Der AI Act: Das erste Gesetz zur Regulierung von KI

Ziele

- vertrauenswürdige KI: Schutz von Grundrechten, Demokratie, Gesundheit, Umwelt
- Innovationsförderung und Harmonisierung
- Risikobasierte Regulierung

Risikoklassen

Die KI-Verordnung unterscheidet vier Risikostufen – je höher das Risiko, desto strenger die regulatorischen Anforderungen.

1.Verbotene KI-Systeme/inakzeptables Risiko
- Unterschwellige Beeinflussung
- KI-Systeme, die Schutzbedürftige ausnutzen
- „Social Scoring“ mit Benachteiligung

→ Verboten

2. Hochrisiko-KI-Systeme
- Biometrische Fernidentifizierungssysteme
- KI in Bildung, Arbeit und Recruiting
- Kritische Infrastruktur
- Maschinen, Aufzüge, Seilbahnen, Sportboote, Kraftfahrzeuge

→ Strenge Anforderungen und Registrierungspflicht

3. Begrenztes Risiko
Interaktion mit KI-System muss erkennbar sein, generierte Inhalte müssen maschinenlesbar gekennzeichnet sein
- Chatbots
- Deepfakes
- Generiete Texte, Videos, Stimmen

→ Transparenzpflicht

4. Minimales Risiko
Keine spezifischen Verpflichtungen (z. B. KI in Videospielen oder Spamfilter)

→ keine Auflagen

Der AI Act: Das erste Gesetz zur Regulierung von KI

Anforderungen an Hochrisiko-Systeme (Kap. III, Absatz 2):

- Risikomanagementsystem
- Daten und Daten-Governance
- Dokumentation und Protokollierung
- Transparenz, Handbücher für Betreiber
- Menschliche Aufsicht
- Genauigkeit, Robustheit, Cybersicherheit
- Registrierung in EU-Datenbank

→ Qualitätsmanagementsystem stellt Einhaltung des AI Acts sicher

Sanktionen (Kap. XII):

- Verwarnungen oder Geldbußen bis zu 35 Millionen Euro oder 7% Jahresumsatz
- Für KMU und Startups gelten jeweils niedrigere Beträge

Zeitlicher Rahmen (Art. 113):

- **Seit 02.02.25:** Kapitel I und II: verbotene KI-Systeme
- **Ab 02.08.25:** Modelle mit allgemeinem Zweck, Governance, Sanktionen, Vertraulichkeit
- **Ab 02.08.25:** Gesamte Verordnung, auch KI-Reallabore
- **Bis 31.12.2030:** bestehende KI-Modelle mit allgemeinem Zweck müssen in Einklang gebracht werden
- AI Act gilt nur begrenzt für Systeme, die vor 02.08.26 in Verkehr gebracht/in Betrieb genommen wurden