Article 6 – Classification of AI Systems

1. Four Risk Levels

AI systems are classified into **four categories** based on their risk potential:

- **Prohibited AI Systems:** AI practices that pose unacceptable risks and are banned.
  - Examples: Subliminal manipulation, exploitation of vulnerabilities, and social scoring systems by public authorities.
- **High-Risk AI Systems:** Systems that significantly impact individuals' safety or fundamental rights.
  - Examples: AI used in biometric identification, critical infrastructure, education, employment, credit scoring, or healthcare.
- **Limited-Risk AI Systems:** Systems requiring transparency obligations but not as tightly regulated as high-risk systems.
  - Examples: Chatbots, recommendation systems, and virtual assistants.
- **Minimal-Risk AI Systems:** Systems with negligible risk, which are largely unregulated.
  - Examples: Entertainment AI, spam filters, and AI-powered games.

---

2. Criteria for Classification

The classification process considers:

- **Sector of Application:** The domain where the AI is deployed (e.g., education, healthcare).
- **Impact on Rights and Safety:** How the AI affects individuals' privacy, safety, or fundamental rights.
- **Severity of Harm:** The potential damage caused by incorrect or biased outcomes.
- **Autonomy of AI Decision-Making:** The level of human involvement or oversight in the AI's decisions.

---

3. High-Risk Systems in Education

Educational AI systems are explicitly listed under **Annex III** of the EU AI Act as high-risk if they:

- Determine **student access to education** (e.g., AI used in admissions).
- Influence **learning outcomes** (e.g., AI-powered grading systems).
- Assess **skills or competencies** that significantly affect career prospects.

These systems must comply with strict regulations, including documentation, risk management, and transparency requirements.

4. Obligations for High-Risk Systems

For AI systems classified as high-risk, the following obligations apply:

- **Conformity Assessments:** Systems must pass pre-deployment evaluations for compliance.
- **Risk Management:** Continuous risk assessment throughout the system's lifecycle.
- **Data Requirements:** High-quality, representative, and bias-free datasets.
- **Human Oversight:** Mechanisms to ensure human intervention when needed.
- **Monitoring and Reporting:** Continuous monitoring of performance and reporting of incidents.

Implications for Stakeholders

AI Developers

- Must evaluate whether their system falls under high-risk categories.
- Implement safeguards like robust testing and documentation.

Educational Institutions

- Ensure AI tools for admissions, grading, or skill assessment meet high-risk criteria.
- Conduct regular audits to verify compliance with the EU AI Act.

Regulators

- Monitor the deployment of AI systems in high-risk areas, especially those influencing fundamental rights.
- Enforce penalties for non-compliance.

Revision #2
Created 22 December 2024 11:26:10 by Admin
Updated 22 December 2024 11:31:11 by Admin