eco Initiative KI in der Praxis

Arbeitsgruppe Rahmenbedingungen

AI Act – Worüber reden wir eigentlich?

Mit freundlicher Unterstützung von

Inhalt

Einleitung **3**

1. AI Act im Kontext des EU-Datenwirtschaftsrechts **6**

2. Was regelt der AI Act? **10**

2.1 Was ist ein „KI-System“? **11**

Beispiel für ein KI-System: KI-gestütztes Ratsinformationssystem der Stadt Freiburg *12*

2.2 Was ist ein „KI-Modell mit allgemeinem Verwendungszweck“ **13**

2.3 Was ist die Risiko-Taxonomie des AI Act und was ist die Konsequenz? **16**

2.3.1 Inakzeptables / Unannehmbares Risiko *18*

2.3.2 Hochrisiko-KI *18*

2.3.3 Begrenztes Risiko *20*

2.3.4 Minimales oder kein Risiko *20*

2.4 Die unterschiedlichen Verpflichteten und Akteure des AI Acts? **21**

2.5 Darstellung von Ziffer 2.1 bis 2.3 ..... **23**

2.6 Zeitlicher Anwendungsbeginn des AI Act zwischen 02.02.2025 und 02.08.2027 ..... **24**

3. „Governance-Gedankenmodell“ anhand IBM watsonx.governance **25**

4. Etablierte Begriffe und ihr Bezug zum AI Act **27**

Impressum **33**

Einleitung

Autoren:
Martin Batz, Lukas Rybok, IONOS SE |
Kai Meinke, deltaDAO AG |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Künstliche Intelligenz (KI) hat sich in den vergangenen Jahren zu einer der einflussreichsten Technologien entwickelt[1], die nahezu jeden Aspekt von Wirtschaft und Gesellschaft berührt. In der Wirtschaft revolutioniert KI Geschäftsmodelle, steigert die Effizienz und stärkt die Innovationskraft. Unternehmen nutzen maschinelles Lernen und intelligente Algorithmen, um Prozesse zu verbessern, Kundenbedürfnisse genauer zu verstehen und innovative Produkte und Dienstleistungen zu entwickeln. Branchen wie das Gesundheitswesen profitieren von KI durch präzisere Diagnosen und personalisierte Medizin, während der Finanzsektor KI zur Risikobewertung und Betrugsprävention einsetzt.

Die rasante technologische Entwicklung und der zunehmende Einsatz von KI bergen nicht nur erhebliche Potenziale, sondern bringen auch komplexe Herausforderungen mit sich[2]. Eine rechtliche Regulierung von KI-Systemen ist unerlässlich, um die Chancen bestmöglich zu nutzen und zugleich die Risiken zu managen. Ohne einen klaren rechtlichen Rahmen besteht die Gefahr, dass die Technologie unbeabsichtigte negative Auswirkungen hat. Dazu gehören Risiken für den Datenschutz und die Privatsphäre, da KI-Systeme große Mengen personenbezogener Daten verarbeiten können und mit ihnen trainiert werden. Eine Regulierung kann sicherstellen, dass Daten verantwortungsvoll genutzt werden und die Rechte der Individuen gewahrt bleiben. Zudem sind KI-Systeme anfällig für Bias und Diskriminierung, besonders wenn die zugrundeliegenden Algorithmen oder Daten unzureichend geprüft wurden. Gesetzliche Vorgaben, die Transparenz und Fairness in der KI-Entwicklung sicherstellen, können dabei helfen, solche Diskriminierungen zu minimieren.

Darüber hinaus trägt eine Regulierung dazu bei, in der Bevölkerung das Vertrauen in KI und die Akzeptanz der Technologie zu fördern. Klare gesetzliche Richtlinien bieten Unternehmen und anderen Organisationen einen Orientierungsrahmen, der Innovationen ermöglicht und gleichzeitig für die Sicherheit und Zuverlässigkeit von KI-Systemen sorgt. Indem sie Verantwortlichkeiten und Haftungsfragen klärt, reduziert eine Regulierung rechtliche Unsicherheiten und schafft Investitionssicherheit für Unternehmen.

[1] "The impact of Artificial Intelligence on productivity, distribution and growth: Key mechanisms, initial evidence and policy challenges", OECD Publishing, https://doi.org/10.1787/8d900037-en, 16.04.2024
[2] "The impact of the General Data Protection Regulation (GDPR) on artificial intelligence.", European Parliamentary Research Services, https://dx.doi.org/10.2861/293, Juni 2020

In diesem Kontext stellt die Verabschiedung des Artificial Intelligence Act (AI Act) der EU[3] einen bahnbrechenden Schritt in der Regulierung von KI dar.[4] Die Verordnung hat weitreichende Auswirkungen auf Unternehmen, öffentliche Institutionen, Anbieter und Nutzende von KI-Systemen in der Europäischen Union, aber auch darüber hinaus[5]. Als weltweit erster umfassender Rechtsrahmen für KI zielt der AI Act darauf ab, eine vertrauenswürdige KI-Entwicklung zu fördern und gleichzeitig potenzielle Risiken anzugehen und dabei die Achtung von Grundrechten im digitalen Raum zu gewährleisten. Die Auswirkungen des Gesetzes reichen über die Grenzen der EU hinaus und werden die globale KI-Governance beeinflussen und für andere Länder als Präzedenzfall dienen.

Angesichts der rasanten Verbreitung von KI-Lösungen ist es für die Akteure aller Branchen von entscheidender Bedeutung, die praktischen Auswirkungen des europäischen KI-Gesetzes zu verstehen, um sich in der neuen Regulierungslandschaft zurechtzufinden.

Die KI-Verordnung birgt Chancen für Unternehmen, stellt die Wirtschaft aber auch vor Herausforderungen[6,7]. Einerseits können Anbieter sich Wettbewerbsvorteile verschaffen, indem sie die Vorschriften des Gesetzes einhalten und „kugelsicherer" KI-Lösungen schaffen. Andererseits gibt es Bedenken hinsichtlich höherer Kosten, Innovationsbeschränkungen und der Abwanderung von Talenten in weniger regulierte Märkte wie die USA, Asien und den Nahen Osten.

Wenn die Vorschriften des AI Act anwendbar sind, was in mehreren Stufen zwischen dem 2. Februar 2025 und dem 2. August 2027 erfolgt, werden mehr Vertrauen und Sicherheit in KI-Anwendungen sowie geringere Kosten bei der Suche nach rechtskonformen Lösungen erwartet. Der AI Act ist Mitte 2024 in Kraft getreten, ist aber erst zu einem späteren Zeitpunkt anwendbar (siehe unten Ziffer 2.6). Allerdings könnten wir auch mit höheren Preisen für KI-Produkte und -Dienstleistungen konfrontiert werden, die auf die Einhaltung der Vorschriften zurückzuführen sind. Die Strafen für Unternehmen, wenn sie gegen das KI-Gesetz verstoßen, sind beträchtlich. Sie reichen von 7,5 Millionen bis 35 Millionen Euro oder ein bis sieben Prozent des weltweiten Jahresumsatzes, je nach Schwere des Verstoßes[8].

[3] VERORDNUNG (EU) 2024/1689 DES EUROPÄISCHEN PARLAMENTS UND DES RATES vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung der Verordnungen (EG) Nr. 300/2008, (EU) Nr. 167/2013, (EU) Nr. 168/2013, (EU) 2018/858, (EU) 2018/1139 und (EU) 2019/2144 sowie der Richtlinien 2014/90/EU, (EU) 2016/797 und (EU) 2020/1828 (Verordnung über künstliche Intelligenz) (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689)

[4] "AI Act", European Commission, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai, 22.09.2024; "Einheitliche Regeln für künstliche Intelligenz", Bundesregierung, https://www.bundesregierung.de/breg-de/themen/digitalisierung/kuenstliche-intelligenz/AI Act-2285944, 22.09.2024

[5] Aufgrund des sog. Marktortprinzips muss der AI Act nicht nur durch in der EU niedergelassene Unternehmen beachtet werden, sondern auch von Unternehmen, die – vereinfacht gesagt – aus Drittstaaten heraus in der EU aktiv sind (Art. 2 Abs. 1 AI Act bspw. "in der Union KI-Systeme in Verkehr bringen oder in Betrieb nehmen" (Art. 2 Abs. 1 lit. a AI Act) und "Anbieter und Betreiber von KI-Systemen, die ihren Sitz in einem Drittland haben oder sich in einem Drittland befinden, wenn die vom KI-System hervorgebrachte Ausgabe in der Union verwendet wird;" (Art. 2 Abs. 1 lit. c AI Act).

[6] "European AI Act: Opportunities and Challenges", Jochen Ditsche und Maria Makhaylenko, Roland Berger, https://www.rolandberger.com/en/Insights/Publications/European-AI Act-Opportunities-and-challenges.html, 22.09.2024

[7] "EU AI Act: Europäische Regulierung und ihre Umsetzung", PwC, https://www.pwc.de/de/risk-regulatory/responsible-ai/europaeische-ki-regulierung-und-ihre-umsetzung.html, 22.09.2024

[8] "Article 99: Penalties", EU Artificial Intelligence Act, https://artificialintelligenceact.eu/article/99/, 22.09.2024

Mit diesem Papier wollen wir eine Navigationshilfe bieten, die den Einstieg in das Thema erleichtern soll. Indem sich die Akteure frühzeitig auf den AI Act ausrichten und bei der Entwicklung, dem Betrieb und der Nutzung von KI-Systemen gemäß der Regulierung richtig vorgehen, können sie die Chancen besser nutzen und Risiken vermeiden.

1. AI Act im Kontext des EU-Datenwirtschaftsrechts

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |
Martin Batz, Mark Neufurth, IONOS SE |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Als Ausfluss der Datenstrategie der Europäischen Union (EU) wurden und werden in schneller Folge verschiedene Gesetzesvorhaben umgesetzt, die Europa in die Lage versetzen sollen, Daten wirtschaftlich nutzbar zu machen, ohne dass die Sorgfalt beim Umgang mit Daten auf der Strecke bleibt oder die berechtigten Interessen der Dateninhaber leiden. Ziel der Union ist es, Europa in Zeiten des digitalen Wandels wettbewerbsfähig zu halten.

Als 'Datenwirtschaftsrecht' begreifen sich alle Regeln zu Datenzugang, -austausch und -nutzung. Dabei setzt sich 'Datenwirtschaftsrecht' offenkundig aus drei Begriffen zusammen: (digitale) Daten, Wirtschaft und Recht. Als Daten bezeichnet man die Darstellung von Informationen in **formalisierter** Art. (Digitale) Daten sind zur **Kommunikation**, **Interpretation** und **Verarbeitung** geeignet. Dabei werden die Informationen durch **Zeichenfolgen** repräsentiert, deren Aufbau einer **vereinbarten Syntax** folgt. Zu Informationen werden die Zeichen, wenn sie in einem Bedeutungskontext interpretiert werden.

Wirtschaft verortet an dieser Stelle den Real- bzw. Lebensbereich, für den Daten von Relevanz sind - für Wirtschaftssubjekte. Recht schließlich verdeutlicht die rechtswissenschaftliche Komponente, also die rechtlich-regulatorische Steuerung eben der Datenwirtschaft.

Diese Definition ist ganz wichtig, denn EU-Kommission, EU-Parlament und EU-Rat haben in ausführlichen Gesetzgebungsverfahren Normen geschaffen, die miteinander zusammenhängen. Sie können nicht getrennt voneinander betrachtet werden, wenn künftig über den Umgang mit Daten geredet werden soll. Zu den Normen zählen:

- der Digital Services Act / Gesetz über digitale Dienste,
- der Digital Markets Act / Gesetz über digitale Märkte,
- der Data Governance Act / Gesetz über Daten-Governance,
- **der Artificial Intelligence Act / Verordnung über künstliche Intelligenz,**
- der Data Act / das Datengesetz.

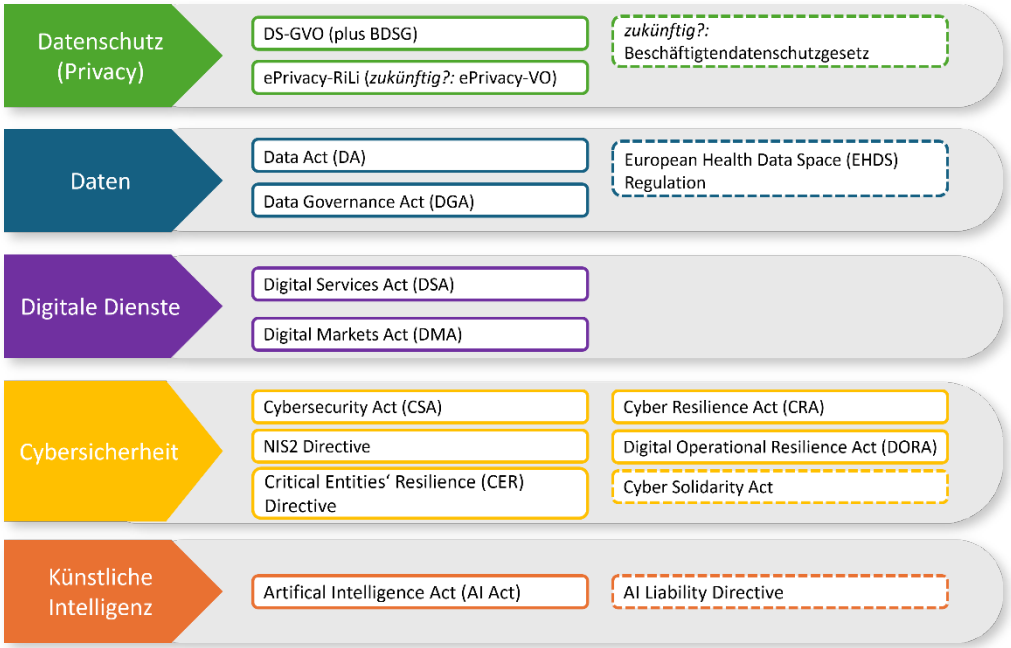

Abb. 1: Regelungsziele Datenwirtschaftsrecht

Die vorstehende Übersicht soll die Regelungsziele der Rechtsakte verdeutlichen. Die Datenschutzbestimmungen schützen die betroffenen natürlichen Personen vor der und gegen die Verarbeitung von deren Daten, wobei die Gesetze diese Verarbeitung nicht per se verbieten, sondern deren Rechtsrahmen regeln. Diese Bestimmungen stehen damit in einem natürlichen Spannungsverhältnis zu den Regelungen des Data Act und des Data Governance Act, die das Teilen und Bereitstellen von Daten an andere als den Erzeuger der Daten regeln. Während der Data Act (DA) neben dem sogenannten Cloud Switching auch Pflichten zum Teilen und zur Bereitstellung von Daten des IoT-Umfeldes einschließlich der Rahmenbedingungen der Bereitstellung regelt, sieht der Data Governance Act (DGA) keine Pflicht zum Teilen vor, sondern definiert nur die Rahmenbedingungen, falls Daten zur Förderung der wirtschaftlichen Entwicklung „freiwillig“ geteilt werden.

Der Digital Services Act (DSA) und der Digital Markets Act (DMA) regeln nicht den Umgang mit Daten, sondern die Bereitstellung von Digitalservices. Sie haben noch stärker als der DA eine wettbewerbsregulierende Funktion. Die Regelungen, die der Cybersecurity zugeordnet werden, tragen den veränderten Cyber- und IT-Risiken Rechnung. Neben der Einführung neuer Regelungen für Digitale Produkte durch den Cyber Resilience Act (CRA) erweitern insbesondere die Network and Information Systems Directive 2 (NIS2-Richtlinie) und die Digital Operational Resilience Act (DORA)bereits bestehende Regelungen. Der Schwerpunkt liegt dabei vor allem darauf, die Resilienz sowie die Betrachtung des ITK-Drittparteienrisikos in den Maßnahmen zur Sicherheit stärker zu betonen und zu verankern. Der AI Act überschneidet sich insofern mit dem CRA, weil beide dem Produktsicherheitsrecht zuzuordnen sind. Der AI Act ist kein „Datenschutz 2.0“. Somit zielt er nicht auf den Schutz der Daten natürlicher Personen ab. Er soll die Gefahren adressieren, die sich aus KI-Produkten ergeben.

Festzustellen ist, dass diese Rechtsakte weder zeitgleich in Kraft getreten sind noch ein einheitliches Regelungsziel haben, sich aber in der praktischen Anwendung überschneiden.

Wir sehen also: In jüngerer Zeit hat die Regulierung der Datenwirtschaft erheblich an Fahrt aufgenommen. Der umfassende Datenregulierungsrahmen soll nicht nur den Datenaustausch befördern

soll, sondern auch den Datenzugang regeln und flexibilisieren, sogenannte Datenintermediäre schaffen und Datenaltruismus in den Fokus rücken. Dieser Regulierungsrahmen umfasst nicht zuletzt den viel besprochenen Artificial Intelligence Act (kurz: AI Act).

Arbeitet man den Wesenskern aller europäischen Normen zur Datenwirtschaft heraus, so zielt das Datenwirtschaftsrecht darauf ab, die sichere und vertrauenswürdige Bereitstellung von Daten (heutzutage oft von Cloud-Anbietern gehostet) zu regeln und ihre Einbindung in (Software-)Anwendungen zu organisieren, gleichzeitig die Beteiligten zu Schutz und Transparenz zu verpflichten, zusätzlich gleiche Wettbewerbsbedingungen für digitale Unternehmen zu garantieren, die Nutzung von Daten in Wirtschaft und Öffentlichem Sektor allgemein zu fördern sowie einen regelrechten EU-Binnenmarkt für Daten zu schaffen.

Zentrale Norm ist hier der Data Act, der eng mit dem AI Act zusammenarbeitet. Bevor auf Letzteren eingegangen werden kann, muss zunächst der Data Act betrachtet werden.

Der EU-Gesetzgeber hat sich auf die finale Version des Data Act geeinigt. Der Data Act ist bereits in Kraft getreten und sieht Übergangsfrist bis zu seinem Anwendungsbeginn im Jahr 2025 vor (Art. 50 Data Act).[9] Für Verträge über die Bereitstellung von Daten gelten damit künftig neue Bedingungen. Der Data Act regelt aber die Bereitstellung von Daten nicht vollumfänglich, sondern nur in seinem sachlichen Anwendungsbereich. Dieser lässt sich – grob und vereinfacht – als „Industrial Internet of Things“ (IIoT) umschreiben (siehe Art. 1 Data Act). Er hält aber auch – wenngleich das mit Blick auf die Bezeichnung überrascht – Regelungen zum sog. Cloud Switching.

Die zugrundeliegende Verordnung schafft erstmals einheitliche Regelungen für den Datenaustausch zwischen Unternehmen, Behörden und anderen Akteuren der Datenwirtschaft. Somit verfolgt Data Act das Ziel, eine Konzentration von Daten bei wenigen großen Unternehmen zu verhindern bzw. rückgängig zu machen. Daten sollen auch anderen Akteuren zugänglich gemacht werden. Im Lichte dessen soll so die Entwicklung neuer datengetriebener Geschäftsmodelle gefördert werden. Profiteure dieser Liberalisierung könnten vor allem kleine und mittelständische Unternehmen sein; in jedem Fall kommt den Betreibern von Internetdiensten hierbei eine besondere Rolle zu. Kein Internet ohne den Austausch von Daten. Ziehen doch Protokolle wie TCP/IP ihre Daseinsberechtigung aus dem Datentransfer. Wirtschaftsunternehmen und Internetdienstanbieter erhalten mit dem Data Act also die Gelegenheit, deutlich einfacher auf Daten zuzugreifen und diese selbst zu nutzen bzw. ihren Austausch zu fördern.

Nutzer vernetzter Produkte und Dienste (Internet of Things, kurz IoT) können demzufolge vom Hersteller eines Investitionsguts die Bereitstellung der Nutzungsdaten verlangen, um diese Daten für eigene Zwecke zu nutzen – etwa für Produktentwicklung oder das Trainieren von KI-Algorithmen. Überall dort aber, wo Daten massenhaft bereitgestellt werden müssen, sind künftig Bereitstellungsverträge einzugehen und die Maßgaben des Data Act zu beachten.

[9] *„Diese Verordnung tritt am zwanzigsten Tag nach ihrer Veröffentlichung im Amtsblatt der Europäischen Union in Kraft. Sie gilt ab dem 12. September 2025. Die Verpflichtung gemäß Artikel 3 Absatz 1 gilt für vernetzte Produkte und die mit ihnen verbundenen Dienste, die nach dem 12. September 2026 in Verkehr gebracht wurden. Kapitel III gilt nur in Bezug auf Datenbereitstellungspflichten nach dem Unionsrecht oder nach im Einklang mit Unionsrecht erlassenen nationalen Rechtsvorschriften, die nach dem 12. September 2025 in Kraft treten. Kapitel IV gilt für Verträge, die nach dem 12. September 2025 geschlossen wurden. Kapitel IV gilt ab dem 12. September 2027 für Verträge, die am oder vor dem 12. September 2025 geschlossen wurden, sofern a) sie unbefristet sind oder b) ihre Geltungsdauer frühestens 10 Jahre nach dem 11. Januar 2024 endet.“* (Art. 50 Data Act)

Der Data Act kennt drei zwingende allgemeine Grundsätze für Verträge über die Bereitstellung von Daten.

Diskriminierungsverbot: Die Daten müssen zu fairen, angemessenen und nichtdiskriminierenden Bedingungen bereitgestellt werden. Unzulässig ist etwa, die Daten einem Partnerunternehmen zu besseren Konditionen bereitzustellen als vergleichbaren Mitbewerbern.

Gegenleistung: Für Datenbereitstellung darf eine angemessene Gegenleistung beansprucht werden. Sonderregeln gelten allerdings für kleine und mittelständische Unternehmen: Für sie ist die Gegenleistung von vornherein begrenzt auf die Selbstkosten des Dateninhabers im Rahmen der Bereitstellung.

Verbot missbräuchlicher Klauseln: Bestimmte Klauseln sind in Bereitstellungsverträgen für Daten schlicht unzulässig. Der Dateninhaber darf z.B. nicht vorschreiben, wie die Daten genutzt werden dürfen.

Klar wird an dieser Stelle: Das wirtschaftliche Potenzial des Data Act ist erheblich. Laut EU-Kommission werden rund 80 Prozent der in der EU entstehenden Industriedaten kaum oder noch gar nicht wirtschaftlich verwertet; gleichzeitig steigen die erzeugten Datenmengen enorm an. Die wirtschaftliche Wertschöpfung wird auf 270 Milliarden Euro allein in den ersten fünf Jahren nach Inkrafttreten des Data Act geschätzt. Datensilos marktbeherrschender Unternehmen könnten nun erstmals aufgebrochen werden. Potenziale für die Entwicklung datengetriebener Geschäftsmodelle liegen auf der Hand.

Auf der Gegenseite tauchen neue Herausforderungen auf. So etwa haben Dateninhaber das Interesse, die eigenen Geschäftsgeheimnisse zu schützen. Kurzum: Die EU sieht die Konkurrenz in der Nutzung und Verwertung von Daten als Motor für Innovation.

Und hier kommt der AI Act ins Spiel. Datenmengen wachsen stündlich weiter an, wie ein Hefepilz in alle Richtungen. Wenn die Frage der Art und Weise der Nutzung auf Basis des Data Act geklärt ist, wenn gleichzeitig rechtliche Fragen aus den Bereichen Urheberrecht, Schutz persönlicher Daten, IT-Datenschutz, NIS2 und CRA-Produkthaftung gegenüber dem Data Act abgegrenzt sind, kann Künstliche Intelligenz ihre Stärke ausspielen.

Denn KI-Systeme saugen Daten gierig auf und verarbeiten sie in bislang unbekannter Geschwindigkeit zu Informationen. Sie legen Einblicke offen, die es bislang nicht gab, finden buchstäblich die goldene Nähnadel der Erkenntnis im digitalen Heuhaufen und sparen beteiligten Menschen erheblichen Aufwand. Nicht zuletzt ermöglicht KI im Zusammenspiel mit Daten autonome Handlungen von Maschinen. Somit knüpft der AI Act nahtlos an die Datenstrategie der EU als Treibsatz für das moderne Datenwirtschaftsrecht an.

2. Was regelt der AI Act?

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Regelungsgegenstand des AI Act ist es, „die Einführung einer auf den Menschen ausgerichteten und vertrauenswürdigen künstlichen Intelligenz (KI) zu fördern und gleichzeitig ein hohes Schutzniveau in Bezug auf Gesundheit, Sicherheit und die in der Charta verankerten Grundrechte, einschließlich Demokratie, Rechtsstaatlichkeit und Umweltschutz, vor schädlichen Auswirkungen von KI-Systemen in der Union zu gewährleisten und die Innovation zu unterstützen“ (Art. 1 Abs. 1 AI Act).

Der AI Act regelt nicht den Schutz personenbezogener Daten, sondern die Anforderungen an den Einsatz von KI-Systemen und KI-Modellen mit allgemeinem Verwendungszweck. Er lässt sich als „Risikobeherrschungsgesetz“ umschreiben. Durch den AI Act wird das von KI ausgehende Risiko grundsätzlich akzeptiert. Der Gesetzgeber geht dabei davon aus, dass durch die Einhaltung der Pflichten des AI Act dieses akzeptierte Risiko beherrscht werden kann.

Die Idee dahinter lässt sich auch so ausdrücken: Kommt es trotz Einhaltung der Pflichten des AI Act zu einem Schaden, dann sollte eine Haftung grundsätzlich ausscheiden, weil der Akteur keinen Rechtsverstoß begangen hat. Wie dies in der Praxis gelebt werden wird, ist noch offen. Zumal aktuell eine AI Liability Directive – also eine grundlegende Regelung zur Haftung – in der Diskussion ist.

Die zentralen Begriffe, mit denen der Regelungsgegenstand des AI Act beschrieben werden, sind „KI-System“ und „KI-Modell mit allgemeinem Verwendungszweck“. Nur wenn ein KI-System gegeben ist, kommt der AI Act zur Anwendung (Art. 2 Abs. 1 AI Act).

Der AI Act nimmt eine autonome Begriffsbestimmung vor. Dies erfordert es, die Anwendung des Gesetzes hieran auszurichten und die bisherigen Technologien anhand dieser Begriffsbestimmungen einzuordnen.

Wie in den Kapiteln zuvor beschrieben und allgemein wahrnehmbar, ist KI eine der faszinierendsten und wichtigsten Technologien der Gegenwart. KI bietet uns Möglichkeiten, menschliche Fähigkeiten zu erweitern und Prozesse zu beschleunigen, birgt aber auch Risiken und Herausforderungen. Diese Risiken und Herausforderungen sind es, die eine staatliche Regulierung notwendig machen und in der EU zum EU AI Act geführt haben.

Diese KI-Verordnung der Europäischen Union ist das erste umfassende Gesetz, welches die Rahmenbedingungen zum Umgang mit künstlicher Intelligenz gestalten soll. Gestartet im Jahr 2021, wurde das Gesetz im Frühjahr 2024 in einer finalen Abstimmung im EU-Parlament verabschiedet und im Juli darauf im Amtsblatt veröffentlicht und ist zum 1. August 2024 offiziell in Kraft getreten.

Damit hat eine Übergangsfrist von 24 Monaten zur Umsetzung begonnen. In dieser Zeit werden die Vorgaben stufenweise bis zum 2. August 2027 verpflichtend (siehe unten Ziffer 2.6).

2.1 Was ist ein „KI-System“?

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Der Begriff „KI-System“ wird in Art. 3 Nr. 1 AI Act definiert als „ein maschinengestütztes System, das für einen in unterschiedlicher 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“.

Der Definition lässt sich entnehmen, dass eine Ergebnisausgabe erforderlich ist und das System über eine hinreichende Autonomie verfügen muss. Diese Autonomie hat die Definition mit der Diskussion um den Begriff „Künstliche Intelligenz“ gemeinsam.

Dem Erwägungsgrund 12 AI Act sind weitergehende Konkretisierungen zu entnehmen, wie der Begriff „KI-System“ auszulegen ist. Dieser Erwägungsgrund 12 AI Act wird nachfolgend ausschnittsweise wiedergegeben:

- „Darüber hinaus sollte die Begriffsbestimmung auf den wesentlichen Merkmalen der KI beruhen, die sie **von einfacheren herkömmlichen Softwaresystemen und Programmierungsansätzen abgrenzen**, und sollte sich **nicht auf Systeme beziehen, die auf ausschließlich von natürlichen Personen definierten Regeln für das automatische Ausführen von Operationen beruhen**.“
- *„Ein wesentliches Merkmal von KI-Systemen ist **ihre Fähigkeit, abzuleiten**. Diese Fähigkeit **bezieht sich auf den Prozess der Erzeugung von Ausgaben, wie Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen, die physische und digitale Umgebungen beeinflussen können, sowie auf die Fähigkeit von KI-Systemen, Modelle oder Algorithmen oder beides aus Eingaben oder Daten abzuleiten.** Zu den Techniken, die während der Gestaltung eines KI-Systems das Ableiten ermöglichen, gehören **Ansätze für maschinelles Lernen, wobei aus Daten gelernt wird, wie bestimmte Ziele erreicht werden können, sowie logik- und wissensgestützte Konzepte, wobei aus kodierten Informationen oder symbolischen Darstellungen der zu lösenden Aufgabe abgeleitet wird**. Die Fähigkeit eines KI-Systems, abzuleiten, **geht über die einfache Datenverarbeitung hinaus, indem Lern-, Schlussfolgerungs- und Modellierungsprozesse ermöglicht werden**.“*
- „Die **Bezeichnung ‚maschinenbasiert‘ bezieht sich auf die Tatsache, dass KI-Systeme von Maschinen betrieben werden**. Durch die Bezugnahme auf explizite oder implizite Ziele wird betont, dass KI-Systeme gemäß explizit festgelegten Zielen oder gemäß impliziten Zielen arbeiten können. Die Ziele des KI-Systems können sich — unter bestimmten Umständen — von der Zweckbestimmung des KI-Systems unterscheiden. Für die Zwecke dieser Verordnung sollten Umgebungen als Kontexte verstanden werden, in denen KI-Systeme betrieben werden, während die von einem KI-System erzeugten Ausgaben verschiedene Funktionen von KI-Systemen widerspiegeln, darunter Vorhersagen, Inhalte, Empfehlungen oder Entscheidungen.“

- ***„KI-Systeme sind mit verschiedenen Graden der Autonomie ausgestattet, was bedeutet, dass sie bis zu einem gewissen Grad unabhängig von menschlichem Zutun agieren und in der Lage sind, ohne menschliches Eingreifen zu arbeiten.*** *Die Anpassungsfähigkeit, die ein KI-System nach Inbetriebnahme aufweisen könnte, bezieht sich auf seine Lernfähigkeit, durch sie es sich während seiner Verwendung verändern kann. KI-Systeme können* ***eigenständig oder als Bestandteil eines Produkts*** *verwendet werden, unabhängig davon, ob das System physisch in das Produkt integriert (eingebettet) ist oder der Funktion des Produkts dient, ohne darin integriert zu sein.“*

Beispiel für ein KI-System: KI-gestütztes Ratsinformationssystem der Stadt Freiburg

Beschreibung:
Das KI-basierte Ratsinformationssystem der Stadt Freiburg ermöglicht eine einfache und effiziente Suche nach Informationen zu Gemeinderatsbeschlüssen. Es analysiert und liefert relevante Daten zu politischen Entscheidungen sowie deren Kontext. Dadurch werden mühsame Recherchen reduziert und die Arbeit der Stadtverwaltung sowie der Zugang zu Informationen optimiert.

Zielgruppe:
Die Hauptnutzer: innen sind Mitarbeitende der Stadtverwaltung, Bürger: innen und politische Entscheider: innen. Sie können mit Hilfe des Systems schnell und gezielt Informationen zu politischen Entscheidungen und Beschlüssen abrufen. Dies fördert Transparenz und erleichtert den Zugang zu kommunalen Informationen.

Ergebnisse und fachliches Ziel:
Das System bietet einen schnellen Zugang zu Daten und verbessert den Service durch vereinfachte Informationsbereitstellung. Es unterstützt die politische Teilhabe, indem es Transparenz schafft und historische Entwicklungen nachvollziehbar macht. Zusätzlich wird Open Government durch die transparente Bereitstellung von Informationen gefördert.

Weiterverwendung:
Das System kann durch die Integration eines Chatbots in weitere Verwaltungsprozesse ausgeweitet werden, z. B. für Bürgerbeteiligung oder Umweltfragen. Die Anbindung an weitere Datenquellen ermöglicht eine noch umfassendere Nutzung. Zudem können Antworten individuell an Nutzerrollen und -präferenzen angepasst werden, um die Effizienz weiter zu steigern.

Quelle: https://www.freiburg.de/pb/datenraum/daten_raum_freiburg/anwendungsfaelle.html

Die Definition zeigt, dass der AI Act eine autonome Begriffsbestimmung vornimmt und nicht versucht zu definieren, was „Künstliche Intelligenz“ ist. Was unter KI zu verstehen ist, wurde zwar schon 1956 auf einer Wissenschaftskonferenz erstmals umrissen, ist aber im Detail nie konkret festgelegt worden und bleibt bis heute umstritten. Ebenso wenig greift diese Definition im AI Act direkt auf bereits gebräuchliche Begriffe der IT-Welt zurück.

Die Regulierung von KI-Systemen ist der Schwerpunkt des AI Act.

2.2 Was ist ein „KI-Modell mit allgemeinem Verwendungszweck“

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

„KI-Modelle mit allgemeinem Verwendungszweck“ haben in Kapitel V des AI Act eine eigenständige Regelung gefunden. Sie sind selbst keine KI-Systeme. Vielmehr handelt es sich bei KI-Modellen um Bestandteile von KI-Systemen. Die Unterscheidung ist von grundlegender Bedeutung.

Ein „KI-Modell mit allgemeinem Verwendungszweck“ ist nach der Definition in Art. 3 Nr. 63 AI Act „ein KI-Modell — einschließlich der Fälle, in denen ein solches KI-Modell mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert wird —, das eine erhebliche allgemeine Verwendbarkeit aufweist und in der Lage ist, unabhängig von der Art und Weise seines Inverkehrbringens ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen, und das in eine Vielzahl nachgelagerter Systeme oder Anwendungen integriert werden kann, ausgenommen KI-Modelle, die vor ihrem Inverkehrbringen für Forschungs- und Entwicklungstätigkeiten oder die Konzipierung von Prototypen eingesetzt werden“.

Insbesondere die Erwägungsgründe 97 bis 100 AI Act enthalten weitere Anhaltspunkte zum Begriffsverständnis, die auszugsweise wiedergegeben werden:

- „(97) Der **Begriff „KI-Modelle mit allgemeinem Verwendungszweck“** sollte […] **vom Begriff der KI-Systeme abgegrenzt werden**, um Rechtssicherheit zu schaffen. Die Begriffsbestimmung sollte auf den wesentlichen funktionalen Merkmalen eines KI-Modells mit allgemeinem Verwendungszweck beruhen, **insbesondere auf der allgemeinen Verwendbarkeit und der Fähigkeit, ein breites Spektrum unterschiedlicher Aufgaben kompetent zu erfüllen**. Diese Modelle werden **in der Regel mit großen Datenmengen** durch **verschiedene Methoden, etwa überwachtes, unüberwachtes und bestärkendes Lernen**, trainiert. […] **Obwohl KI-Modelle wesentliche Komponenten von KI-Systemen sind, stellen sie für sich genommen keine KI-Systeme dar. Damit KI-Modelle zu KI-Systemen werden, ist die Hinzufügung weiterer Komponenten, zum Beispiel einer Nutzerschnittstelle, erforderlich**. KI-Modelle sind in der Regel in KI-Systeme integriert und Teil davon. […]“
- „(98) Die allgemeine Verwendbarkeit eines Modells könnte zwar unter anderem auch durch eine bestimmte Anzahl von Parametern bestimmt werden, doch sollten **Modelle mit mindestens einer Milliarde Parametern, die mit einer großen Datenmenge unter umfassender Selbstüberwachung trainiert werden, als Modelle gelten, die eine erhebliche allgemeine Verwendbarkeit aufweisen** und ein breites Spektrum unterschiedlicher Aufgaben kompetent erfüllen.“
- „(99) **Große generative KI-Modelle** sind ein **typisches Beispiel für ein KI-Modell mit allgemeinem Verwendungszweck**, da sie eine flexible Erzeugung von Inhalten ermöglichen, etwa in Form von Text- Audio-, Bild- oder Videoinhalten, die leicht ein breites Spektrum unterschiedlicher Aufgaben umfassen können.“

- „(100) Wenn ein KI-Modell mit allgemeinem Verwendungszweck in ein KI-System integriert oder Teil davon ist, sollte dieses System als KI-System mit allgemeinem Verwendungszweck gelten, wenn dieses System aufgrund dieser Integration in der Lage ist, einer Vielzahl von Zwecken zu dienen. "Ein KI-System mit allgemeinem Verwendungszweck kann direkt eingesetzt oder in andere KI-Systeme integriert werden.“

Was ein KI-Modell ist, wird in Art. 3 AI Act nicht definiert. Aktuell stehen insbesondere Large-Language-Modelle (LLM) im Fokus. Sie fallen in den Bereich der Generativen-KI, einer Teildisziplin der KI, bei der beispielsweise Texte, Bilder oder Software-Code erzeugt werden. LLM zählen zu den Foundation-Modellen (FM), die im Allgemeinen durch selbstüberwachtes Lernen (Self-supervised Learning) erstellt werden. FM bieten den Vorzug, dass sie im Unterschied zu klassischen KI-Modellen direkt genutzt werden können und nicht erst langwierig mit großen Datenmengen trainiert werden müssen. Foundation-Modellen können nicht nur genutzt werden, um Texte oder Bilder zu erzeugen. Auch für Deep Learning und allgemein zur Erstellung von KI-Modellen lassen sie sich einsetzen. Sie erlauben einen schnellen Start und können die Anfangsphase von KI-Projekten verkürzen.

Vereinfacht lässt sich ein KI-Modell als ein Programm beschreiben, das anhand einer bestimmten Menge von Daten dazu trainiert wurde, Muster zu erkennen oder Entscheidungen ohne weiteres menschliches Eingreifen zu treffen.

Abschließend noch einige Beispiele für Sprachmodelle (LLMs):

- ChatGPT von Open AI (Microsoft)
- DALL-E von OpenAI
- Gemini von Google
- Imagen von Google
- Jasper von Jasper AI
- Claude von Anthropic
- Llama von Meta (Open Source)
- Granite von IBM (Open Source)
- Mistral von Mistral AI (Open Source)
- Stable Diffusion von Stability

Der AI Act definiert ein „KI-System mit allgemeinem Verwendungszweck“ in Art. 3 Nr. 66 AI Act als „ein KI-System, das auf einem KI-Modell mit allgemeinem Verwendungszweck beruht und in der Lage ist, einer Vielzahl von Zwecken sowohl für die direkte Verwendung als auch für die Integration in andere KI-Systeme zu dienen“.

Die Regelungen im Kapitel V des AI Act zu den „KI-Modellen mit allgemeinem Verwendungszweck“ erfassen im Schwerpunkt nur „KI-Modelle mit allgemeinem Verwendungszweck“ mit einem sogenannten systemischen Risiko. Ein „systemisches Risiko“ ist nach der Definition in Art. 3 Nr. 65 AI Act „ein Risiko, das für die Fähigkeiten mit hoher Wirkkraft von KI-Modellen mit allgemeinem Verwendungszweck spezifisch ist und aufgrund deren Reichweite oder aufgrund tatsächlicher oder vernünftigerweise vorhersehbarer negativer Folgen für die öffentliche Gesundheit, die Sicherheit, die öffentliche Sicherheit, die Grundrechte oder die Gesellschaft insgesamt erhebliche Auswirkungen auf den

Unionsmarkt hat, die sich in großem Umfang über die gesamte Wertschöpfungskette hinweg verbreiten können“.

Der AI Act nimmt auch insoweit eine autonome Begriffsbestimmung vor. Dies erfordert es, die Anwendung des Gesetzes hieran auszurichten und die bisherigen Technologien anhand dieser Begriffsbestimmungen einzuordnen.

2.3 Was ist die Risiko-Taxonomie des AI Act und was ist die Konsequenz?

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Review:
Jamal Lammert, Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Die Risikotaxonomie wird durch die Europäische Kommission[10] wie folgt dargestellt. Die Pyramide soll wohl die Einschätzung der EU-Kommission zum Umfang der Anwendung der Regelungen des jeweiligen Risikobereichs zum Ausdruck bringen:

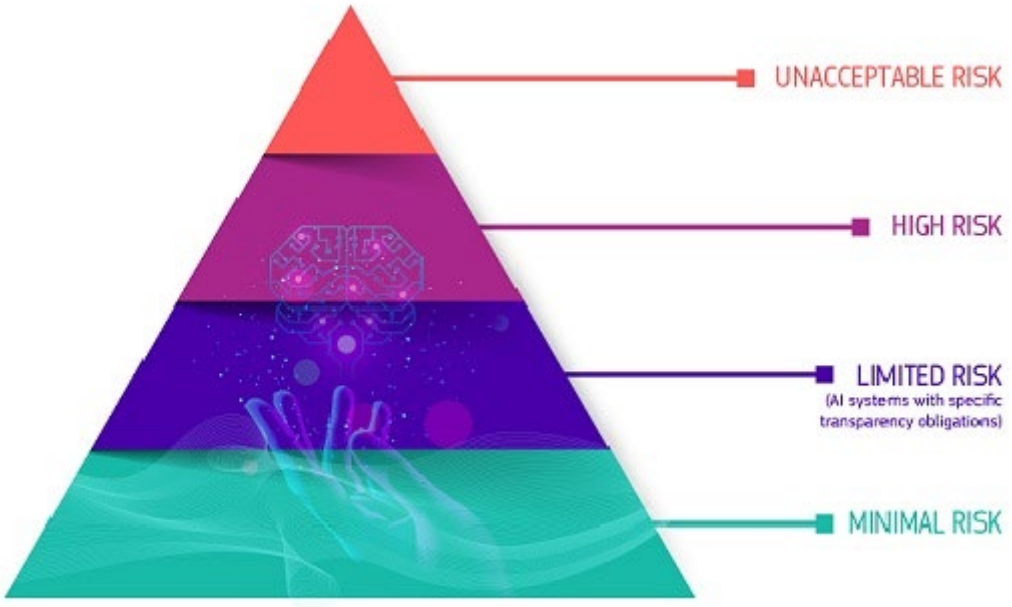

Abb. 2: Risikoeinstufung der EU-Kommission zum AI Act

[10] "A risk-based approach", Europäische Kommission, https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai, 22.09.2024

Regelungsmechanik des AI Act

Verbotene KI-Systeme	Art. 5 AI Act (Verboten Praktiken)	bspw. unterschwellige Technologien zur (negativen) Beeinflussung	
Hochrisiko KI-Systeme	Festlegung anhand Art. 6 mit Anhängen I und III Regelungen in Art. 6 ff. AI Act	bspw. Personalmanagement, Kreditwürdigkeit, Vertragsabschlüsse,	Hauptregelungs-Gegenstand des AI Act
KI-Modelle mit allgemeinem Verwendungszweck (General Purpose AI)	Regelungen in Kapitel V „außerhalb" der (im Übrigen) Risikotaxonomie	bspw. Modelle wie GPT-4	Erfassung im Gesetzgebungsverfahren umstritten
geringes Risiko KI-System	Art. 50 AI Act	bspw. Chat-Bots	
Minimales Risiko KI-System	Aufsicht nach AI Act	Alle anderen KI-Systeme	

Es werden im AI Act vier Risikostufen unterschieden:[11]

1. Inakzeptables / Unannehmbares Risiko: KI-Systeme, die hierunter fallen, sind in der EU seit dem 02.02.2025 verboten.
2. Hohes Risiko: Derart klassifizierte Systeme obliegen einer Vielzahl von Auflagen, die erfüllt werden müssen
3. Begrenztes Risiko: Solche Systeme unterliegen Transparenzpflichten
4. Minimales oder kein Risiko

Im Nachfolgenden werden diese vier Stufen anhand von Beispielen konkretisiert. Zudem wird erläutert, was die Einordnung für KI-Systeme bedeutet, d.h. welche Anforderungen sie auf der jeweiligen Stufe erfüllen müssen (sofern sie nicht zu den verbotenen Systemen zählen).

[11] EU AI Act: https://artificialintelligenceact.eu/de/high-level-summary/

2.3.1 Inakzeptables / Unannehmbares Risiko

KI-Systeme dieser Klasse werden allgemein als zu gefährlich für die Grundwerte und Grundrechte innerhalb der EU angesehen. Art. 5 regelt diese sog. Verbotenen Praktiken und verbietet sie. Rein formal wird nicht das KI-System verboten, sondern die Nutzung von KI-Systemen für solche Praktiken („Folgende Praktiken im KI-Bereich sind verboten:“, Art. 5 AI Act). Diese Verbote gelten ab dem 02.02.2025 (siehe auch unten Ziffer 2.6).

Nachfolgend erfolgt ein Überblick über die adressierten Themen, wobei es für das Verbot jedoch auf die konkrete(!) Formulierung der Praktiken und der Ausnahmen ankommt:

- Verwendung eines KI-Systems, das Techniken der unterschwelligen Beeinflussung außerhalb des Bewusstseins einer Person oder absichtlich manipulative oder täuschende Techniken mit dem Ziel oder der Wirkung einsetzt, das Verhalten einer Person oder einer Gruppe von Personen wesentlich zu verändern
- Verwendung eines KI-Systems, das eine Vulnerabilität oder Schutzbedürftigkeit einer natürlichen Person oder einer bestimmten Gruppe von Personen aufgrund ihres Alters, einer Behinderung oder einer bestimmten sozialen oder wirtschaftlichen Situation mit dem Ziel oder der Wirkung ausnutzt, das Verhalten dieser Person oder einer dieser Gruppe angehörenden Person in einer Weise wesentlich zu verändern
- KI-Systemen zur Bewertung oder Einstufung von natürlichen Personen oder Gruppen von Personen über einen bestimmten Zeitraum auf der Grundlage ihres sozialen Verhaltens oder bekannter, abgeleiteter oder vorhergesagter persönlicher Eigenschaften oder Persönlichkeitsmerkmale
- KI-Systems zur Durchführung von Risikobewertungen in Bezug auf natürliche Personen, um das Risiko, dass eine natürliche Person eine Straftat begeht,
- KI-Systemen, die Datenbanken zur Gesichtserkennung durch das gezielte Auslesen von Gesichtsbildern aus dem Internet oder von Überwachungsaufnahmen erstellen oder erweitern
- KI-Systemen zur Ableitung von Emotionen einer natürlichen Person am Arbeitsplatz und in Bildungseinrichtungen
- Systemen zur biometrischen Kategorisierung, mit denen natürliche Personen individuell auf der Grundlage ihrer biometrischen Daten kategorisiert werden, um ihre Rasse, ihre politischen Einstellungen, ihre Gewerkschaftszugehörigkeit, ihre religiösen oder weltanschaulichen Überzeugungen, ihr Sexualleben oder ihre sexuelle Ausrichtung zu erschließen oder abzuleiten
- biometrischer Echtzeit-Fernidentifizierungssysteme in öffentlich zugänglichen Räumen zu Strafverfolgungszwecken

2.3.2 Hochrisiko-KI

Das Kapitel III des AI Act regelt die sog. Hochrisiko KI-Systeme.

Art. 6 AI Act regelt im Zusammenspiel mit den Anhängen I und III, in welchen Konstellationen ein KI-System als hochrisikoreich eingeordnet wird.

Die Regelung Art. 6 Abs. 1 AI Act stellt – vereinfacht – darauf ab, ob das KI-System

- als Sicherheitsbauteil eines unter die in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union fallenden Produkts verwendet werden oder das KI-System ist selbst ein solches Produkt; oder
- das Produkt, dessen Sicherheitsbauteil gemäß Buchstabe a das KI-System ist, oder das KI-System selbst als Produkt muss einer Konformitätsbewertung durch Dritte im Hinblick auf das Inverkehrbringen oder die Inbetriebnahme dieses Produkts gemäß den in Anhang I aufgeführten Harmonisierungsrechtsvorschriften der Union unterzogen werden.

Diese Regelung gilt – anders als Art. 6 Abs. 2 AI Act (siehe nachfolgend) - erst ab dem 02.08.2027 (Art. 113 AI Act).

Art. 6 Abs. 2 AI Act bezieht sich auf die in Anhang II des AI Act genannten Einsatzbereiche. Die folgenden Themen werden adressiert, wobei es für ein Verbot jedoch auf die konkrete(!) Formulierung der Praktiken und der Ausnahmen ankommt:

- Biometrische Informationssysteme
- Kritische Infrastruktur
- Allgemeine und berufliche Bildung
- Beschäftigung, Personalmanagement und Zugang zur Selbstständigkeit
- Zugänglichkeit und Inanspruchnahme grundlegender privater und grundlegender öffentlicher Dienste und Leistungen:
- Strafverfolgung,
- Migration, Asyl und Grenzkontrolle
- Rechtspflege und demokratische Prozesse

Art. 6 Abs. 2 AI Act gilt ab dem 2. August 2026 (Art. 113 AI Act).

KI-Systeme dieser Stufe dürfen nur, wenn strenge Pflichten und Anforderungen eingehalten werden, entwickelt und betrieben werden. Hier liegt der inhaltliche Regelungsschwerpunkt des AI Act.

Im Kern geht es bei den Pflichten um eine regulatorisch kontinuierliche Nachweisbarkeit der Entwicklung und des Betriebs der KI-Anwendung. Dies bedeutet, dass hierfür folgende Anforderungen erfüllt sein müssen:

1. Einführung eines Risiko- und Qualitätsmanagement Systems
2. Erstellung einer technischen Dokumentation
3. Durchführung von Konformitätsbewertungen
4. Registrierung in einer EU-Datenbank
5. Kontinuierliche Überwachung und kontinuierliche Berichterstattung

Um einen Eindruck hierfür zu vermitteln: Anbieter haben insbesondere die Regelungen in Art. 8 bis 22 AI Act zu erfüllen. Betreiber haben die Anforderungen der 12 Absätze des Art. 26 AI Act zu erfüllen. Das macht auch deutlich, dass die Anbieter und Betreiber die Primäradressaten der Regelungen sind.

2.3.3 Begrenztes Risiko

Art. 50 AI Act sieht für KI-Systeme in bestimmten Anwendungsbereichen besondere Transparenzpflichten vor. Aufgrund dieser Nutzung der KI-Systeme geht der Gesetzgeber von einem erhöhten Risiko im Vergleich zu sonstigen Verwendungen von KI-Systemen, aber nicht von einem hohen Risiko wie nach Art. 6 AI Act aus. Diese Regelungen gelten ab dem 02. August 2026 (siehe unten Ziffer 2.6).

Art. 50 Abs. 1 bis 4 AI Act legen die Anwendungsbereiche fest, wobei Art. 50 AI Act weitgehend die Modalitäten zur Erfüllung der Transparenzpflichten regeln:

- „Die Anbieter stellen sicher, dass KI-Systeme, die für die direkte Interaktion mit natürlichen Personen bestimmt sind, so konzipiert und entwickelt werden, dass […]“ (Art. 50 Abs. 1 AI Act)
- „Anbieter von KI-Systemen, einschließlich KI-Systemen mit allgemeinem Verwendungszweck, die synthetische Audio-, Bild-, Video- oder Textinhalte erzeugen, stellen sicher, dass […]“ (Art. 50 Abs. 2 AI Act)
- „Die Betreiber eines Emotionserkennungssystems oder eines Systems zur biometrischen Kategorisierung informieren die davon betroffenen natürlichen Personen […]“ (Art. 50 Abs. 3 AI Act)
- „Betreiber eines KI-Systems, das Bild-, Ton- oder Videoinhalte erzeugt oder manipuliert, die ein Deep Fake sind, müssen offenlegen, dass […]“ (Art. 50 Abs. 4 UAbs. 1 AI Act)
- Betreiber eines KI-Systems, das Text erzeugt oder manipuliert, der veröffentlicht wird, um die Öffentlichkeit über Angelegenheiten von öffentlichem Interesse zu informieren, […]“ (Art. 50 Abs. 4 UAbs. 2 AI Act)

2.3.4 Minimales oder kein Risiko

Alle anderen KI-Systeme werden ebenfalls durch den AI Act erfasst und unterstehen damit insbesondere der behördlichen Aufsicht nach dem AI Act. Für diese sind jedoch nicht umfassend ausgeprägte Regelungen im AI Act enthalten.

2.4 Die unterschiedlichen Verpflichteten und Akteure des AI Acts?

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |
Kai Meinke, deltaDAO AG |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Der EU AI Act definiert verschiedene Akteure mit unterschiedlichen Verpflichtungen im Zusammenhang mit KI-Systemen. Der Oberbegriff für die durch den AI Act Verpflichteten ist „Akteur“. Der Begriff „Akteur“ umfasst nach Art. 3 Nr. 8 AI Act Anbieter, Produkthersteller, Betreiber, Bevollmächtigten, Einführer und Händler. Diese sind wiederum in Art. 3 Nr. 3, 4, 5, 6 und 7 AI Act definiert. Der Begriff Produkthersteller ist nicht in Art. 3 AI Act definiert.

Die wichtigsten Verpflichteten sind der Anbieter und der Betreiber von KI-Systemen. Genauer:

- „Anbieter“ ist 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. (Art. 3 Nr. 3 AI Act)
- „Betreiber“ ist 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 (Art. 3 Nr. 4 AI Act)

Die genaue Beachtung des Wortlauts der Definitionen ist entscheidend für das Rollenverständnis und damit auch für den Umfang der Pflichten. Diese sind bspw. im Bereich der Hochrisiko-KI stark unterschiedlich ausgeprägt. Auch die Regelungen für begrenzte Risiken in Art. 50 AI Act unterscheidet zwischen diesen.

Von besonderer Bedeutung ist das Risiko des Rollenwechsels zum Anbieter. Akteure, die originär nicht Anbieter des KI-Systems sind, werden durch bestimmte Verhaltensweisen zu Anbietern und „rutschen“ in deren Pflichtenkatalog.

Für Hochrisiko-KI-Systeme regelt Art. 25 AI Act den Rollenwechsel wie folgt. Nach Art. 25 Abs 1 AI Act gelten Händler, Einführer, Betreiber oder sonstige Dritte als Anbieter eines Hochrisiko- KI-Systems für die Zwecke dieser Verordnung und unterliegen den Anbieterpflichten gemäß Artikel 16

- wenn sie ein bereits in Verkehr gebrachtes oder in Betrieb genommenes Hochrisiko-KI-System mit ihrem Namen oder ihrer Handelsmarke versehen, unbeschadet vertraglicher Vereinbarungen, die eine andere Aufteilung der Pflichten vorsehen;
- wenn sie eine wesentliche Veränderung eines Hochrisiko-KI-Systems, das bereits in Verkehr gebracht oder in Betrieb genommen wurde, so vornehmen, dass es weiterhin ein Hochrisiko-KI-System gemäß Artikel 6 bleibt;

- wenn sie die Zweckbestimmung eines KI-Systems, einschließlich eines KI-Systems mit allgemeinem Verwendungszweck, das nicht als hochriskant eingestuft wurde und bereits in Verkehr gebracht oder in Betrieb genommen wurde, so verändern, dass das betreffende KI-System zu einem Hochrisiko-KI-System im Sinne von Artikel 6 wird.

Im Unternehmen muss durch geeignete Maßnahmen sichergestellt werden, dass es nicht unbeabsichtigt zu einem solchen Rollenwechsel kommt! Diese Klarstellung kann auch ein Bestandteil der nach Art. 4 AI Act jeder mitarbeitenden Person zu vermittelnden KI-Kompetenz (siehe nachfolgend) sein.

Für **Anbieter** und **Betreiber** besteht unabhängig von der Risikotaxonomie ab dem 02.02.2025 (siehe unten Ziffer 2.6) die Pflicht zur Sicherstellung der **KI-Kompetenz nach Art. 4 AI Act**: 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.

Einführer und **Händler** unterliegen Sorgfaltspflichten und müssen ebenfalls die Konformität und Kennzeichnung von Systemen prüfen, bevor Systeme in den Verkehr gebracht werden. Damit gemeint sind alle Unternehmen oder Personen, welche KI-Anwendungen von außerhalb der EU in den Verkehr bringen oder betreiben.

Alle europäischen Mitgliedsstaaten benennen **zuständige nationale Behörden**[12], die für die Überwachung und Durchsetzung des AI Act verantwortlich sind. Bis zum 2. August 2025 müssen die zuständigen nationalen Behörden benannt sein.

Das KI-Büro (AI-Office) der Europäischen Kommission[13], mit rund 140 Mitarbeitenden, soll die Regeln auf europäischer Ebene durchsetzen. Das KI-Büro besteht aus fünf Referaten und zwei Beratenden:

- Referat „Exzellenz in KI und Robotik“
- Referat „Regulierung und Einhaltung“
- Referat „KI-Sicherheit“
- Referat „KI-Innovation und Politikkoordinierung“
- Referat „KI für das Gemeinwohl“
- Leitende wissenschaftliche Beratung
- Beratung für internationale Angelegenheiten

Das KI-Büro hat weitreichende Vollmachten zur Untersuchung und Sanktionierung von Verstößen gegen die Regulierung und ist verantwortlich für die Risikoeinstufung von KI-Anwendungen.

[12] “KI-Regulierung: Nationale Aufsicht muss verbraucherfreundlich sein”, Bundesverband Verbraucherzentrale, https://www.vzbv.de/publikationen/ki-regulierung-nationale-aufsicht-muss-verbraucherfreundlich-sein, 22.09.2024

[13] “Europäisches Amt für künstliche Intelligenz”, Europäische Kommission, https://digital-strategy.ec.europa.eu/de/policies/ai-office, 22.09.2024

Bevollmächtigte von Anbietern von KI-System werden durch KI-Anbietende in Drittländern benannt und dienen als Kontaktpunkte für europäische Behörden. Sie müssen mit entsprechenden Kompetenzen und Mitteln ausgestattet werden, um Ihre Aufgaben erfüllen zu können.

2.5 Darstellung von Ziffer 2.1 bis 2.3

Autoren:
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |
Kai Meinke, deltaDAO AG |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Der AI Act enthält Regelungen für verschiedene Konstellationen. Nicht alle Regelungen gelten für alle Konstellationen. Die Anwendbarkeit hängt sowohl von der Risikoklassifizierung als auch von der Rolle im Sinne des AI Act ab. Daher ist es entscheidend, die relevanten Regelungen zu identifizieren. Wir schlagen dazu folgendes Vorgehen vor:

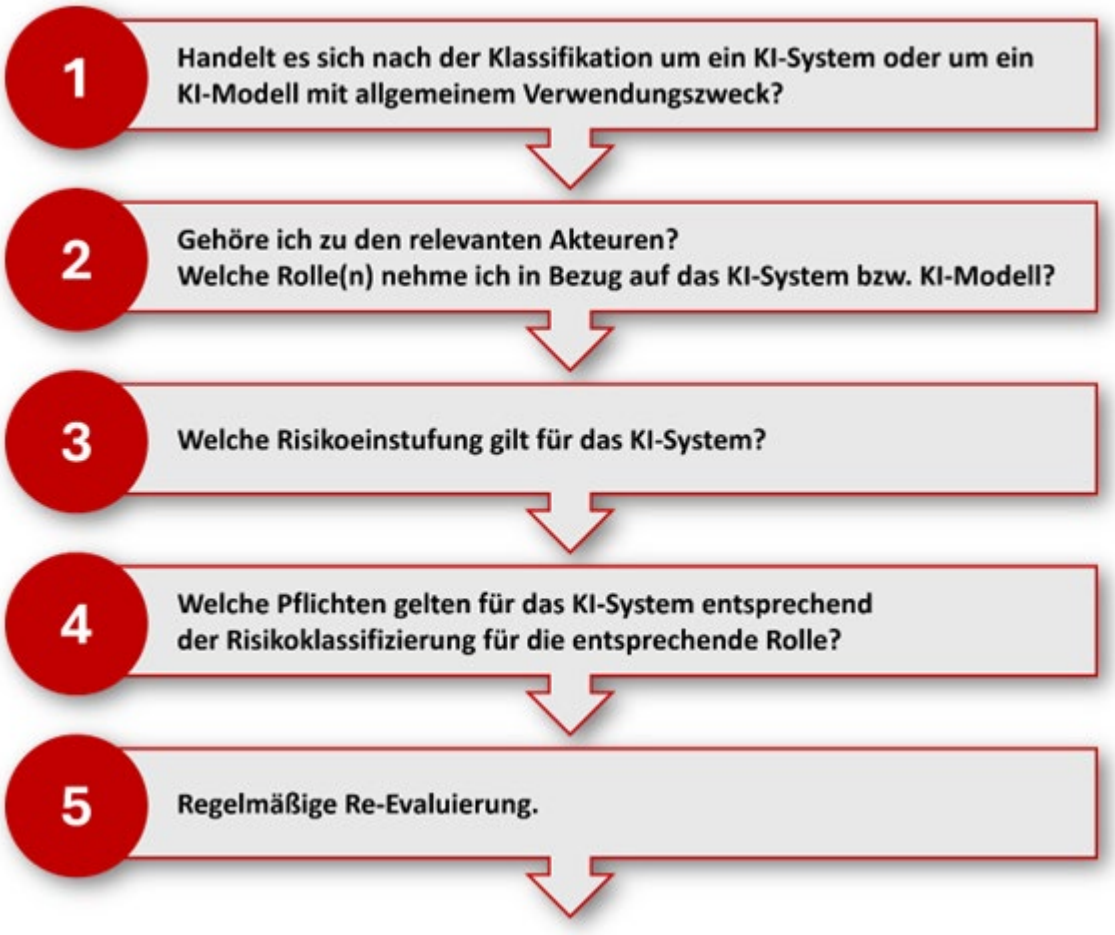

Abb. 4: Identifizierung der relevanten Regelung

2.6 Zeitlicher Anwendungsbeginn des AI Act zwischen 02.02.2025 und 02.08.2027

Der AI Act ist Mitte 2024 in Kraft getreten. Art. 113 AI Act regelt jedoch unterschiedliche Zeitpunkte zur Anwendung der Regelungen des AI Act. Damit soll den Verpflichteten die Zeit gegeben werden, die Anforderungen des AI Acts bis zum Anwendungsbeginn der Regelungen umzusetzen.

Die praktische Relevanz besteht nicht zuletzt darin, dass erst ab dem Anwendungsbeginn der jeweiligen Regelungen Geldbußen wegen ihrer Missachtung verhängt werden dürfen. Grundsätzlich gilt dasselbe für die zivilrechtliche Haftung (bspw. auf Schadensersatz), wobei insoweit Vorsicht und eine differenziertere Betrachtung geboten ist.

Anwendungsbeginn	Regelungen des AI Act
ab 02.08.2026, aber:	Genereller Anwendungsbeginn des AI Act, **aber:**
ab 02.02.2025	Artt. 1 bis 5 AI Act – also insbesondere auch: • Art. 4 AI Act: KI Kompetenz (siehe Ziffer 2.4) • Art. 5 AI Act: Verbotene Praktiken (siehe Ziffer 2.3)
ab 02.08.2025:	Kapitel III Abschnitt 4, Kapitel V, Kapitel VII und Kapitel XII sowie Artikel 78, mit Ausnahme des Artikels 101 – also insbesondere: • Regelungen zu KI-Modellen mit allgemeinem Verwendungszweck (Kapitel V)
02.08.2026	Genereller Anwendungsbeginn, insbesondere: • Hochrisiko-KI im Sinne des Art. 6 Abs. 2 AI Act (siehe Ziffer 2.3) • KI mit geringem Risiko (Art. 50 AI Act) (siehe Ziffer 2.3)
02.08.2027	Hochrisiko-KI im Sinne des Art. 6 Abs. 1 AI Act (siehe Ziffer 2.3) und die entsprechenden Pflichten gemäß des AI Act
Achtung – Stichwort „Bestandsschutz“: Weitere Übergangsregelung in Art. 111 AI Act für bereits in Verkehr gebrachte oder in Betrieb genommene KI-Systeme und bereits in Verkehr gebrachte KI-Modelle mit allgemeinem Verwendungszweck.	

3. „Governance-Gedankenmodell“ anhand IBM watsonx.governance

Autoren:
Stephan Schnieber, IBM Deutschland GmbH |
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Die zuvor beschriebenen Regelungen des EUAI Act und die Anforderungen, die daraus für die Erstellung und den Betrieb von KI-Systemen folgen, haben IBM dazu veranlasst, ein Lösungsangebot für AI Governance zu entwickeln. IBM bietet dazu das offene Produkt watsonx.governance an. Mit dessen Hilfe ist nicht nur eine AI Governance für die eigene KI-Entwicklungsplattform watsonx.ai möglich, sondern auch für KI-Entwicklungsumgebungen anderer Anbieter (zum Beispiel für AWS Sagemaker).

Die nachfolgenden Diagramme orientieren sich an den Pflichten des Anbieters einer Hochrisiko-KI (Art. 6 ff. AI Act)(zu den Definitionen siehe oben). Denn dies ist der umfangreichste Regelungskomplex des AI Act. Die Regelung für KI-Modelle (Kapitel V des AI Act) mit allgemeinem Verwendungszweck und für KI mit geringem Risiko (Art. 50 AI Act) werden in weiteren Papieren beleuchtet werden.

Die folgende Grafik stellt den allgemeinen Zyklus von Erstellung und Betrieb von KI-Änderungen dar:

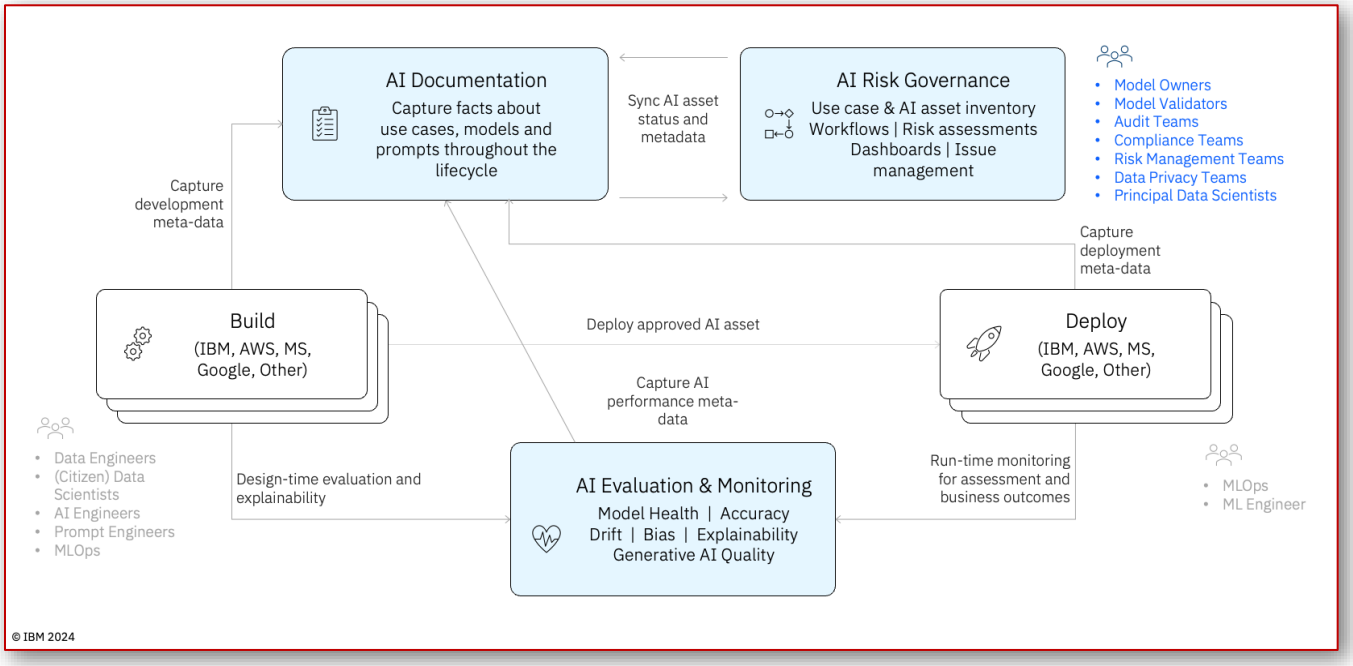

Abb. 5: The general cycle of creation and operation of AI modifications © IBM 2024

Im nachfolgenden Bild ist ein Mapping der Vorgaben des EU-AI Act auf die einzelnen Phasen und Schritte vorgenommen worden:

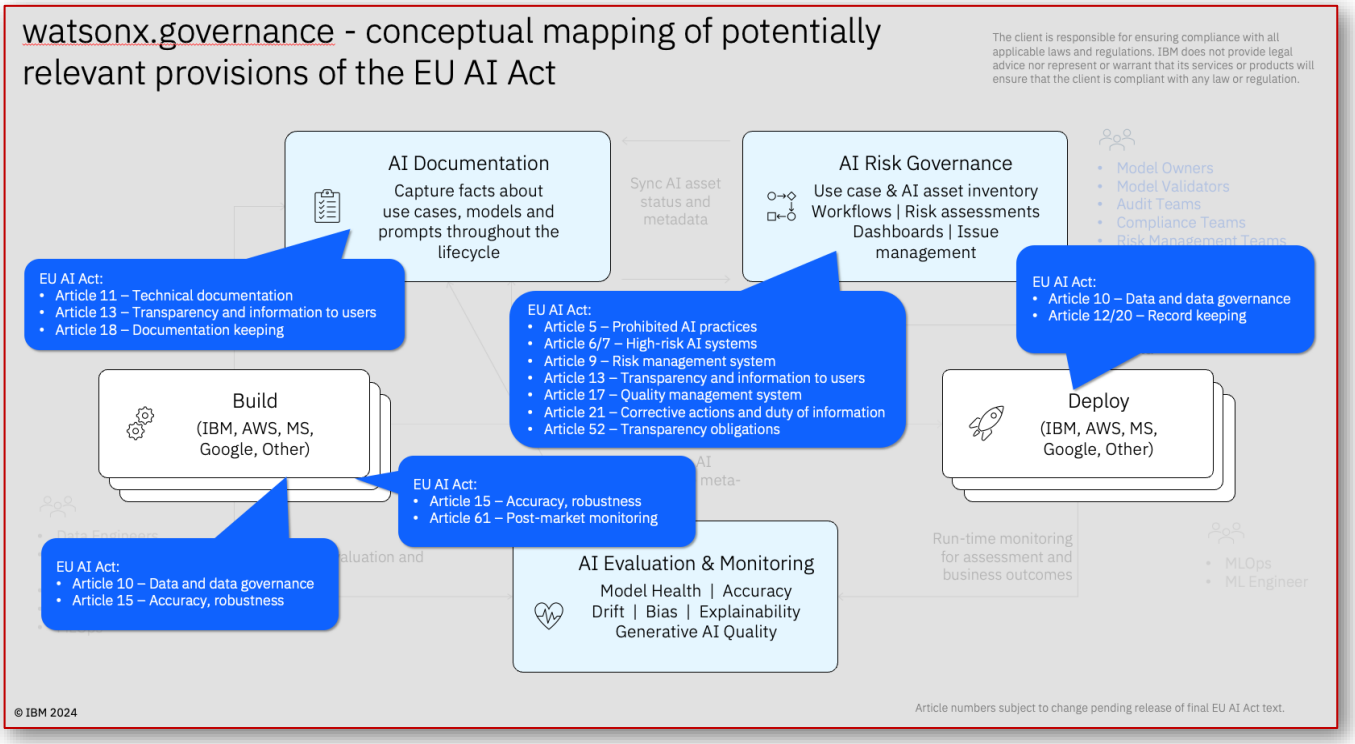

Abb. 6: watsonx.governance - conceptual mapping of potentially relevant provisions of the AI Act © IBM 2024

Das dargestellte Mapping ist nur exemplarisch und dient der Orientierung. Dies bedeutet, dass weder die gesamte Funktionalität noch alle Regelungen des AI-Act umfänglich dargestellt sind. Zu beachten bleibt, dass dieses Angebot von IBM als Beispiel einer Unterstützung im Kontext der Nutzung von KI im Unternehmen zu sehen ist. Ob sich damit alle Anforderungen des EU-AI-Act und die des jeweiligen Unternehmens abdecken lassen, muss von Fall zu Fall evaluiert werden. Weitere Details sind in einem AI Risiko Atlas[14] abgebildet.

Die Pflichten des Betreibers einer Hochrisiko-KI (zu den Definitionen siehe oben) ergeben sich im Kern aus den 12 Absätzen des Art. 26 AI Act. Diese knüpfen an die Pflichten des Anbieters der Hochrisiko-KI bzw. bauen darauf auf. Dies gilt beispielsweise für die menschliche Aufsicht und die bestimmungsgemäße Verwendung. Daher erscheint es auch für Betreiber sinnvoll, sich (zumindest in der Übersicht) mit den Pflichten des Anbieters zu befassen.

[14] https://ibm.biz/AI-RiskAtlas

4. Etablierte Begriffe und ihr Bezug zum AI Act

Autoren:
Dr. Jens Eckhardt, pitc legal Eckhardt Rechtsanwälte Partnerschaft mbB |
Christine Neubauer, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Review:
Michael Hase, eco - Verband der Internetwirtschaft e.V. |
Christian Weber, GasLINE CP Customer Projects GmbH |

Die gängigen Begriffe, die in der KI-Praxis benutzt werden, finden sich in den Definitionen des AI Act oftmals nicht wieder. Denn die EU-Verordnung greift diese Begriffe kaum auf, sondern verwendet ihre eigene Terminologie. In der Praxis müssen die Akteure ein „Matching“ zwischen den Begriffen aus dem technologischen Kontext und denen aus dem rechtlichen Kontext vornehmen. Eine sinnvolle Hilfestellung wäre es daher, die Bezüge zwischen den wichtigsten technologischen Termini und den maßgeblichen Begriffen des AI Act herzustellen.

Eine solche Zuordnung ist aber nicht ganz einfach, da sich die Begriffe nicht immer eins zu eins „übersetzen“ lassen. Dieser Aufgabe wollen wir uns noch im ersten Quartal 2025 auf unserer Website unter (https://www.eco.de/ki-in-der-praxis/rahmenbedingungen/begriffe/) widmen. Als Basis dafür haben wir nachfolgend ein Glossar erstellt. Es greift Begriffe der „KI-Welt“ auf und erklärt sie anhand frei verfügbarer Quellen. Für diese Begriffe soll in einem nächsten Schritt der Bezug zur Terminologie des AI Act hergestellt werden.

Online finden Sie die gebräuchlichen Begriffe und die Liste lässt sich auch erweitern. Teilen Sie uns hier Ihre Vorschläge mit: christine.neubauer@eco.de

Diese Begriffe und deren Erläuterung sind relevant, um die technologischen Aspekte von KI zu verstehen. Sie sind aber nicht maßgeblich für die Auslegung des AI Act. Wie mehrfach ausgeführt, hat die EU-Verordnung eine autonome Begriffswelt, und die ist für die Rechtsanwendung allein verbindlich. Daraus folgt zugleich, dass vertraute technologische Begriffe nicht einfach in den AI Act „hineingelesen“ werden dürfen.

Bitte beachten Sie: In der Praxis weicht das Verständnis ein und desselben Begriffs nicht selten zwischen verschiedenen Quellen voneinander ab. Auch wenn zu jeder Begriffserklärung die jeweilige Quelle genannt wird, ist sie nicht allgemeinverbindlich.

Begriff	Quelle	Definition
Artificial Intelligence bzw. Künstliche Intelligenz	Lernende Systeme	Verordnung der Europäischen Union (EU), die Regeln für die Entwicklung und den Einsatz von KI-Systemen festlegt und einen einheitlichen, verbindlichen Rechtsrahmen schafft. Ziel ist es, vertrauenswürdige KI zu gestalten, die entsprechend den europäischen Wertvorstellungen eingesetzt wird. So sollen KI-Systeme, die in der EU verwendet werden, sicher, transparent, ethisch, unparteiisch und unter menschlicher Kontrolle sein. Gleichzeitig verfolgt die Verordnung den Anspruch, KI-Technik und Forschung innerhalb der EU wettbewerbsfähig zu halten und Innovationen zu ermöglichen. Der 2024 verabschiedete AI Act ist das weltweit erste transnationale KI-Regelwerk.
Basis Modelle (Foundation Models)	Lernende Systeme	Modelle des maschinellen Lernens, die mit Hilfe von Deep Learning (z. B. Transformer) auf umfangreichen Datensätzen (z.B. Text, Bilder, Videos) aus Internet, Social Media oder anderen Quellen vortrainiert wurden. Hierfür benötigen Basismodelle derzeit eine hohe Rechenleistung. Charakteristisch für Basismodelle sind u.a. ihre Wiederverwendbarkeit und Anpassungsfähigkeit für spezifische Aufgaben und Domänen (siehe: Finetuning). Ein weiteres Kennzeichen ist, dass nach dem Trainieren Fähigkeiten des Modells entdeckt werden, an die zuvor nicht explizit gedacht wurde (sog. Emergenz). Basismodelle können sowohl Sprachmodelle als auch multimodale Modelle sein. Basismodelle bilden die Grundlage für viele Anwendungen, die auf generativer KI beruhen, wie Chatbots oder Bild- und Videogeneratoren. Bekannte Basismodelle sind etwa GPT-4 (Open AI), Llama (Meta) oder Gemini (Google). Der Begriff "Foundation Models" wurde 2021 vom Stanford Institute for Human-Centered Artificial Intelligence's (HAI) geprägt.
Basismodell	ZVKI	KI-Modelle, die für viele verschiedene Aufgaben infrage kommen. Basismodelle (Foundation Models) bilden die Grundlage – die Basis – für viele weitere KI-Anwendungen. Sie zeichnen sich vor allem dadurch aus, dass sie eine Vielzahl von verschiedenen Aufgaben lösen können, beispielsweise Texte oder Bilder erstellen, Musik erzeugen oder Programmiercode ausgeben. Solche Anwendungen werden auch als „General Purpose AI" bezeichnet. ChatGPT basiert zum Beispiel auf einem Foundation Model. Damit die KI-Anwendungen gut funktionieren, müssen sie von Menschen feinjustiert werden. Bei Programmen zur Texterstellung bewerten Menschen beispielweise, ob eine Aussage Sinn ergibt oder nicht. Basismodelle können immense Datenmengen und Parameter – Knotenpunkte für Gewichtungen von Daten – in ihre Berechnungen einbeziehen. Sie verarbeiten unter anderem Bilder, Videos und Texte. In Abgrenzung zu Basismodellen erfüllen bisher gängige KI-Systeme eine bestimmte und eindeutig vorgegebene Aufgabe und können nicht einfach angepasst werden. Bei Foundation Models handelt es sich dennoch NICHT um sogenannte starke KI. Ihre Einsatzmöglichkeiten sind zwar vielfältig, aber weiterhin klar begrenzt.
bias	ZVKI	Verzerrung, Vorurteil oder Voreingenommenheit. Im Zusammenhang mit Künstlicher Intelligenz meint der englische Begriff bias, dass das Ergebnis bzw. die Ausgabe einer KI-Anwendung verzerrt ist. Die Ursachen hierfür sind vielfältig. Beispielsweise können subjektive Sichtweisen von Entwickler:innen oder die generellen strukturellen Ungleichheiten unserer Gesellschaft (beabsichtigt oder unbeabsichtigt) in die Programmierung von Algorithmen einfließen. So können etwa die ausgewählten Trainingsdaten fehlerhaft und/ oder unvollständig sein, weil bestimmte (Bevölkerungs-)Gruppen unterrepräsentiert sind. Die sich darin spiegelnden Diskriminierungen und Rassismen reproduzieren das algorithmische System. Die Beispiele für algorithmenvermittelte Diskriminierung sind zahlreich. So hat unter anderem Google ein KI-Modell zur Bilderkennung entwickelt, das Schwarze Menschen deutlich schlechter erkannte als weiße Menschen. Ein anderes Beispiel ist ein von Amazon entwickeltes KI-Modell, das dabei helfen sollte, geeignete Bewerber:innen für offene Stellen auszuwählen. Das Programm benachteiligte systematisch Frauen, weil Amazon in der Vergangenheit vor allem Männer eingestellt hatte und sich dies in den Trainingsdaten für das System widerspiegelte.
Chat Bot	Lernende Systeme	Virtuelle Dialogsysteme, die zunehmend im Kundenservice und für Benutzerschnittstellen im Allgemeinen eingesetzt werden. Über eine Textein- und Textausgabemaske (z.B. ein Dialogfenster auf einer Website) kommunizieren sie in natürlicher Sprache mit dem Menschen. Durch Methoden des maschinellen Lernens können Chatbots aus Eingaben ständig dazu lernen – um etwa die Stimmlage des Menschen zu interpretieren oder personalisierte Antworten zu geben.
Chatbot	ZVKI	Computerprogramme, die per Text- oder Spracheingabe Daten verarbeiten. Nutzer*innen können mit einem Chatbot interagieren, indem sie Eingaben tätigen – zum Beispiel, indem sie schriftliche oder mündliche Fragen stellen – und im Gegenzug Antworten bzw. Aussagen des Systems lesen oder hören. Chatbots werden beispielsweise in der Kund*innenbetreuung eingesetzt, in der sie mitunter helfen, Fragen gezielt an die entsprechenden Mitarbeiter:innen weiterzuleiten. Sprachbasierte Chatbots sind digitale Assistenten wie Siri, Alexa und Google Assistant. Dank stetiger Fortschritte beim natural language processing nimmt die Qualität von Chatbots rasant zu. Den Turing-Test bestehen sie aber noch nicht.

Deep Fake	ZVKI	Realistisch anmutende, aber künstlich erzeugte bzw. manipulierte Foto-, Video- oder Tonaufnahme.Der Begriff setzt sich aus Deep Learning und Fake (auf Deutsch: Fälschung) zusammen. Erstellt werden Deep Fakes mithilfe von künstlichen neuronalen Netzen. Dafür werden beispielsweise Videoaufnahmen einer Person analysiert, um ihre Gesichtsbewegungen simulieren zu können. Mit natural language processing können auch Aussagen stimmlich nachgebildet werden. Auf diese Weise kann man Menschen Dinge sagen lassen, die sie in Wirklichkeit nie von sich gegeben haben. Deep Fakes können ein Mittel sein, um gezielt Desinformationen zu verbreiten oder jemanden in ein schlechtes Licht zu rücken. Für ein überzeugendes Ergebnis sind allerdings sehr viele Bild- oder Tonaufnahmen des Opfers erforderlich.
Deep Learning	Lernende Systeme	Methode des maschinellen Lernens in künstlichen neuronalen Netzen. Diese umfassen mehrere Schichten – typischerweise eine Eingabe- und Ausgabeschicht sowie mehr als eine „versteckte“ dazwischenliegende Schicht. Die einzelnen Schichten bestehen aus einer Vielzahl künstlicher Neuronen, die miteinander verbunden sind und auf Eingaben von Neuronen aus der jeweils vorherigen Schicht reagieren. In der ersten Schicht wird etwa ein Muster erkannt, in der zweiten Schicht ein Muster von Mustern und so weiter. Je komplexer das Netz (gemessen an der Anzahl der Schichten von Neuronen, der Verbindungen zwischen Neuronen sowie der Neuronen pro Schicht), desto höher ist der mögliche Abstraktionsgrad – und desto komplexere Sachverhalte können verarbeitet werden. Angewendet wird Deep Learning bei der Bild-, Sprach- und Objekterkennung sowie dem verstärkenden Lernen.
Deep learning	ZVKI	(auf Deutsch: tiefes Lernen) Komplexe künstliche neuronale Netze, die mehr als drei Schichten künstlicher Neuronen umfassen. Der Begriff „deep“ bzw. „tief“ bezieht sich auf die Menge der Schichten künstlicher Neuronen. Mit jeder zusätzlichen Schicht erhöht sich der Abstrahierungsgrad des Systems und die Fähigkeit, komplexere Aufgaben zu bewältigen. Dazu gehört die Erstellung von realistischen Deep Fakes. Gleichzeitig verstärkt sich dabei allerdings das Problem von Künstlicher Intelligenz als Blackbox.
Generative KI	Lernende Systeme	KI-Systeme, die mit großen Datensätzen trainiert wurden und in der Lage sind, Inhalte zu erzeugen (z. B. Text, Programmcode, Videos, Bilder, Proteinstrukturen, Bauteile). Sie stützen sich dabei auf große Rechenleistung und spezielle Algorithmen, die unter anderem auf dem so genannten Transformer-Modell basieren. Die Anwendungsbereiche generativer KI-Systeme sind breit, zu den bekanntesten Systemen zählen das Ende 2022 veröffentlichte Sprachmodell ChatGPT (Open AI) sowie BARD (Google) und LLaMA (Meta). Im Fall von ChatGPT erfolgt das Training der generativen KI in einem mehrstufigen Prozess auf Basis von Deep Learning und anderen Lernmethoden wie folgt: - Der Transformer lernt auf Grundlage von Wahrscheinlichkeiten vorherzusagen, welches Wort auf die vorangehende Wortfolge folgt. Das KI-System lernt hier selbstüberwacht – und benötigt keine vom Menschen vorgenommene Datenetikettierung (Labeling). - Menschen erstellen idealtypische Texte für bestimmte Arbeitsanweisungen und vermitteln dem KI-System so, welche Texte erwünscht sind. Dann stellen Menschen Anfragen und ordnen die Antworten der KI gemäß ihrer Qualität in Rangfolgen an. - Auf dieser Basis lernt das KI-System, seine eigenen Ausgaben zu bewerten und sich für gute Ausgaben zu belohnen (verstärkendes Lernen). So nähert es sich den Qualitätsansprüchen des Menschen an und verbessert sich stetig. Neben Wörtern können generative KI-Systeme auch Pixel, Töne, Programm- oder DNA-Code verarbeiten.
Generative KI	ZVKI	KI-Modelle, die in der Lage sind, Inhalte zu erstellen. Modelle generativer KI (Generative AI) können beispielsweise Texte, Bilder, Audiodateien, Videos oder Programmiercode erzeugen. Zu den verbreiteten Anwendungen zählen die Bildgeneratoren Stable Diffusionund Midjourney oder der Textgenerator ChatGPT. Viele generative KI-Systeme basieren auf Basismodellen. Einfachen Chatbots, die etwa bestimmte Serviceanfragen beantworten können, liegen aber weniger komplexe KI-Modelle zugrunde.
Halluzination	Uni Siegen	Eine KI-Halluzination tritt auf, wenn ein KI-Modell falsche oder irreführende Informationen generiert. Diese generierten Inhalte können plausibel erscheinen, da die Modelle darauf ausgelegt sind, kohärente und flüssige Ausgaben zu produzieren, auch wenn sie nicht notwendigerweise der Realität entsprechen. KI-Halluzinationen können in verschiedenen Formen auftreten, wie beispielsweise falsche Behauptungen, inkonsistente Logik oder unrealistische Szenarien. Es ist wichtig, sich der Möglichkeit von Halluzinationen bewusst zu sein und die Ergebnisse von KI-Modellen kritisch zu hinterfragen.

KI-Modell	HPE	KI-Modelle oder Modelle für künstliche Intelligenz sind Programme, die anhand einer Sammlung von Datensätzen bestimmte Muster erkennen. Es ist die Darstellung eines Systems, das Dateneingaben erhält und Schlussfolgerungen ziehtbzw. basierend auf diesen Schlussfolgerungen agieren kann. Ist ein KI-Modell einmal trainiert, kann es beispielsweise Zukunftsprognosen erstellen oder basierend auf zuvor überwachten Daten handeln. KI-Modelle können für zahlreiche Aktivitäten verwendet werden: in der Bild- und Videoerkennung, Natural Language Processing (NLP), Anomalieerkennung, für Empfehlungssysteme, zur vorausschauenden Modellerstellung und für Prognosen sowie in der Robotik und in Kontrollsystemen.
Künstliche Intelligenz	KUKA	Maschinen als intelligente Partner. Künstliche Intelligenz (KI) ist der Schritt zur Realisierung der vierten Stufe der von KUKA postulierten Robotik-Revolutionen. Sie setzt voraus, dass Maschinen, Informationssysteme und Roboter in der Lage sind, noch sehr viel intelligenter und reaktiver zu werden. In den Bereichen Servicerobotik und Home Assisted Living werden diese intelligenten Maschinen mit ihren kognitiven und sensitiven Fähigkeiten als Helfer des Menschen immer mehr Bedeutung erlangen. Heute hängen diese Systeme noch vollständig von der Programmierung durch den Menschen ab. Je mehr jedoch der Autonomiegrad der Systeme steigt, desto dringlicher wird sich die Frage nach einem verantwortungsvollen Umgang mit künstlicher Intelligenz stellen.
Künstliche Intelligenz	ZVKI	(auf Englisch: artificial intelligence, AI) Teilgebiet der Informatik, das sich damit beschäftigt, menschliche Intelligenz technisch nachzubilden. In diesem Teilgebiet werden Methoden entwickelt, mit denen Computerprogramme oder Maschinen automatisiert Aufgaben erfüllen sollen. Deswegen sprechen wir von verschiedenen KI-Methoden, die in KI-Systemen angewendet werden, anstatt von der „einen“ KI. Bisher fehlt eine einheitliche Definition für Künstliche Intelligenz. In der Regel werden als Künstliche Intelligenz aber insbesondere Methoden des maschinellen Lernens bezeichnet. Zudem wird zwischen starker und schwacher KI unterschieden.
Künstliche Intelligenz (KI)	Lernende Systeme	Eine allgemein akzeptierte Definition zu Künstlicher Intelligenz (KI) gibt es nicht. KI ist zum einen ein Teilgebiet der Informatik, das versucht, mit Hilfe von Algorithmen kognitive Fähigkeiten wie Lernen, Planen oder Problemlösen in Computersystemen zu realisieren. Begründet wurde der Begriff Artificial Intelligence im Zuge des Dartmouth Workshops (1956), der auch heute noch die moderne KI-Forschung prägt. Das internationale Standardlehrbuch für Künstliche Intelligenz von Russel/Norvig behandelt folgende Forschungsfelder: - Problemlösen - Wissensrepräsentation und Schlussfolgern - Unsicherheit und Schlussfolgern - Maschinelles Lernen - Wahrnehmung und Sehen - Verstehen und Generieren von natürlicher Sprache - Interaktion - Robotik Der Begriff KI steht zugleich für Systeme, die ein Verhalten zeigen, für das gemeinhin menschliche Intelligenz vorausgesetzt wird. Ziel moderner KI-Systeme (Lernende Systeme) ist es, Maschinen, Roboter und Softwaresysteme zu befähigen, abstrakt beschriebene Aufgaben und Probleme eigenständig zu bearbeiten und zu lösen, ohne dass jeder Schritt vom Menschen programmiert wird. Dabei sollen sich die Systeme auch an veränderte Bedingungen und ihre Umwelt anpassen können. In diesem Sinne schafft Künstliche Intelligenz die Voraussetzungen für Lernende Systeme. Die Lernfähigkeit der Systeme wurde bereits zu Beginn der KI-Forschung als grundlegende kognitive Fähigkeit definiert. Es ist jedoch schwierig, abschließend zu bestimmen, was als „intelligent“ gilt. Abhängig vom jeweiligen Stand der Technik entwickelt sich daher stets das Verständnis darüber weiter, was als KI bezeichnet wird.
Künstliches Neuronales Netz (KNN)	CLICKWORKER	Ein künstliches neuronales Netz (KNN) (EN: Artificial Neural Network, ANN) ist ein Rechenmodell nach dem Vorbild des Gehirns. Es besteht aus vielen miteinander verbundener Verarbeitungsknoten (Neuronen), die zusammenarbeiten, um bestimmte Probleme zu lösen. KNNs werden verwendet, um zukünftige Ereignisse vorherzusagen, Muster zu erkennen und Entscheidungen zu treffen. Sie werden häufig für Aufgaben eingesetzt, die für herkömmliche Computersysteme zu schwierig sind – zum Beispiel Bild- und Spracherkennung. Im Vergleich zum Gehirn haben KNNs eine Reihe von Vorteilen. - Sie können enorme Datenmengen viel schneller verarbeiten als das Gehirn. -Sie sind nicht durch das Kurzzeitgedächtnis eingeschränkt. - Darüber hinaus können künstliche neuronale Netze so konzipiert werden, dass sie fehlertolerant sind. Sie funktionieren auch dann, wenn einige ihrer Neuronen beschädigt oder zerstört werden. Trotz dieser Vorteile haben KNNs immer noch eine Reihe von Einschränkungen. Es kann schwierig sein, sie zu entwerfen und zu trainieren, und sie benötigen oft eine große Menge an Daten, um effektiv zu sein. Außerdem können falsche Trainingsdaten die KNNs zu ungenauen Ergebnissen führen.

Large Language Modell (LLM)	ZVKI	Sprachmodell, das sich durch seine Größe auszeichnet. Large Language Models sind Basismodelle und können für viele verschiedene Aufgaben der Sprachverarbeitung eingesetzt werden, beispielsweise um Texte zu übersetzen, zusammenzufassen oder zu analysieren oder um Texte und andere Inhalte (z. B. Code) zu generieren. Verbreitete LLMs sind etwa GPT-3und GPT-4 von OpenAI oder LaMDA von Google. Die Entwicklung von Large Language Models ist ein Teilbereich der Forschung rund um Natural Language Processing.
Large Language Modelle (LLM)	Lernende Systeme	Sie hierzu Sprachmodelle
Machine Learning	KUKA	Wissen aus Erfahrung. Intelligente Maschinen schöpfen ihr Wissen aus Erfahrung. Bei vernetzten Maschinen ist es dabei unerheblich, ob sie die Erfahrung selbst gemacht haben oder ob die Erfahrung aus der Schwarmintelligenz stammt. Dabei lernt ein künstliches System immer aus dem Abgleich zwischen dem angestrebten Ziel und auftretenden Anomalien. Es kann Korrelationen, Muster und Gesetzmäßigkeiten erkennen, daraus Schlüsse ziehen und sein zukünftiges Verhalten verändern – diesen synthetischen Prozess bezeichnet man als Machine Learning. Speziell in unstrukturierten Umgebungen und bei hochflexiblen Prozessen wie Industrie 4.0 ist Machine Learning im Schwarm oder in der Cloud eine effektive Methode, um Produktionsprozesse nahezu in Echtzeit intelligent und autonom an die jeweiligen Rahmenbedingungen anzupassen.
Machine Learning Operations (MLOps)	Lernende Systeme	beziehen die Development Operations (DevOps) auf das maschinelle Lernen mit dem Ziel, maschinelles Lernen, Softwareentwicklung und den laufenden Betrieb der Systeme zusammenzubringen. Angestrebt wird insbesondere eine effiziente, zuverlässige und qualitativ hochwertige Gestaltung der Entwicklung, Bereitstellung, Verwaltung und Überwachung von KI-Modellen.
Maschinelles Lernen	Gaia-X Hub Germany	Mithilfe von Lernalgorithmen sowie der Anwendung statischer Modelle entwickelt eine Maschine auf Basis von Lerndaten ein komplexes Modell und erkennt Muster, aus denen Klassifikationen und Vorhersagen abgeleitet werden können. Genutzt wird diese Technologie beispielsweise in alltäglichen Anwendungen wie dem Spamfilter, in Mailprogrammen oder Übersetzungsprogrammen.
Maschinelles Lernen	Lernende Systeme	Maschinelles Lernen ist eine grundlegende Methode der Künstlichen Intelligenz (KI). Sie zielt darauf, dass Maschinen ohne explizite Programmierung eines konkreten Lösungswegs automatisiert sinnvolle Ergebnisse liefern. Spezielle Algorithmen lernen aus den vorliegenden Beispieldaten Modelle, die dann auch auf neue, zuvor noch nicht gesehene Daten angewendet werden können. Dabei werden drei Lernstile unterschieden: überwachtes Lernen, unüberwachtes Lernen und verstärkendes Lernen. Maschinelles Lernen mit großen neuronalen Netzen wird als Deep Learning bezeichnet. Maschinelle Lernverfahren kommen zum Einsatz beim Data Mining, beim Generieren von Smart Data und in praktisch allen modernen KI-Systemen.
maschinelles Lernen	ZVKI	(auch machine learning) Grundlegende Methode im Bereich Künstliche Intelligenz. Bei diesem Verfahren programmieren die Entwickler*innen den Algorithmus, der eine Aufgabe lösen soll, nicht selbst. Stattdessen legen die Entwickler*innen Zielvorgaben fest und bauen ein Computerprogramm, das den besten Lösungsalgorithmus selbstständig findet. Unterschieden werden drei grundlegende Lernmethoden: überwachtes Lernen, unüberwachtes Lernen und bestärkendes Lernen. Inzwischen werden immer häufiger künstliche neuronale Netze eingesetzt.
Multimodale Modelle	Lernende Systeme	KI-Modelle, die auf einem Datensatz trainiert werden, der verschiedene Modalitäten (z.B. Text, Bild, Ton, Programmcode, Video) enthält. Durch deren Verknüpfung lassen sich bessere Ergebnisse erzielen und neue Aufgaben (z.B. Text-zu-Bild oder Text-zu-Code) erledigen. Die Integration von Daten aus verschiedenen Modalitäten kann komplex sein. Da Daten unterschiedliche Formate, Skalen und Bedeutungen haben können, sind spezielle Techniken der Datenkuratierung und -verarbeitung erforderlich, um qualitativ hochwertige, kohärente und nützliche Ergebnisse zu erzielen. GPT-4 ist ein multimodales KI-Modell, das sowohl Text als auch Bilder verarbeiten kann.
natural language processing	ZVKI	maschinelle Verarbeitung natürlicher Sprache. Damit ist zum einen gemeint, dass mithilfe algorithmischer Systeme der Inhalt einer gesprochenen oder geschriebenen Aussage ermittelt wird, damit eine Software diesen verarbeiten kann. Dafür müssen Computerprogramme zunächst mittels maschinellen Lernens und unzähliger Trainingsdaten umfangreich trainiert werden. Zum anderen ist damit die künstliche Erstellung von Aussagen in natürlicher Sprache gemeint, beispielsweise bei Vorleseprogrammen. Bei digitalen Sprachassistenten wie Siri, Alexa und Google Assistant kommen beide Formen von natural language processing zur Anwendung: Sie können unsere Befehle verarbeiten und Antworten ausgeben. Dass das nicht immer gelingt, verdeutlicht, wie anspruchsvoll diese Aufgabe ist.

Prompt	Uni Siegen	Prompts sind textliche Anfragen oder Anweisungen, die von Nutzer:innen verwendet werden, um Antworten, Informationen oder Texte von Chat-KIs oder Bilder und Grafiken von Bild-KIs zu erhalten. Ist das Ergebnis nicht zufriedenstellend, muss die Anfrage optimiert oder erweitert werden. In der dialogischen Interaktion mit dem KI-System entsteht so nach und nach ein immer besseres Ergebnis. Umfangreichere Eingabeaufforderungen oder Anfragen an ein KI-Modell, um komplexe Aufgaben oder Texte zu erstellen, werden Mega-Prompts genannt. Diese Prompts sind oft ausführlicher und detaillierter als gewöhnliche Anfragen, um hochwertige und präzise Ergebnisse zu erzielen. Sie dienen dazu, das Modell gezielt zu steuern und sicherzustellen, dass es den gewünschten Text oder das Bild generiert (Bsp.: „Ein Apfel auf einem Tisch“ – besser: „Ein roter Apfel in einer hölzernen, ovalen Schale auf einem braunen Holztisch, gemalt im impressionistischen Stil“, s. auch nachfolgend ‚Prompt-Labor‘). Das sog. Prompt Engineering beinhaltet das Experimentieren mit verschiedenen Formulierungen, Strukturen und Kontexten, um die gewünschten Antworten oder Ausgaben zu optimieren. Es ist eine Methode, um die Leistung von KI-Modellen in natürlicher Sprachverarbeitung gezielt zu verbessern. Tipp: Im Modul „Materialien zur Einführung“ des Selbstlernkurses „Prompt-Labor: Generative KI in der Hochschullehre“ der Lernplattform KI-Campus werden verschiedene Empfehlungen und Techniken vorgestellt, um durch geschicktes Vorgehen beim Prompt-Design bessere Ergebnisse zu erzielen (Anm.: Kostenlose Registrierung erforderlich).
Sprachmodelle	Lernende Systeme	KI-Modelle, die mit maschinellen Lernverfahren auf Textdatensätzen trainiert wurden, beispielsweise mit Transformern oder rekurrenten neuronalen Netzwerken. Solche Modelle sagen z.B. basierend auf Wahrscheinlichkeiten voraus, welche Wörter auf eine vorhandene Reihe an Wörtern folgt. Moderne Sprachmodelle werden oft als große Sprachmodelle (engl. Large Language Models) bezeichnet, da sie sehr große Parameterzahlen aufweisen und auf umfangreichen Textdaten trainiert wurden. Solche Modelle erkennen, produzieren, übersetzen und verarbeiten natürliche Sprache und können für eine Vielzahl von Aufgaben eingesetzt werden – etwa das Erstellen und Zusammenfassen von Text, das Beantworten von Fragen, das Generieren von Programmcode und vieles mehr. Aufgrund ihrer Einsatzbreite gelten große Sprachmodelle wie GPT 4, Gemini, Claude oder Mistral als Schlüsseltechnologie der Künstlichen Intelligenz. Große Sprachmodelle werden auch als Basismodelle bzw. Foundation Models bezeichnet, wenn sie an bestimmte Zwecke und Aufgaben angepasst werden können.

Impressum

eco – Verband der Internetwirtschaft e. V.
Lichtstr. 43h,
50825 Köln
Fon + 49 (0) 221-70 00 48-0
Fax +49 (0) 221-70 00 48-111
info@eco.de
www.eco.de