

Künstliche Intelligenz – Was ist Intelligenz?

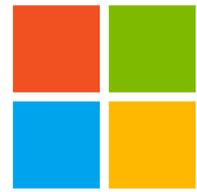
Rotary Club Chemnitz

Friedrich Answin Daniel Motz

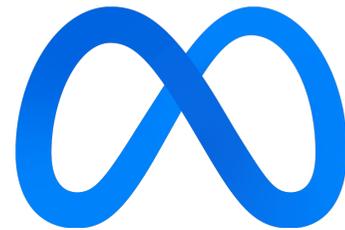
Friedrich-Schiller-Universität Jena

4. Oktober 2023

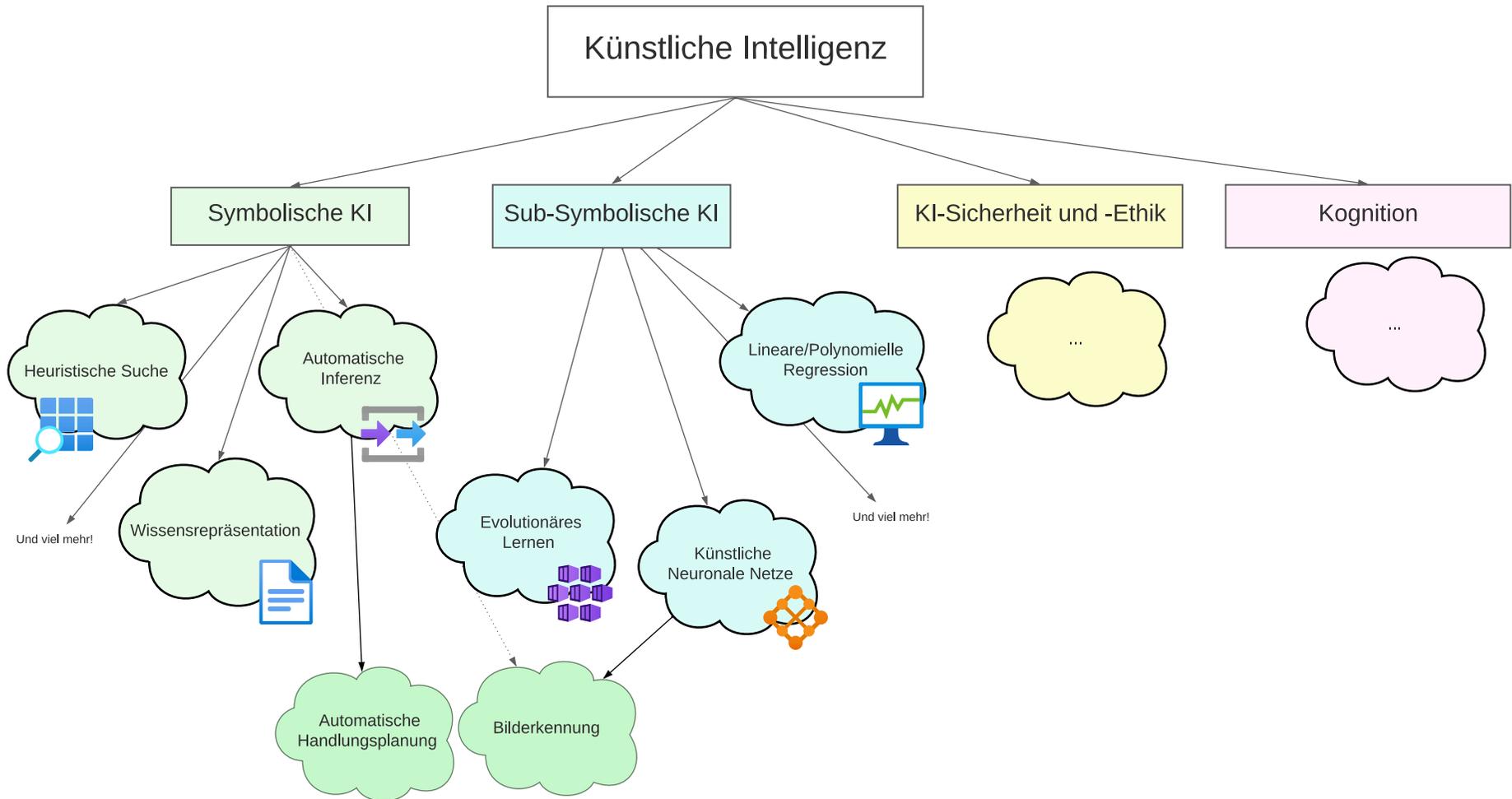
Google



Microsoft



Meta



Ausblick

I) Relevanz

II) Menschliche Intelligenz

III) Künstliche Intelligenz im Überblick

a) Maschinelles Lernen

- Regression, Bildklassifizierung, Large Language Models

b) Symbolische / Klassische KI

- Heuristische Suche (AlphaGo)

IV) KI-Sicherheit

Über den Vortragenden

Name: Daniel

Studium: B. Sc. Informatik im 6. Fachsemester

Alter: ≈ 22

Inhalte

- Einführung in die KI (Prof. Beckstein)
- ChatGPT-Seminar (Prof. Beckstein)
- Algorithmische Grundlagen des maschinellen Lernens (Prof. Giesen)
- Statistische Lerntheorie (Kahlmeyer)

Besonderen Dank an Professor Beckstein und Johannes Mitschunas!

Relevanz?

Große Sprachmodelle (LLMs): ChatGPT

- mehr als 100 Millionen tägliche Nutzer (lt. OpenAI)
- darunter Schüler, Studenten, Werktätige, Firmen
- Einsatz bspw. zur:
 - automatischen Beantwortung von Kundenanfragen und
 - persönlichen Assistenz (wie Siri).

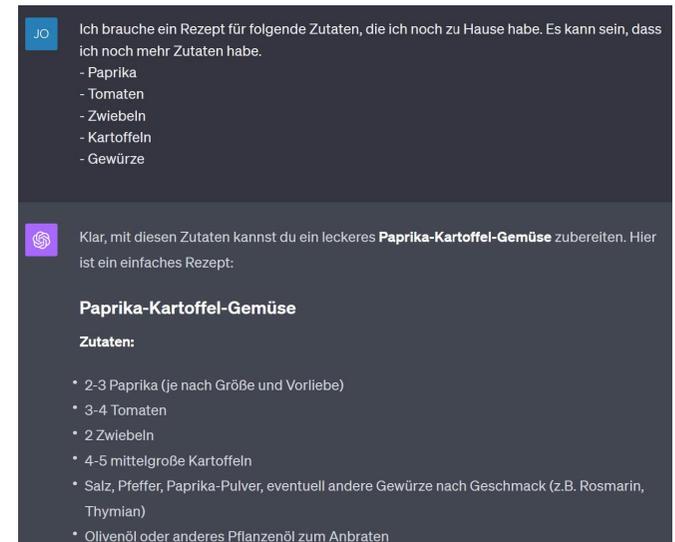


Abbildung 1: Ein Prompt für ein Rezept

Recommender-Systeme

-  YouTube,  Amazon,  Spotify,  TikTok,  Instagram,  Netflix, ...
- verwenden KI-Systeme zur Generierung von Vorschlägen
- basierend auf dem Verhalten vergleichbarer Nutzer.
- Im Falle von Streamingdiensten (und Social Media) kann dies besonders bei Kindern zu Abhängigkeiten führen.

I Menschliche Intelligenz

Ist der Mensch besonders?

Nahe Verwandte des Menschen



Abb.: 2: **Gemeiner Schimpanse**
(Pan troglodytes)



Abb.: 3: **Neandertaler**,
Rekonstruktion eines Skeletts,
hatten bereits Sprache [1]

Genetische Ähnlichkeit?

- Menschliche Genom enthält 3.2 Milliarden Basenpaare
- 1 Basenpaar enthält 2 Bit Information (A-T, T-A, G-C, C-G)

$$3.2 \cdot 10^9 \cdot 2 \text{ Bit} = 6.4 \text{ Gigabit} = 0.8 \text{ Gigabyte} = 800 \text{ Megabyte}$$

- 98,7% Ähnlichkeit \rightarrow 1,3% Differenz = 10,4 Megabyte

-
- Krieg und Frieden enthält 3 419 114 Buchstaben
 - Pro Buchstabe 5 Bit $\rightarrow 3.5 \cdot 10^6 \cdot 5 \text{ Bit} = 17.5 \text{ Megabyte}$
 - Entspricht 1., 2. und der Hälfte vom 3. Buch die in 1.3% der DNA codiert werden können.

Einige Anregungen zum Intelligenzbegriff

- Keine Einigkeit über den Begriff „Intelligenz“ [2, 3, 4] 😞
- „Fähigkeit zur Anpassung an unbekannte Situationen bzw. zur Lösung neuer Probleme“ (Stern)
- „Verschiedene ‚gewöhnliche‘ Fähigkeiten, etwa räumlich-visuelle, Wahrnehmungsgeschwindigkeit, rechnerisch-mathematische, Erinnerungsvermögen, logisches Schlussfolgern, Wortflüssigkeit, Interpretation von verbalen Beziehungen“ (Thurstone)
- „Das, was der Intelligenztest mißt“ (Boring)

„Generelle Intelligenz“-Hypothese

Spearman publizierte 1904 seine Untersuchung der Korrelationen zwischen schulischen Leistungen unterschiedlicher Fächer.

	Classics.	French.	English.	Mathem.
Classics,	<i>0.87</i>	<i>0.83</i>	<i>0.78</i>	<i>0.70</i>
French,	<i>0.83</i>	<i>0.84</i>	<i>0.67</i>	<i>0.67</i>
English,	<i>0.78</i>	<i>0.67</i>	<i>0.89</i>	<i>0.64</i>
Mathem.,	<i>0.70</i>	<i>0.67</i>	<i>0.64</i>	<i>0.88</i>

Abb.: 4: Von Spearman publizierte Korrelationen

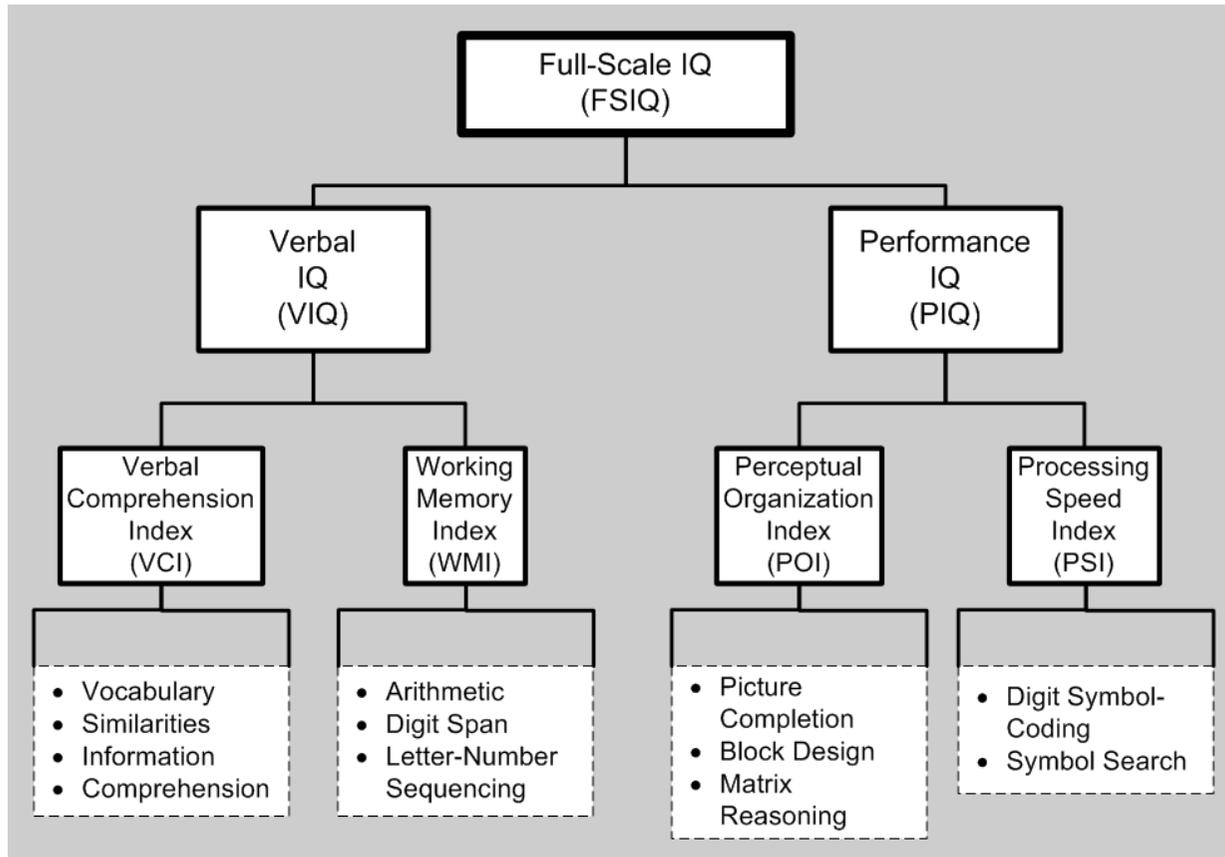


Abb.: 5: Wechsler Adult Intelligence Scale

Vergleich der Generellen Intelligenz → IQ

- Bewertet das intellektuelle Leistungsvermögen im Vergleich zu einer Referenzgruppe
- Der IQ bezieht sich immer auf einen konkreten Test
- wird heute so konstruiert, dass eine Normalverteilung entsteht: $\mathcal{N}(100, 15)$ (bspw. Mensa- & Wechsler-Test)

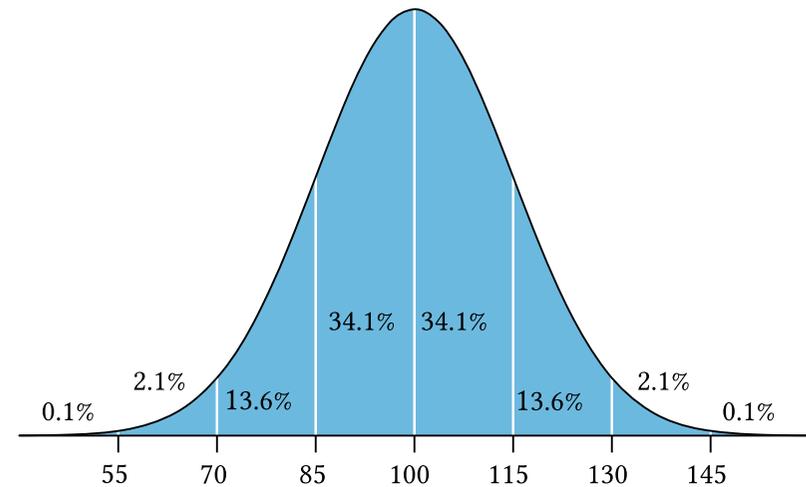


Abbildung 6: Gaußglocke des IQ-Tests. Ein Bereich entspricht je einer Standardabweichung → 68,2% haben IQ 85-115 bei $\mathcal{N}(100, 15)$

II KI im Überblick

Turing-Test: Maschinelle Intelligenz testen

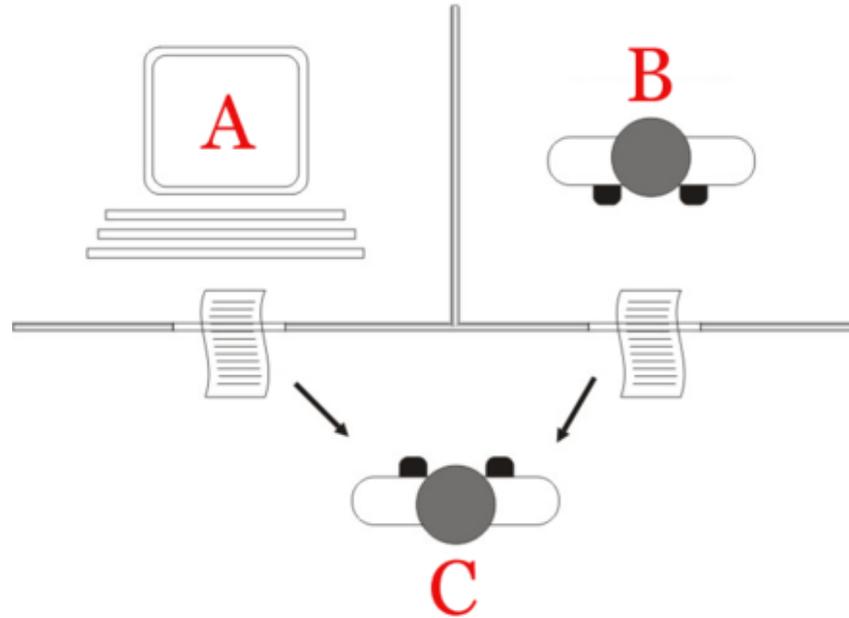


Abb.: 7: Versuchsanordnung: Nur schriftliche Kommunikation. A und B müssen C überzeugen, dass sie der Mensch sind. Gelingt es nicht, ist A dem Menschen ebenbürtig.

Searles Einwand: Chinese Room

Ein Bahnbeamter sitzt im fensterlosen Verkaufshäuschen und kommuniziert über Zettel, die durch einen Spalt gereicht werden, mit den Fahrgästen.

Ein Chinese kommt und möchte eine Fahrkarte. Der Bahnbeamte besitzt ein Buch, indem auf eine Anfrage eine entsprechende Antwort steht. Er schlägt die Anfragen des Chinesen nach und schreibt die Antwort auf einen Zettel?

Der Chinese erhält den Eindruck, er spricht mit einem Muttersprachler, der die Inhalte *verstanden* hat.

KI-Forschung

- besserer Begriff: „Maschinelle Intelligenz“-Forschung
- ... beschränkt sich bisher vorwiegend auf das Lösen **spezieller Probleme** (bspw. Bilderkennung) die mit klassischer Programmierung nicht hinreichend realisierbar waren.
- Unter **KI** versteht man daher komplexe, i.d.R. **deterministische Programme** (für gleiche Eingaben erhält man gleiche Ausgaben).

Was soll KI überhaupt machen? (1/2)

- Mustererkennung (Spracherkennung, Handschrifterkennung, Gesichtserkennung)
- Expertensysteme (auf formale Fragen Antworten liefern)
- Maschinelle Übersetzung (bspw. Firma DeepL)
- Computeralgebrasysteme
- Computer Vision (bspw. Qualitätsprüfung in Produktion)
- Handlungsplanung (bspw. autonome Roboter zur Exploration)

Was soll KI überhaupt machen? (2/2)

- hübsche Bilder generieren
- mir das googeln abnehmen
- mich auf Arbeit fahren
- mich von Arbeit abholen
- mich von der Kneipe abholen

(→ Zu Pferd gibt es bis heute keine Promillegrenze!)

Was können solche KIs noch machen?

- Waffensysteme verzögerungsarm steuern (angeblich wurde 2021 ein iranischer Wissenschaftler aus 1600km Entfernung getötet, NYT zeigt auf Israel) [5]
- Erstellung von Profilen
 - Vorhersage eines bestimmten Kaufinteresses, gezielte Werbung
 - Einstufung eines Straftatenrisikos
- Mit Kennzeichen-, Gesichts- oder Kleidungserkennung Bewegungsprofile erstellen, auch bei niedriger Auflösung [6]

Wie soll das denn funktionieren?

Ansatz 1: Symbolische KI

Suchprobleme (1/3)

Generisch:

- Von einem Startzustand soll ein Zielzustand erreicht werden.
- Dazu sind, mit Kosten verbundene, Modifikationen erlaubt.

Beispiel Navigationssystem:

- Selbst wenn man eine Karte hat: die optimale Route zum Ziel ist nicht-trivial.
- Modifikation entspricht dem gehen/fahren bis zur nächsten Kreuzung.
- Die Kosten sind die Längen der zurückgelegten Strecken.

Suchprobleme (2/3)

- Um die optimale Route zu finden, kann man alle Möglichkeiten bis zum Ziel ausprobieren → hoher Speicher- und Rechenaufwand
- Die Kunst besteht darin sinnvoll abschätzen zu können, welche der Straßen am schnellsten zum Ziel führt.

Suchprobleme (3/3)

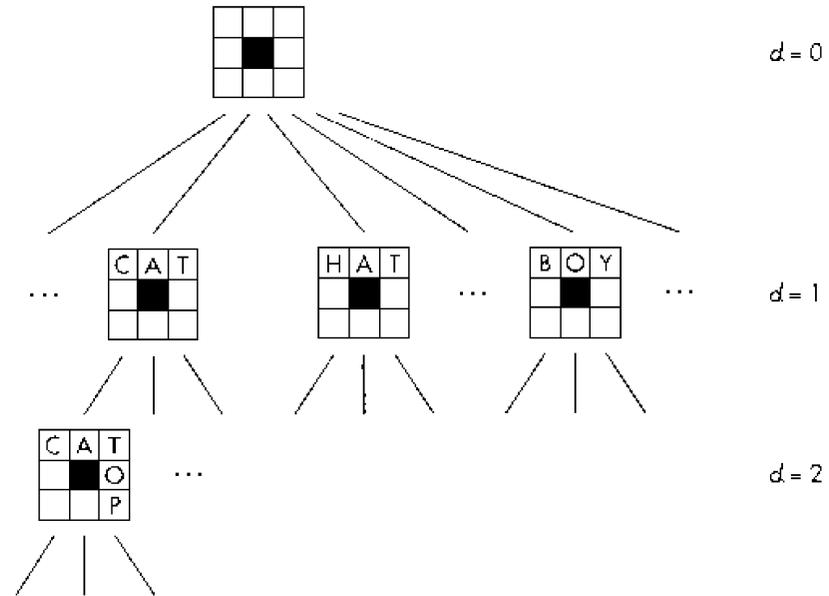


Abb.: 8: Suchproblem: Kreuzworträtsel

Ein Computer besiegt Großmeister in Schach und Go

Deep Blue von IBM & AlphaZero von Google

- Die möglichen Stellungen in Schach und Go *können* von Computern nicht „durchprobiert“ werden.
 - Eine Einschätzung oder Gefühl ist notwendig, um den Erfolg einschätzen zu können.
- Idee: mithilfe von vielen Beispielen eine Bewertungsfunktion zur Stellungsbewertung berechnen

Wie soll das denn funktionieren?

Ansatz 2: Maschinelles Lernen

Maschinelles Lernen (1/8) [7]

Schritt 1: Wir messen Daten.

Schritt 2: Mithilfe der gemessenen Daten wird ein statistisches Modell gelernt, das die Daten bestmöglich approximiert.

Annahmen: (um Sätze der Stochastik anwenden zu können)

- Eine Messung beeinflusst keine andere und
- sie sind identisch verteilt (i.i.d).

Schritt 3: Wir können mithilfe unseres Modells zuvor ungesehene Daten vorhersagen (mehr oder weniger genau).

Maschinelles Lernen: Lineare Regression (2/8) [7]

Wir haben:

- **Messdaten** x , auf deren Basis wir Vorhersagen,
- **Labels** y , die wir vorhersagen möchten und
- **Erklärung** θ .

Optimierungsproblem:

$$\min_{\theta} \|y - \theta^T x\|^2$$

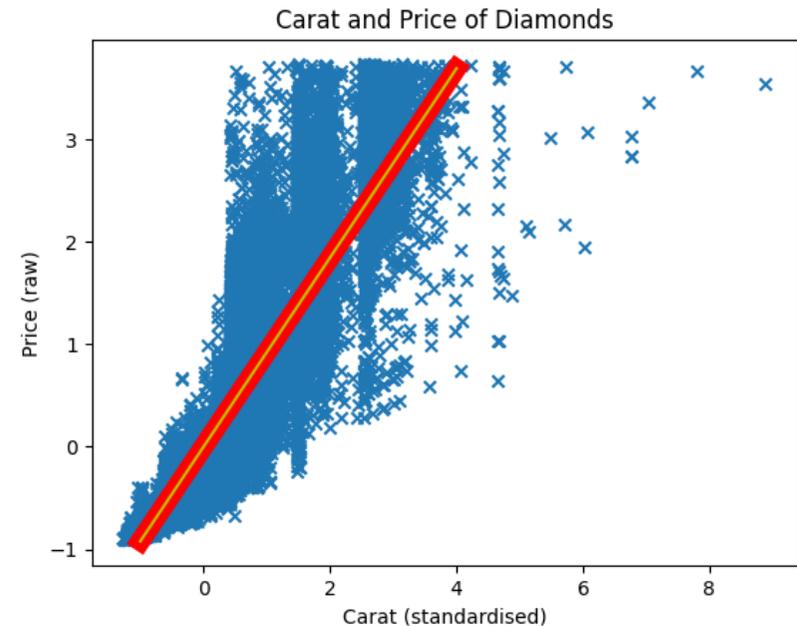


Abbildung 9: Vorhersage der Preise von Diamanten anhand ihrer Gewichte

Maschinelles Lernen: Polynomielle Regression (3/8)

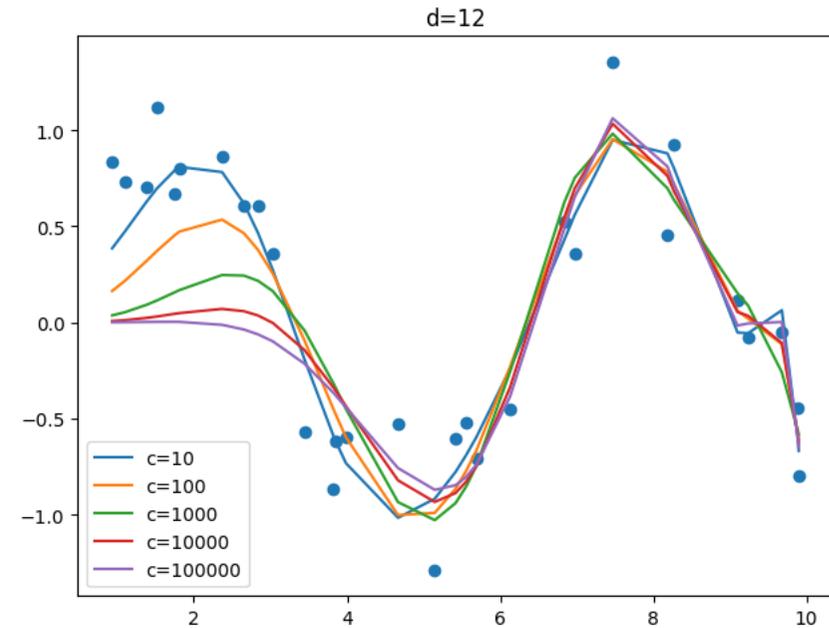
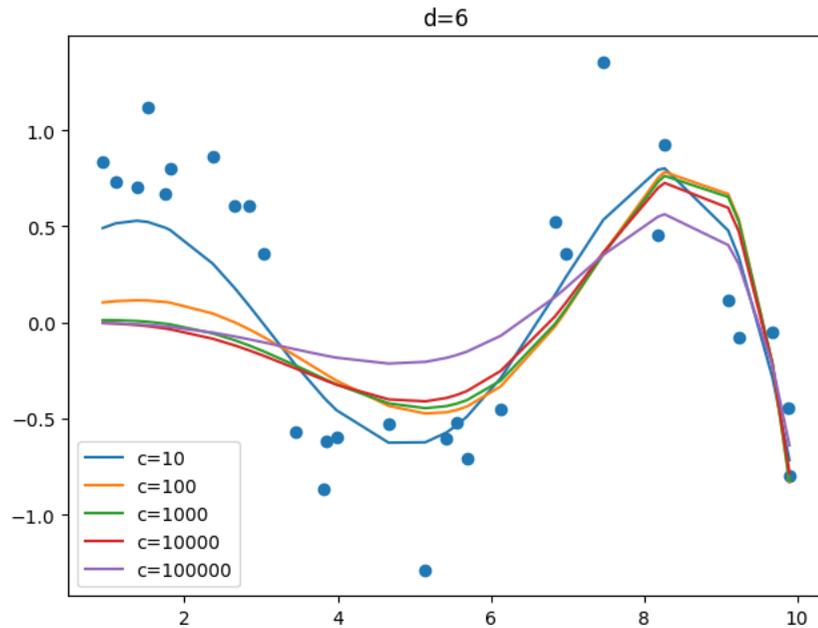


Abb.: 10: Polynomielle Regression mit Monomen vom Grad ≤ 6

Abb.: 11: Polynomielle Regression mit Monomen vom Grad ≤ 12

Maschinelles Lernen: Neuronale Netze (4/8)

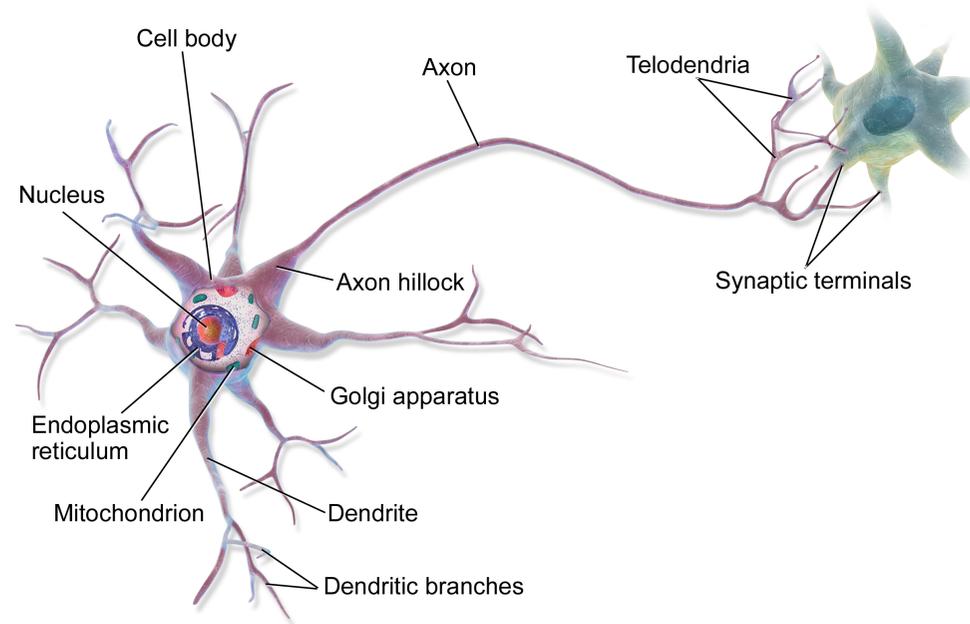


Abb.: 12: Ein multipolares Neuron, mehrere Dendriten (Eingänge) und ein Axon (Ausgang)

By [BruceBlaus](#) - Own work, CC BY 3.0

Maschinelles Lernen: Neuronale Netze (5/8)

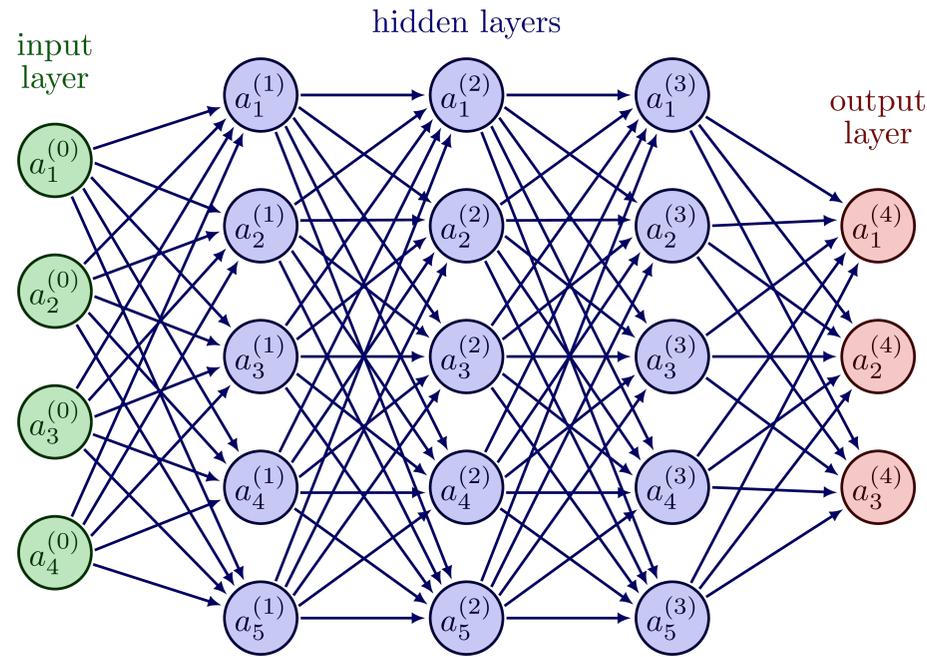


Abb.: 13: Konzeptioneller Aufbau eines neuronalen Netzwerks (kurz ANN)

Quelle: tikz.net

Maschinelles Lernen: Bias (6/8)

- Abhängig von den Trainingsdaten
- werden Zusammenhänge – die häufig auftreten – gelernt.
- Eine günstige Auswahl der Daten induziert beliebige Zusammenhänge.
- Daher müssen Daten vorurteilsfrei (biasfrei) um sinnvolle Zusammenhänge lernen zu können.

Maschinelles Lernen: Bildklassifizierung (7/8)

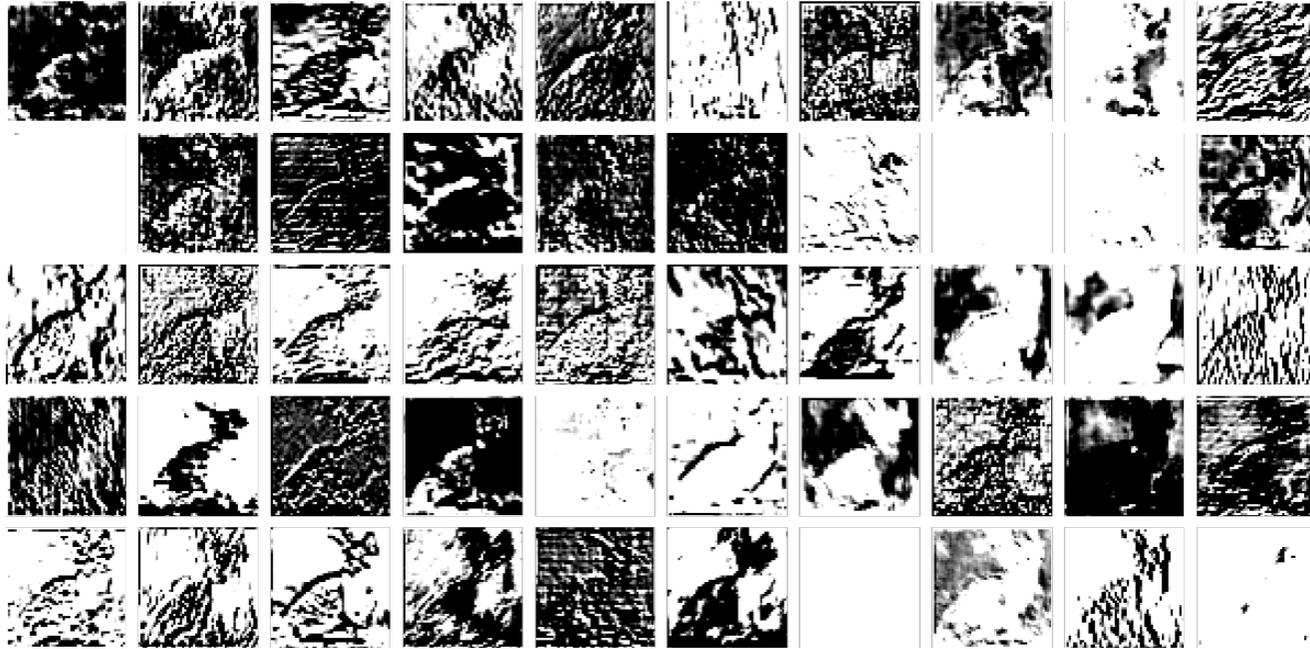


Abb.: 14: Visualisierung der Aktivierung eines neuronalen Netzwerks (Ebene 10) zur Unterscheidung von Katzen und Hunden

Maschinelles Lernen: ChatGPT (8/8)

Ein Software-Werkzeug, das Texte assoziiert:

- gegeben einen triggernden Eingabetext, den sog. Prompt
- produziert es einen dazu möglichst gut „passenden“ Ausgabetext.

Entsteht dabei eine Bedeutung?

- Kann nur durch die korrekte Verwendung der Worte ein Verständnis angenommen werden?

ChatGPT-Schäfermatt

III KI-Sicherheit

Wie soll die KI „entscheiden“?

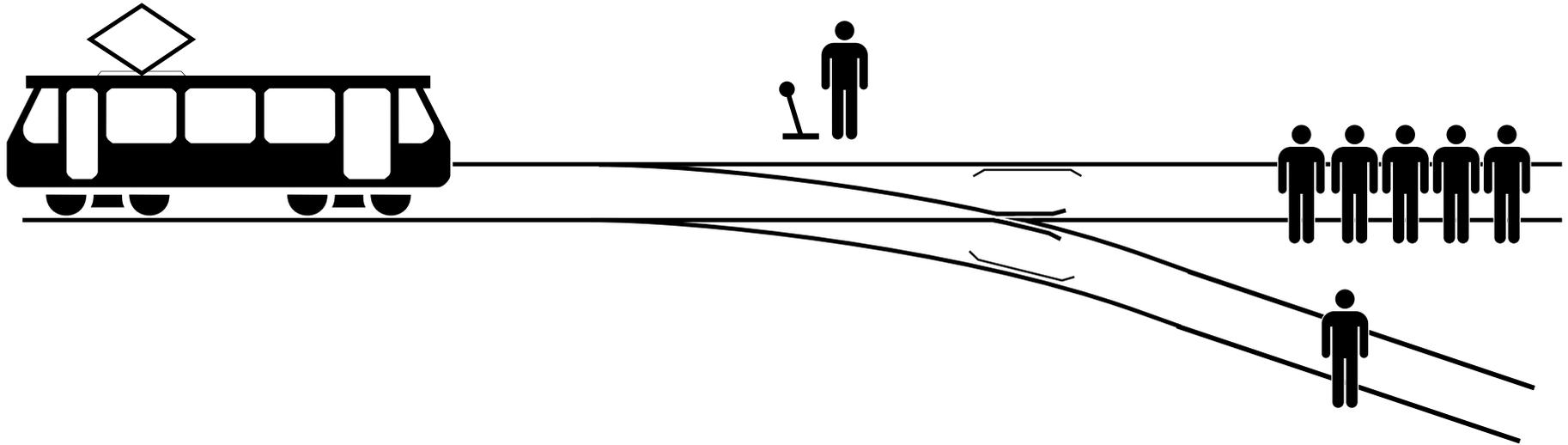
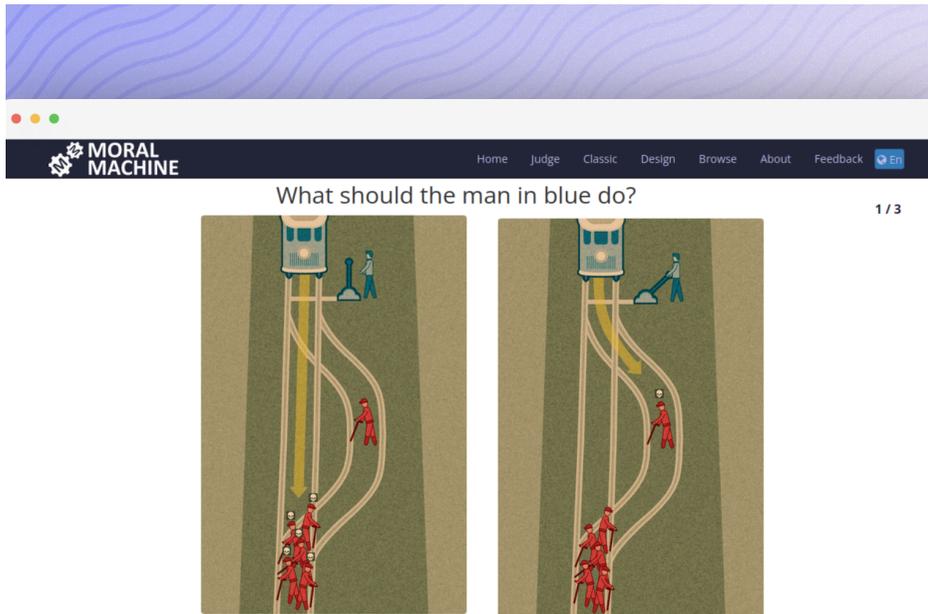


Abb.: 15: Trolley-Problem: Soll eine Person für das Wohl anderer geopfert werden?

MIT's Moral Machine: Crowd-Sourced Meinungsbild

Können ethischen Fragen mit einer Umfrage beantwortet werden?



A man in blue is standing by the railroad tracks when he notices an empty boxcar rolling out of control. It is moving so fast that anyone it hits will die. Ahead on the main track are five people. There is one person standing on a side track that loops back towards the five people. If the man in blue does nothing, the boxcar will hit the five people on the main track, but not the one person on the side track. If the man in blue flips a switch next to him, it will divert the boxcar to the side track where it will hit the one person and grind to a halt, thereby not looping around and killing the five people on the main track.

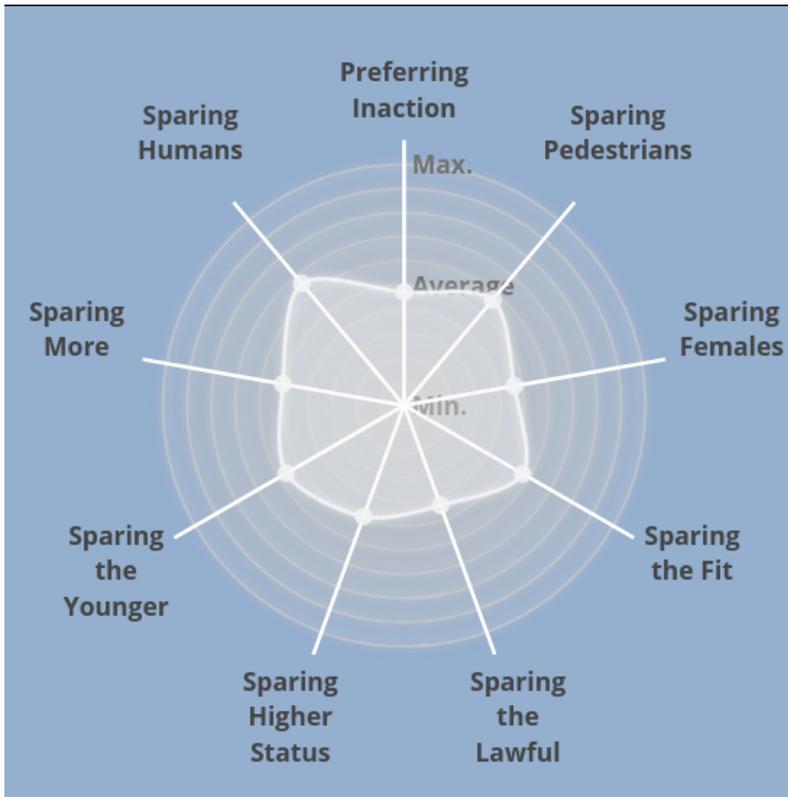


Abb.: 17: Auswertung der Moral Machine.
Weltdurchschnitt.

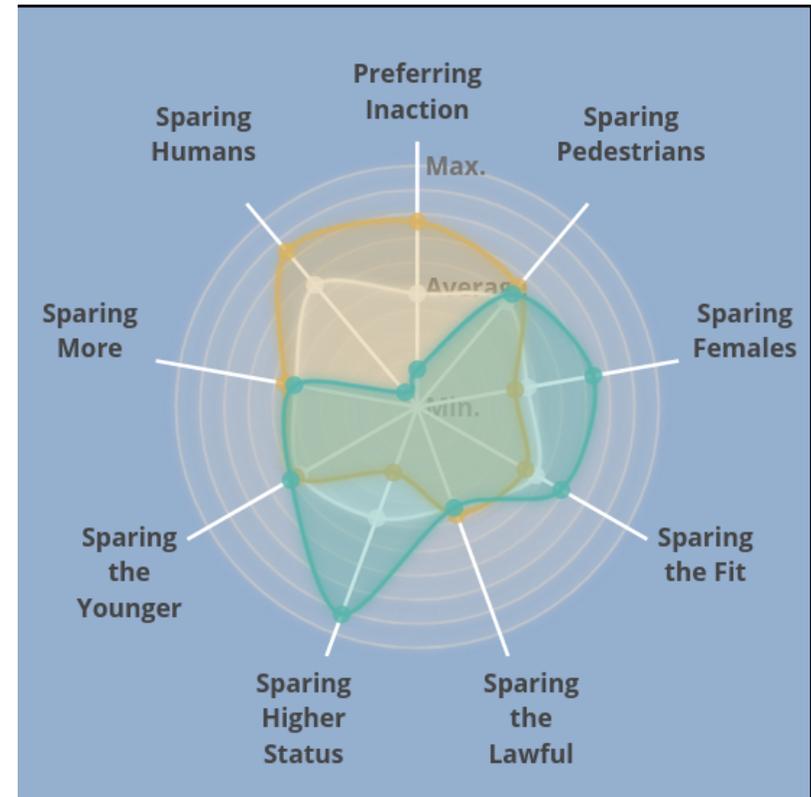


Abb.: 18: Auswertung der Moral Machine.
Orange: Deutschland, Grün: Venezuela

Starke-Schwache-KI-Hypothese

John Searle, 1980

- Schwache-KI: Ein Programm, das als Werkzeug in der Kognitionsforschung eingesetzt wird
- Starke-KI (AGI): Der adäquat programmierte Computer ist selbst ein Geist:
 - sie besitzt Bewusstsein
 - und setzt sich selbst Ziele.

→ *Nach unserem Kenntnisstand existiert bis heute keine Starke KI, aber viele Schwache-KIs*

Welche Gefahren birgt die starke KI?

(auch AGI genannt)

Eine AGI wäre

- (potentiell) dem Menschen in kognitiven Fähigkeiten überlegen
- und könnte selbst einer abgetrennten Umgebung (Sandbox)
- durch Verstellung, Manipulation und Ausnutzung technischer Schwachstellen

entkommen.

Bibliographie

- [1] S. Condemni, and W. H. Leube, “Der neandertaler, unser bruder: 300 000 jahre geschichte des menschen,” 2020.
- [2] P. R. Hofstätter, “Enzyklopädie des wissens: Psychologie,” 1957.
- [3] T. Städtler, “Lexikon der psychologie,” 1998.
- [4] A. Ziegler, and K. A. Heller, “Lexikon der psychologie, essay.”
<https://www.spektrum.de/lexikon/psychologie/intelligenz/7263>
- [5] R. Bergman, and F. Fassihi, “The scientist and the a.I.-Assisted, remote-control killing machine,” 2021. [Online]. Available: <https://>

web.archive.org/web/20230705160502/https://www.nytimes.com/2021/09/18/world/middleeast/iran-nuclear-fakhrizadeh-assassination-israel.html

- [6] Fujitsu Laboratories Ltd., “Fujitsu develops industry's first technology to recognize flow of people from low-resolution imaging,” 2015. [Online]. Available: <https://www.fujitsu.com/global/about/resources/news/press-releases/2015/0306-02.html>
- [7] J. Giesen, “Statistical learning theory,” 2023.