Überblick KI-Verordnung

Regelungsgegenstand

Die KI-VO teilt KI-Systeme entsprechend ihrem Verwendungszweck in **drei Kategorien** ein und macht Vorgaben für ihre Verwendung

1. **verbotene KI-Systeme (Art. 5 KI-VO)** z.B. Social Scoring, predictive policing, Emotionserkennung am Arbeitsplatz / in Bildungseinrichtungen, Manipulation von Personen, Ausnutzen von Schutzbedürftigen
2. **Hochrisiko-KI → strengen Regeln unterworfen**
3. **allgemeiner Verwendungszweck (general purpose AI, GPAI) → grundsätzlich frei zu verwenden**

Spezielle, nicht-hochriskante KI ist nicht vom Anwendungsbereich der KI-VO erfasst.

Hochrisiko-Systeme

„Risiko“ die Kombination aus der Wahrscheinlichkeit des Auftretens eines Schadens und der Schwere dieses Schadens

Hochrisikosysteme im Bildungsbereich: Art. 6 Abs. 2 i.V.m. Anhang III Nr. 3 KI-VO: Allgemeine und berufliche Bildung

1. **Zulassung zu Bildungseinrichtungen** (KI-Systeme, die bestimmungsgemäß zur Feststellung des Zugangs oder der Zulassung oder zur Zuweisung natürlicher Personen zu Einrichtungen aller Ebenen der allgemeinen und beruflichen Bildung verwendet werden sollen)
2. **Prüfungsbewertung** (KI-Systeme, die bestimmungsgemäß für die Bewertung von Lernergebnissen verwendet werden sollen, einschließlich des Falles, dass diese Ergebnisse dazu dienen, den Lernprozess natürlicher Personen in Einrichtungen oder Programmen aller Ebenen der allgemeinen und beruflichen Bildung zu steuern)
3. **Bewertung des Bildungsniveaus einer Person** (KI-Systeme, die bestimmungsgemäß zum Zweck der Bewertung des angemessenen Bildungsniveaus, das eine Person im Rahmen von oder innerhalb von Einrichtungen aller Ebenen der allgemeinen und beruflichen Bildung erhalten wird oder zu denen sie Zugang erhalten wird, verwendet werden sollen)
4. **Überwachung verbotenen Verhaltens im Rahmen von Prüfungen** (KI-Systeme, die bestimmungsgemäß zur Überwachung und Erkennung von verbotenem Verhalten von Schülern bei Prüfungen im Rahmen von oder innerhalb von Einrichtungen aller Ebenen der allgemeinen und beruflichen Bildung verwendet werden sollen)

Ausnahmen

Ausnahme hiervon in Art. 6 Abs. 3 Uabs. 1 KI-VO:

Ein KI-System ist dann nicht hochriskant, wenn es **kein erhebliches Risiko** der Beeinträchtigung in Bezug auf die Gesundheit, Sicherheit oder **Grundrechte** natürlicher Personen birgt, indem es unter anderem **nicht das Ergebnis der Entscheidungsfindung wesentlich beeinflusst**.

Das ist der Fall, wenn

a) das KI-System ist dazu bestimmt, eine **eng gefasste Verfahrensaufgabe** durchzuführen;
b) das KI-System ist dazu bestimmt, das Ergebnis einer **zuvor abgeschlossenen menschlichen Tätigkeit** zu verbessern;
c) das KI-System ist dazu bestimmt, **Entscheidungsmuster** oder Abweichungen von früheren Entscheidungsmustern zu **erkennen**, und ist nicht dazu gedacht, die zuvor abgeschlossene menschliche Bewertung ohne eine angemessene menschliche Überprüfung zu ersetzen oder zu beeinflussen; oder
d) das KI-System ist dazu bestimmt, eine **vorbereitende Aufgabe für eine Bewertung durchzuführen**, die für die Zwecke der in Anhang III aufgeführten Anwendungsfälle relevant ist.

Adressat*innen

KI-Verordnung richtet sich an Anbieter und Betreiber von KI

➔ Also an Hochschulen, nicht an Anwender*innen

Sagt nichts darüber, wie Studierende KI z.B. in Prüfungen einsetzen dürfen

➔ allgemeine Regeln gelten

➔ Konkretisierung durch Hochschulen erforderlich (PO, Leitlinien)

Persönlicher Anwendungsbereich: Anbieter vs. Betreiber

„Betreiber“ eine natürliche oder juristische Person, **Behörde**, Einrichtung oder sonstige Stelle, **die ein KI-System in eigener Verantwortung verwendet**, es sei denn, das KI-System wird im Rahmen einer persönlichen und nicht beruflichen Tätigkeit verwendet;

„Anbieter“ eine natürliche oder juristische Person, **Behörde**, Einrichtung oder sonstige Stelle, die ein **KI-System** oder ein KI-Modell mit allgemeinem Verwendungszweck **entwickelt oder entwickeln lässt** und es unter ihrem **eigenen Namen** oder ihrer Handelsmarke **in Verkehr bringt oder** das KI-System unter ihrem eigenen Namen oder ihrer Handelsmarke **in Betrieb nimmt**, sei es entgeltlich oder unentgeltlich;

Achtung: **Betreiber können Anbieter werden**, wenn sie die **Zweckbestimmung** eines KI-Systems, [..] das nicht als hochriskant eingestuft wurde [...], so **verändern**, dass das betreffende KI-System zu einem Hochrisiko-KI-System wird.

Fallbeispiel

Die Lehrenden setzen ChatGPT zur Prüfungsbewertung ein.

Folge (1): ChatGPT wir von GPAI zum Hochrisiko-KI-System

Folge (2): Hochschule wird von der Betreiberin zur Anbieterin

Das bedeutet: Hochschule muss nun selbst alle Pflichten aus Art. 16 ff KI-VO erfüllen (insb. Informations- und Transparenzpflichten, Qualitätsmanagement)

Vorbeugen:

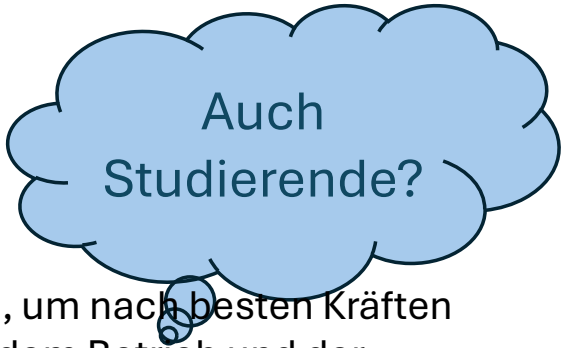
(1) Prüfenden die Verwendung von GPAI zur Prüfungsbewertung generell untersagen.

ODER

(2) Betreiber-Lizenz vom Anbieter erwerben (→ dann nur Betreiber-Pflichten aus Art. 26 KI-VO und Art. 27 KI-VO erfüllen – insb. TOM, Kompetenz, menschl. Aufsicht, Grundrechtefolgeabschätzung)

→ im Fall von ChatGPT etc. wird dies aber nicht möglich sein.

Allgemeine Pflicht für die Hochschulen: KI-Kompetenz vermitteln



Art. 4 KI-VO: Die Anbieter und **Betreiber** von KI-Systemen **ergreifen Maßnahmen**, um nach besten Kräften sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, **über ein ausreichendes Maß an KI-Kompetenz verfügen**, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, zu berücksichtigen sind.

Art. 3 Nr. 56 KI-VO: „**KI-Kompetenz**“ die Fähigkeiten, die Kenntnisse und das Verständnis, die es Anbietern, Betreibern und Betroffenen unter Berücksichtigung ihrer jeweiligen Rechte und Pflichten im Rahmen dieser Verordnung ermöglichen, KI-Systeme **sachkundig** einzusetzen sowie sich der **Chancen und Risiken** von KI und möglicher **Schäden**, die sie verursachen kann, bewusst zu werden.

➔ Hochschule muss insbesondere sicherstellen, dass ihr Personal weiß, dass KI nur als Hilfsmittel eingesetzt werden soll, nicht aber als Ersatz für eigenverantwortliche Entscheidungen betrachtet werden soll.

Kennzeichnungspflichten für Betreiber

- Art. 50 Abs. 4 Uabs. 1 KI-VO: Wer **Bild-, Ton- oder Videoinhalte** erzeugt oder manipuliert, die ein **Deepfake** sind
  - Deepfake = ein durch KI erzeugten oder manipulierten Bild-, Ton- oder Videoinhalt, der wirklichen Personen, Gegenständen, Orten, Einrichtungen oder Ereignissen ähnelt und einer Person fälschlicherweise als echt oder wahrheitsgemäß erscheinen würde (Art. 3 Nr. 60 KI-VO)
- Art. 50 Abs. 4 Uabs 2 KI-VO: Wer **Text** erzeugt oder manipuliert, der veröffentlicht wird, **um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren**,
  - Ausnahme: **Menschlichen Überprüfung** oder **redaktionellen Kontrolle**

Ab wann gilt KI-VO?

Inkrafttreten am 01.08.2024

- Verbote besonders riskanter KI-Systeme gelten schon nach 6 Monaten
- Regeln für generative KI gelten nach 12 Monaten
- Verpflichtungen für „eingebettete KI-Systeme“ erst nach 36 Monaten
- Restliche Regelungen gelten 24 Monate nach Inkrafttreten

Wie kennzeichnen?

- spätestens zum Zeitpunkt der ersten Interaktion oder Aussetzung
- in klarer und eindeutiger Weise
- barrierefrei

- Ist Inhalt Teil eines offensichtlich künstlerischen, kreativen, satirischen, fiktionalen oder analogen Werks oder Programms → dann kann Hinweis so geschehen, dass er nicht den Werkgenuss beeinträchtigt

- ErwG 133: Diese Techniken und Methoden sollten – soweit technisch möglich – hinreichend **zuverlässig, interoperabel, wirksam und belastbar** sein, wobei verfügbare Techniken, wie Wasserzeichen, Metadatenidentifizierungen, kryptografische Methoden zum Nachweis der Herkunft und Authentizität des Inhalts, Protokollierungsmethoden, Fingerabdrücke oder andere Techniken, oder eine Kombination solcher Techniken je nach Sachlage zu berücksichtigen sind.

Use cases

- Prüfungsbewertung
- Anrechnung und Anerkennung
  - https://blog.his-he.de/2024/04/25/ki-technologien-in-der-hochschulverwaltung-studie-zur-analyse-der-potenziale-von-ki-in-anerkennungs-und-anrechnungsprozessen/
- Täuschungsversuche
- Personalmaßnahmen
- Gestaltung der Lehre
- Chatbots zur Studierendenberatung