Entwicklung künstlicher Intelligenz in Zeiten von Regularien

Inhalt

Regularien zu künstlicher Intelligenz und ihre Gründe

Spätestens mit dem Erscheinen von OpenAIs ChatGPT und dem anschließenden KI-Hype ist künstliche Intelligenz auch in der breiten Gesellschaft angekommen. Auf künstlicher Intelligenz basierende Systeme haben eine Reihe an Risikopotentialen. Die Nutzung von Nutzeranfragen zur Verbesserung der Systeme ist gängige Praxis. Hier könnten personenbezogene Daten verarbeitet werden. Des Weiteren können die Technologien für schändliche Zwecke wie Betrug oder Manipulation eingesetzt werden. Das Verwenden von voreingenommenen Daten zum Training von KI-Modellen führt zu voreingenommenen Systemen. Dass diese Risiken früher oder später durch Verordnungen und Gesetze reguliert werden würden, ist eine logische Folge.

Die Entwicklung von KI-basierten Anwendungen und Systemen unterliegt in der Europäischen Union (EU) Regulierungen wie dem EU AI Act und der Datenschutzgrundverordnung (DSGVO). Diese sind in der EU bindend und Verstöße resultieren in Strafen. Die Regelungen gelten sowohl für die Entwicklung also auch für den Betrieb und die Nutzung der Systeme.

EU AI Act

Der EU AI Act bildet seit August 2024 einen einheitlichen, rechtlichen Rahmen für die Entwicklung, den Einsatz und die Vermarktung von KI-Systemen in der Europäischen Union.

Im Rahmen der Verordnung werden KI-Systemen eine Reihe an Eigenschaften zugesagt (siehe Schaubild).

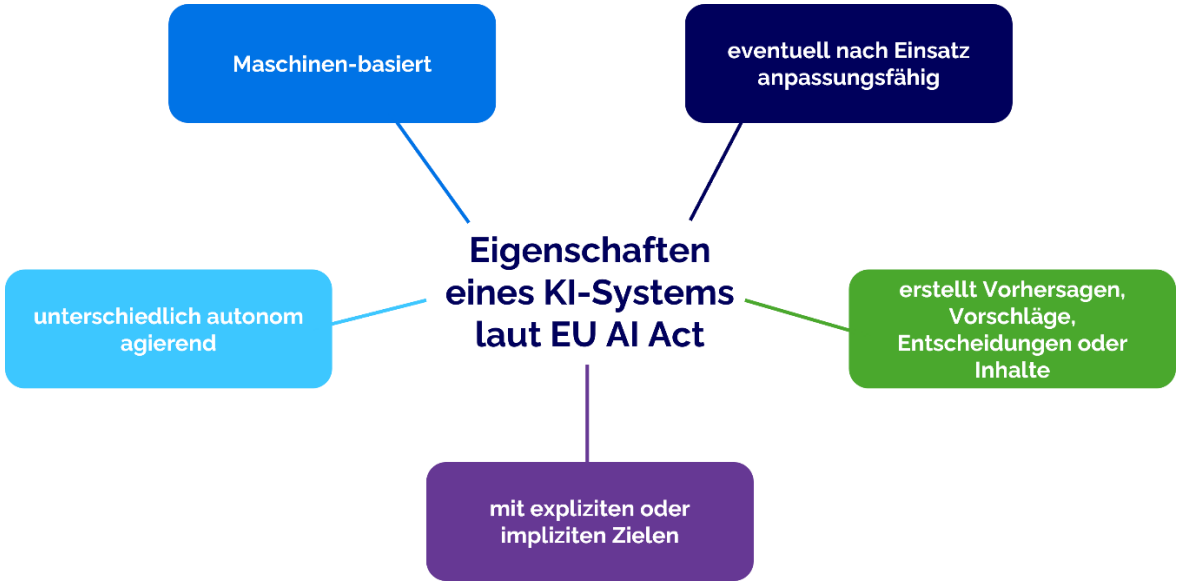


Abbildung 1: KI-Systeme laut EU AI Act

Für die Bewertung von KI-Systemen legt der EU AI Act vier Kategorien fest: „minimales“, „begrenztes“, „hohes“ und „unakzeptables“ Risiko. Je höher das Risiko ist, desto größer ist auch das Maß an Verantwortung und Kontrolle, das für die Entwicklung und den Betrieb des Systems aufgewendet werden muss.
KI-Systeme von unakzeptablem Risiko sind prinzipiell verboten. Hierzu zählen Social Scoring und Anwendungen mit dem Zweck der Manipulation.

Der Fokus der Verordnung liegt auf Systemen mit hohem Risiko, da diese Strukturen mit dem höchsten Maß an Fürsorge umgesetzt werden müssen. Hierzu zählen die automatische Verarbeitung von personenbezogenen Daten und die biometrische Identifizierung von Personen.

Wird das Risiko als begrenzt eingestuft, gelten leichtere Transparenzanforderungen und es muss sichergestellt werden, dass Endnutzer über den KI-Einsatz Bescheid wissen. In diese Kategorie werden Chatbots und Deepfake-Anwendungen eingestuft.

Anwendungen, deren Risiko als minimal kategorisiert wird, bleiben laut dem EU AI Act unreguliert. Hierzu zählen KIs in Videospielen und Spam-Filter.

Unakzeptables Risiko

Social Scoring
Manipulation

Hohes Risiko

Verarbeitung personenbezogener Daten
Biometrische Indentifizierung

Begrenztes Risiko

Chatbots
Nutzung von Deepfakes

Minimales Risiko

Spam-Filter
Videospiel-KI

Abbildung 2: Risikokategroien des EU AI Acts

Der EU AI Act sieht den Großteil der Erfüllungspflichten bei den Entwicklern eines KI-Systems, aber auch die Bereitsteller einer Lösung sind in der Verantwortung. Zu deren Pflichten gehören ausreichendes Risikomanagement, Data Governance, Erstellen von technischen Dokumentationen und Nutzungsanweisungen, Datenhaltungsanforderungen und Einhaltung von Standards für Genauigkeit, Robustheit, Cybersecurity und Qualitätsmanagement für KI-Systeme. [1]

[1] Regulation (EU) 2024/1689 (2024) *OJ L 13.7.2024*, S. 1 ff.

DSGVO

Die Datenschutzgrundverordnung (DSGVO) trat im Mai 2018 in Kraft und regelt den Umgang mit personenbezogenen Daten in der EU. Die Daten-Verarbeitung muss sieben Grundsätzen entsprechen.

1 **Rechtmäßigkeit, Verarbeitung nach Treu und Glauben**
Die Verarbeitung muss rechtmäßig und nachvollziehbar durchgeführt werden.

2 **Zweckbindung**
Daten dürfen nur für den Zweck verwendet werden, für den sie erhoben wurden.

3 **Datenminimierung**
Es dürfen nur für den konkreten Zweck benötigte personenbezogene Daten erfasst werden.

4 **Richtigkeit**
Personenbezogene Daten sollen aktuell und richtig sein. Falsche Daten müssen korrigiert oder gelöscht werden.

5 **Speicherbegrenzung**
Daten müssen gelöscht werden, sobald sie nicht mehr benötigt werden. Hierbei gibt es Ausnahmen für Archivfälle.

6 **Integrität und Vertraulichkeit**
Die Daten müssen sicher aufbewahrt werden und vor dem Zugriff durch Unbefugte geschützt werden.

7 **Rechenschaftspflicht**
Die Einhaltung der Grundsätze muss nachgewiesen werden können..

Diese Grundsätze finden ebenfalls bei der Entwicklung von KI-Systemen, die personenbezogene Daten verarbeiten Einsatz.[2]

[2] Verordnung (EU) 2016/679 (2016) *OJ L 119*, S. 1-88

Weitere Regularien

Zusätzlich zu den beiden genannten relevanten Regularien gibt es noch Regelwerke und Frameworks aus weiteren Quellen. Folgend wird ein Auszug relevanter Quellen genannt.

So hat die *Organisation for Economic Co-operation and Development* (OECD) Richtlinien für den Einsatz von KI festgelegt. Hier sind ebenfalls Fairness, Transparenz, Erklärbarkeit, Robustheit, Sicherheit und eine Rechenschaftspflicht als Kriterien für die Entwicklung von KI festgelegt.[3]

Das amerikanische *National Institute of Standards and Technology* veröffentlichte 2022 ein Framework für Risikomanagement für die Entwicklung künstlicher Intelligenz und erweiterte das Framework 2024 um spezifische Informationen für generative KI. Das Framework unterstützt bei der Evaluierung von Risiken und umfasst Merkmale wie Validität und Zuverlässigkeit, Sicherheit, Fairness und Bias-Management und Erklärbarkeit und Interpretierbarkeit.[4,5]

Die Regularien der Länder, in denen ein System und seine Ergebnisse erstellt, genutzt oder verwendet werden, können ebenfalls Anwendung finden.

---

[3] OECD, Empfehlung des Rates zu künstlicher Intelligenz, OECD/LEGAL/0449

[4] Tabassi, E. (2023), Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.AI.100-1, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=936225 (Accessed March 3, 2025)

[5] Autio, C. , Schwartz, R. , Dunietz, J. , Jain, S. , Stanley, M. , Tabassi, E. , Hall, P. and Roberts, K. (2024), Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, NIST Trustworthy and Responsible AI, National Institute of Standards and Technology, Gaithersburg, MD, [online], https://doi.org/10.6028/NIST.AI.600-1, https://tsapps.nist.gov/publication/get_pdf.cfm?pub_id=958388 (Accessed March 4, 2025)

Regularien in Unternehmenskontext

Unternehmen sind in der Verantwortung, die relevanten Regularien umzusetzen. Gleichzeitig müssen sie sich strategisch aufstellen, um die Potenziale von künstlicher Intelligenz als Technologie zu nutzen.

KI-Strategien

Die Regularien haben so einen Einfluss auf die Strategien eines Unternehmens. Spezielle KI-Ansätze haben zum Ziel, die Nutzung der Technologien und die Standpunkte des Unternehmens zu diesen festzuhalten und beinhalten folgende Punkte:

Wie will das Unternehmen künstliche Intelligenz nutzen?

Für diesen Punkt befasst man sich mit Bestandsaufnahmen aktueller Technologien sowie ihrer Stärken und Schwächen. Man erfasst Pilotprojekte und Kriterien für die Erfolgsmessung. Hier bietet es sich an, die unternehmenseigene Datenstrategie zu integrieren, um Datenmanagement und Datenqualitätssicherungsmaßnahmen anzugleichen.

Wie interagiert künstliche Intelligenz mit der Unternehmensphilosophie und -Kultur?

Hier definiert ein Unternehmen seine Stellung zur ethischen Verwendung künstlicher Intelligenz in Bezug auf ihre eigenen Werte. Dabei sind unter anderem die Einflüsse auf Umweltschutzthemen zu berücksichtigen. Die Auswahl der KI-Partner beeinflusst ebenfalls die Unternehmenskultur. Es macht Sinn, Partner mit kompatiblen Werten zu berücksichtigen.

Welchen Einfluss hat die Nutzung von künstlicher Intelligenz auf die Unternehmensmission und -Vision?

Die zuvor beschriebenen Umstände münden in die Evaluierung der Einflüsse auf das Unternehmen und seine mittel- und langfristigen Ziele. Es ist festzuhalten, wo Synergien und wo Konflikte entstehen können.

KI-Policen

Der Geltungsbereich und das Ziel einer Unternehmenspolice ist es, für jeden Mitarbeitenden einen einheitlichen Rahmen für den Umgang mit KI-Technologie zu bilden. Die Police ist ein Leitfaden für den Mitarbeitenden und beinhaltet alle Verpflichtungen aus geltenden Regularien, ausgearbeitet für den Unternehmenskontext. Rollen, Verantwortlichkeiten und Ansprechpartner werden definiert. Zulässige und nichtzulässige Werkzeuge werden festgehalten und die entsprechenden Kriterien werden genannt. Eine KI-Police umfasst die Prozesse zur Prüfung von Anwendungsfällen und zur Einhaltung der Regularien durch die Mitarbeitende.

Im Gegensatz zur Strategie betrifft eine Police direkt die Gegenwart und den Alltag der Mitarbeitenden.

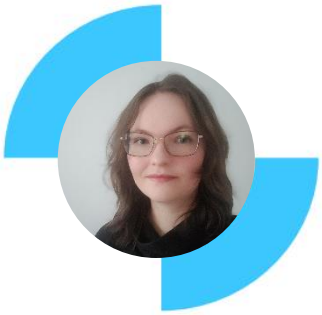
Betriebsvereinbarungen

Betriebsratsvereinbarungen werden zwischen den Mitbestimmungsgremien und dem Arbeitgeber getroffen und regeln Arbeitsbedingungen und betriebliche Belange. In Bezug auf künstliche Intelligenz können hier Informationspflichten und Schulungen geregelt und Datenschutz- und Sicherheitsmaßnahmen erfasst werden. Die Risiken von künstlicher Intelligenz für Mitarbeitende kann erfasst und bewertet werden.

Die relevanten Regularien zu künstlicher Intelligenz bilden ein bindendes Rahmenwerk für den Umgang mit KI und sollen in unternehmenseigene Policen und Strategien einfließen, um Mitarbeitenden einen einheitlichen Startpunkt für das Auseinandersetzen mit den Technologien geben und verlässliche Anwendungen zu entwickeln.

Autorin

Autorin



Lea Elina Kleemann
AI Application Developerin
Atos Deutschland

Über Atos

Atos ist ein weltweit führender Anbieter für die digitale Transformation mit ca. 82.000 Mitarbeitern und einem Jahresumsatz von zirka 10 Milliarden Euro. Als europäischer Marktführer für Cybersecurity sowie Cloud und High Performance Computing bietet die Atos Gruppe maßgeschneiderte, ganzheitliche Lösungen für sämtliche Branchen in 69 Ländern. Als Pionier im Bereich nachhaltiger Dienstleistungen und Produkte arbeitet Atos für seine Kunden an sicheren, dekarbonisierten Digitaltechnologien. Atos ist eine SE (Societas Europaea), die an der Börse Euronext Paris notiert ist.

Das Ziel von Atos ist es, die Zukunft der Informationstechnologie mitzugestalten. Fachwissen und Services von Atos fördern Wissensentwicklung, Bildung sowie Forschung in einer multikulturellen Welt und tragen zu wissenschaftlicher und technologischer Exzellenz bei. Weltweit ermöglicht die Atos Gruppe ihren Kunden und Mitarbeitern sowie der Gesellschaft insgesamt, in einem sicheren Informationsraum nachhaltig zu leben, zu arbeiten und sich zu entwickeln.

Weitere Informationen finden Sie unter www.atos.net

Über Tech Foundations

Tech Foundations umfasst den Managed-Services-Geschäftsbereich der Atos Gruppe mit Fokus auf Hybrid Cloud Infrastructure, Employee Experience und Technology Services. Mit seinen dekarbonisierten, automatisierten und KI-gestützten Lösungen ist Tech Foundations führend in diesem Bereich und treibt mit seinen 41.000 Mitarbeitern Themen voran, die Unternehmen, Institutionen und die Gesellschaft weltweit am dringendsten beschäftigen. Das Unternehmen ist in 69 Ländern vertreten und erzielt einen Jahresumsatz von ca. 5 Milliarden Euro.

Atos