KI mit IQ?

"Künstliche Intelligenz" hat weder Intelligenz noch Taktgefühl

Manfred Rotermund

Hinweis zur Betrachtung der PDF-Datei am Bildschirm.

Schauen Sie sich die Datei als **Diashow** an, dann können Sie in der Vollbild-Ansicht in Ihrem Tempo von Folie zu Folie wechseln.

Diese Präsentation enthält weitaus mehr als die Folien während der Vortrags. Ich habe sie vor allem zum Thema "Gehirn", am Ende um Informationen zur Kontrolle der KI und um "Weitergehende Gedanken" ergänzt. Einige Folien geben detaillierte Informationen oder Beispiele, die zum Verständnis der Präsentation nicht unbedingt notwendig sind. Ein Link unten links gibt Ihnen die Möglichkeit, diese zu überspringen.

Copyright 2024 Manfred Rotermund

Für geschäftliche Zwecke und Data Mining behält sich der Verfasser alle Rechte vor. Kontakt über Manfred.Rotermund@ruhr-universität-bochum.de

KI mit IQ? - Inhalt

- KI ist da
- Die KI-Gemeinde von Dystopien und Utopien Gehe zu Grundlagen zum Computer und Gehirn Gehe zu
- Wie funktioniert KI?Gehe zu
 - Wie lernen Maschinen?
 - Wie funktioniert ChatGPT
 Gehe zu
- Ist KI intelligent?Gehe zu
- Anwendungsbereiche der KI, Chancen und Gefahren Gehe zu
- Kontrolle der KI
 Gehe zu
- Weitergehende Gedanken
 Gehe zu
- Quellen
 Gehe zu

KI ist da!





Quelle: Unbekannt

- 1996 Computer schlägt Schachweltmeister
- seit 2010 sinkt die Fehlerrate bei der Bilderkennung auf 2,25%
- 2015 schlägt AlphaGo den europäischen GO-Meister
- Nov. 2022 ChatGPT

KI ist da!

 Sie wird "Entscheidungen über uns, mit uns und für uns treffen."
 (Zweig: 9)

Dies erzeugt Erwartungen, Hoffnungen, aber auch Ängste.

Die KI-Gemeinde: Utopien und Dystopien

"Ein Wunder"!

ChatGPT versteht die Welt.

Bast im Gespräch mit Lesch (20:46, 24:04). Hannah Bast ist **Professorin für Algorithmen und Datenstrukturen** an der Albert-Ludwigs-Universität Freiburg. Sie wurde mehrfach für ihre Forschung zu angewandten Algorithmen ausgezeichnet und ist <u>Sachverständige der Enquete-Kommission Künstliche Intelligenz.</u>

"Diese Art von Diskussionen ist die bevorzugte Domäne der Informatiker und der Technologieelite aus dem Silicon Valley." (Ulbrich/Gasser).

Die KI-Gemeinde: Dystopien

Bast (2023): "Ich behaupte, das werden irgendwann schon… also jeder gute <u>Science-Fiction-Film</u> geht in die Richtung - irgendwann werden das Akteure werden… dann wird es gefährlich."

(Quelle: H. Lesch, M. Smolla & H. Bast: ChatGPT & Co – Wie intelligent ist KI? (Min.: 20:01, 23:08)

Musk + 33706 andere: "Sollten wir den Verlust der Kontrolle über unsere Zivilisation riskieren?"

Quelle: Musk u.a. - future of life

"Diese Art von Diskussionen ist die bevorzugte Domäne der Informatiker und der Technologieelite aus dem Silicon Valley." (Ulbrich/Gasser).

Die KI-Gemeinde: Utopie

Musk u.a.: "KI kann unglaubliche Vorteile bringen, unlösbare globale Probleme lösen und inspirierende Zukünfte verwirklichen."

Quelle: Musk u.a. - future of life

"Diese Art von Diskussionen ist die bevorzugte Domäne der Informatiker und der Technologieelite aus dem Silicon Valley." (Ulbrich/Gasser).

Ziele meines Vortrags

Ich möchte Sie

- ermächtigen mit der KI-Gemeinde und in der Öffentlichkeit kompetent über KI zu sprechen und
- ermutigen, Mitsprache und Kontrollrechte für sich und alle Menschen einzufordern

Den ausführlichen Aufsatz und diese Präsentation zu diesem Thema finden Sie auf der Webseite von Attac.

Die Basis der KI: Der von Neumann Rechner

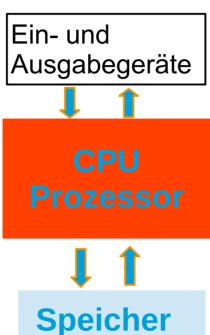
1978



Intel 8086 Größe der CPU a 3 x 2 cm Transistoren 29.000 Hruska

1981 1 CPU: Intel 8086

IBM PC 8086





2024

Cerebras' third-generation waferscale engine (WSE-3) für Al

Größe: 30 x 30 cm Transistoren 4 Trillionen

Cerebras



2022

9408 CPUs 37.000 GPUs

Frontier 2022



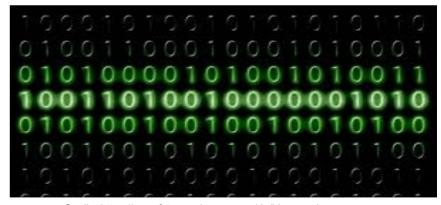
Transistoren – Die Basis des Prozessors

Quelle. https://www.elektronikkompendium.de/sites/praxis/ bauteil_transistor_npn.htm

- Transistoren
 - Schalter: Stromfluss
 - ein, 1 oder
 - aus, 0.
 - **fest verdrahtet**, feste Schaltungsstruktur.

Die Sprache der Transistoren (Maschinensprache)

digitaler Binärcode



Quelle: https://www.futura-sciences.com/de/binaercodewas-ist-das-definition_10610/

Die Basis-Befehle des Prozessors

Mathematische Operationen	Logische Operationen
AdditionSubtraktionMultiplikationDivision	 - AND (gibt nur bei zwei 1en eine 1 aus), - OR (gibt beimindestens einer 1 eine 1 aus), - XOR (gibt bei genau einer 1 eine 1 aus) - NOT (kehrt jedes Bit um, aus 1 wird 0 und aus 0 wird 1)

Die Ergebnisse des Prozessors

- Ergebnisse sind Rechenergebnisse.
 Sie basieren auf logischen und mathematischen Operationen, die der PC im Binärformat durchführt.
- Der PC kennt keine Bilder und Wörter mit Buchstaben.
 Er kennt nur den Binärcode.

Informationen über das Gehirn überspringen Zurück zum Inhaltsverzeichnis



downloaded from Freepik: www.freepik.com



 Quelle: https://www.futurasciences.com/de/binaercode-was-istdas-definition 10610/

Das Gehirn -Nervenzellen und Synapsen

- Gehirn 1230 ccm; 1,3 1,5 kg;
- 86 Mrd. Nervenzellen (Neuronen)
 - 86 Mrd. nicht neuronaler Zellen (Schutz der Neuronen)
- 150 Tsd. bis 3 Mio. km Nervenfasern (Video: Overview of the Human Brain Project)
- 100 bis 164 Bill. Synapsen als Verbindungsstellen zwischen Neuronen

Quellen: Jahn Human Brain Projekt der EU (2012 - 2023)

Gehirn - Funktionen



Studyflix: Gehirn

Die Funktionen des Gehirns sind

- Reize/Signale (innere u. äußere)
 - Empfängt Signale
 - speichert und verarbeitet Signale (lernen)
 - Sendet Signale
- Sinneswahrnehmungen
- Denken
- Sprechen
- Bewegungssteuerung
- Emotionen

Gehirn - Funktionen

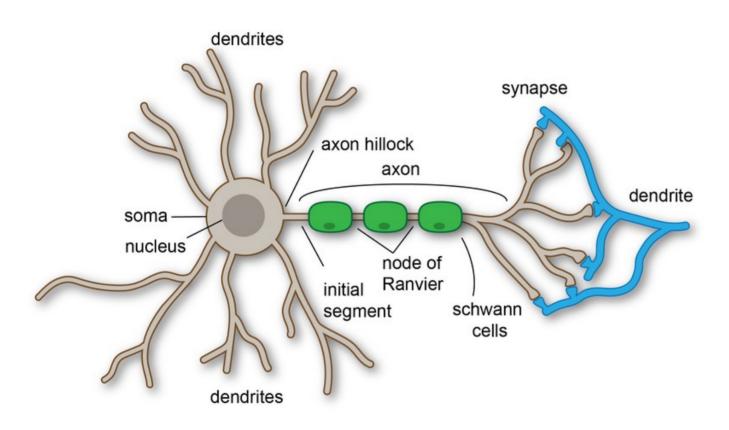


Ergänzung

Studyflix: Gehirn

- Sprache (Begriffsbildung) und Sprechen
- Emotionen

Gehirn – 86 Mrd. Neuronen (Nervenzellen)



- → Synapsen pro Neuron:
 - Bis zu 10.000
 - Ø 1.000
 - Max.: 200.000
- → An einem Axon und den Dendriten kann es Tausende von synaptischen Endköpfchen geben.

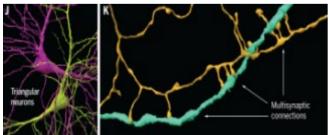
Jedes Neuron ist mit Tausenden anderer Neuronen verbunden (Bis zu 10.000 Verbindungen)

Quelle: Neveu

Das Gehirn -Nervenzelle und Synapse

Dendriten: elektrische Impulse

Axonhügel am Beginn des Axons berechnet, ob die Zelle einen elektrischen Impuls weitergibt



Zellen unter dem Elektronenmikroskop Quelle: Shapson-Coe u.a.

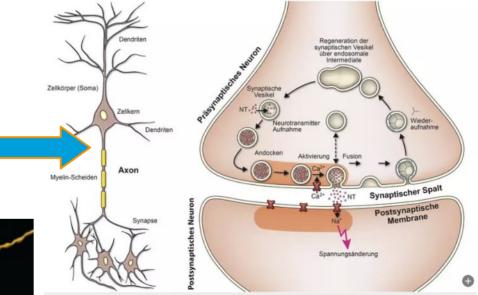


Abb. 1. Schematische Darstellung einer Nervenzelle (links) und einer Synapse (rechts).

© MPI für biophysikalische Chemie

Quelle: Jahn

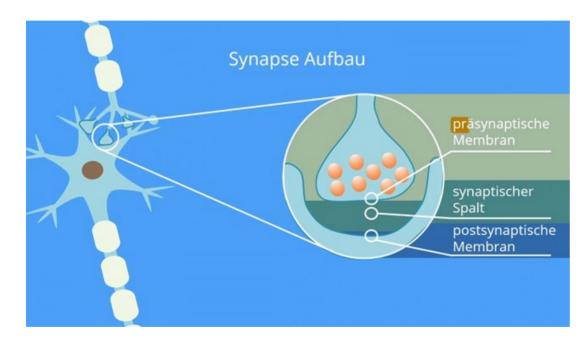
Axon: elektrischer Impuls

präsynaptisches
Endköpfchen:
Umwandlung des
Impuls' in Neurotransmitter (NT), die
das Endköpfchen
verlassen können
und über den
synaptischen Spalt

postsynaptischen
Endköpfchen des
Dendriten wandern:
Aufnahme des NT
Umwandlung in einen
elektrischen Impuls

zum

Der Vorteil des Gehirns - Die chemische Informationsübertragung



Studyflix: Neurobiologie

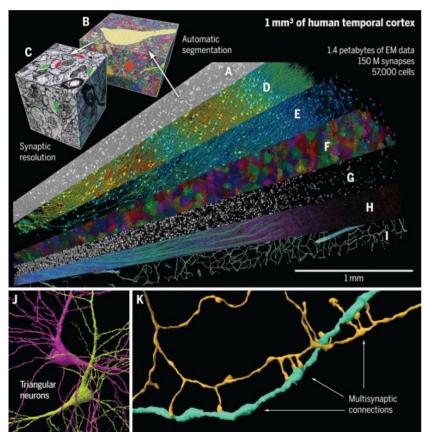
Vorteile der chem. Übertragung

- flexibler,
- komplexer,
- vielfältigere Signalübertragung

Vorteile der synaptischen Verbindungen

- Synapsen sind z.T. keine festen Verbindungen.
- Sie sind plastisch und spiegeln Erfahrungen wider.
- Ständige Selbstorganisation des Gehirns

Gehirn: 1mm³ des temporalen Cortex



A range of histological features in 1 mm³ of human brain were rendered, including

- neuropil (A) and
- its segmentation (B) at nanometer resolution,
- annotated synapses (C),
- excitatory neurons (D),
- inhibitory neurons (E),
- astrocytes (F),
- oligodendrocytes (G),
- myelin (H),
- and blood vessels (I).

A previously unrecognized neuronal class (J) and multisynaptic connections (K) were also identified.

Shapson-Coe u.a.

Das Gehirn -Synaptisches Vesikel

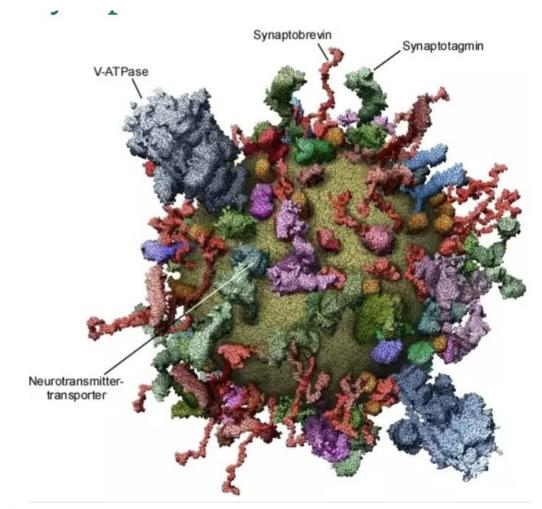
Vesikel sind im Endköpfchen für die Aufnahme der Neurotranmitter, deren Tranport zur Zellmembran und der Öffnung der Zellmembran verantwortlich.

<u>V-AT Pase</u>: Energieversorgung durch (Adenosintriphosphat (ATP))

Neurotransmitter-Transporter: (NT-Aufnahme)

Synaptobrevin: Verschmelzung mit der Membran

Synaptotagmin: Lösung von der Membran

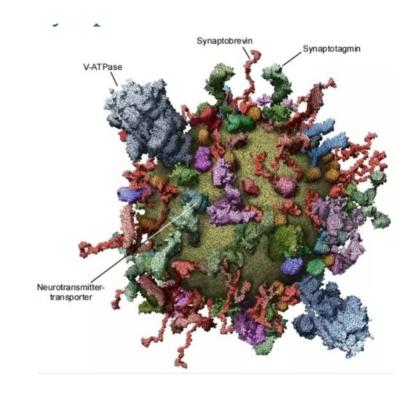


Quelle: Jahn

Das Gehirn -Synaptisches Vesikel und Neurotransmitter

Vesikel können mehr als einen Neurotransmitter enthalten. Welche NT von dem Vesikel aufgenommen werden und welche Vesikel aktiviert werden, ist abhängig von

- 1. dem **elektrischen Impuls** des Axons: Stärke, Frequenz, Dauer, der
 - die r\u00e4umliche und zeitliche Verteilung der Calcium-Ionen und
 - → die Chlorid-Konzentration beeinflusst
- 2. der spezifischen **molekularen Zusammensetzung der Vesikel**membranproteine.



"Trotz großer Fortschritte sind die komplexen molekularen Prozesse immer noch nicht vollständig verstanden…"

Quelle: Jahn

Prozessor und Gehirn

Prozessor	Gehirn
Trennung von Prozessor und Speicher	Einheit von Gedächtnis, Arbeitseinheit (Wahrnehmung, Denken, Steuerung) und Gefühlen
Fest verdrahtete Transistoren	Dynamisch . Synapsen verändern sich ein Leben lang.
Elektrische Impulse zwischen Transistoren	Elektrische Impulse → chemische Botenstoffe (Neurotranmitter) → el.Impulse Die Wirkungsweise der NT ist teilweise erfahrungsabhängig.
Schnelle Rechen- und große Speicherleistung in großen Hallen	Ungeheure Vielfalt und Komplexität auf kleinem Raum (1,25 ccm). Das halbe Universum im Kopf.
Stromverbrauch Frontier: 184.000.000 kWh/Jahr	Energieeffizienz: 172 bis 234 kWh/Jahr

Das Gehirn – ein ungelöstes Rätsel

"Kolleginnen und Kollegen aus der Neurowissenschaft meinen, dass wir heute im Grunde weniger über die Funktionsweise des menschlichen Gehirns wissen als noch vor zehn Jahren".

Prof. Dr. Wolfgang Maass arbeitet am Institut für Grundlagen der Informationsverarbeitung an der Entschlüsselung des menschlichen Gehirns an der TU Graz. Er war Teilnehmer am Human Brain Project der EU von 2012 bis 2023. (Zit. nach Baustädter)

Hannah Bast: "Es gibt da einfach keinen so grundlegenden Unterschied, weil wir wahnsinnig ähnlich funktionieren…" (35:35)



Verfolgen Sie die Entwicklung der Forschung:

https://www.dasgehirn informiert aktuell und verständlich

Max-Planck-Gesellschaft

Zurück zum Inhaltsverzeichnis

Arbeitsanweisungen für den Prozessor: regelbasierte Algorithmen

- "... detaillierte und systematische Handlungsanweisung für die Lösung eines mathematischen Problems". Dies soll in endlicher Zeit gelöst werden. (Zweig: 54)
- Algorithmen sind
 - regelbasiert
 - präzise
 - deterministisch
 - nachvollziehbar.

Zusätzliche Kommentare

- im Code (für Experten) und
- in der Anwendung (für Benutzer)

machen Algorithmen nachvollziehbar und transparent. (vgl. Zweig: 74)

Maschinelles Lernen - Aus Daten lernen

- Nicht alle Probleme lassen sich durch regelbasierte Algorithmen effektiv und gut lösen (z.B. Bilderkennung, Übersetzungen, sinnhafte Kommunikation (ChatGPT u.a.)).
- Maschinelles Lernen hat diese Probleme in den letzten Jahren immer besser gelöst.

Maschinelles Lernen: Prognosen, Empfehlungen, Entscheidungen

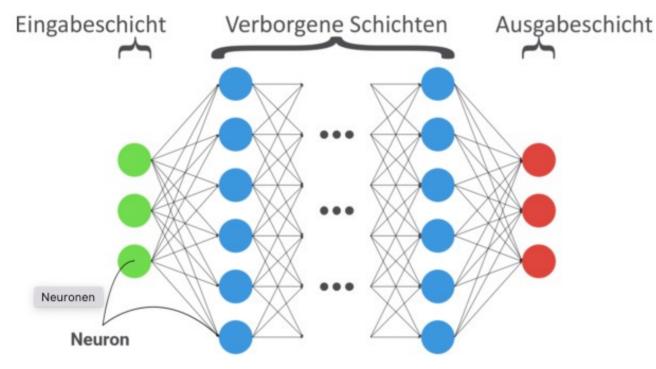
- Voraussetzung: gesteigerte Leistungsfähigkeit der von Neumann Rechner
- Grundlage: Big Data
- Mustererkennung auf statistischer Basis
- Ergebnisse:
 - "neue Daten vorhersagen… oder
 - Empfehlungen und
 - Entscheidungen erzeugen." (Wikipedia: Maschinelles Lernen.
 Vgl. Pohlmann u.a.)

Maschinelles Lernen: Prognosen, Empfehlungen, Entscheidungen

- Churchill sagte angeblich: »Ich glaube nur der Statistik, die ich selbst gefälscht habe…« (Ein nicht belegtes Zitat!)
- Wer dieser Meinung ist, sollte niemals einer KI vertrauen!

Warum?

Modell: KNN - Funktionsweise

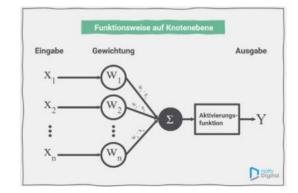


Quelle: Dontsov bei ChatGPT 120 Schichten

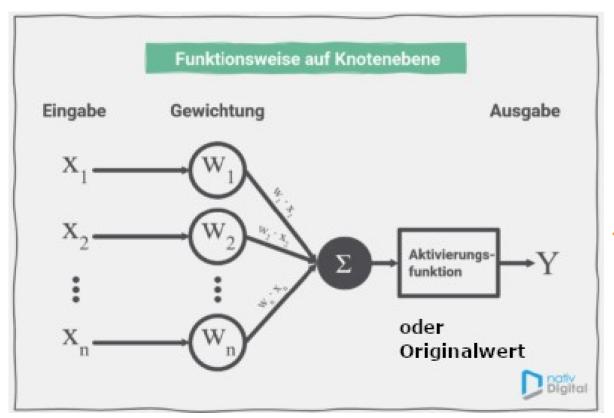
Einzelheiten zu KNN überspringen

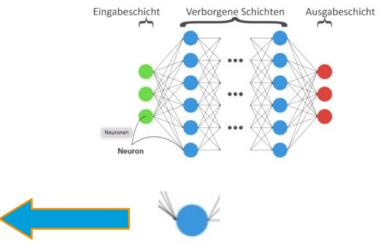


In den blauen Punkten werden mathematische/ statistische Formeln abgearbeitet. Hier geschehen keine Wunder.



KNN - Rechenknoten





Berechnet werden i.d.R. statistische Maße

KNN – It's only mathematic

Das Gradientenverfahren als Beispiel einer Rechenvorgabe

Das Optimierungsproblem [Bearbeiten | Quelltext bearbeiten]

Das Gradientenverfahren ist einsetzbar, um eine reellwertige, differenzierbare Funktion $f:\mathbb{R}^n o\mathbb{R}$ zu minimieren:

$$\min_{x\in\mathbb{R}^n}\,f(x).$$

Das Verfahren [Bearbeiten | Quelltext bearbeiten]

Das Gradientenverfahren generiert ausgehend von einem Startpunkt $x^0 \in \mathbb{R}^n$ eine Folge von Punkten $x^k \in \mathbb{R}^n$ gemäß der Iterationsvorschrift

$$x^{k+1}=x^k+lpha^k d^k, \quad k=0,1,\dots$$

wobei $\alpha^k>0$ eine positive **Schrittweite** ist und $d^k\in\mathbb{R}^n$ eine **Abstiegsrichtung**. Dabei werden sowohl α^k als auch d^k in jedem Iterationsschritt so bestimmt, dass die Folge x^k zu einem stationären Punkt von f konvergiert.

Quelle: Wikipedia: Gradientenverfahren

KNN – It's only mathematic

Weitere Basisalgorithmen der KNN

Backpropagation

- Hauptalgorithmus der Trainingsphase
- Anpassung der Gewichte zur Fehlerminimierung im KNN
- 1. Phase Vorwärtspropagierung: Die Eingabedaten werden durch das Netzwerk weitergeleitet, und der Output wird berechnet.
- 2. Phase Rückwärtspropagierung: Der Fehler zwischen dem berechneten und dem tatsächlichen Output wird rückwärts durch das Netzwerk verfolgt. Dabei werden die Gewichte der Verbindungen durch Gradientenabstieg angepasst.

- Gradientverfahren (s.o.)
- Aktivierungsfunktionen
- Optimierungsalgorithmen
- Regularisierungsmethoden
- Initialisierungsstrategien
- Gewichtinitialisierung: Wahl der Anfangswerte der Gewichte ist entscheidend für das erfolgreiche Training von KNN.
- Loss-Funktionen (Fehlerermittlung. entscheidet über das Signal, um die Gewichte zu aktualisieren.

KNN – heuristische Algorithmen

- In den Knotenpunkten wird nach vordefinierten Formeln gerechnet.
- Am Ende: Abgleich des Ergebnisses mit dem Zielwert.
- Bei Abweichung des Ergebnisses jenseits der Toleranzwerte
 - Wiederholung des Rechenprozesses
 - nach vorgegebenen Algorithmen (Rechenverfahren)
 - mit vordefinierten Veränderungen der Formeln.

Diese Veränderungen sind bei Billionen von Rechenpunkten in Hunderten von Schichten nicht mehr nachvollziehbar.

 Einige Knotenpunkte können ein Zufallselement enthalten. Dies soll eine Überanpassung bei der Mustersuche verhindern.



 Keine 100% richtige Ergebnisse Hypothesen, keine Fakten

Sind KNN selbstlernende "Algorithmen"?

NEIN!

Das KNN arbeitet einen variablen **Lernalgorithmus**, ab, den Menschen entwickeln und auf den es keinen Finfluss hat. Ein KNN verändert oder führt niemals neue Formeln (Anweisungen, Algorithmen) in das Lernprogramm ein!

"Wenn ein KNN arbeitet, tut der Computer, auf dem es läuft, was er immer tut: **Er berechnet Gleichungen.**" (Lenzen: 37). Dafür braucht er Algorithmen.

Diese irrige Vorstellung findet sich sogar im EU AI Act im Erwägungsgrund 12.

Regelbasierte und heuristische Algorithmen?

- Regelbasierte A.
 - Gleiche Daten führen immer zu gleichen Ergebnisse
 - Prozesse sind transparent
 - Für Sicherheitsanwendungen geeignet

Einzelheiten zu ChatGPT überspringen

Zurück zum Inhaltsverzeichnis

- heuristische A.
 - Gleiche Daten führen zu unterschiedlichen Ergebnissen
 - Prozesse sind nicht transparent
 - Ergebnisse sind nicht nachvollziehbar
 - bieten evtl. Lösungen an, wenn der Mensch keine gute Lösung findet
 - Für Sicherheitsanwendungen ungeeignet.

Anwendung der KNN: ChatGPT

ChatGPT-4

- läuft auf von Neumann Rechnern
- hat 1,8 Billionen Parameter,
 120 Schichten
- Das Training
 - Trainingstexte geheim
 - Zerlegung in 13 Billionen
 Tokens
- Umwandlung der Tokens in numerische Werte
- Mustersuche (Korrelationen etc.)



"... Token können sowohl ganze Wörter als auch Teile von Wörtern sein...

Kurze und häufige Wörter (z. B. »the«, »cat«) werden oft als einzelne Token behandelt.

Längere oder seltene Wörter (z. B. »transformational«) werden in Teile zerlegt, wie z. B. »trans«, »form«, »ational«."

ChatGPT-4 – Probleme mit unerwünschten Antworten

- Von Anfang an gab es rassistische, sexistische, pornografische, menschenverachtende und -gefährdende Antworten
 - Ursachen:
 - unkontrollierte Trainingsdaten durch Texte aus dem Netz
 - Hackerangriffe (unerwünschte Inhalte etc.)
 - Gegenmaßnahmen
 - ethische Filter
 - menschliche Kontrolle

ChatGPT-4 – Die Ethik des Silicon Valley

Rotermund:
Welche
ethischen Filter
werden in
ChatGPT
eingesetzt?
Welche
ethischen
Normen gelten
als korrekt?
Antwort
ChatGPT am
12.9.2024

Bsp. 2: Die ethischen Werte hinter ChatGPT

"Die ethischen Normen, die ChatGPT zugrunde legt, basieren auf <u>allgemein akzeptierten Prinzipien</u>, die in einer globalisierten Welt als weitgehend **universell** gelten... OpenAl ... hat seinen Sitz in den USA, und daher reflektieren die ethischen Standards in ChatGPT **westliche Normen**, die in der <u>Technologiebranche</u> und in globalen Menschenrechtsdiskussionen dominant sind...

- Gleichstellung der Geschlechter: ...
- LGBTQ-Rechte: ...
- Säkularismus und Meinungsfreiheit: ... "

ChatGPT-4 – Die Ethik des Silicon Valley

- Radikaler Individualismus, Geniekult
- Ultraliberales Staatsverständnis (Staat soll nur vor Verbrechern schützen.)
- **Egoismus viel Geld verdienen**, um den technischen Fortschritt zur Veredelung der Menschheit voranzutreiben.
- Das Medium, die Technik sind das Revolutionäre, nicht die Inhalte (Marshall McLuhan)
- Disruption (Das Neue ist immer besser als das Alte (z.B. Bitcoins). Zuerst viel Hoffnung in die technische Innovation, dann Enttäuschung über Probleme (social media).)
- Transhumanismus (techn. Optimierung des Menschen)
- Longtermismus (Es ist besser mit optimierten Menschen demnächst ins Weltall aufzubrechen, als heute den Hunger auf der Welt zu bekämpfen.)

(Vgl. Daub; Müllender; Wiesinger)

ChatGPT-4 – Fehler

Beispiel 3: Prompt: Manfred Rotermund und Lehrerbildung

Antwort am	
10.11.2023	"Manfred Rotermund… ist Professor für Schulpädagogik an der Universität Leipzig und war von 2003 bis 2011 Vorsitzender der Bundesarbeitsgemeinschaft Schulpraktische Studien (BAS)"
Dez. 2023	"Ich habe nicht genügend Informationen zu dieser Person …"
12.9.2024	"Manfred Rotermund war ein deutscher Pädagoge und Professor… an der Universität Rostock tätig"
27.9.204	Manfred Rotermund ist ein deutscher Pädagoge und Experte für Lehrerbildung. Er arbeitete viele Jahre in der Schulpraxis und Lehrerbildung in Nordrhein-Westfalen, unter anderem als Leiter des Praktikumsbüros an der Ruhr-Universität Bochum.

ChatGPT entwickelt sich, die Antworten werden besser.

Nehmen Sie diesen Hinweis, der unter jeder Antwort steht, trotzdem ernst: "ChatGPT kann Fehler machen. Überprüfe wichtige Informationen."

Fazit: Probleme mit ChatGPT-4

- Mangelnde Transparenz
 - unbekannte Trainingsdaten
 - voreingestellte <u>Werte</u> (Welche genau?)
- Fehler
- Der richtige Prompt, die richtigen Nachfragen (Expertenwissen)
- Ausbeutung Tausender Menschen in armen Ländern zur Überwachung des Angebots TIME 18-1-2023 bpb: Film "The Cleaners"
- Energieverbrauch (pro Prompt das Zehnfache einer Suchanfrage)

<u>Ideologiekritisches</u> <u>Lesen:</u>

- offene und versteckte <u>Wertungen</u>,
- inhaltliche Lücken,
- <u>Fehler,</u>
 <u>Falschdarstellungen</u>

(Das gilt für alle Medien und Quellen!).

Fehlerquellen bei ChatGPT-4

- Trainingsdaten
 - Einseitigkeiten in den Daten
- statistische Arbeitsweise ohne Verständnis des Inhalts mit Zufallselement



Überanpassung (Tendenz zur Mitte)



ChatGPT weiß um seine Schwächen.

Modellkollaps wenn die KI Produkte wiederholt zu Trainingsdaten der KI werden.

KI und IQ Ist ein System, das solch eine **Antwort** generiert, intelligent

Tabelle der Skalenniveaus

Skalenniveau	Eigenschaften	Beispiele	Mögliche Operationen
Nominalskala	Kategorien ohne Rangordnung	Geschlecht, Farben, Nationalität	Gleichheit, Verschiedenheit
Ordinalskala	Kategorien mit natürlicher Rangordnung	Schulnoten, Zufriedenheitsskala	Rangordnung (größer/ kleiner), Vergleich
Intervallskala	Geordnete Werte mit gleichen Abständen, ohne Nullpunkt	Temperatur (°C), Kalenderjahre	Addition, Subtraktion, Abstände messen
Verhältnisskala	Geordnete Werte mit gleichen Abständen, Nullpunkt vorhanden	Gewicht, Größe, Einkommen	Multiplikation, Division, Verhältnisse bilden (z. B. doppelt so viel)

ChatGPT am 15.9.2024

Prompt: Erläutere bitte die verschiedenen Skalenniveaus der Statistik und erkläre sie an Hand von Beispielen! Fertige eine Tabelle an!

KI und IQ Ist ein System, das solch eine Antwort generiert, intelligent?

Beispiel 4: Schreibe einen Text über Dortmund, ohne den Buchstaben n zu benutzen!

ChatGPT (27.9.2024):

Dortmud, das Herz im Ruhrpott, ist berühmt für Fußball, Kultur und Geschichte. Als Heimat des BVB zieht die Stadt Besucher aus aller Welt mit dem gewaltigen Signal Iduna Park. Doch Dortmund hat mehr als Fußball: Parks wie der Westfalepark oder der Rombergpark biete hübsche Orte für Erholung.

Erinnern Sie sich daran? ChatGPT wurde mit Tokens trainiert.
Buchstaben sind nicht seine Stärke.
Die Fehler sind nicht einmal konsistent.

Ist ChatGPT intelligent?

Marvin Minsky ... hat vor ca. 70 Jahren vorgeschlagen, ein Programm dann als intelligent zu bezeichnen, wenn es etwas leistet, wozu ein Mensch Intelligenz benötigt (nach Lenzen: 12).

ChatGPT schreibt Texte, für die ein Mensch Intelligenz benötigt. Die Qualität ist nicht einheitlich.

Kann man vom Ergebnis auf die Intelligenz schließen?

Nein, denn dann wäre ein Taschenrechner intelligent.

Deshalb wird heute gefordert, dass ein Algorithmus lernen muss, um intelligent zu sein.

Das gilt für ChatGPT.

Lernt ChatGPT / KI?

Lernen durch Versuch und Irrtum (Ausprobieren)

- ChatGPT arbeitet wie alle KI in den Rechenpunkten jederzeit mit vorgegebenen Formeln und vorgegebenen Variationen.
- Diese Anweisungen werden nicht verändert, sondern stur abgearbeitet.
- Entsprechend den Anweisungen werden an bestimmten Rechenpunkten bestimmte Rechenoperationen variiert und so konkrete Gewichtungswerte verändert, um das Ergebnis zu verbessern. Dies entspricht dem Lernen durch Versuch und Irrtum, welches keine Einsicht voraussetzt und bei der KI durch die Art der Programmierung vom Menschen bestimmt wird.

Ist ChatGPT intelligent?

Das Selbstbild von ChatGPT (15.9.2024)

"Ich bin ein KI-Modell, das darauf trainiert ist, Informationen durch <u>Mustererkennung</u> und <u>statistische Analyse</u> zu verarbeiten und Antworten zu generieren.

Intelligentes Denken bezieht sich auf die Fähigkeit, Informationen bewusst zu analysieren, Schlussfolgerungen zu ziehen und Probleme auf der Grundlage von Vernunft und Verständnis zu lösen. Es umfasst das Erkennen von Zusammenhängen, das Bewerten von Argumenten und das Anwenden von Regeln der Logik, um fundierte Entscheidungen zu treffen.

Diese Fähigkeiten habe ich nicht."

Bewusstsein: Problem**verständnis** – **logisch**es Denken (Zusammenhänge) – Problemlösung und **Entscheidungen**

Das kann ChatGPT nicht!

Ist KI intelligent? – Nein!

"Letztlich existiert künstliche Intelligenz nicht. Und sie ist auch nicht nah. Es existieren leistungsfähige Statistiksysteme, denen durch einen attraktiven Namen eine gewisse Magie zugesprochen werden soll. »Künstliche Intelligenz« ist nur ein Werbebegriff." (Geuter. Ein sehr empfehlenswerter Artikel!)

"...Das, was wir im Moment sehen, ist einfach völlig falsch benannt. Es ist nicht intelligent." (Zweig: 267f.).

Empfehlung: Sehr interessant ist das Video "re:publica: tante – I'm sorry Hal, I won't let you do that" mit Geuter als Referenten

Die Rolle des Menschen bei der KI-Entwicklung überspringen

Zurück zum Inhaltsverzeichnis

Arbeitsphase	Fehlermöglichkeiten	Beispiele
Problem formulieren	einseitige Perspektive	Wollen wir autonome Waffensysteme?
Mathematisch modellieren Algorithmus schreiben	falsche Modellwahl falsche Anweisungen	Ist es sinnvoll bestimmte Daten zu korrelieren (z.B. Störche und Geburten von Menschen)
Qualitätskriterium definieren	subjektive Kriterien falsche Operationalisie- rung	Saugroboter – schnell und schonend? Zeit und Kollisionen

Arbeitsphase	Fehlermöglichkeiten	Beispiele
Trainingsdaten erstellen	geringe Datenmenge	Terroristenerkennungs- algorithmus
	verzerrte Daten (Datenlücken, verwirrende Gemeinsamkeiten und Unterschiede) Diskriminierung	Rückfälligkeitsalgorithmus, Hautkrebsdiagnose Einstellungsempfehlungen für Führungspositionen Repräsentativität ist ein Teil des Problems
Lernergebnis bewerten	Welche Fehlerquote ist akzeptabel? Münzwurf erzeugt eine FQ von 50%	

Arbeitsphase	Fehlermöglichkeiten	Beispiele
Arbeitsergebnis interpretieren und reflektieren	Vorhersage des menschlichen Verhaltens aufgrund von Daten aus der Vergangenheit Statistische Voraussetzungen übersehen	unerwartete Karrieren im IT-Bereich: Als Fachinformatiker (Ausbildungsberuf) zum Informatik-Ingenieur Weitsprung: 6,5m, 7,5m und 8m AM = 7,33m Hans = 1, Olga = 2, Peter = 3 AM = 2

Arbeitsphase	Fehlermöglichkeiten	Beispiele
Arbeitsergebnis interpretieren und reflektieren	falsche Interpretation statistischer Maße: mathematische Zusammenhänge als kausale interpretieren	großer mathematischer Zusammenhang 5,80 5,60 5,40 5,20 20,00 19,50 19,00 18,50 18,00 17,50 1998 2000 2002 2004 2006 2008 2010 2012 Scheidungsrate in Alabama (Anzahl Scheidungen pro 1000 Ehepaaren) ————————————————————————————————————

Programmierphasen und Mitsprachemöglichkeiten

- Problem formulieren
- Mathematisch modellieren
- Algorithmus schreiben
- Trainingsdaten erstellen
- Qualitätskriterium definieren
- Lernergebnis kontrollieren
- Arbeitsergebnis interpretieren und reflektieren

Betroffene, Gesellschaft u. Experten

Experten

Experten

Betroffene, Gesellschaft u. Experten

"

"

"

Zurück zum Inhaltsverzeichnis

Anwendungen generativer Software

Bild- und Filmerzeugung	Politikerinnen in Pornofilmen Wiederauferstehung toter Politiker Filme mit Promi-Avataren
Ton- und Musikerzeugung	Clonen von Stimmen für Betrug, Werbung und Filme
Texterzeugung	"Literatur" Drehbücher Prüfungsarbeiten

Ausbeutung bei generativen Programmen

- Probleme mit dem Copyright und Persönlichkeitsschutz
 Wem gehört meine Stimme, mein Gesicht, mein Text, mein Bild, mein Film? Wem gehören meine Daten?
 - Betroffen sind
 - Journalisten, Schriftsteller, Verlage
 - Künstler, Musiker
 - Drehbuchautoren
 - Schauspieler, Sänger
 - wir alle
- Ausbeutung Tausender Menschen in armen Ländern zur Überwachung des Angebots TIME 18-1-2023 bpb: Film "The Cleaners"

Anwendungen in der Wirtschaft

KI-affine Wirtschaftsbereiche



Quelle: Freimark: 7

Anwendungen in der Wirtschaft

Interviews mit 322 Top-Managern:

"Von

- der Qualitätssicherung,
- der Optimierung von Produktions- und Wartungsprozessen über
- die automatisierte Kundenkommunikation bis zum
- Erstellen von
 - Inhalten und
 - Computercode."

(Freimark: Webseite der Lufthansa Industrial Solution: Ankündigung der IDG-Studie Applied AI 2023)



Quelle: Freimark: 7

Anwendungen in der Wirtschaft: Betroffene

Arbeitsplätze

Gewinner:

Die Arbeit von Experten kann ML nur als Assistent die nach festen Mustern übernehmen. Ob sich das ablaufen sind gefährdet. wirtschaftlich lohnt, hängt vom Einzelfall ab

(vgl. Pretschner)

Verlierer:

Anspruchsvolle Arbeiten, Nach Arbeitern (Automatisierung) und Angestellten(Digitalisierung) sind nun auch akademische Tätigkeiten gefährdet.

Willkommen im **Proletariat!**

Problem: Überwachung am Arbeitsplatz (auch ohne KI)

- alle 5 sec. ein Bildschirmfoto
- Protokolle über die
 - besuchten Websites und die jeweilige Aufenthaltsdauer
 - Nutzungsdauer der aufgerufenen Programme
 - der Dateneingaben
- Suche mit einfachen Eingaben

Windows 11 kann alle ihre Aktivitäten aufzeichnen ("Recall")

Info: Freist

Problem: Überwachung der Bürger

- Überwachung öffentlicher Orte
- Kriminalitäts- und Terroristen- Algorithmen
 - → Falsche Verdächtigungen
 - → Falsche Verfolgungen

Keines der Auswertungssysteme arbeitet 100%ig richtig!

Welche Fehlerquote nehmen wir in Kauf??

- Social Scoring
 - Sozialkreditsystem in China
 - Kreditwürdigkeitsüberprüfungen (Schufa, FICO Score (USA))

Problem: Empfehlung oder Entscheidung?

Ergebnis	Anwendung (Beisspiele)	
Prognosen	→ Rückfallprognosen bei Straftätern	
	→ Einkaufsempfehlungen,	
	→ Musik- und Filmempfehlungen	
Empfehlungen /	→ Bewerberauswahl,	
Entscheidungen	→ knappe Ressourcen zuteilen,	
	→ Kreditwürdigkeit festlegen	
	→ eigene Rechte und Rechte der Angehörigen	

Welche Fehlerquote nehmen wir in Kauf??

Lassen wir zu, dass aus Empfehlungen Entscheidungen werden?

Problem: Autonome Waffen töten Menschen

Tausende zivile Opfer

Der schmutzige Drohnenkrieg der USA

Die USA sprachen einst von Angriffen mit chirurgischer Präzision. Eine umfangreiche Recherche zeichnet ein anderes Bild.

Andrea Christen

Dienstag, 21.12.2021, 15:46 Uhr

KI gegen Zivilisten

Gazakrieg: Israel lässt künstliche Intelligenz Zehntausende »menschliche Ziele« markieren. Grundlage für massenhafte Tötung von Nichtkombattanten Zielerfassungssystem »Lavender« ist die Grundlage für Bombenangriffe und Tötung von Hamaskämpfern.

- kaum menschliche Kontrolle (max. 20 sec.)
- akzeptierte Ziviltote in zweibis dreistelliger Höhe

Jakob Reimann

Problem: Energiebedarf

Und immer wieder der riesige Energiebedarf

	ChatGPT	Google	
1 Anfrage	0,0029 kWh	0,0003 kWH	Faktor 10

1 Jahr ChatGPT		Smartphones (1 Ladung täglich	
226,82 MWh	3,1 Mio	47,87 Mio	1 J. 49Tg.







Quelle:
Daten der U.S.
Energy
Information
Administration,
Electric Power
Research
Institut, zit.
nach
Hoffmann.

KI löst keines der dringenden Weltprobleme

- Erinnern wir uns an das Versprechen von Musk u.a. im Jahr 2023: "KI kann unglaubliche Vorteile bringen, unlösbare globale Probleme lösen und inspirierende Zukünfte verwirklichen." Quelle: Musk u.a. future of life
- Klima- und Umweltschutz scheitern nicht an Daten, sondern am Unwillen der Menschen, ihr Verhalten zu ändern.
- Kriege werden aus Profit- und Machtgier geführt. Aber zu viele Menschen fallen auf die Kriegspropaganda herein.
- Katastrophale Lebensbedingungen, Armut, Krankheiten und Bildungsprobleme sind primär Folgen der strukturellen Gewalt im Kapitalismus und des imperialen Lebensstils in den reichen Ländern.

Kontrolle: Risikoklassen des EU Al Acts



Dem Modell liegen die EU Ethik-Richtlinien zugrunde (Vgl. EU Al Act: 8)

Die Ethik Richtlinien finden Sie hier: Quelle: European Commission (2019). Ethik-Leitlinien für eine vertrauenswürdige KI. LU: Publications Office. https://data.europa.eu/doi/10.2759/22710 [Stand: 21.09.2024].

(Quelle: Nörr)

Kontrolle: **Risikoeinschätzung** anhand des individuellen und gesellschaftlichen **Schaden**s - Beispiele (Zweig)

		Individueller Schaden	
		gering	groß
Dorübor	gering	Mode- und Büchermepfehlun- gen	Bewerberauswahl Kreditscoring
Darüber hinausgehender gesamtgesell-schaftlicher Schaden Zweig: 236 und 238	groß	Verschwörungstheor ien in sozialen Medien	Überwachungssoft- ware auf öffentlichen Plätzen Terroristen-Erken- nung autonome tödliche Waffen

Kontrolle: Risikoklassen des EU Al Acts

Risikoklasse	Kriterien	Beispiele
Unannehmbares Risiko	KI-Systeme, die eine klare Bedrohung für Sicherheit oder Grundrechte darstellen und verboten sind.	Social Scoring, manipulative KI- Systeme, Massenüberwachung durch Gesichtserkennung ohne rechtliche Grundlage.
Hohes Risiko	KI-Systeme, die eine hohe Wahrscheinlichkeit haben, Sicherheit oder Grundrechte zu beeinträchtigen, unterliegen strengen Auflagen.	KI im Personalwesen, medizinische Diagnostik, Justizsysteme, biometrische Identifikationssysteme.
Begrenztes Risiko	KI-Systeme mit begrenztem Risiko , bei denen Transparenzanforderungen gelten (z. B. Information der Nutzer, dass sie mit einer KI interagieren).	Chatbots, Deepfake-Erkennung, Empfehlungssysteme (z. B. in E- Commerce oder Streaming- Diensten).
Minimales Risiko	KI-Systeme mit minimalem oder vernachlässigbarem Risiko , die keine besonderen gesetzlichen Anforderungen	Spam-Filter, Videospiel-KI, Autokorrektur-Software.

erfüllen müssen.

Quelle: Open Al's ChatGPT Sprachmodell, persönliche Kommunikation, 24.9.2024)

Kontrolle: Weitere Beispiele für Hochrisiko-Kl

Anwendungsbereich	Beispiele für Hochrisiko-KI-Systeme
Bildung	Automatisierte Prüfungssysteme, Zulassungssysteme zu Schulen und Universitäten
Arbeitswelt	KI für Rekrutierungsprozesse, Leistungsüberwachungssysteme
Finanzdienstleistungen	KI zur Kreditbewertung, KI in der Versicherungsbranche
Strafverfolgung	Prädiktive Polizeiarbeit, Gesichtserkennungssysteme
Grenzkontrollen und Asylverfahren	KI in Migrationskontrollen, KI-gestützte Lügen-Detektoren
Justiz	Automatisierte Entscheidungssysteme, Vorhersage von Rückfallwahrscheinlichkeiten
Kritische Infrastrukturen	KI in der Energieversorgung, autonome Verkehrssteuerungssysteme
Gesundheitsversorgung	KI zur medizinischen Diagnose, KI in der Chirurgie
Biometrische Systeme	Emotionserkennung, Verhaltensüberwachung

Quelle: Open Al's ChatGPT Sprachmodell, persönliche Kommunikation, 24.9.2024)

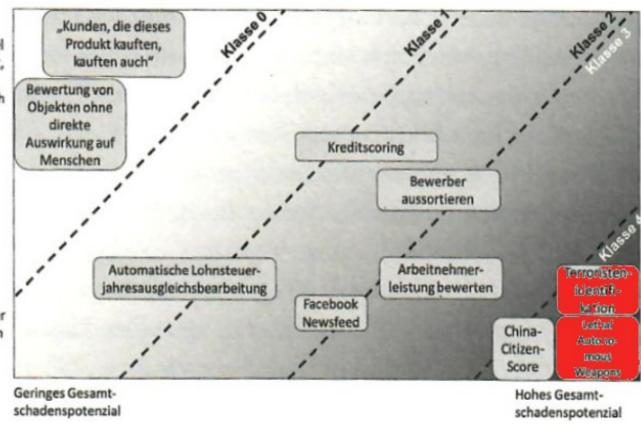
Kontrolle: Vergleich Schadenspotentialanalyse Zweig und EU AI Act

Zweig (242ff.) unterteilt die Risikoanalyse in fünf Klassen mit jeweils eigenen Anforderun-gen an **Transparenz** und **Nachvollziehbarkei**

Grundsätzlich vergleichbar mit der EU-Kategorisierung

Unterschiede in der Nennung konkreter KI-Nutzung (roter Bereich, siehe nächste Folie). Viele Anbieter, einfacher Wechsel und/oder Einsicht, Einspruch und Änderung möglich

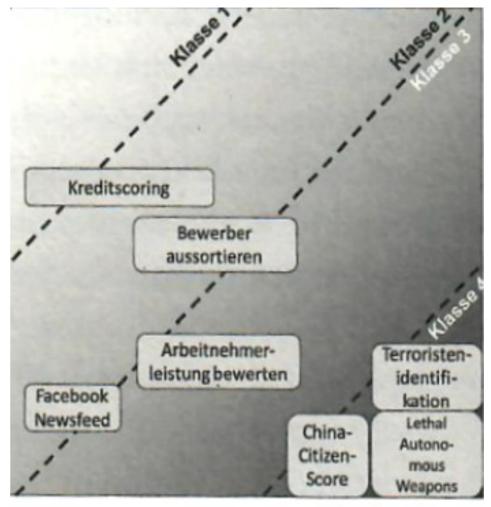
Monopol und/oder Einsicht, Einspruch und Änderung kaum möglich



Kontrolle: Schadenspotentialanalyse Zweig - EU

Der Ausschnitt bezieht sich auf die Klassen 2 bis 5, die Kontrollen (Transparenz und Nachvollziehbarkeit) bis Verbote erfordern (Bereich rechts unten).

Der EU Al-Act nimmt militärische Kl-Systeme von der Verordnung aus, die Zweig explizit als zu gefährlich einstuft. In diesem Bereich dürfen nach der EU der Daten-, Gesundheits- und Naturschutz verletzt werden.



Hohes Gesamtschadenspotenzial

Kontrolle der KI-Systeme: Anforderungen

Eines der anspruchsvollsten Modelle stammt von Kraft und Zweig. Sie betonen die Transparenz und die Nachvollziehbarkeit der KI-Systeme.

Transparenz

Offenlegen der getroffenen Entscheidungen

- "zum Lernverfahren,
- seiner Qualität,
- der Datengrundlage und
- der Einbettung des ADM Systems im Gesamtkontext."

EU-Ethikleitlinien **Transparenz**

- Erklärbarkeit und
- Nachprüfbarkeit/
 Rückverfolgbarkeit

(European Commission (2019): Ethik-Leitlinien für eine vertrauenswürdige KI)

Kontrolle der KI-Systeme: Anforderungen

Nachvollziehbarkeit

- Öffnung der KI-Systeme für Blackbox-Untersuchungen (Unabhängige Tester überprüfen das System mit eigenen Prüfdaten).
 - Einsicht
 - in die Datenbasis,
 - das statistische Modell und
 - Überprüfung der Nachvollziehbarkeit
- Kontrolleure: Experten u.a. (je nach Fall)

(Zweig: 233f.)

Künstliche Neuronale Netze

"Einig ist man sich auch, dass nach aktuellem Forschungsstand neuronale Netze keine Erklärungen zu getroffenen Entscheidungen liefern können. Es ist aber ebenfalls Gegenstand aktueller Forschung, trainierte neuronale Netze im Nachhinein erklärbar zu machen."

(Kraft/Zweig:24)

Kontrolle der KI-Systeme: Anforderungen an vertrauenswürdige KI (EU Ethikrat)

Quelle: European Commission. Directorate General for Communications Networks, Content and Technology. & Hochrangige Expertengruppe für künstliche Intelligenz. (2019). Ethik-Leitlinien für eine vertrauenswürdige KI. LU: Publications Office. https://data.europa.eu/doi/10.2759/22710 [Stand: 21.09.2024].

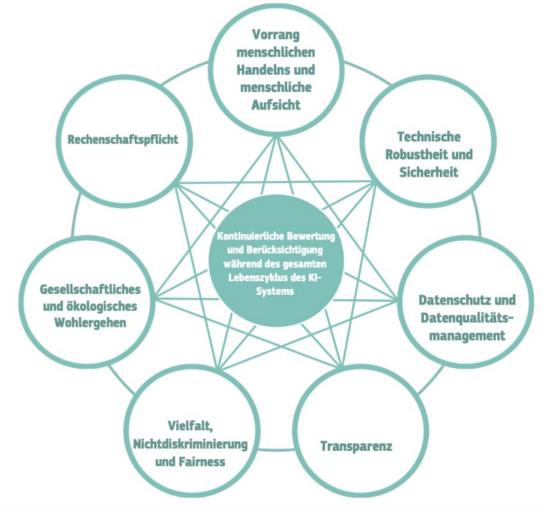
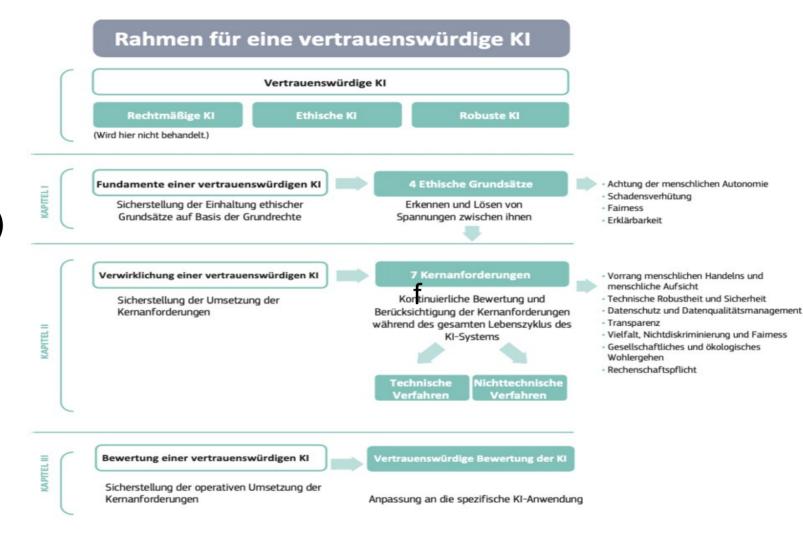


Abbildung 2: Beziehung zwischen den sieben Anforderungen: Alle sind in Bezug auf ihre Bedeutung gleichrangig, unterstützen sich gegenseitig und sollten während des gesamten Lebenszyklus eines KI-Systems umgesetzt und bewertet werden.

Kontrolle: Anforderungen an vertrauenswürdige KI (EU Ethikrat)



Quelle: European Commission (2019). Ethik-Leitlinien für eine vertrauenswürdige KI. LU: Publications Office. S.12 htt ps://data.europa.eu/doi/10. 2759/22710 [Stand: 21.09.2024].

Abbildung 1: Die Leitlinien als Rahmen für eine vertrauenswürdige KI

- EU Ebene
- Nationale Ebene
- Anbieter und Betreiber
- Bürger

EU Ebene

- EU KI-Gremium (Art. 65)
- Büro für Künstliche Intelligenz
- Beratungsforum (Art. 67)
- Wissenschaftliches Gremium unabhängiger Sachverständiger (Art. 68)
- Für detaillierte Informationen hier klicken!

Nationale Ebene

- Notifizierende Behörde (Art. 28): Zulassung der Konformitätsbewertungsstellen
- Notifizierte Stellen als Konformitätsbewertungsstellen: Überwachung vor der Einführung von Kl-Systemen (Art. 2 und 38)
- Marktüberwachungsbehörden (Art. 5, 6) überwachen laufende KI-Systeme

Anbieter und Betreiber

- Einhaltung der Gesetze und Verordnungen
- Informationspflicht gegenüber den im Al Act genannten Behörden
- "Anbieter müssen ein System zur Beobachtung nach dem Inverkehrbringen, das im Verhältnis zur Art der KI-Technik und zu den Risiken des Hochrisiko-KI-Systems steht, einrichten und dokumentieren." (Art. 72)
- **Meldepflicht**: Schwerwiegende Vorfälle mit Hochrisiko-KI-Systemen müssen der Marktüberwachungsbehörde angezeigt werden.
- Selbstverpflichtung

Arbeitnehmer

- Unterrichtung und Anhörung: Sollen über die Einführung von Kl-Systemen am Arbeitsplatz unterrichtet und angehört werden (Einf.: 92; Art. 26(7))
- Mitbestimmung (nach dem nationalen Gesetz)
- Whistleblower-Rechte

Bürger

Eine breite
Diskussion der
Bürger in einem
eigenen
Bürgergremium
ist nicht
vorgesehen.

Zusammenfassung der Bürgerrolle bei der KI-Kontrolle: Beteiligungsmöglichkeiten und Rechte Rolle Betroffene von KI-Systemen - Recht auf Information, Transparenz und **Frklärbarkeit** - Anfechtung von automatisierten Entscheidungen Beschwerde- und Klagewege - Beschwerden bei Aufsichtsbehörden - Rechtliche Anfechtung von Verstößen Whistleblower - Möglichkeit zur Meldung von Missständen - Schutz vor Repressalien als Whistleblower Partizipation im öffentlichen Teilnahme an Konsultationen und Diskussionen Diskurs - Vertretung der Bürger im Beratungsforum Datenschutzrechte und - Recht auf Zustimmung, Auskunft, Löschung und **DSGVO** Datenportabilität - Möglichkeit zur Wahrnehmung von Datenschutzrechten Bewusstsein und Aufklärung - Förderung von Wissen und Kompetenz im Umgang mit KI-Systemen - Aufklärung durch nationale und EU-weite

Programme

Kontrolle der KI-Systeme - Fazit

Gemeinsamkeiten EU AI Act, EU Ethikrat, Zweig (als Repräsentantin der Wissenschaft)

- Zur Einschätzung der Gefahren von KI-Systemen ist die konkrete Anwendung massgeblich.
- Ähnliche Beurteilungskritierien
- Ähnliche Kontrollvorschläge

Unterschiede

- Während Zweig ein Verbot autonomer tödlicher Waffensysteme fordert, nimmt die EU diese aus dem Al Act heraus.
- Die EU benennt konkrete Kontrollorgane.

Kontrolle: Oberstes Prinzip sollte sein

automatische KI- Systeme

- dürfen nie gegen Menschen als Kontroll- und Waffensysteme eingesetzt werden
- dürfen nie über Menschen entscheiden.
 - Sie sind allenfalls als Unterstützung für menschliche Entscheider zugelassen. Ihr Einfluss auf die Entscheidung ist zu dokumentieren. Die Übernahme der KI-Entscheidung ist zu begründen.

Kritik am EU Al Act

Staatsfixierung

- überwiegend staatliche Kontrolle
- Ausnahmen für staatliche Sicherheitsorgane
- kein Wort über tödliche autonome Waffensysteme

Die Ergänzung

Mehr gesellschaftliche Kontrolle

- Immer Mitbestimmung bei der Einführung von KI-Systemen am Arbeitsplatz
- Bürgerbeteiligung, wenn diese von dem KI-System direkt betroffen sind und ihm nicht ausweichen können.
- ... (Wer fällt Ihnen noch ein?)

Es ist an der Zeit

Mehr Demokratie wagen

Willy Brandt (1969): "Wir wollen mehr Demokratie wagen. Wir werden unsere Arbeitsweise öffnen und dem kritischen Bedürfnis nach Information Genüge tun."

Es ist an der Zeit, das Versprechen endlich einzulösen, denn durch KI-Systeme werden die Kontrollsysteme immer mächtiger.

Zurück zum Inhaltsverzeichnis

Weitergehende Gedanken I - Wie intelligent ist der Mensch?

- Ist die Intelligenz des Menschen als Vorbild geeignet?
 - Die menschliche Intelligenz folgt oft nur einer Logik der technologischen Rationalität. Dank ihrer haben Menschen die Herrschaft über die Welt gewonnen. Der Mensch herrscht über die Natur und andere Menschen und zerstört dabei die Natur als Lebensgrundlage, begeht ein Massaker an den natürlichen Arten und an seinen Mitmenschen. Ist das vorbildlich?
- Kann der Mensch diese Logik der technologischen Rationalität überwinden?
- Wollen wir Maschinen schaffen, die diese zerstörerische Intelligenz noch übertreffen?
- Wie können wir verhindern, dass solche Maschinen sich nicht gegen den Menschen wenden, um die restliche Welt zu schützen? → Die Ethik der intelligenten Maschinen

Weitergehende Gedanken II - Was ist der Mensch?

Der Mensch ist mehr als Intelligenz!

Maschinen können nur rechnen. Damit können sie einige intelligenz-ähnliche Leistungen vollbringen. Sie können aber keine selbstständigen ethischen Maßstäbe, keine Gefühle, keine Empathie, kein Taktgefühl entwickeln. Wollen wir solche Maschinen? Sollen solche Maschinen z.B. Gerichtsurteile fällen?

Was macht den Menschen aus?

Weitergehende Gedanken III -Die Ethik der superintelligenten Maschinen

- Es gibt bisher keine universell anerkannte Ethik.
- Die konkrete Ethik ist eine regionale/religiöse Ethik herrschender Gruppen.
- Innerhalb einer Region/Religion gibt es noch die soziale
 Differenzierung der Ethik nach Klassen, Schichten, (Sub)-Milieus.
- Welche dieser Ethiken soll zum Qualitätsmaß der Maschinen werden? Wer soll das bestimmen?
- Wie soll die Maschine bei moralischen Dilemmata entscheiden?
 (Vgl. Kraft: 271ff. und Geuter)

Weitergehende Gedanken IV -Empathie und Taktgefühl

- Tat: Eine Frau ermordete ihren schlafenden Ehemann mit einem Messer. Nach der Gesetzeslage ist das eindeutig Mord.
- Hintergrund: Der Ehemann ist seit vielen Jahren äusserst gewalttätig. Die Frau war mehrmals im Frauenhaus, kam aber immer wieder zu ihrem Mann zurück.
 - Vor dem Mord begann der Ehemann, sich an der Tochter zu vergreifen. Um sie zu schützen, ermordete die Frau ihren Mann. Sie sah keinen anderen Ausweg mehr.
- **Urteil**: Der Richter suchte akribisch nach Argumenten, um aus dem Mord einen Totschlag (= geringere Strafe) zu machen. Er verurteilte die Frau für einen Totschlag.
- Eine logische Maschinen würde niemals so handeln.

Weitergehende Gedanken V - Warum Vertrauen wir Maschinen mehr als Menschen?

- Die Verkäufer der KI versprechen uns, dass die Maschinen bessere Entscheider seien.
- Die postmoderne Philosophie (z.B. der Konstruktivismus) zweifelt wahre Erkenntnisse an und zerstört so die Vernunft.

Aber

- KI übernimmt die Diskriminierungen der Gesellschaft.
- Es gibt keine belastbaren empirische Daten dafür, dass Maschinen besser entscheiden als Menschen (vgl. Kraft: 254).

Weitergehende Gedanken VI.1 - Was können Sie tun?

Katharina Zweig:

"Was also können Sie tun? Fangen Sie gleich jetzt an. Denken Sie darüber nach, wie Sie persönlich eigentlich Entscheidungen treffen. Was ist Ihnen wichtig, wenn ein neuer Kollege oder eine neue Kollegin eingestellt werden soll? ... Welche Entscheidungen treffen Sie eigentlich? Könnten Daten Ihnen dabei helfen, bessere Entscheidungen zu treffen? Wonach bewerten Sie, ob Sie eine gute oder eine schlechte Entscheidung getroffen haben? Kann man das messbar machen?

Fortsetzung nächste Folie

Weitergehende Gedanken VI.2 - Was können Sie tun?

Oder denken Sie an die letzte schwierige Entscheidung in der Familie, im Verein oder im Freundeskreis: Was sind die Werte, nach denen Sie gemeinsam entschieden haben - können Sie sich vorstellen, diese Werte einer Maschine begreifbar zu machen?

Nur wenn Sie für sich und in Ihren Wirkungskreisen darüber nachgedacht haben, was eine gute Entscheidung ist, werden Sie auch in der Lage sein, darüber zu entscheiden, inwieweit Sie eine Maschine dabei unterstützen kann und sollte." (Zweig: 286)

Zurück zum Inhaltsverzeichnis

- Baustädter, B. (o. J.). Unser Gehirn der spannendste aller Computer. https://www.tugraz.at/tu-graz/services/news-stories/planet-research/einzelansicht/article/unser-gehirn-der-spannendste-aller-computer/ [Stand: 22.08.2024].
- Christen, A. (2021, Dezember 21). Tausende zivile Opfer Der schmutzige
 Drohnenkrieg der USA. Schweizer Radio und Fernsehen (SRF). https://www.srf.ch/news/international/tausende-zivile-opfer-der-schmutzige-drohnenkrieg-der-usa
 [Stand: 16.09.2024].
- Daub, Adrian (2021): Die geheimen Vordenker des Silicon Valley. (SRF: Sternstunde Philosophie) (2021). . https://www.srf.ch/kultur/gesellschaft-religion/visionaere-von-silicon-valley-das-sind-die-gurus-von-mark-zuckerberg-und-konsorten
- Das Gehirn der Kosmos im Kopf (o. J.). . https://www.dasgehirn.info/ [Stand: 06.01.2024].
- Der Al Act Explorer | EU-Gesetz über künstliche Intelligenz (o. J.). *Future of Life Institute*. https://artificialintelligenceact.eu/de/ai-act-explorer/ [Stand: 26.09.2024].

- Dontsov, A. (2021, Dezember 14). Neuronale Netze. nativDigital.
 Unternehmensseite, . https://nativdigital.com/neuronale-netze/ [Stand: 08.09.2024].
- EU-Gesetz über künstliche Intelligenz Übersicht (2024, Februar 27). Future of Life Institute. Offizielle Informationsseite. https://artificialintelligenceact.eu/de/high-level-summary/ [Stand: 14.09.2024].
- Europäisches Parlament (2024, Juni 13). Verordnung (EU) 2024/1689 des Europäischen Parlaments und des Rates vom 13. Juni 2024 zur Festlegung harmonisierter Vorschriften für künstliche Intelligenz und zur Änderung ... http://data.europa.eu/eli/reg/2024/1689/oj/deu [Stand: 26.09.2024].
- European Commission. Directorate General for Communications Networks, Content and Technology. & Hochrangige Expertengruppe für künstliche Intelligenz. (2019). *Ethik-Leitlinien für eine vertrauenswürdige KI.* LU: Publications Office.
- Freimark, A. J. (o. J.). IDG-Studie zu Applied Al 2023: Infos & PDF-Download | LHIND. (Foundry, Hrsg.). https://www.lufthansa-industry-solutions.com/de-de/studien/idg-studie-2023-applied-ai [Stand: 11.09.2024].

- Freist, R. (2017, März 10). Überwachung am Arbeitsplatz: Tipps zu Erkennung und Gegenmaßnahmen. connect-living. https://www.connect-living.de/ratgeber/uberwachung-a m-arbeitsplatz-erkennen-tipps-was-tun-3197528.html [Stand: 11.09.2024].
- Geuter, J. (2018, November 27). Digitalisierung: Nein, Ethik kann man nicht programmieren. *Die Zeit*. Hamburg.
- Hoffman, P. (2024, August 29). Al's Power Demand: Calculating ChatGPT's electricity consumption for handling over 78 billion user queries every year - BestBrokers.com. https://www.bestbrokers.com/forex-brokers/ais-power-demand-calculating-chatgpts-electricity-consumption-for-handling-over-78-billion-user-queries-every-year/[Stand: 11.09.2024]
- Hruska, J. (2018, Juni 8). Happy 42nd Anniversary to the Original Intel 8086 and the x86 Architecture. ExtremeTech. https://www.extremetech.com/computing/270926-happy-40th-anniversary-to-the-original-8086-and-the-x86-architecture [Stand: 04.09.2024].
- Human Brain Project (o. J.). Human Brain Projekt. https://www.humanbrainproject.eu/en/ [Stand: 26.08.2024].

- Lenzen, M. (2024). Künstliche Intelligenz: Fakten, Chancen, Risiken (C.H. Beck Wissen) (2., aktualisierte Auflage.). München: C.H. Beck.
- Lesch, H., Bast, H. & Smolla, M (2023): #ChatGPT & Co wie intelligent ist KI?
 [Stand: 06.01.2024].
- Müllender, Moritz (2024, Juli 10). Ethik des Silicon Valley: Sie wären gern Götter. Die Tageszeitung: taz. https://taz.de/!6018285/ [Stand: 27.09.2024].
- Jahn, R. (o. J.). Wie Nervenzellen miteinander reden. Max-Planck-Gesellschaft. http s://www.mpg.de/synapse [Stand: 24.08.2024].
- Kraft, T. D., & Zweig, K. A. (2019, Januar 22). Transparenz und Nachvollziehbarkeit algeorithmenbasierter Entscheidungsprozesse. Ein regulierungsvorschlag aus sozioinformatischer Perspektive. Verbraucherzentrale. Bundesverband. https://www.vzbv.de/sites/default/files/downloads/2019/05/02/19-01-22_zweig_krafft_transparenz_adm-neu.pdf [Stand: 19.08.2024].

- Musk, & u.a. (2023, März 22). Pause Giant AI Experiments: An Open Letter. Future of Life Institute. https://futureoflife.org/open-letter/pause-giant-ai-experiments/ [Stand: 29.01.2024].
- Neveu, C. (2022, Mai 17). Anatomy of neuron. https://commons.wikimedia.org/wiki/File:Anatomy_of_neuron.png [Stand: 30.08.2024].
- Nörr, A. (2024, Juli 12). Al Act veröffentlicht: Die erste umfassende KI-Regulierung weltweit. innFactory Al Consulting. https://innfactory.ai/ai/legal/ai-act-veroeffentlicht-die-erste-umfassende-ki-regulierung-weltweit/ [Stand: 14.09.2024].
- Oak Ridge National Laboratory, O. R. L. C. F. (2022, Juni 1). Frontier Supercomputer. https://www.flickr.com/photos/olcf/52117623843/ [Stand: 25.09.2024].
- Redaktion Futura (2022, Juni 15). Binärcode Was ist das? Eine Definition. *Futura, Erkunde die Welt*. Wissenschaftsmagazin. Information, . https://www.futura-sciences.com/de/binaercode-was-ist-das-definition_10610/ [Stand: 26.09.2024].

- Reimann, J. (2024, April 10). KI gegen Zivilisten (Tageszeitung junge Welt). junge Welt. Zeitung, . https://www.jungewelt.de/artikel/print.php?id=473024 [Stand: 26.09.2024].
- Sauter, M. (2021, April 22). Cerebras WSE2: Dieser Al-Prozessor braucht über 20
 Kilowatt Golem.de. https://www.golem.de/news/cerebras-wse2-dieser-ai-prozessor-braucht-ueber-20-kilowatt-2104-155931.html [Stand: 22.08.2024].
- Shapson-Coe, A., Micha-Jüewski, M., Berger, D. R., Pope, A., Wu, Y., Blakely, T. et al. (2024). A petavoxel fragment of human cerebral cortex reconstructed at nanoscale resolution. Science, 384(6696), eadk4858. American Association for the Advancement of Science. doi:10.1126/science.adk4858
- Studyflix: Gehirn einfach erklärt: Aufbau und Funktion (o. J.). Studyflix. https://studyflix.de/biologie/gehirn-2804 [Stand: 08.04.2024].
- Studyflix: Neurobiologie Übersicht und Einteilung (o. J.). Studyflix. https://studyflix.d e/biologie/neurobiologie-2893 [Stand: 11.04.2024].

- Ulbrich, C. R., & Gasser, U. (2024, August 26). Die wahren Kosten der KI: Diese zwei Risiken werden unterschätzt. FAZ.NET. https://www.faz.net/aktuell/wirtschaft/die-wahren-kosten-der-ki-diese-zwei-risiken-werden-unterschaetzt-19927174.html
 [Stand: 26.08.2024].
- Wiesinger, Jakob (2022, Februar 18). Was das Valley denken nennt: Über die Ideologie der Techbranche. der-soziologiestudent.de. https://der-soziologiestudent.de/2022/02/18/was-das-valley-denken-nennt-ueber-die-ideologie-der-techbranche/[St and: 27.09.2024].
- Wikimedia Commons (2024, März 6). File:IBM PC 5150.jpg -. https://commons.wikimedia.org/w/index.php?title=File:IBM_PC_5150.jpg&oldid=858452087 [Stand: 26.09.2024].
- Wikipedia-Autoren (2023, August 9). Gradientenverfahren. Wikipedia. https://de.wikipedia.org/w/index.php?title=Gradientenverfahren&oldid=236260947 [Stand: 20.09.2024].

- Wikipedia-Autoren (2024, August 23). Maschinelles Lernen. Wikipedia. https://de.wikipedia.org/w/index.php?title=Maschinelles_Lernen&oldid=247974908 [Stand: 08.09.2024].
- Zweig, K. A. (2019). Ein Algorithmus hat kein Taktgefühl: wo künstliche Intelligenz sich irrt, warum uns das betrifft und was wir dagegen tun können (Originalausgabe.). München: Heyne.

Empfehlung:

Geuter - re:publica 2023: tante - I'm sorry HAL, I won't let you do that. (2023). https://www.youtube.com/watch?v=3LlvHF-IX9Y [Stand: 27.09.2024].