# Künstliche Intelligenz konkret

- Was wir von einer KI-Anwendung lernen können -

# KI – Lösung aller Probleme?

- Die **Veröffentlichung von ChatGPT** hat die Diskussion über die Anwendung und Folgen der Künstlichen Intelligenz angeheizt.
- Sehen die einen die KI als einen **Beitrag zur Lösung** vieler unserer aktuellen Probleme (Gesundheit, Klima), glauben die anderen, dass eine wildgewordene KI möglicherweise die **Weltherrschaft** an sich reißen könnte.
- Tatsächlich geben schon die ersten Schüler und Studenten ihre von ChatGPT verfassten Arbeiten ab und einige Zeitungen veröffentlichen z.B. Sport- und Wirtschaftsberichte, die ebenfalls mit Hilfe von ChatGPT erstellt wurden.
- Wir “normalen” Beobachter müssen also versuchen, wenigstens im Ansatz zu verstehen, wie KI-Anwendungen funktionieren, damit wir uns zu diesem Thema eine halbwegs angemessene Meinung bilden können.

# Wir untersuchen eine KI-Anwendung

- Wir nehmen heute eine spezielle KI-Anwendung etwas genauer unter die Lupe und schauen der Künstlichen Intelligenz bei der Arbeit zu.
- Wir verfolgen, wie eine KI-Anwendung verschiedene Arten von Schwertlilien erkennt. Dazu verwenden wir den öffentlich zugänglichen **Iris-Datensatz** und eine Software-Bibliothek von **Scikit-learn**.
- Bei Scikit-learn handelt es sich um eine Software-Bibliothek für maschinelles Lernen. Die Software ist frei erhältlich und für die Programmiersprache Python vorgesehen. Für das maschinelle Lernen stehen verschiedene Algorithmen wie Clustering-, Regressions- oder Klassifizierungsalgorithmen zur Verfügung. Die Bibliothek zeichnet sich durch ihre robusten und gut dokumentierten Funktionen aus.
- Aus unseren Beobachtungen werden wir versuchen einige allgemeine Aussagen über Möglichkeiten, Grenzen und Auswirkungen von KI-Anwendungen zu gewinnen und dann zu diskutieren.

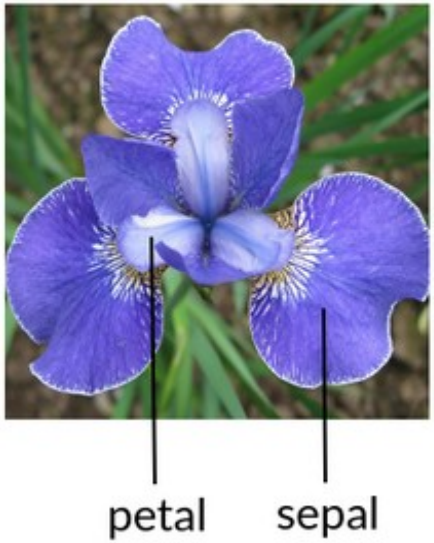
# Was ist Künstliche Intelligenz?

- Unter dem Begriff **“Künstliche Intelligenz”** (KI) versteht man Programme, die aufgrund von Mustern und Wahrscheinlichkeiten, die sie in einer Lernphase in Trainingsdaten gefunden haben, **eigene Entscheidungen** treffen.
- Auch die Programmierer wissen nicht, wie das Programm genau zu einer bestimmten Entscheidung gekommen ist, der Entscheidungsprozess findet in einer **Black Box** statt.
- Konnte man nach der Erfindung der Dampfmaschine verstehen, was im Inneren dieser Maschinen passiert, wird bei KI-Anwendungen der **Prozess der Entscheidungsfindung selbst nicht vollständig verstanden**, sondern er kann nur durch Trainingsdaten und den Aufbau des Algorithmus beeinflusst werden.
- Was ist ein Programm? (Demo)

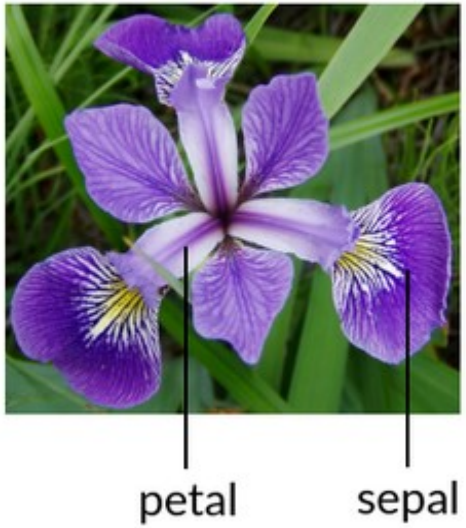
# Welche Schwertlilie haben wir?

Einer KI-Anwendung wird die Aufgabe gestellt, eine von drei Arten von Schwertlilien (Iris) zu erkennen. Wir Menschen würden uns die Form und Farben der Blätter einprägen und nach einiger Übung die richtige Art erkennen.

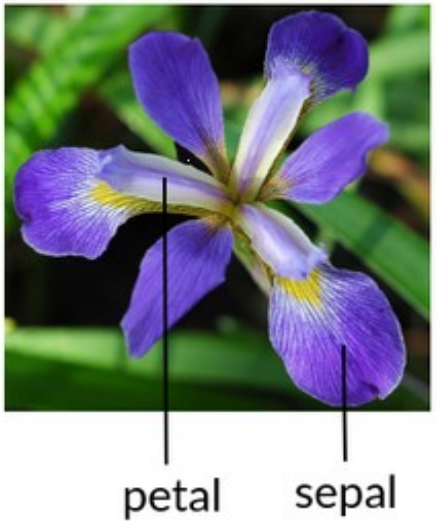
**iris setosa**




**iris versicolor**

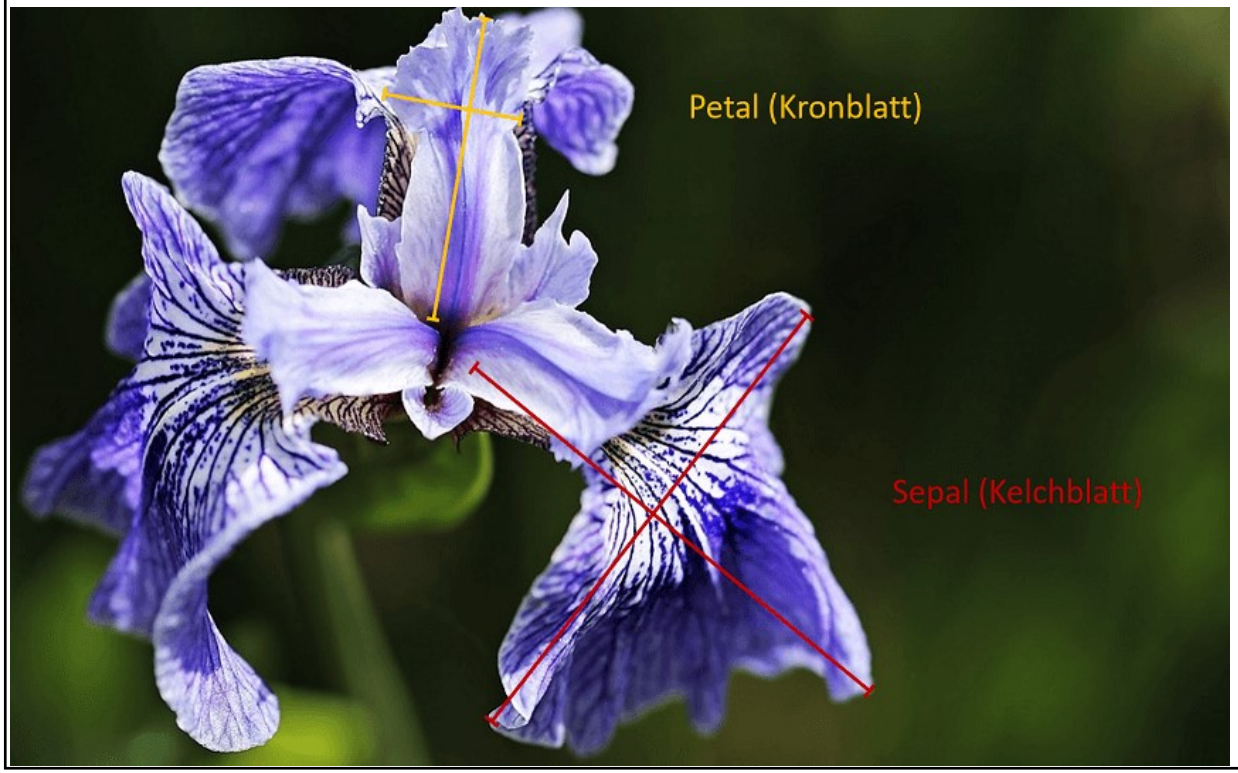



**iris virginica**



# Welche Schwertlilie haben wir?

Der Computer hingegen braucht für diese Aufgabe Merkmale, die sich in Zahlen ausdrücken lassen: Hier sind das die Länge und die Breite eines Blattes in cm.




Petal = Kronblatt
Sepal = Kelchblatt

Iris-Setosa
Iris-versicolor
Iris-virginica

# Der Iris-Datensatz

Im frei verfügbaren Iris-Datensatz sind 150 Beobachtungen von je vier Eigenschaften von Schwertlinien zusammengestellt. Diese Daten sind die Grundlage für das Training der KI-Anwendung.

| Sepal-Länge (cm) | Sepal-Breite (cm) | Petal-Länge (cm) | Petal-Breite (cm) | Species |
|---|---|---|---|---|
| 5.1 | 3.5 | 1.4 | 0.2 | Iris-setosa |
| 4.9 | 3.0 | 1.4 | 0.2 | Iris-setosa |
| 4.7 | 3.2 | 1.3 | 0.2 | Iris-setosa |
| 4.6 | 3.1 | 1.5 | 0.2 | Iris-setosa |
| 5.0 | 3.6 | 1.4 | 0.2 | Iris-setosa |
| 5.4 | 3.9 | 1.7 | 0.4 | Iris-setosa |
| 4.6 | 3.4 | 1.4 | 0.3 | Iris-setosa |
| 5.0 | 3.4 | 1.5 | 0.2 | Iris-setosa |
| 4.4 | 2.9 | 1.4 | 0.2 | Iris-setosa |
| 4.9 | 3.1 | 1.5 | 0.1 | Iris-setosa |
| 5.4 | 3.7 | 1.5 | 0.2 | Iris-setosa |
| 4.8 | 3.4 | 1.6 | 0.2 | Iris-setosa |
| 4.8 | 3.0 | 1.4 | 0.1 | Iris-setosa |

# Trainings- und Testdaten

Die Irisdaten werden aufgeteilt in Trainings- und Testdaten. Die **Trainingsdaten** enthalten die Längen- und Breitenangaben der Blätter sowie den Namen der Art. Diese Daten werden von der KI-Anwendung zum Lernen benutzt.

| | | | | |
|---|---|---|---|---|
| 7.0 | 3.2 | 4.7 | 1.4 | Iris-versicolor |
| 6.4 | 3.2 | 4.5 | 1.5 | Iris-versicolor |
| 6.9 | 3.1 | 4.9 | 1.5 | Iris-versicolor |
| 5.5 | 2.3 | 4.0 | 1.3 | Iris-versicolor |
| 6.5 | 2.8 | 4.6 | 1.5 | Iris-versicolor |
| 5.7 | 2.8 | 4.5 | 1.3 | Iris-versicolor |

Die **Testdaten** enthalten den Namen der Art nicht mehr. Die Testdaten dienen dazu, die Erkennungsrate der KI-Anwendung nach dem Training zu prüfen.

| | | | | |
|---|---|---|---|---|
| 5.2 | 2.7 | 3.9 | 1.4 | |
| 5.0 | 2.0 | 3.5 | 1.0 | |
| 5.9 | 3.0 | 4.2 | 1.5 | |
| 6.0 | 2.2 | 4.0 | 1.0 | |
| 6.1 | 2.9 | 4.7 | 1.4 | |

# Qualität der Trainingsdaten

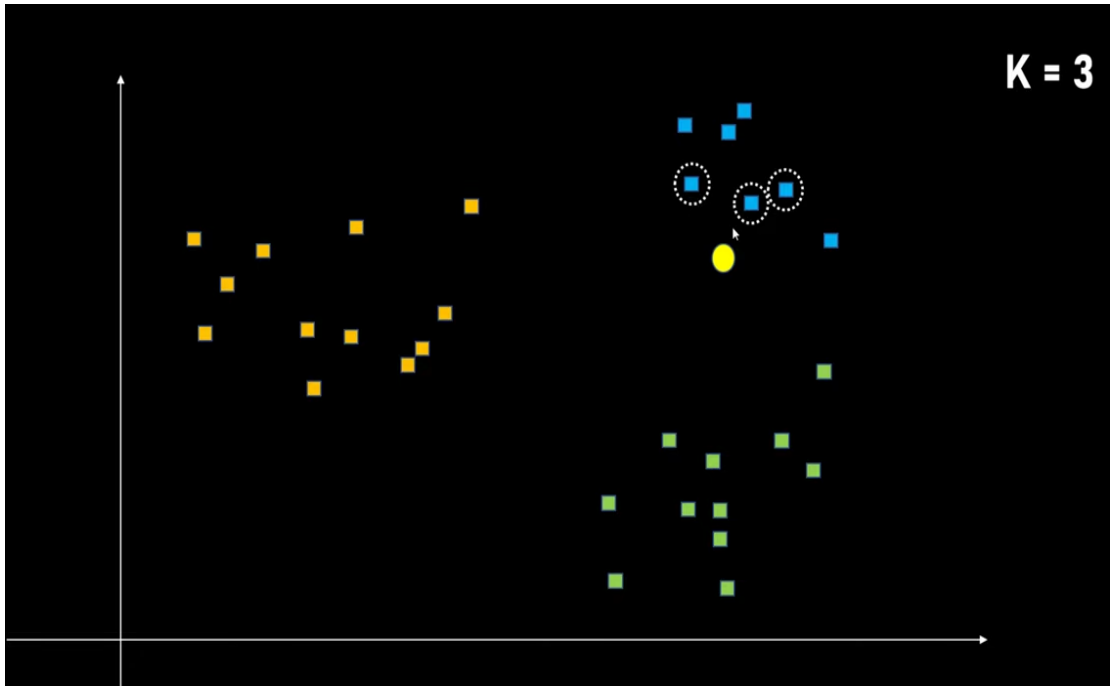
- Die **Erkennungsrate** einer KI-Anwendung hängt ganz wesentlich von der Qualität der Trainingsdaten ab.
- So müssen die Maße der Blätter stimmen, d.h. sie müssen richtig gemessen und richtig eingetippt worden sein. Die Benennung der Art muss den Messwerten entsprechen. Die Daten jeder Iris-Art müssen etwa gleich häufig vorkommen.
- In **unserem Beispiel** wird der Iris-Datensatz nach Zufall aufgesplittet in **70% Trainings-** und **30% Testdaten.** Damit stehen 105 Datensätze zum Training und 45 Datensätze zum Testen zur Verfügung.
- **105 Trainingsdatensätze** sind in einer realen KI-Anwendung viel **zu wenig.**
- **Für die Diskussion:** Wer mit KI-Anwendungen Erfolge erzielen will, braucht eine goße Menge hochwertiger Trainingsdaten. (Demo)

# Erkennungsrate

- Unsere Anwendung erreicht nach dem Training mit den 105 Trainingsdatensätzen eine **Erkennungsrate von 97,8 %.**
- Von 100 Versuchen werden also knapp 98 richtig erkannt , was für die wenigen Trainingsdaten und die kurze Trainingsphase auf meinem billigen PC schon recht ordentlich ist.
- Wie lässt sich die **Erkennungsrate verbessern**?
  - Die Zahl der Trainingsdatensätze erhöhen
  - Die Zahl der Trainingsdurchgänge variieren
  - Den Algorithmus optimieren
- Es bleibt aber ausgesprochen schwierig, unter allen Umständen eine Erkennungsrate von 100% zuverlässig zu erreichen. Das ist auch ein Grund, warum die Autoindustrie sich von der Stufe 5 des autonomen Fahrens verabschiedet hat.

# Wie funktioniert die Erkennung?

Der **KNeighborsClassifier** ist ein Algorithmus, der einen bestimmten Datenpunkt nach den Trends kategorisiert, die in den am nächsten liegenden Datenpunkten (Nachbarn, K=3) gefunden werden.




Wenn ein neuer Datenpunkt (gelb), bestimmt werden soll, wird die Entfernung zu seinen Nachbarn berechnet. In diesem Beispiel ist es sehr wahrscheinlich, dass der neue Datenpunkt den blauen Quadraten zugewiesen wird.

# Grenzen des Algorithmus

- Die **Erkennungsrate von 98%** wurde schon genannt; von 100 Anfragen sind 2 falsch, wir wissen aber nie welche.
- Es werden offensichtlich **unsinnige Anfragen**, wie z.B. [0.0, 0.0, 100.00, 100.00] als blaue virginica erkannt. Der Algorithmus weiß also nichts über Schwertlilien, während der Mensch sofort erkennen würde, dass es ein Blatt mit diesen Maßen nicht gibt.
- Der Algorithmus wird alle Maßanfragen mit seinem Rechenweg bearbeiten, auch wenn es sich nicht um Maße von Schwertlinien, sondern z.B. um die von Geldscheinen handelt. Auch hier gilt wieder: **Der Algorithmus weiß nichts von der Welt.**
- Der Algorithmus beruht auf dem intensiven Berechnen von Abständen und deren Bewertung. Wir haben es hier mit einem **mathematischen Modell** zu tun, das **berechnet mit welcher Wahrscheinlichkeit Daten zusammenpassen**. Für die Diskussion: Es entsteht die Frage, ob dies schon als intelligentes Verhalten bezeichnet werden kann.

# Einige alltägliche KI-Anwendungen

- **Empfehlungsalgorithmen sammeln** Daten über Ihr Nutzungsverhalten und zeigt Ihnen dann auf Basis Ihrer Präferenzen **Empfehlungen** an. Der Empfehlungsalgorithmus beruht auf der Anwendung Künstlicher Intelligenz.
- **Optimierung von Suchmaschinen:** Künstliche Intelligenz hilft dabei, die Ergebnisse von Suchanfragen zu analysieren und dabei nicht nur **natürliche Sprache** zu berücksichtigen, sondern auch Verknüpfungen zwischen verschiedenen Suchinhalten herzustellen, was die Passgenauigkeit der Suchergebnisse immer weiter verfeinert.
- Wenn Sie an Ihrem Smartphone die Bildschirmsperre durch **Gesichtserkennung** aufheben, werden Sie sich vermutlich schon gefragt haben, wie die Software Ihr Gesicht auch mit wechselndem Make-Up oder bei Dämmerlicht erkennt. Dahinter steckt ein **KI-Algorithmus**, der stetig dazulernt und jedes Mal, wenn Sie Ihr Gesicht vor die Kamera halten, besser darin wird, dieses zu erkennen.
- **Spielberichte, Börsennachrichten oder Wettervorhersagen:** Meldungen, die stark **datenbasiert** sind, werden nicht selten von einer Künstlichen Intelligenz verfasst. Betreiber von Webseiten oder Zeitungsverlage können diese Meldungen bei Unternehmen bestellen, deren KI innerhalb kürzester Zeit eine Vielzahl von unterschiedlichen Berichten produzieren kann, die allesamt auf Daten basieren.
- Auch **Übersetzungsprogramme –** das bekannteste von ihnen vermutlich Google Translate – basieren auf Künstlicher Intelligenz. Sicher haben Sie bemerkt, dass die Übersetzungen von Google Translate und Co.. in den letzten Jahren immer besser geworden sind.

# Auswirkungen: Bedeutung der Trainingsdaten

- Der vorgestellte Algorithmus hat zugegebenermaßen schon einige Jahre auf dem Buckel. Im Zuge der rasanten technischen Entwicklung werden sicher eine Reihe von “Noch-Problemen” bald verschwinden.
- Wir haben die **Bedeutung der Trainingsdaten** genannt. Schon heute muss es darum gehen, offen zu legen mit welchen Daten eine KI trainiert wurde.
- Das Anhäufen riesiger Datenmengen in privater Hand verschafft den Besitzern **machtpolitischen Einfluss**. Dieser muss begrenzt und die Datenbestände sollten Allgemeingut werden.
- Das Nutzen offen zugänglicher Daten als Trainingsmateral für kommenzielle Anwendungen kann private **Urheberrechte verletzen.** Die «New York Times» hat als erste große Zeitung die Software-Unternehmen Open AI und Microsoft wegen Chat-GPT verklagt.
- “**Data-Colonialismus**” meint die Ausbeutung von prekären Klickworkern bei der Datenbeschaffung und Erfassung.

# Auswirkungen auf die Bildung

- Wir sollten weiterhin alles über Schwertlilien lernen, denn die Auskünfte der KI sind ja nicht immer richtig. Letztendlich müssen wir **Menschen entscheiden**, was stimmt und was nicht.
- Immer mehr an Bedeutung gewinnen wird die Fähigkeit, die Quelle, den Verfasser, dessen Interesse an seiner Veröffentlichung und **den Wahrheitsgehalt seiner Information zu prüfen**.
- Der Deutsche Ethikrat befürchtet einen **Verlust menschlicher Fähigkeiten**, wenn wir zuviele kognitive Tätigkeiten an die KI delegieren (siehe die Beispiele Taschenrechner, Naivigationssysteme und ChatGPT).
- Wenn wir z.B. weniger **schreiben** würden, weil das ChatGPT für uns erledigt, hätte das negative Auswirkungen auf unser Denken, da wir beim Schreiben unsere Gedanken sammeln, ordnen und ausdrücken (es grüßt herzlich PISA 2023!).

# Auswirkungen auf die Demokratie

- Durch die mit Hilfe der KI leicht zu erzeugenden Texte und Meldungen haben wir eine neue Qualität von Informationen, bei denen wir ebenso wie bei den “sozialen” Medien, dem Privatfernsehen und der Bild-Zeitung den **Wahrheitsgehalt nur schwer prüfen können.** Geoffrey Hinton, ein Pate der KI, sieht das Risiko, dass die Menschen “nicht mehr in der Lage sein werden, zu wissen, was wahr ist. “
- Ein Beispiel: https://youtu.be/UL6dOto-yQM?si=N0q9AgEbKBVwyOGz
- Immer mehr Fälschungen untergraben bei den Menschen den Glauben an überprüfbare Tatsachen, der Zerfall der Öffentlichkeit in Filterblasen und Glaubensgemeinschaften wird weiter gefördert.
- Die **“Weltanschauung” der KI** gründet sich auf die verwendeten Algorithmen und auf ihre Trainingsdaten. Wir müssen also wissen, wessen Geistes Kind eine KI-Anwendung und ihre Trainingsdaten sind.
- KI schafft “verbesserte” Möglichkeiten personalisierter Manipulation.

# Auswirkungen auf die Gesellschaft

- Folgt man **Armin Nassehi** und seinem Buch **„Muster“**, so helfen Digitalisierung und KI der Gesellschaft, d.h. Unternehmen, Staaten, Verwaltungen, Strafverfolgungsbehörden, Wissenschaft usw. beim Erkennen von vorhandenen, aber **bisher unsichtbaren Mustern.**
- Auch wenn das Erkennen von Mustern heute mit Hilfe einer wesentlich höher entwickelten digitalen Technik geschieht als in der vordigitalen Welt, der Zweck bleibt der gleiche: Das Erkennen dieser Muster verspricht einen **großen gesellschaftlichen Nutzen.**
- Die Komplexität der gesellschaftlichen Moderne wurde durch die fortschreitende Digitalisierung überhaupt erst sichtbar. Nur so sei es möglich, die Funktionsweisen der modernen Gesellschaft sehr viel präziser nachzuvollziehen als davor. Zum Beispiel lässt sich **menschliches Verhalten** mit Hilfe gefundener Muster **besser voraussagen.**

# Auswirkungen auf die Gesellschaft

- Der Ansatz von Nassehi  blendet aber leider das Machtgefälle aus, dass in der Anwendung von KI liegen kann:
- Wer kommt mit welchen Methoden in den Besitz der zum Training der KI-Anwendungen notwendigen Daten?
- Welche „Weltanschauung“ liegt den eingesetzten Algorithmen zugrunde?
- Das Beispiel China zeigt, das KI massiv zur Überwachung der Bevölkerung eingesetzt werden kann.
- Beispiele hierzu sind Gesichtserkennung – auch mit Atem-schutzmaske -  und Zensur im digitalen Raum.
- Chinesische Firmen bieten Programme an, die das Ausrollen eines Transparents in Reichweite einer Überwachungskamera automatisch erkennen.

# Regulierungs-Maßnahmen (EU AI Act)

- Im Juni 2123 haben die Abgeordneten des EU-Parlaments ihre Verhandlungsposition zum Gesetz über künstliche Intelligenz angenommen. Im Dezember 2023 haben sich das EU-Parlament und die Vertreter der EU-Staaten auf das **weltweit erste KI-Gesetz** geeinigt. Danach haben die Staaten ca. 2 Jahre Zeit, die Regelungen in Nationales Recht umzusetzen. Es bleibt abzuwarten, ob die Lobbyisten von Microsoft, OpenAI und Co. die Umsetzung noch verhindern oder verwässern können.
- Das Europäische Parlament will vor allem sicherstellen, dass die in der EU eingesetzten **KI-Systeme sicher, transparent, nachvollziehbar, nicht diskriminierend und umweltfreundlich** sind. KI-Systeme sollten von Menschen und nicht von der Automatisierung überwacht werden, um schädliche Ergebnisse zu verhindern. Das Parlament möchte außerdem eine technologieneutrale, einheitliche Definition für KI festlegen, die auf zukünftige KI-Systeme angewendet werden könnte.
- Die Anwendung  von KI-Systemen unterliegt bisher in Deutschland keinen spezifischen, d. h. auf KI-Systeme besonders zugeschnittenen, nationalen Gesetzen und Verordnungen (Wissenschaftliche Dienste Deutscher Bundestag, 2023).

# Regulierungs-Maßnahmen (EU AI Act)

| RISIKOKLASSE | BEISPIELE | FOLGEN BZW. ANFORDERUNGEN |
|---|---|---|
| **Verbotene KI-Systeme (Art 5 EU AI Act)** | Systeme zur unterschwelligen bzw. absichtlichen Verhaltensmanipulation, „social-scoring"-Systeme oder biometrische Echtzeit-Fernidentifizierungssysteme. | KI-Systeme mit unannehmbarem Risiko sind von vornherein verboten. |
| **Hochrisiko-KI-Systeme (Art 6 ff EU AI Act)** | KI-Systeme, die als Sicherheitskomponenten in der Verwaltung und dem Betrieb des Straßenverkehrs sowie in der Wasser-, Gas-, Wärme- und Stromversorgung eingesetzt werden, sowie Bewerbertools oder Verfahren zur Kreditwürdigkeitsprüfung. | • Einrichtung eines Risikomanagementsystems,<br>• besondere Verfahren der Datenkontrolle und Daten-Governance,<br>• technische Dokumentation,<br>• Aufzeichnungspflichten,<br>• Sicherstellung von Transparenz und die Bereitstellung von Informationen für die Nutzer,<br>• Sicherstellung der menschlichen Aufsicht und<br>• Einhaltung das Gebot von Genauigkeit, Robustheit und Cyber-Sicherheit. |
| **Geringes Risiko (Art 52 EU AI Act)** | Chatbots, Anwendungen, die Bild-, Ton bzw. Videoinhalte erzeugen und manipulieren (zB Deepfakes). | Transparenz- bzw. Offenlegungsanforderungen. |
| **Minimales Risiko (Art 69 EU AI Act)** | „Auffangnetz" für Anwendungen, von denen grundsätzlich nur minimale Risiken ausgehen, wie zB Spam-Filter oder KI-gestützte Videospiele. | Selbstregulierung durch freiwillige Verhaltenskodizes. |

# Künstliche Intelligenz konkret

"Wir neigen dazu, die kurzfristigen Auswirkungen einer Technologie zu überschätzen und die langfristigen Auswirkungen zu unterschätzen." (Roy Amara)

Ich bedanke mich für Ihre Aufmerksamkeit.

Literatur-Empfehlungen:

Philip Häusser: Natürlich alles künstlich – Was KI kann und was (noch) nicht, Droemer, 2021

Armin Nassehi: Muster – Theorie der digitalen Gesellschaft, bpb Bonn, 2020