

UNESCO AI Ethics Recommendations & EU AI Act

Luís Macedo

University of Coimbra

September 23, 2024

- Growing global efforts to ensure the ethical and responsible development of AI.
- **UNESCO's AI Ethics Recommendations** provide a global framework for ethical AI.
- The **EU AI Act** sets legal standards for AI in Europe, focusing on risk-based regulation.
- These frameworks aim to balance innovation with human rights and societal well-being.

UNESCO AI Ethics Recommendations

- Adopted in 2021, the **UNESCO Recommendation on the Ethics of AI** is a global framework for AI development.
- It focuses on the following core areas:
 - **Human Dignity:** Ensuring AI systems respect human rights and dignity.
 - **Inclusive Development:** Reducing inequality and promoting diversity.
 - **Environmental Sustainability:** Minimizing the environmental impact of AI technologies.
 - **Data Privacy and Security:** Safeguarding personal data and ensuring user control.
- Encourages nations to incorporate ethical AI principles into their laws and policies.

Core Ethical Principles of UNESCO's AI Guidelines

- **Human Rights and Freedom:** AI must respect international human rights laws.
- **Fairness and Non-discrimination:** AI should avoid bias and promote equality.
- **Accountability and Transparency:** Developers and users must be accountable, and AI decisions must be explainable.
- **Cultural and Gender Diversity:** Promote diversity in AI development and usage.
- **Sustainability:** AI systems should be developed in line with sustainable development goals.

The EU AI Act: Overview

- The **EU AI Act** is the first comprehensive legal framework for AI.
- Focuses on regulating AI based on its level of risk:
 - **Unacceptable Risk**: AI applications that violate human rights (e.g., social scoring).
 - **High Risk**: AI used in critical sectors like healthcare, education, and law enforcement.
 - **Limited and Minimal Risk**: Consumer-facing AI systems with less regulatory scrutiny.
- Aims to ensure safety, fairness, and compliance with EU fundamental rights.

Risk-Based Approach of the EU AI Act

- **Unacceptable Risk:** These AI systems are banned, such as those involved in social scoring or mass surveillance.
- **High-Risk AI Systems:** These include:
 - Biometric identification and critical infrastructure management.
 - AI used in job recruitment, credit scoring, and education.
- High-risk AI systems must comply with strict rules on transparency, accountability, and data security.

Obligations for High-Risk AI Systems

- High-risk AI systems must comply with:
 - **Data Governance:** Ensuring high-quality datasets free from bias.
 - **Transparency:** Clear documentation and explainability of AI decisions.
 - **Accountability:** Continuous monitoring and human oversight.
 - **Risk Management:** Regular assessment of potential harms and implementing mitigation strategies.

Transparency and Accountability in AI (EU AI Act)

- **Transparency:** AI systems must be explainable, providing clear insights into how decisions are made.
- **Accountability:** Organizations developing or using AI are responsible for ensuring ethical and legal compliance.
- Ongoing audits and assessments of AI system performance and risk.
- Transparency also ensures users know when they are interacting with AI systems.

Challenges and Future Directions

- Implementing global AI ethics standards while fostering innovation.
- Ensuring compliance with the EU AI Act without stifling technological progress.
- Continuous adaptation of policies to address emerging AI risks.
- Fostering international cooperation to address AI challenges and opportunities.

Conclusion

- The UNESCO Recommendations and EU AI Act are critical in shaping the future of ethical and legal AI.
- Ensuring that AI respects human rights, promotes fairness, and remains transparent is key to its responsible deployment.
- A balance between innovation and regulation will determine the future trajectory of AI.