g., "Your resume lacked keywords in X skill").
  - Allow applicants to **opt out** of AI screening and request human review.
- **Privacy Protection:**
  - Follow **GDPR/CCPA compliance**—anonymize data and allow deletion requests.
  - Store data securely with **encryption** and **access controls**.

### 3. Long-Term Ethical Commitment:

- **Continuous Monitoring:** Regularly re-audit the model post-deployment.
- **Stakeholder Feedback:** Involve HR teams and job seekers in evaluating fairness.

- **Public Transparency:** Publish an **AI Ethics Report** detailing bias checks and corrective actions.