

Faire Algorithmen: Rechtlicher Rahmen für Künstliche Intelligenz



MANAGEMENT SUMMARY

Künstliche Intelligenz (KI) kann für viele Zwecke genutzt werden. Sie ist dabei wie jede andere Technologie zunächst wertneutral und kann - falsch eingesetzt - auch für Mensch und Umwelt schädliche Effekte hervorrufen. Die Europäische Union hat daher im Sommer 2024 das erste Gesetz für die Entwicklung und den Einsatz von KI verabschiedet. Dieses hat das Potential - ähnlich wie die Verordnungen im Bereich Datenschutz - durch globale Vernetzung und Übernahme in Länder außerhalb der EU zum weltweiten De-Facto-Standard für ethische KI zu werden. Wir fassen den aktuellen Stand (Februar 2025) kompakt zusammen.



FAKTEN

Entwicklung der KI-Gesetzgebung in der EU

Während KI in den USA von großen Konzernen und in China von staatlicher Seite gelenkt wird, hinkt Europa in diesem Bereich bisher eher hinterher. Die Ursachen hierfür sind vielfältig. Moralische Vorbehalte, rechtliche Bedenken und daraus resultierendes Misstrauen gegenüber der Technologie tragen in jedem Fall dazu bei, dass öffentliche und privatwirtschaftliche Initiativen nur gebremst oder gar nicht voranschreiten und KI nur sehr zögerlich Eingang in Produkte und Prozesse findet.

Die EU möchte die Entwicklung von KI fördern, gleichzeitig aber auch öffentliche Interessen schützen und Vertrauen in KI schaffen. Zu diesem Zweck sollen Standards definiert werden, die rechtliche und ethische Klarheit schaffen und KI-Entwicklern

sinnvolle Rahmenbedingungen geben. Aus diesem Grund erarbeitete eine internationale, unabhängige Expertengruppe eine 2019 veröffentlichte Leitlinie für vertrauenswürdige KI, die als Basis für den offiziellen Entwurf für ein europäisches KI-Gesetz¹ diene. Der im Mai 2024 von den EU-Mitgliedsstaaten verabschiedete AI-Act soll den Rechtsrahmen zum Thema Künstliche Intelligenz in der EU bilden (als Teil des sogenannten „Koordinierten Plans für KI“²). Das Gesetz beinhaltet Verbote bestimmter KI-Anwendungen (Eintritt erster solcher Verbote im Februar 2025³), Anforderungen und Pflichten für KI-Systeme mit hohem Risiko, sowie darüber hinaus Transparenzpflichten für KI-Systeme. Außerdem wurden, vor dem Hintergrund des Green Deal der EU, ökologische und soziale Nachhaltigkeitsaspekte gestärkt. In vollem Umfang soll das KI-Gesetz ab dem 2. August 2026 greifen.

Ebenfalls 2019 erschienen ähnliche Grundsätze für Künstliche Intelligenz auch seitens der OECD als Empfehlung für Regierungen, Investoren und Entwickler von KI. Rechtlich bindend sind diese OECD-Grundsätze jedoch nicht. Das europäische KI-Gesetz hingegen verpflichtet die EU-Mitgliedsstaaten weitgehend zur Umsetzung entsprechender Rahmenbedingungen. Aufgrund der wirtschaftlichen Bedeutung Europas und der globalen Verflechtung von Märkten hat das EU-Recht für KI das Potential, sich zum weltweiten Standard zu entwickeln (ähnlich wie dies z.B. im Bereich Datenschutz durch GDPR bereits geschehen ist).



Grundsätzliche Vorgaben für vertrauenswürdige KI

Die Leitlinie und das Gesetz der EU machen einige grundlegende Vorgaben. So müssen KI-Systeme alle geltenden Gesetze, aber auch ethische Grundsätze befolgen, Menschenrechte achten und Schäden – speziell für Menschen – vermeiden. Sie müssen zuverlässig gegen nachteilige Effekte gesichert sein, egal ob diese aus unbeabsichtigten Fehlern des Systems oder aus externen Angriffen resultieren. Sicherheit und Robustheit gegenüber Fehlern beinhalten auch Aspekte wie Genauigkeit, Zuverlässigkeit und Reproduzierbarkeit.

Fairnessgebot und Diskriminierungsverbot mögen selbstverständlich klingen, waren aber in den letzten Jahren bereits in der öffentlichen Diskussion, da sich in manche KI bereits menschliche Vorurteile eingeschlichen haben. Algorithmen lernen auf der Basis von eingespeisten Daten. Sind diese unausgewogen oder beinhalten sie Zusammenhänge, die sich aus faktischer Diskriminierung in der analogen Welt ergeben, dann führt dies zum vieldiskutierten KI-Bias (Voreingenommenheit der KI).

Die geforderte Menschenzentrierung für KI findet auch darin Ausdruck, dass KI erklärbar sein muss. Transparente Rückverfolgbarkeit, Nachvollziehbarkeit, Überprüfbarkeit

¹ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689>

² <https://digital-strategy.ec.europa.eu/en/policies/plan-ai>

³ https://germany.representation.ec.europa.eu/news/ki-gesetz-erste-vorschriften-kraft-getreten-2025-02-03_de

und Verständlichkeit von Empfehlungen und Entscheidungen sind hierbei essenziell. Entsprechende Kommunikation, z.B. mit Behörden und Betroffenen, muss gewährleistet sein. Dies gilt auch für das Auftreten und die Wiedergutmachung von Schadensfällen (Rechenschaftspflicht).

Neben den Algorithmen selbst stellt das Gesetz auch Anforderungen an die Qualität und Sicherheit der zugrundeliegenden Daten. Natürlich gelten die Prinzipien des Schutzes personenbezogener Daten (Datenschutz gemäß GDPR) in vollem Umfang. Idealerweise werden nur anonyme Daten verwendet, und wo doch personenbezogene Daten erforderlich sind, sind diese auf ein Minimum und den vorab vereinbarten Zweck zu beschränken und nach Erfüllung dieses Zwecks zu löschen.

Für das ordnungsgemäße und legale Funktionieren eines KI-Systems muss aber auch sichergestellt sein, dass die Daten korrekt sind und nicht durch schlechte Datenqualität fehlerhafte KI-Modelle und Entscheidungen entstehen. Das Thema Datenqualität hat bekanntlich über triviale Fehler hinaus viele Facetten. Probleme fallen aufgrund des hohen Automatisierungsgrades innerhalb von KI-Systemen nicht immer sofort auf. Intensive Prüfmaßnahmen – auch und gerade im laufenden Betrieb – können erforderlich sein.

Die genannten Vorgaben gelten für alle Phasen des Lebenszyklus eines KI-Systems, also bei Konzeption und Entwicklung, während der produktiven Nutzung, aber auch danach (z.B. in Archiven). Nach Verabschiedung des Gesetzes sind sie für jegliche Nutzer und Anbieter von KI-Lösungen innerhalb der EU verbindlich, unabhängig von Herkunft und Firmensitz. Hierin liegt der potenziell weltweite Hebel der EU-Gesetzgebung.



Konkretisierung für die Praxis

Um die obigen allgemeinen Grundsätze praktikabel zu machen, klassifizieren Richtlinie und EU-Gesetz die verschiedenen KI-Systeme hinsichtlich ihres Risikos in vier Risikostufen.

In die Kategorie des inakzeptablen (unannehmbaren) Risikos fallen all diejenigen Anwendung, die eine Bedrohung für die Sicherheit, den Lebensunterhalt oder die Rechte des Menschen darstellen. Hierzu gehören u.a. Social-Scoring-Applikationen, Systeme, die menschliches Verhalten unterbewusst manipulieren, aber auch biometrische Fernerkennung (z.B. Gesicht, Stimme, Gang) in Echtzeit. Sie alle sind im Geltungsbereich des EU-Gesetzes verboten⁴.

Als Hochrisikoanwendungen gelten Applikationen, die ein erhebliches Risiko für Gesundheit, Sicherheit oder Grundrechte aufweisen.

⁴ <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32024R1689>

Die Liste solcher Hochrisikoapplikationen ist lang: kritische Infrastruktur (wozu auch autonomes Fahren gezählt wird) und Sicherheitskomponenten innerhalb von (z.B. medizinischen) Produkten gehören genauso dazu wie jegliche Anwendungen, die Menschen bewerten und durch diese Bewertung benachteiligen können (beispielsweise Kreditrisikobewertung oder automatisches Scoring von Bewerbungen oder Prüfungen). Auch hoheitliche Aufgabengebiete (z.B. in der Strafverfolgung oder bei Grenzkontrollen) und erlaubte Anwendungen für biometrische Identifizierungslösungen werden als hochriskant eingestuft. Für alle diese Applikationen gelten besonders strenge Vorschriften hinsichtlich Robustheit, Sicherheit, Risikomanagement, Transparenz, Datenqualität und anderer Kriterien. Um Risiken zu minimieren, müssen sie in angemessenem Umfang unter menschlicher Aufsicht stehen.

Bei KI-Systemen mit geringem Risiko steht Transparenz als wesentliches Gebot im Vordergrund. So müssen beispielsweise Nutzer von Chatbots erkennen können, dass sie mit einer Maschine interagieren, oder Anbieter von generativer KI sicherstellen, dass KI-generierte Inhalte klar zu erkennen sind. Der Einsatz von KI-Systemen mit minimalem oder keinem Risiko (z.B. Spamfilter und viele andere KI-gestützte Funktionen innerhalb von größeren Applikationen) soll auch in dieser Hinsicht weitgehend uneingeschränkt möglich sein. Hier greift letztlich die normale Produkthaftung.



FAZIT

Trotz mancher noch offener Themen in vielen Anwendungsgebieten, z.B. der Frage nach den Rechten an den zugrundeliegenden Daten, Haftungsfragen oder auch nach den konkreten zivil- und strafrechtlichen Folgen von Verstößen, bietet der Rechtsrahmen eine sichere Grundlage für Anbieter und Anwender von KI. Durch den Abschluss des Gesetzgebungsverfahrens ist zu hoffen, dass auch die im „Koordinierten Plan für KI“ angelegte systematische Förderung zeitnah umgesetzt wird und die Aufholjagd Europas in diesem Gebiet mit vollem Schwung vorangeht.



Autor

Dr. Marcus Dill ist mehrfacher erfolgreicher Unternehmensgründer und langjähriger Management Berater mit den Schwerpunkten Data, Analytics und Sustainability.

Kontakt: Ingdilligence GmbH

www.ingdilligence.com

sustainability@ingdilligence.com

