

# Rechtliche Aspekte der KI als industrielle „Zukunftstechnik“



05.09.2024

# Zukunftstechnik I

Art 3 Z 14 IE-RL (2010): „**Zukunftstechnik**: eine **neue** Technik für eine industrielle Tätigkeit, die bei gewerblicher Nutzung entweder ein **höheres allgemeines Umweltschutzniveau** oder zumindest das gleiche Umweltschutzniveau und **größere Kostenersparnisse** bieten **könnte** als bestehende beste verfügbare Techniken“.

ErwGrund (1) AI-Act (2024): „...um die **Einführung menschenzentrierter und vertrauenswürdiger KI** zu fördern und gleichzeitig ein **hohes Schutzniveau** in Bezug auf Gesundheit, Sicherheit und Grundrechte, einschließlich [...] **Umweltschutz** sicherzustellen“.

# Zukunftstechnik II

- **Entwicklungsrisiko** als rechtliche Herausforderung
- Wie regulieren wir (noch) ungewisse, unsichere Entwicklungen?
  - ...nach welchen Maßstäben, wenn es noch keine Maßstäbe gibt?
  - ...mit welchen Privilegien für die Forschungsfreiheit (Ausnahmen von Grenzwerten)?
  - ...und bis zu welcher Grenze gehen wir Risiken ein?
- **Grunddilemma der „regulierten Innovation“: Lenkung = Hemmung?**

# Rechtliche Instrumente zur Beherrschung von Entwicklungsrisiken

- **präventive Schadensvermeidung** (Verbote, Prüfvorbehalte, Überwachung; freiwillige Systeme – Compliance – flexible Anpassungsmechanismen je nach Neubewertung der Gefahrenlage): **AI-Act**
- **reparativer Schadensausgleich** (Haftung; Versicherung): **KI-Haftungs-RL** (Entwurf)

# „Ur-Angst“ des Rechts gegenüber automatisierten/autonomen Prozessen

- **Kontrollverlust / Unterwerfung** des Menschen unter die Maschine

Art 22 Abs 1 DSGVO: „Die betroffene Person hat das Recht, **nicht** einer **ausschließlich** auf einer **automatisierten** Verarbeitung – einschließlich Profiling – beruhenden **Entscheidung unterworfen zu werden**, die ihr gegenüber rechtliche Wirkung entfaltet oder sie in ähnlicher Weise erheblich beeinträchtigt“.

(„Kafka-Klausel“)

# Verbote

## Verbotene Praktiken im KI-Bereich (Art 5 AI-Act), zB

- *Unterschwellige **Beeinflussung** oder **absichtlich manipulative** oder täuschende Techniken + Veranlassung einer andernfalls nicht getroffenen Entscheidung + erheblicher Schaden*
- *Ausnutzung der Vulnerabilität oder Schutzbedürftigkeit + wesentliche Verhaltensänderung + erheblicher Schaden*
- ***Social scoring***
- *Gefährdungsbeurteilung einer Person ausschließlich aufgrund des Profiling*
- *Ableitung von **Emotionen** am Arbeitsplatz, es sei denn aus Sicherheitsgründen*
- *Biometrische Identifizierung in Echtzeit, außer wenn zur Terrorismus-Abwehr erforderlich*

**Schutzgut: **Autonomie menschlichen Entscheidens und Verhaltens****

# Hochrisikosysteme I

## **Einstufung** (Art 6 Abs 1 iVm Anh III)

- Sicherheitsbauteile bestimmter harmonisierter Sektoren (Maschinensicherheit, Medizinprodukte, etc)
- Kritische Infrastrukturen (Digital-, Verkehrs- Energieinfrastruktur)
- Personalmanagement, etc

Liste in Anhang III ist **durch delegierten Rechtsakt der Kommission änderbar (Flexibilität)**

# Hochrisikosysteme II

**Ausstufung** (Ausnahme gem Art 6 Abs 3), wenn

- eng gefasste Verfahrensaufgabe oder
- Verbesserung einer zuvor abgeschlossenen menschlichen Tätigkeit
- Entscheidungsvorbereitung oder Erkennen von Entscheidungsmustern, **ohne menschliche Bewertung zu ersetzen oder zu beeinflussen**

Parameter für Änderungen des Anh III, u.a.

- das Ausmaß in dem das **KI-System autonom** handelt, und die Möglichkeit, dass ein **Mensch eine Entscheidung, die zu einem potenziellen Schaden führen kann, außer Kraft** setzt

# Hochrisikosysteme III

## Anforderungen (Art 8 ff)

- Einrichtung eines Risikomanagementsystems
- Qualitätskriterien für Trainingsdatensätze
- Technische Dokumentation
- Automatische Aufzeichnung (zB von Störungen)
- Transparenz und Gebrauchsanweisung
- **Ermöglichung menschlicher Aufsicht (Mensch-Maschinen-Schnittstelle)**
- Genauigkeit, Robustheit, Cybersicherheit
- Qualitätsmanagementsystem – Verantwortlichkeit in der Wertschöpfungskette

**Konformitätsbewertung > Registrierung EU-Datenbank > Konformitätserklärung**

# Mensch-Maschine-Schnittstelle

## Risikoadäquate Aufsichtsmaßnahmen (Art 14):

- menschl Aufsicht soll in der Lage sein, Betrieb des KI-System ordnungsgemäß zu überwachen, einschließlich in Bezug auf das **Erkennen und Beheben von Anomalien, Fehlfunktionen und unerwarteter Leistung**;
- Kenntnis über mögliche **Automatisierungsbias** (übermäßiges Vertrauen in KI)
- Möglichkeit des menschlichen Eingriffs und der **Stopptaste**

## Menschliche Kontrolle mittels zweiter KI?

# Nachhaltigkeit I

## KI als Öko-Wirkungsursache (=Ursache von Umweltauswirkungen)

- Beurteilung: Energie + Ressourcenverbrauch für KI (ggf mittelbar) als zurechenbarer Wirk- und Emissionsfaktor einer Industrieanlage
  - Konsequenz: Vermeidungs- und Minimierungsstrategien für Projektdesign = KI-Design
- **optimiertes Öko-Design für KI als Umweltauflage?**

# Nachhaltigkeit II

## KI als Instrument der Nachhaltigkeit

### medienübergreifende Emissionssteuerung:

- cross-media-optimierte Grenzwerte
- Lösung des *trout-murder*-Dilemmas
- **Situationselastische Grenzwertfindung** innerhalb vorgegebener Bandbreiten?

# Nachhaltigkeit III

- **energie- und ressourceneffizienten** Prozesssteuerung
- **Lieferketten**management
- **integrierten** Nachhaltigkeitsberichterstattung

# Freiraum zur Weiterentwicklung I

- Freiräume für Forschung in Weiterentwicklung werden nur **befristet** und **unter Aufsicht** eröffnet (Reallabore – Art 57f AI-Act, Versuchsbetriebe, Pilotprojekte)
- Zukunftstechniken sind rechtlich betrachtet „instabile Zustände“, **Verfestigung zum Stand der Technik erwünscht** – **stabile Geltung für angemessene Perioden (7-10 Jahre)**
- **Überforderung des Rechts durch kontinuierlich lernende und autonom fortschreitende Systeme?**

# Freiraum zur Weiterentwicklung II

Notwendigkeit eines rechtlichen Paradigmenwechsels:

- vom **befristeten Versuchsrahmen**
- zur **unbefristeten, entwicklungsoffenen Rahmenbewilligung**

erster Ansatz: GewO-Novelle 2024 ermöglicht Black-Box-Einreichung

# Fazit

Regulierte Innovation braucht **innovative Regulierung**

**Platz für Fragen...und ein Danke!**

