IDW KNOWLEDGE PAPER

EU AI ACT

EU-VERORDNUNG ÜBER KÜNSTLICHE INTELLIGENZ

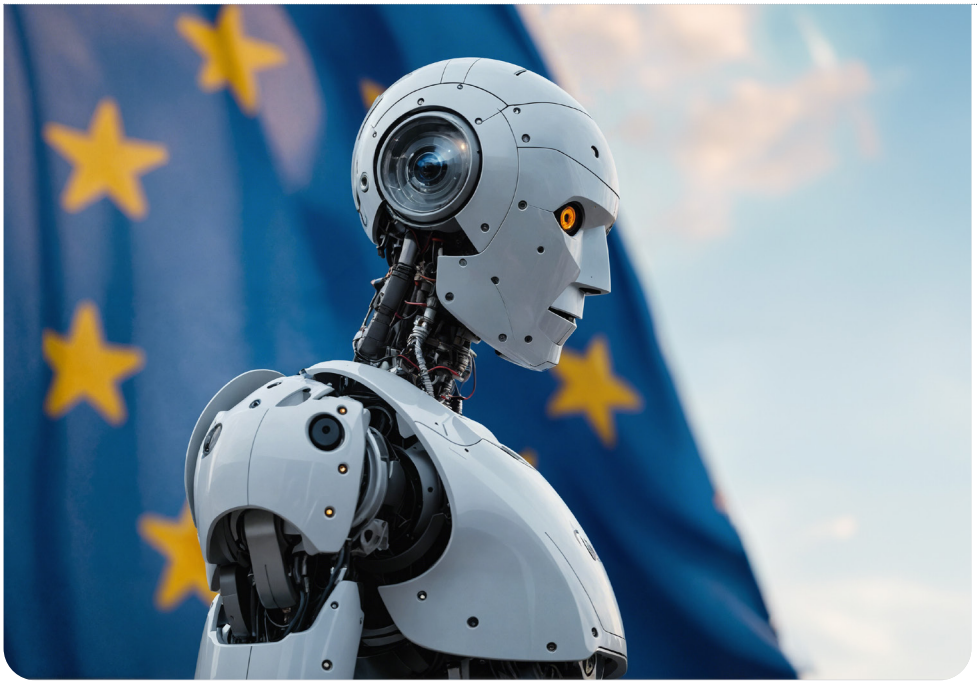
STAND 15.08.2024

IDW
INSTITUT DER WIRTSCHAFTSPRÜFER

INHALT

EINLEITUNG

Die EU hat mit der Verordnung über künstliche Intelligenz (Artificial Intelligence Act (AI Act)) einen umfassenden Rechtsrahmen für künstliche Intelligenz (KI) verabschiedet. Die EU-weit einheitliche Regelung soll den Einsatz einer auf den Menschen ausgerichteten und vertrauenswürdigen KI fördern und gleichzeitig ein hohes Schutzniveau bei Grundrechten, Demokratie und Rechtsstaatlichkeit sicherstellen, sie soll Innovationen unterstützen und der EU zu einer führenden Rolle im Bereich KI verhelfen. Um diese Ziele zu erreichen, verfolgt der AI Act einen risikobasierten Ansatz bei der Einstufung der KI. Die in diesem Knowledge Paper genannten Verweise auf Kapitel, Abschnitte und Artikel beziehen sich sämtlich auf den AI Act.



1. WAS VERSTEHT DIE EU UNTER KÜNSTLICHER INTELLIGENZ?

Der AI Act definiert in Art. 3 Nr. 1 ein KI-System als „ein maschinengestütztes System, das für einen in unterschiedlichem Grade autonomen Betrieb ausgelegt ist und das nach seiner Betriebsaufnahme anpassungsfähig sein kann und das aus den erhaltenen Eingaben für explizite oder implizite Ziele ableitet, wie Ausgaben wie etwa Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen erstellt werden, die physische oder virtuelle Umgebungen beeinflussen können." Hinzu kommen in Art. 3 Nr. 63 und Nr. 66 Definitionen zu KI-Modellen und KI-Systemen mit allgemeinem Verwendungszweck.

2. AB WANN SIND DIE REGELUNGEN IM AI ACT ANWENDBAR?

Der AI Act wurde am 12. Juli 2024 im EU-Amtsblatt veröffentlicht und trat 20 Tage nach seiner Veröffentlichung in Kraft, d.h. am 2. August 2024 (Art. 113). Die vollständige Umsetzung der Regelungen vollzieht sich gestaffelt über einen Zeitraum von drei Jahren. Innerhalb dieses Zeitraums sind die folgenden Einzelvorschriften zu beachten (Art. 113 lit. a)-c)):

VORSCHRIFTEN	ANWENDBARKEIT NACH INKRAFTTRETEN
• Allgemeine Bestimmungen (Kap. I) • Verbotene Praktiken im KI-Bereich (Kap. II)	6 Monate (2. Februar 2025)
• Vorschriften für KI-Modelle mit allgemeinem Verwendungszweck (Kap. V) • Notifizierende Behörden und notifizierte Stellen (Kap. III, Abschnitt 4) • Governance auf EU- und nationaler Ebene (Kap. VII) • Vertraulichkeit (Art. 78) • Sanktionen (Kap. XII) mit Ausnahme der Geldbußen für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck (Art. 101)	12 Monate (2. August 2025)
• Alle Vorschriften, mit Ausnahme der Pflichten für Hochrisiko-KI-Systeme, die als Sicherheitsbauteil in einem von der EU regulierten Produkt verwendet werden oder selbst ein solches Produkt sind und einer Konformitätsbewertung durch Dritte unterzogen wurden (Art. 6 Abs. 1 iVm. Anhang I)	24 Monate (2. August 2026)
• Pflichten für Hochrisiko-KI-Systeme, die als Sicherheitsbauteil in einem von der EU regulierten Produkt verwendet werden oder selbst ein solches Produkt sind und einer Konformitätsbewertung durch Dritte unterzogen wurden (Art. 6 Abs. 1 iVm. Anhang I)	36 Monate (2. August 2027)

3. WER FÄLLT UNTER DIE REGELUNGEN DES AI ACT?

In den Anwendungsbereich des AI Act fallen die folgenden Einheiten (Art. 2 Abs. 1):

- Anbieter, die in der EU oder einem Drittland niedergelassen sind, und in der EU
  - KI-Systeme in Verkehr bringen oder in Betrieb nehmen, oder
  - KI-Modelle mit allgemeinem Verwendungszweck in Verkehr bringen,
- Betreiber von KI-Systemen, die ihren Sitz in der EU haben oder sich in der EU befinden,
- Anbieter und Betreiber von KI-Systemen aus einem Drittland, wenn die von deren KI-System erzeugten Ergebnisse in der EU verwendet werden,
- Einführer und Händler von KI-Systemen, Produkthersteller, Bevollmächtigte und betroffene Personen.

4. WELCHE EINSTUFUNGEN VON KI-SYSTEMEN GIBT ES?

Dem AI Act liegt ein risikobasierter Ansatz zu Grunde. Je höher das mit dem KI-System einhergehende Risiko ist, umso weitreichender sind die Verpflichtungen, die die Unternehmen erfüllen müssen. Es werden die im Folgenden dargestellten fünf Kategorien unterschieden.

4.1. Verbotene Praktiken im Bereich der künstlichen Intelligenz

Das Inverkehrbringen, die Inbetriebnahme oder die Verwendung eines KI-Systems mit besonders hohem Risiko ist verboten (Art. 5). Darunter fallen die folgenden Praktiken:

- unterschwellige Beeinflussung von Personen oder Personengruppen, so dass diesen erheblicher Schaden zugefügt wird oder zugefügt werden kann;
- Ausnutzung der Vulnerabilität oder Schutzbedürftigkeit einzelner Personen oder Personengruppen, so dass diesen erheblicher Schaden zugefügt wird oder zugefügt werden kann;
- Bewertung oder Einstufung von Personen oder Personengruppen aufgrund ihres sozialen Verhaltens oder ihrer Persönlichkeitsmerkmale (sog. Social Scoring), wenn diese zu ungerechtfertigter oder unverhältnismäßiger Schlechterstellung oder Benachteiligung führt;
- Wahrscheinlichkeitsermittlung der Durchführung von Straftaten durch Personen ausschließlich auf der Grundlage des Profiling mit Ausnahmen für Unterstützungsleistungen im unmittelbaren Zusammenhang mit kriminellen Aktivitäten;
- Erstellung von Datenbanken zur Gesichtserkennung durch ungezieltes Auslesen von Gesichtsbildern aus dem Internet;
- Ableitung von Emotionen am Arbeitsplatz und in Bildungseinrichtungen, außer aus medizinischen Gründen oder Sicherheitsgründen; sowie
- biometrische Kategorisierung und biometrische Echtzeit-Fernidentifizierungssysteme mit Ausnahmen für die Strafverfolgung.

4.2. Hochrisiko-KI-Systeme

DEFINITION: Ein KI-System gilt als Hochrisiko-KI-System, wenn es als Sicherheitskomponente eines in Anhang I aufgeführten regulierten Produkts genutzt wird oder selbst ein solches Produkt ist (Art. 6 Abs. 1). Dazu zählen u.a. Funkanlagen, Druckgeräte, Medizinprodukte, Fahrzeuge, Luftfahrzeuge. Zusätzlich gelten die in Anhang III genannten KI-Systeme als hochriskant, sofern diese ein erhebliches Risiko in Bezug auf die Gesundheit, Sicherheit oder Grundrechte von Personen bergen (Art. 6 Abs. 2, 3). Dazu zählen KI-Systeme, die u.a. für Kritische Infrastrukturen, Feststellung des Zugangs zu Bildungseinrichtungen, Einstellung oder Auswahl von Bewerbenden, Entscheidungen zu Beförderungen oder Kündigungen von Arbeitsverhältnissen eingesetzt werden.

PFLICHTEN: Für Hochrisiko-KI-Systeme gilt ein umfassender Pflichtenkatalog.

Die Artikel 8 bis 15 enthalten die Anforderungen an das Hochrisiko-KI-System selbst, vor allem

- die Einrichtung, Anwendung, Dokumentation und Aufrechterhaltung eines Risikomanagementsystems,
- Daten-Governance- und Datenverwaltungsverfahren für die Trainings-, Validierungs- und Testdaten,
- die Erstellung einer technischen Dokumentation,
- Protokollierungspflichten,
- Transparenz- und Informationspflichten,
- die Gewährleistung der menschlichen Aufsicht sowie
- die Erreichung eines angemessenen Maßes an Genauigkeit, Robustheit und Cybersicherheit.

Anbieter von Hochrisiko-KI-Systemen haben zudem u.a. ein Qualitätsmanagementsystem einzurichten und das System vor dem Inverkehrbringen oder der Inbetriebnahme einem Konformitätsbewertungsverfahren zu unterziehen (Art. 16 ff.).

Betreiber von Hochrisiko-KI-Systemen haben u.a. durch geeignete technische und organisatorische Maßnahmen sicherzustellen, dass die Systeme ihrer Zweckbestimmung entsprechen, die Systeme über eine menschliche Aufsicht verfügen und überwacht werden, damit Vorfälle an den Anbieter und die zuständigen Behörden gemeldet werden können (Art. 26f.).

4.3. Transparenzpflichten für bestimmte KI-Systeme

Der AI Act sieht für Anbieter von KI-Systemen, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, besondere Transparenzpflichten vor.

PFLICHTEN: Die betreffenden Personen müssen vor der ersten Nutzung in klarer und eindeutiger Weise darüber informiert werden, dass sie mit einem KI-System interagieren (Art. 50 Abs. 1, 5).

Anbieter und Betreiber von KI-Systemen, die synthetische Audio-, Bild, Video- oder Textinhalte erzeugen, müssen sicherstellen, dass die Ausgaben als künstlich erzeugt oder manipuliert erkennbar sind. Die Kennzeichnungspflicht besteht nicht, wenn von einem KI-System erzeugte Texte einer redaktionellen Kontrolle unterzogen wurden und eine natürliche oder juristische Person die redaktionelle Verantwortung für die Veröffentlichung der Inhalte trägt (Art. 50 Abs. 4).

4.4. KI-Modelle mit allgemeinem Verwendungszweck

Der AI Act sieht besondere Regelungen für KI-Modelle mit allgemeinem Verwendungszweck vor.

DEFINITION: Solche Modelle weisen nach der Definition in Art. 3 Nr. 63 eine erhebliche allgemeine Verwendbarkeit auf und sind in der Lage, unabhängig von der Art und Weise ihres Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben in einer Vielzahl nachgelagerter Systeme oder Anwendungen kompetent zu erfüllen. Hierunter fallen vor allem generative KI-Modelle, die Text-, Audio-, Bild- oder Videoinhalte erzeugen.

PFLICHTEN: Anbieter von KI-Modellen mit allgemeinem Verwendungszweck müssen eine detaillierte technische Dokumentation des Modells erstellen und aktuell halten, die mindestens die in Anhang XI aufgeführten Informationen enthält, damit sie dem Büro für Künstliche Intelligenz und den zuständigen nationalen Behörden auf Anfrage zur Verfügung gestellt werden kann (Art. 53 Abs. 1 lit. a)).

Für nachgelagerte Anbieter, die das Modell in ihr KI-System integrieren, ist ebenfalls eine technische Dokumentation zu erstellen und aktuell zu halten, die aber weniger detailliert ist und mindestens die in Anhang XII aufgeführten Informationen enthalten muss (Art. 53. Abs. 1 lit. b)).

Für KI-Modelle, die im Rahmen einer freien und offenen Lizenz unter bestimmten Voraussetzungen bereitgestellt werden, gelten diese Anforderungen nicht (Art. 53 Abs. 2).

Für Anbieter von KI-Modellen mit allgemeinem Verwendungszweck und systemischem Risiko gelten zusätzliche Anforderungen.

DEFINITION: Systemische Risiken bestehen, wenn das Modell über Fähigkeiten mit einem hohen Wirkungsgrad verfügt (Art. 51 Abs. 1). Dies wird angenommen, wenn für das Training ein sehr hoher Rechenaufwand von mehr als 1025 Gleitkommaoperationen pro Sekunde (sog. FLOPS) benötigt wird (Art. 51 Abs. 2).

PFLICHTEN: Die Anbieter müssen eine Modellbewertung mit standardisierten Protokollen und Instrumenten inkl. Angriffstests durchführen, um Systemrisiken zu ermitteln und zu mindern. Weiterhin sind sie verpflichtet, mögliche systemische Risiken auf EU-Ebene im Zusammenhang mit der Entwicklung, dem Inverkehrbringen oder der Verwendung des KI-Modells zu bewerten und zu mindern, schwerwiegenden Vorfällen nachzugehen und diese zu melden sowie ein angemessenes Maß an Cybersicherheit zu gewährleisten (Art. 55 Abs. 1).

4.5. Freiwillige Anwendung von Anforderungen

Anbieter und Betreiber von KI-Systemen, die in keine der vorgenannten Kategorien fallen, können freiwillig einige oder alle der in Kapitel III Abschnitt 2 genannten Anforderungen für KI-Systeme, die kein hohes Risiko bergen, anwenden (Art. 95 Abs. 1). Hierzu wird die Aufstellung von Verhaltenskodizes empfohlen.

5. AUSWIRKUNGEN FÜR UNTERNEHMEN UND WIRTSCHAFTSPRÜFER

- Da der Anwendungsbereich sehr weit gefasst ist, sollten sich alle Unternehmen und Wirtschaftsprüfer mit den Anforderungen des AI Act vertraut machen.

- Zum Einsatz kommende KI-Systeme müssen als solche identifiziert und kategorisiert werden. Die sich aus der Risikokategorie ergebenden Auswirkungen sind zu bewerten und die entsprechenden Maßnahmen im Rahmen von Governance- und Compliance Strukturen umzusetzen.

- Setzen Wirtschaftsprüfer ein KI-System i.S. des AI Act ein, empfiehlt sich eine Integration der Vorgaben aus dem AI Act in das bestehende Qualitätsmanagement der Wirtschaftsprüferpraxis *(IDW QMS 1 (09.2022))*.

- Zusätzlich spielt auch die im AI Act vorgesehene KI-Kompetenz (Art. 4) eine wesentliche Rolle. Unternehmen und Wirtschaftsprüfer, die KI-Systeme i.S. des AI Act einsetzen, sind verpflichtet, Maßnahmen zu ergreifen, damit ihr Personal über ein ausreichendes Maß an KI-Kompetenz verfügt.

- Wirtschaftsprüfer können zudem dem steigenden Bedarf der Unternehmen nach standardisierten KI-Prüfungen auf der Basis geeigneter Kriterien durch die Anwendung des *IDW Prüfungsstandards: Prüfung von KI-Systemen (IDW PS 861) (03.2023)* nachkommen.

Ihre Notizen:

Dieses Knowledge Paper wurde von der IDW Geschäftsstelle in Zusammenarbeit mit der Arbeitsgruppe „Prüfung von KI“ erarbeitet.

Wir freuen uns über Ihre Anmerkungen. Sie können diese an das Institut der Wirtschaftsprüfer in Deutschland e.V., Postfach 320580, 40420 Düsseldorf oder an grit.baum@idw.de senden.

Bildrechte:

Seite 3, 4,, 5, 6, 10: perstige©Adobe-Stock.com, Seite 3: CoreDESIGN©Adobe-Stock.com,
Seite 4: All Creative Lines©Adobe-Stock.com, Seite 6: tanaonte©Adobe-Stock.com,
Seite 10/11: noah9000©Adobe-Stock.com

INSTITUT DER WIRTSCHAFTSPRÜFER IN DEUTSCHLAND E.V.

Roßstr. 74
40476 Düsseldorf

Postfach 32 05 80
40420 Düsseldorf

Telefon: +49 (0) 211/4561-0
Telefax: +49 (0) 211/4561097

E-Mail: info@idw.de
Web: www.idw.de

IDW
INSTITUT DER WIRTSCHAFTSPRÜFER