



Einführung: Künstliche Intelligenz

Maximilian Koch

Freie Universität Berlin, FUB-IT, Abt. Lehre, Studium und Forschung
Arbeitsbereich E-Learning und E-Examinations (EEE)

VERANSTALTET VON:



IM RAHMEN EINES PROJEKTES VON:



GEFÖRDERT VON:



Einführung: Künstliche Intelligenz

Einblick, Überblick, Ausblick

Programm 27.02.2025; 10-12 Uhr

(Anmoderation und Vorstellung)

1. Was ist Künstliche Intelligenz?

- Geschichte der KI
- Grundlegende Begrifflichkeiten

2. Training von KI-Anwendungen

- Trainingsmethoden
- Probleme und Lösungen(?)

3. Rechtliche & ethische Fragen zu KI

- Rechtlicher Rahmen (AI Act)
- Relevante weitere Gesetze
- Moralische KI

Schlusswort: „Textverständnis“ von KI

(Evaluation und Abmoderation)

Was ist Künstliche Intelligenz?

Einstiegsfrage: Slido

Welche KI-Anwendungen haben Sie in dieser Woche schon bewusst genutzt?

- Inhalt-/Textgenerierung (ChatGPT und Co.)
- Empfehlungssysteme (Youtube, Twitter und Amazon)
- Sprachassistentz (Siri, Alexa und Cortana)

- Integrierte KI:
 - Word: Bildbeschreibung
 - Adobe: Optical Character Recognition
 - Iphone: Bildersuche

Geschichte der Künstlichen Intelligenz

Meilensteine der KI-Forschung

1936

- Mathematischer Grundstein für KI: Kognitive Prozesse durch Algorithmus abbildbar (Alan Turing)

1956

- „Künstliche Intelligenz“ als Begriff für Simulation menschlicher Intelligenz eingeführt (John McCarthy)

1957

- Rosenblatt-Perzeptron als erstes künstliches Neuron (Frank Rosenblatt)

Geschichte der Künstlichen Intelligenz

Meilensteine der KI-Forschung

1997

- „Deep Blue“ (IBM) schlägt Schachweltmeister Garry Kasparov

2006

- „Deep Learning“ erkennt Handschrift mit hoher Genauigkeit (Geoffrey Hinton et al.)

2011

- „Siri“ als erste Spracherkennung erreicht marktreife

Geschichte der Künstlichen Intelligenz

Meilensteine der KI-Forschung

2017

- Selbstfahrende Taxis von Waymo (Google) als erstes kommerzielles Angebot dieser Art in Phoenix (Arizona)

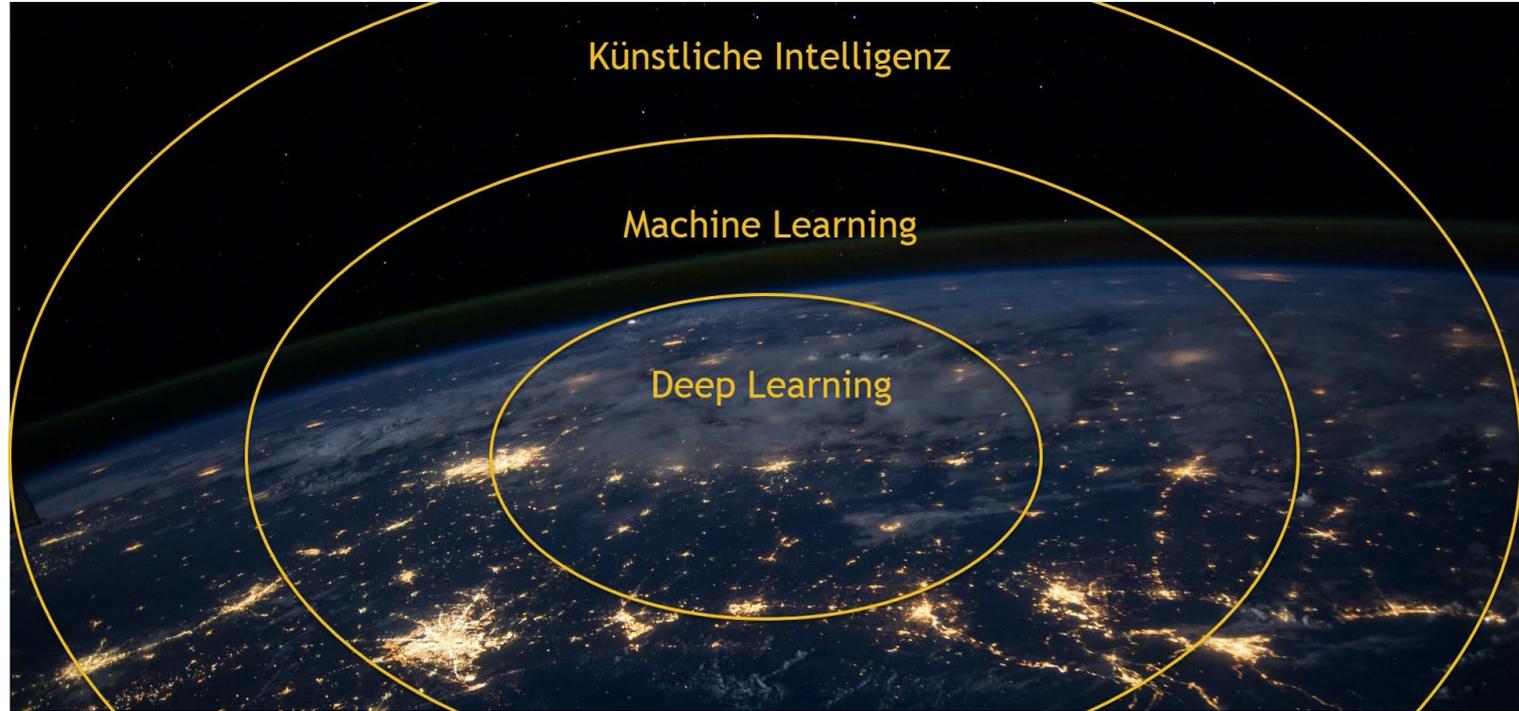
2022

- ChatGPT von Open AI zeigt die Fähigkeiten von Large Language Modellen und macht Textgenerierung für alle verfügbar

2025

- DeepSeek von DeepSeek stellt den enormen Ressourcenaufwand von ChatGPT oder Llama in Frage

Begrifflichkeiten Künstlicher Intelligenz



Eigene Darstellung; Foto: [NASA/Unsplash](#)

Begrifflichkeiten Künstlicher Intelligenz

Schwache KI vs. Starke KI
(Spezifische KI vs. Allgemeine KI)



Bild: [stepintofuture/pixabay.com](#)

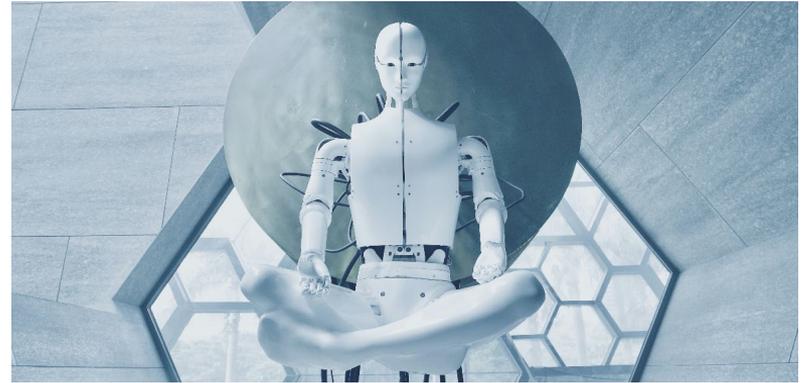
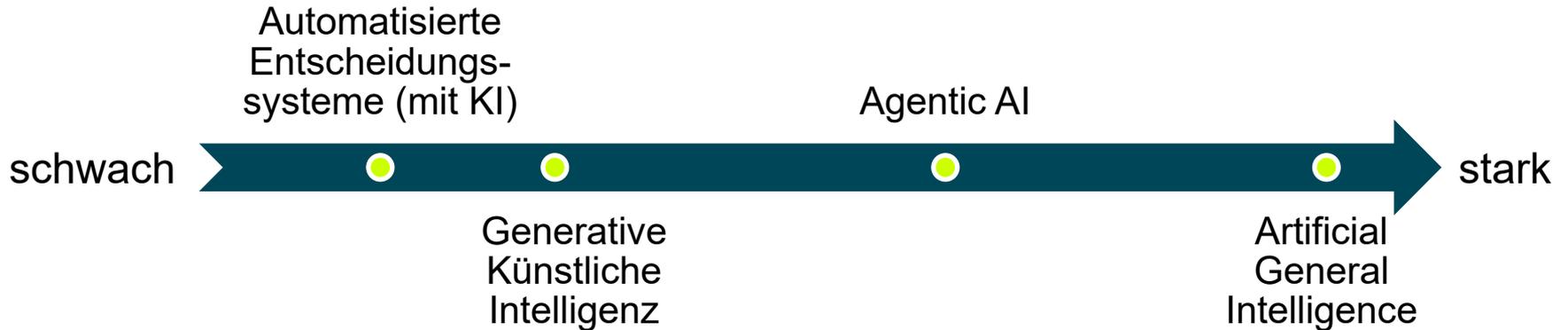


Bild: [Aideal Hwa](#) on [Unsplash](#)

Begrifflichkeiten Künstlicher Intelligenz



Umgang mit Unsicherheit

Begrifflichkeiten Künstlicher Intelligenz: Big Data

- ✓ großer Umfang (Volume)
- ✓ hohe Geschwindigkeit (Velocity)
- ✓ enorme Vielfalt (Variety)

Eigene Darstellung; Photo: [NASA/Unsplash](#)

Q & A: Slido



Bild: [Emily Morter](#) auf [Unsplash](#)

Training von KI-Anwendungen

Einstiegsfrage: Slido

Was sind für Sie die wichtigsten Qualitäten einer KI-Anwendung?

- Genauigkeit vs. Kreativität
- unverzerrte Ergebnisse
- nachvollziehbare Ergebnisse

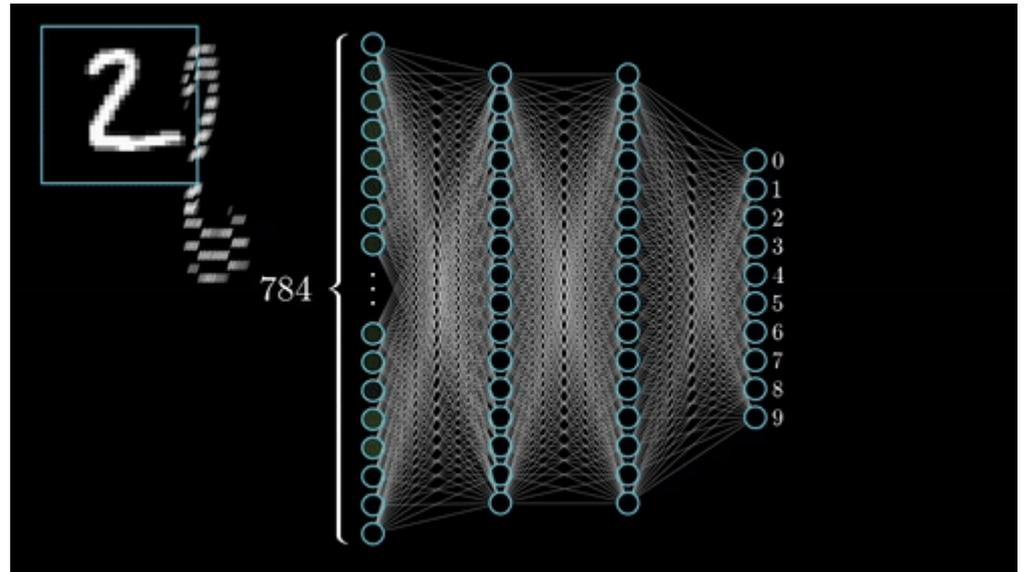
- Schnelligkeit
- Intuitive Bedienoberfläche (UX-Design)

Training von KI: Deep Learning

Künstliche Neuronales Netzwerke

- Knoten-Punkte → Neuronen
- Verbindungen → Gewichtung
- Werte der Verbindung entscheiden über Ergebnis

Training = Anpassung der Gewichte



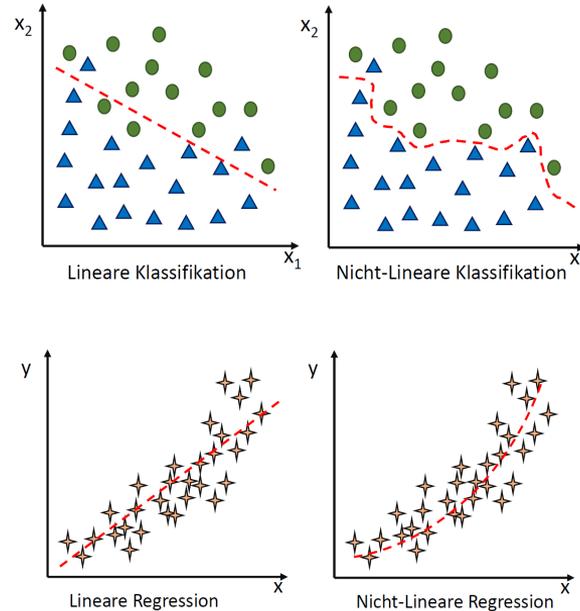
GIF: 3Blue1Brown, Neural networks: <https://youtu.be/aircAruvnKk?si=Y9kiyzYSp8itht7Z>

Training von KI: Supervised Learning

Methode:

Trainingsdaten und Testdaten haben überprüfbares „richtiges“ Ergebnis.

Hauptaufgabe: Erkennen, Vorhersagen.



Quelle: <https://data-science-blog.com/blog/2017/12/20/maschinelles-lernen-klassifikation-vs-regression/>

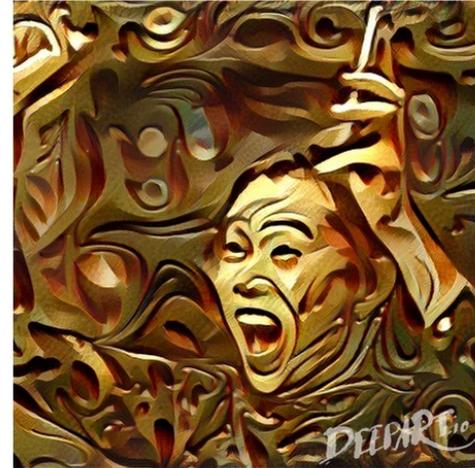


Training von KI: Unsupervised Learning

Methode:

Trainingsdaten haben kein gewünschtes Ergebnis; Maschinen erkennen Muster in Daten und reproduzieren sie.

Hauptaufgabe: Mustererkennung, Generieren.



Quelle: Screenshot/ deepart.io

Training von KI: Reinforcement Learning

Methode:

Wie Konditionierung: Agent (KI) versucht mit handeln die Belohnung zu maximieren.

**Hauptaufgabe: Wegfindung,
Handlungsentscheidungen**



Bitc: DeepMind

„Hausaufgabe“: Bilderkennung trainieren

Aufgabe:

- Trainieren von „Teachable Machine“ (Google)
- 1000 Katzen- und 1000 Hundebilder
- Learning: Erfahren Sie wie sich das Training auf die Qualität der KI auswirkt.

Links:

- [Teachable Machine](#)
- [Trainingsdaten](#)



Bild: [Andrew S on Unsplash](#)

Training von KI: „Garbage in – Garbage out“

Underfitting:

Das Modell hat nicht genug Informationen für eine akkurate Entscheidung.

Beispiel: Erkennt einen Chiwawa als Katze.

Overfitting:

Das Modell orientiert sich zu stark an den Trainingsdaten: „Irrelevante“ Informationen werden in das Modell aufgenommen.

Beispiel: Hintergrund von Hunde-/Katzenbildern.

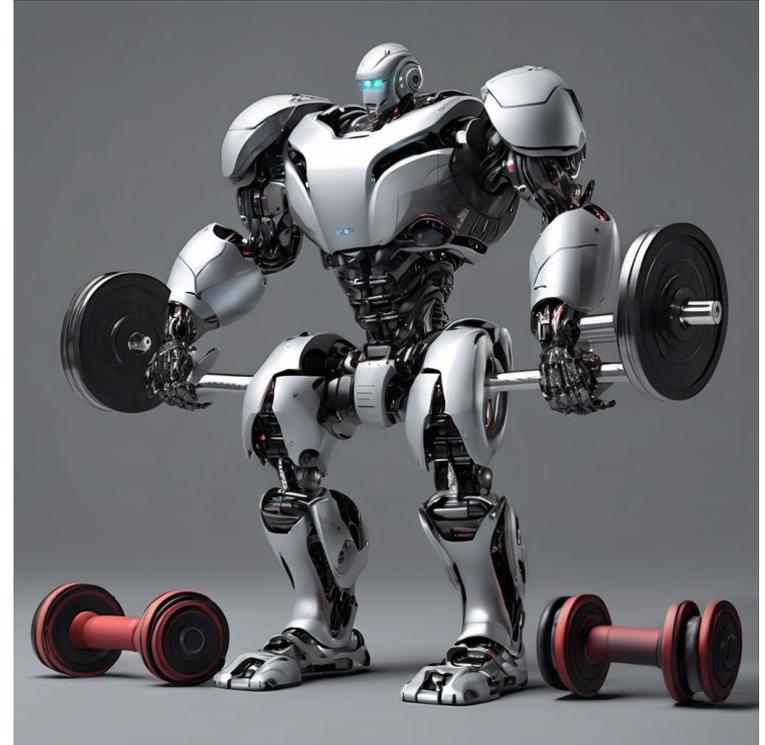


Bild: dreamstudio.ai; Prompt: "Super realistic Roboter lifting weights"

Training von KI: „Bias in – Bias out“

KI-Anwendungen lernen von Menschen: Menschliche Vorurteile werden übernommen.

- [Chatbot Tay](#) wird durch Twitter rassistisch, beleidigend und leugnet Holocaust (2016)
- [Google](#) erkennt keine Gorillas (2018)
- [Amazons](#) Bewerbungsauswahl durch KI sexistisch (2018)
- Bildvorschau bei [Twitter](#) ist rassistisch und sexistisch (2020)



<https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist>

Training von KI: „Explainable AI“ als Lösung?

Blackbox-Problem: Wie entscheidet die KI?



Bilder: <https://de.steadforce.com/how-tos/explainable-object-detection>, <https://cocodataset.org/#explore?id=76937>, <https://github.com/marcotcr/lime>

Q & A: Slido



Bild: [Emily Morter](#) auf [Unsplash](#)

Rechtlicher Rahmen für KI-Anwendungen

Einstiegsfrage: Slido

Welche Bedenken/Ängste haben Sie bezüglich Einsatz und Entwicklung von KI-Anwendungen?

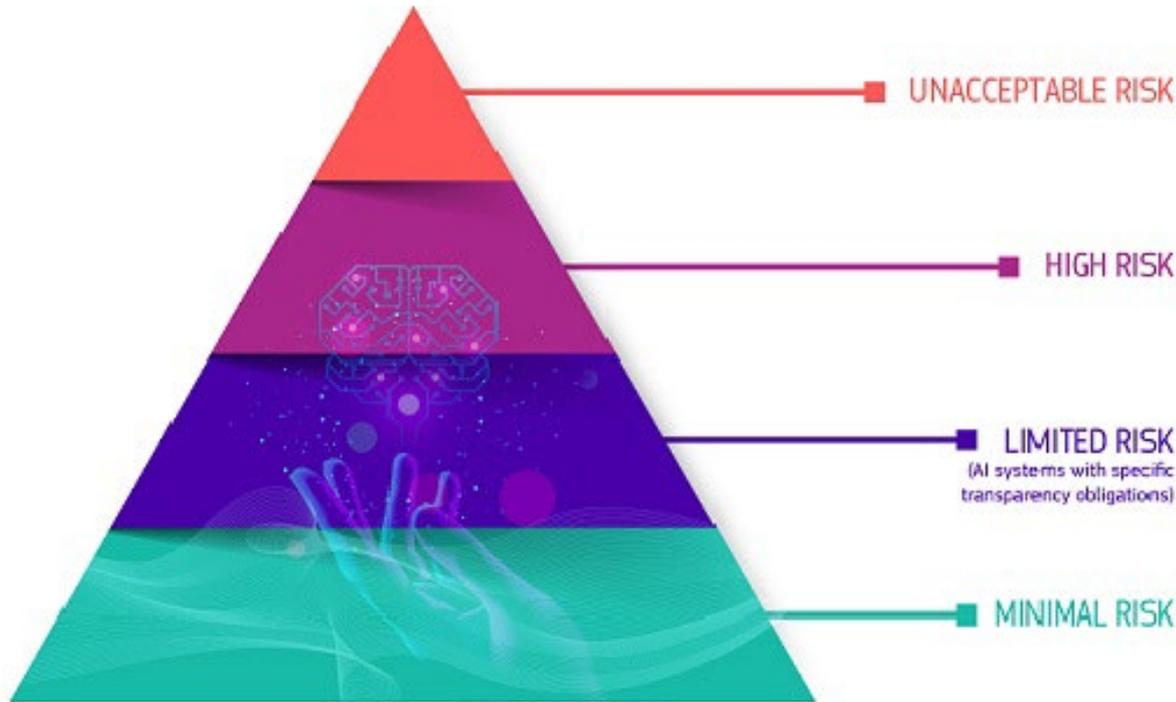
- Grundrechte-Verletzung durch KI-Nutzung → AI Act (EU)
- Nutzung/Missbrauch von sensiblen Daten → Datenschutz
- Nutzung von geistigem Eigentum anderer → Urheberrecht
- Vortäuschen von Leistung → Prüfungsrecht

Regelung in der EU: AI Act



- **AI Act (in Kraft) – 5.2024**
 - Fokus: KI-Anwendungen, LLMs (General Purpose AI)
 - Ziele:
 - Unternehmen verantwortlich machen (Anforderung)
 - Schutz der Nutzer*innen (Verbot, Kennzeichnungspflicht)
- Umsetzung:
 - Anwendung ab 8.2025-8.2027 (Schonfrist)
 - Strafen höher als bei DSGVO-Verstößen

Regelung in der EU: AI Act



Bedrohung von Sicherheit und Grundrechten

Schwerwiegende Risiko für Sicherheit, Gesundheit und Grundrechten

Risiko der Täuschung: Kennzeichnungspflicht

Kein Risiko, keine Regelung

Bild: <https://digital-strategy.ec.europa.eu/de/policies/regulatory-framework-ai>

Regelung in der EU: Was heißt „KI-Kompetenz“?

„Die Anbieter und Betreiber von KI-Systemen ergreifen Maßnahmen, um **nach besten Kräften** sicherzustellen, dass ihr Personal und andere Personen, die in ihrem Auftrag mit dem Betrieb und der Nutzung von KI-Systemen befasst sind, über **ein ausreichendes Maß an KI-Kompetenz** verfügen, wobei ihre technischen Kenntnisse, ihre Erfahrung, ihre Ausbildung und Schulung und der Kontext, in dem die KI-Systeme eingesetzt werden sollen, sowie die Personen oder Personengruppen, bei denen die KI-Systeme eingesetzt werden sollen, **zu berücksichtigen** sind.“ (KI-Verordnung, Art. 4)

https://eur-lex.europa.eu/legal-content/DE/TXT/HTML/?uri=OJ:L_202401689#art_4

Relevante Gesetze: Datenschutz (DSGVO)

Datenschutz wie andere Datenverarbeitungen:

- personenbezogene Daten:
 - Anmelde-Daten
 - Eingabe
- Transparenz
- Recht auf Löschung etc.



Bild: dreamstudio.ai; Prompt: "sculpture of justitia with scale in one hand made up of binary code protecting data"

Relevante Gesetze: Urheberrecht (UrhG)

Juristischer Streit über ChatGPT

OpenAI kritisiert »New York Times« für Urheberrechtsklage

US-Schriftsteller verklagen ChatGPT-Entwickler

SUNO IS A MUSIC AI COMPANY AIMING TO GENERATE \$120 BILLION PER YEAR. BUT IS IT TRAINED ON COPYRIGHTED RECORDINGS?

Ausnahme (§ 60d UrhG):
wissenschaftliches Text- und
Datamining

[→ Urteil LG Hamburg 28.9.2024](#)



Quelle: <https://www.heise.de/news/12-Millionen-Bilder-kopiert-Getty-klagt-auch-in-den-USA-gegen-Stability-AI-7487081.html>

KI und Moral: Fallbeispiel „Auto-Entscheidungen“

Autonomes Fahren in Deutschland

- Ethische Empfehlung der EU ([2020](#))
- Räumlich begrenzt möglich ([Gesetz, 2021](#))
- Entwicklung ethischer Algorithmen ([TUM, 2023](#))

Trolley-Dilemma:

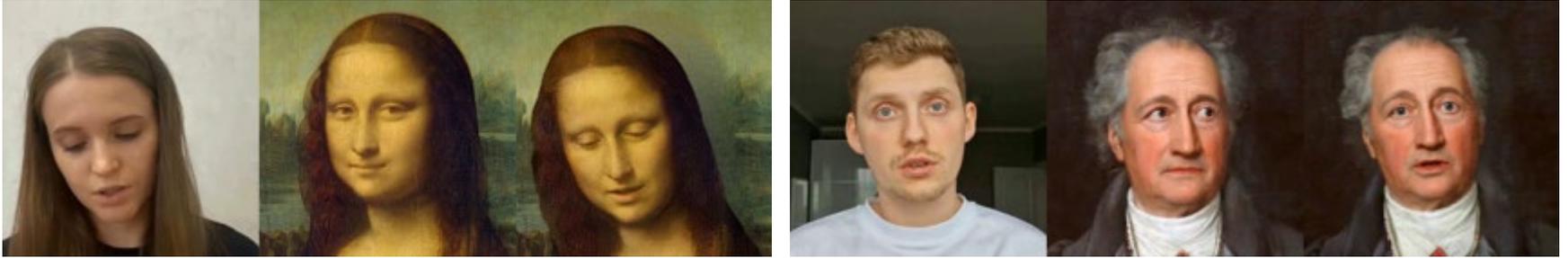
Wie würden Sie entscheiden?

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



Quelle:
https://bmdv.bund.de/SharedDocs/DE/Publikationen/DG/die-zukunft-faehrt-autonom.pdf?__blob=publicationFile

KI und Moral: Fallbeispiel „Deep Fakes“



Quelle: <https://nseek2303.github.io/MegaPortraits/>

„Mit großer Macht, kommt große Verantwortung.“ (Onkel Ben)

- Macht der Anbieter*innen → Was darf KI nicht können?
- Macht der Anwender*innen → Wofür setze ich KI ein?

Q & A: Slido



Bild: [Emily Morter](#) auf [Unsplash](#)

Schlusswort: Textverständnis von KI

„Versteht“ eine KI den Text?

Natural Language Processing

- Text wird in kleine Einheiten (Token) zerlegt
- Beziehungen zu anderen Texteinheiten durch Zahlen
- **Textgenerierung durch „Erraten“ einer menschlichen Antwort**
- Problem: z.B. Halluzinationen

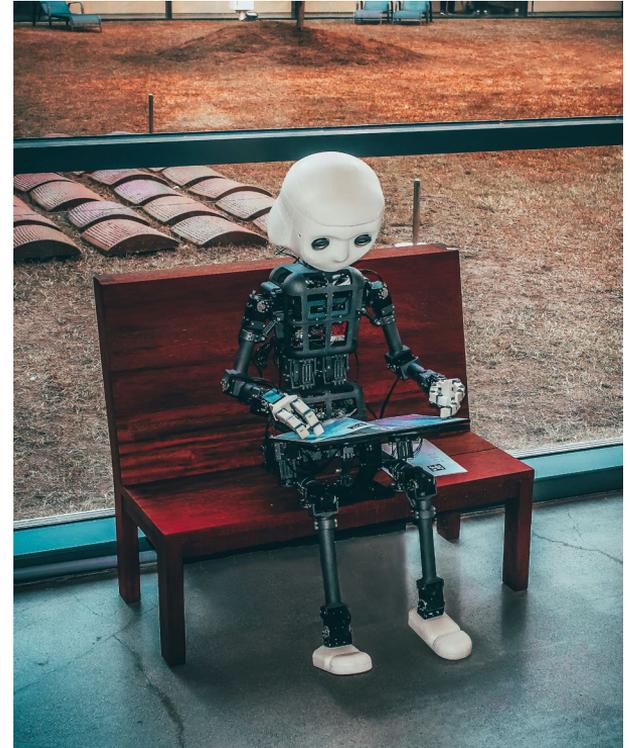


Bild: [Andrea De Santis/Unsplash](#)

Vielen Dank!

Maximilian Koch

Maximilian.koch@fu-berlin.de

Freie Universität

Lehren, Lernen und Prüfen mit
Künstlicher Intelligenz

ki@fu-berlin.de

AI Week: Organisatorisches

e-teaching@cedis.fu-berlin.de

AIQualify - Qualifizierung zu Künstlicher Intelligenz in der Hochschullehre

Koordination und Kontakt
Freie Universität Berlin
FUB-IT. Lehre, Studium, Forschung (ex-CeDiS)
E-Learning und E-Examinations (EEE)
Innestraße 24, D-14195 Berlin
e-teaching@cedis.fu-berlin.de
<https://www.cedis.fu-berlin.de/nel-ki>

VERANSTALTET VON:



IM RAHMEN EINES PROJEKTES VON:



GEFÖRDERT VON:

