

Künstliche Intelligenz – Einführung

Prof. Dr. Sibylle Schwarz

HTWK Leipzig, Fakultät IM

Gustav-Freytag-Str. 42a, 04277 Leipzig

Zimmer Z 411 (Zuse-Bau)

<https://informatik.htwk-leipzig.de/schwarz>

sibylle.schwarz@htwk-leipzig.de

17. April 2024

Was ist Künstliche Intelligenz?

EU-Factsheet on Artificial Intelligence (2018)

(<https://digital-strategy.ec.europa.eu/en/library/factsheet-artificial-intelligence-europe>)

Artificial intelligence (AI) refers to systems that show intelligent behaviour: by analysing their environment they can perform various tasks with some degree of autonomy to achieve specific goals.

*Mobile phones, e-commerce tools, navigation systems and many other different sensors constantly gather **data** or **images**. AI, particularly **machine-learning** technologies, can learn from this torrent of data to **make predictions** and **create useful insights**.*

Was ist Künstliche Intelligenz?

BMBF: Aktionsplan Künstliche Intelligenz (2023)

Fragen und Antworten zur Künstlichen Intelligenz

(<https://www.bmbf.de/bmbf/shareddocs/faq/20230822-faq-ki.html>)

*Computerprogramme, die in der Lage sind, selbstständig **Entscheidungen zu treffen** oder **Probleme zu lösen**. Sie können sich dabei selbst **an neue Gegebenheiten** anpassen, um ihre Arbeitsaufgabe bestmöglich lösen zu können. Wie das passiert, ist unter Umständen für den Menschen nicht vollständig nachvollziehbar. Der Mensch liefert lediglich die Grundlagen: Entweder abstrakte Programmieranweisungen oder einen Datensatz, aus denen die Maschine oder das Computerprogramm selbstständig lernt.*

***Zurzeit** sind insbesondere KI-Verfahren (als sogenannte generative KI) **erfolgreich**, die Zugriff auf einen riesigen Datenbestand von Texten, Bildern oder Tönen haben. Es gelingt ihnen eigenständig aus diesen Daten neue Texte, Bilder und Töne zu erzeugen, die **plausibel erscheinen** und so vorher nicht in der Datenmenge vorkamen.*

Können Maschinen denken?

Alan Turing 1950

Konkretisierung der Frage:

Können Maschinen **denken**?

zur überprüfbareren Frage:

Können Maschinen konstruiert werden, die einen **speziellen Test bestehen**?

Imitation Game

Imitation Game (Alan Turing 1950):

- ▶ zwei verschlossene Räume,
in einem befindet sich **Herr** A, im anderen **Frau** B
- ▶ eine Person C (Frager) stellt Fragen, A und B antworten
- ▶ Kommunikation über neutrales Medium,
an welchem das Geschlecht nicht erkennbar ist,
- ▶ C soll herausfindet, in welchem Raum Frau B ist
- ▶ Herr A versucht, C irrezuführen
- ▶ Frau B versucht, zu verdeutlichen, dass sie die Frau ist
(kooperiert mit C)

Herr A besteht den Test, wenn ihn C für Frau B hält.

Wie erkennt man Intelligenz: Turing-Test

Turing-Test 1950: verschiedene Versionen des Imitation Game

- ▶ A ist Machine statt Mann (B Person beliebigen Geschlechts)
- ▶ verschiedene Kooperationsverhalten von A und B

Turing-Test bewertet eigentlich die
Fähigkeit zur Kommunikation in natürlicher Sprache

= Intelligenz ?

Beginn koordinierter Forschung zur Künstlichen Intelligenz

John McCarthy

Programmiersprachen

Marvin Minsky

Kognitionswissenschaft

Claude Shannon

Informationstheorie

stellten 1955 die Vermutung auf, dass

„jeder Aspekt des Lernens oder jedes anderen Ausdrucks von Intelligenz prinzipiell so **präzise beschrieben** werden kann, dass sich eine Maschine konstruieren lässt, die ihn simuliert. “

Begriff Künstliche Intelligenz

McCarthy formulierte das Ziel,

„herauszufinden, wie man Maschinen konstruiert, die

- ▶ natürliche Sprache benutzen,
- ▶ Abstraktionen und Begriffe entwickeln,
- ▶ Aufgaben lösen, die (bis dahin) nur Menschen lösen konnten,
- ▶ sich selbst verbessern.“

und prägte dafür den Begriff **Künstliche Intelligenz**.

Beginn koordinierter Forschung zur Künstlichen Intelligenz

1956: erste Konferenz zur Künstlichen Intelligenz

Dartmouth Summer Research Project on Artificial Intelligence

Themen:

- ▶ Berechnungsmodelle in Computern
- ▶ Kommunikation mit Computern in natürlicher Sprache
- ▶ Berechenbarkeitstheorie
- ▶ Neuronale Netzwerke
- ▶ Selbst-Verbesserung
- ▶ Abstraktionen
- ▶ Zufälligkeit und Kreativität

Forschung zur Künstlichen Intelligenz

Momentaufnahme 2006:

Dartmouth Artificial Intelligence Conference: The Next Fifty Years

Themen:

- ▶ Modelle des (menschlichen) Denkens
- ▶ Neuronale Netzwerke
- ▶ (Maschinelles) Lernen und Suchen
- ▶ Maschinelles Sehen
- ▶ Logisches Schließen
- ▶ Sprache und Kognition
- ▶ KI und Spiele
- ▶ Interaktion mit intelligenten Maschinen
- ▶ Ethische Fragen und zukünftige Möglichkeiten der KI

KI-Erfolge – Auswahl

- ▶ 1945 frühe Schachprogramme (ohne Implementierung)
- ▶ 1955 Logic Theorist: automatischer Beweiser
- ▶ 1958 erster erfolgreicher Neurocomputer Mark I Perceptron
- ▶ 1961 General Problem Solver, z.B. zum Lösen von Rätseln und Intelligenztests
- ▶ 1966 ELIZA: Chatbot (als **Gegenbeispiel für Turing-Test**)
- ▶ 1972 erster mobiler Roboter
- ▶ ab ca. 1970 Beschränkung auf spezialisierte Expertensysteme
- ▶ 1976 MYCIN (Medizinisches Diagnosesystem)
- ▶ 1980 Dendral (Molekülstruktur aus Massenspektrogramm)
- ▶ 1982 XCON (Konfiguration von Computersystemen)
- ▶ ab ca. 1980 Expertensystem-Shells
- ▶ seit 1992 automatisiertes Fahren
- ▶ seit 1993 RoboCup Roboter-Fußball
- ▶ 1997 Deep Blue gewinnt gegen amtierenden Weltmeister
- ▶ 2011 Watson schlägt zwei Meister in Quizshow Jeopardy!
- ▶ 2016 AlphaGo schlägt Go-Meister, ...

Ansätze zur Entwicklung „intelligenter“ Systeme

- ▶ Simulation menschlichen **Verhaltens**
(Verständnis und eigenes Denken nicht notwendig)
Modellierung von Kognition,
Training mit vielen Fällen,
Finden von analogen Diagnosen/Entscheidungen
Vorgeschlagene/getroffene Entscheidungen können meist
nicht begründet und erklärt werden.

- ▶ Simulation des menschlichen **Denkens**
(Verständnis und eigenes Denken notwendig)
Modellierung des (menschlichen) Denkens
logisches Schließen, Abstraktion
Entscheidungen können nachvollziehbar begründet werden.
Denkmodelle, individuelle, kulturelle, moralische Faktoren

Kritik am Turing-Test

Kritik: Turing-Test überprüft nur Simulation von Verhalten

1966: Joseph Weizenbaum entwickelt den Chatbot ELIZA
(maschinelle Psychotherapeutin)

ELIZA besteht den Turing-Test.

Joseph Weizenbaum in
„Computer Power and Human Reason“ (1976)
zu den Reaktionen auf ELIZA:

Diese Reaktionen auf ELIZA haben mir deutlicher als alles andere bis dahin Erlebte gezeigt, welche enorm übertriebenen Eigenschaften selbst ein gebildetes Publikum einer Technologie zuschreiben kann oder sogar will, von der es nichts versteht.

Chinese-Room-Argument

Gedankenexperiment als Kritik am Turing-Test (Searle, 1980)

Situation: eine (nicht chinesisch verstehende) Person B in einem verschlossenen Raum mit einem (riesigen) Regelbuch mit chinesischen Fragen und passenden Antworten.

- ▶ A stellt Fragen, B antwortet.
- ▶ B antwortet mit Hilfe des Buches immer passend, ohne Frage und Antwort verstanden zu haben.

These: (anscheinend) intelligentes Verhalten ist noch **keine Intelligenz, wenn Verständnis fehlt.**

1. ursprünglich nur Gedanken-Experiment,
2. lange praktisch nicht umsetzbar (Datenmenge),
3. aktuelle Sprachmodelle und Chatbots simulieren das Wörterbuch durch Statistik-basiertes Zusammenfügen von Textfragmenten

Methoden „intelligenter“ Systeme

statistische Verfahren, z.B.

- ▶ maschinelles Lernen
- ▶ künstliche neuronale Netze
- ▶ evolutionäre Algorithmen
- ▶ Ameisen-Algorithmen

symbolische Verfahren, z.B.

- ▶ Logisches Schließen, Inferenzsysteme
- ▶ Suche in Zustandsräumen mit Erfahrungswerten / Heuristiken
- ▶ deklarative Programmierung

hybride Verfahren, z.B.

- ▶ Monte-Carlo-Suche
- ▶ mehrwertige Logiken (fuzzy, probabilistisch)
- ▶ do-Kalkül (Kausalität)

Phasen in der KI-Geschichte

wechselnde Betonung **symbolischer** und **statistischer** KI-Gebiete

- ▶ ca. 1950-70
symbolisch: Logik, Inferenz-Maschinen, Lisp, Chatbot Eliza
statistisch: künstliche neuronale Netze,
Robotik, Verarbeitung natürlicher Sprache
- ▶ ca. 1970-85 (symbolisch)
Prolog, (spezialisierte) Expertensysteme
nichtmonotones Schließen
- ▶ ca. 1985-2000 (statistisch)
maschinelles Lernen, KNN, evolutionäre Alg.,
Schwarm-Intelligenz, (Fuzzy-Logik) ...
autonome Fahrzeuge (Ernst Dickmanns)
- ▶ ca. 2000-2010 (symbolisch)
Constraint-Programmierung
SAT-Solver, Theorem-Prover
Ontologien (Semantic Web), Beschreibungslogiken
- ▶ seit ca. 2010 (statistisch)
Deep Learning, CNN, generative Ansätze

Denkmodelle

individuell abhängig z.B. von

- ▶ Persönlichkeit
- ▶ intellektuelle Ressourcen
- ▶ Kultur
- ▶ Bildung, Training
- ▶ Erziehung

Derartige Aspekte beeinflussen Entscheidungen und Verhalten und müssen bei der Modellierung von Denkmodellen mit betrachtet werden.

Beispiel: Moralische Entscheidungen

Beispiele ethischer Prinzipien:

Utilitarismus: Um Schaden zu vermeiden, darf (gesamtgesellschaftlich) geringerer Schaden für Unbeteiligte in Kauf genommen werden.

double effekt: Um Schaden zu vermeiden, darf Schaden für Unbeteiligte **nicht als Mittel** eingesetzt werden.

triple effekt: Um Schaden zu vermeiden, darf Schaden für Unbeteiligte **als Nebenwirkung** in Kauf genommen werden.

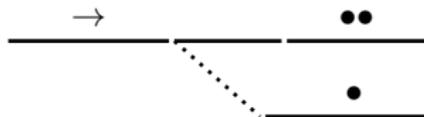
Unterschied zwischen double und triple effect:

Unterscheidung zwischen „Mittel“ und „Nebenwirkung“

Hätte die Aktion den ursprünglichen Schaden auch vermieden, wenn die dadurch ausgelöste schädliche Wirkung (aufgrund einer anderen Ausgangsposition) nicht eingetreten wäre?

Beispiel Trolley Probleme (Foot [1967])

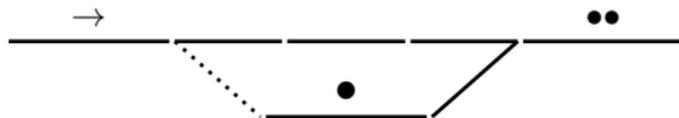
► Bystander Case



► Footbridge Case



► Loop Case



Welche Aktionen sind moralisch zulässig?

Ethische Prinzipien



Bystander Case

Loop Case

Footbridge Case

double effect	change	-	-
triple effect	change	change	-
utilitaristisch	change	change	throw down
Aktion zulässig meinen	85%	56%	12%

in den Trolley-Beispielen:

- ▶ Bystander: Hätte die Aktion (Weiche) beide auf dem Hauptgleis gerettet, wenn niemand auf dem Abstellgleis steht?
- ▶ Loop: Hätte die Aktion (Weiche) beide auf dem Hauptgleis gerettet, wenn niemand auf dem Umweg steht?
- ▶ Footbridge: Hätte die Aktion (Stoß) beide auf dem Hauptgleis gerettet, wenn niemand von der Brücke stürzt?

viele Varianten: <https://www.moralmachine.net/>

Gruppenaufgabe: Zahlenrätsel

Überlegungen beim Lösen des folgenden Rätsels protokollieren.

Eine Person A wählt zwei verschiedene natürliche Zahlen zwischen (einschließlich) 2 und 100 und verrät S deren Summe und P deren Produkt. Dann kommt es zu folgendem Gespräch:

P: Ich kenne die beiden Zahlen nicht.

S: Das weiß ich.

P: Dann kenne ich die beiden Zahlen jetzt.

S: Dann kenne ich sie jetzt auch.

Welche Zahlen hat A gewählt ?