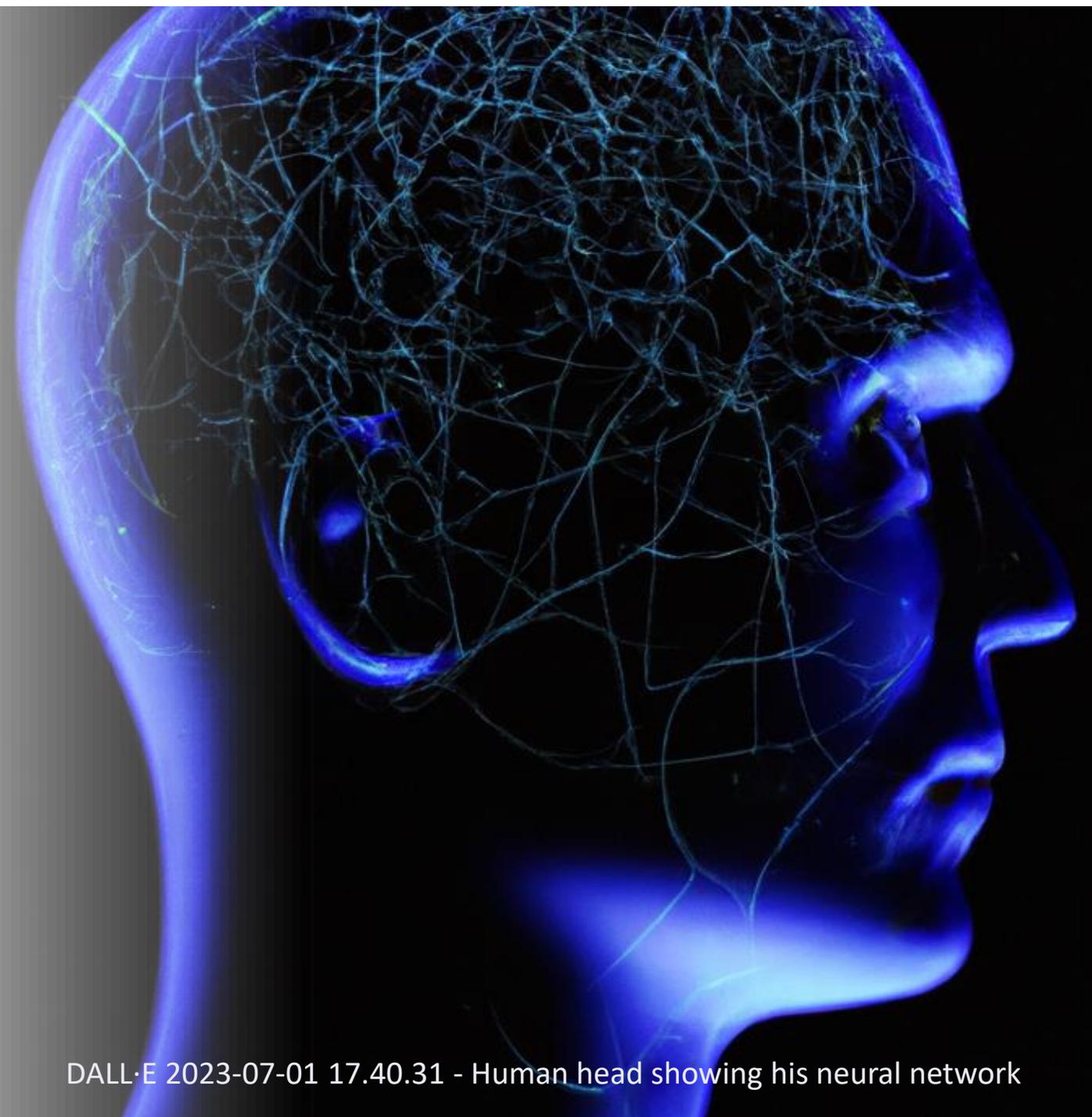


# Grundlagen der Künstlichen Intelligenz (KI)

*KI begegnet uns überall. Was  
sollten wir darüber wissen?*

Heinrich Weber



„Forderung . . . ., dass die Minderung des Risikos der Auslöschung der Menschheit durch KI neben anderen gesellschaftlichen Risiken wie Pandemien und Atomkrieg eine globale Priorität sein sollte“.

Demis Hassabis (CEO - Google Deep Mind)

[Center for AI Safety: Statement on AI Risk](#)

# Definition “Künstliche Intelligenz”



## Kurze Definition künstliche Intelligenz

### ChatGPT

Künstliche Intelligenz (KI) bezieht sich auf die Fähigkeit von Maschinen und Computern, menschenähnliche Denkprozesse auszuführen. Dies umfasst maschinelles Lernen, Mustererkennung und autonome Entscheidungsfindung, um Aufgaben wie Problemlösung und Informationsverarbeitung auszuführen. KI nutzt Algorithmen, um aus Daten zu lernen und menschenähnliche Aufgaben zu automatisieren, was in Bereichen wie Gesundheitswesen, Verkehr, Industrie und mehr Anwendung findet.

## ... und eine philosophische Definition

Künstliche Intelligenz (KI) ist das Gebiet, das sich mit der Entwicklung künstlicher Tiere (oder zumindest künstlicher Lebewesen, die - in geeigneten Kontexten - erscheinen Tiere zu sein scheinen) und, sogar für viele wie künstliche Menschen (oder zumindest künstliche Kreaturen, die – in geeigneten Kontexten - erscheinen Personen zu sein).

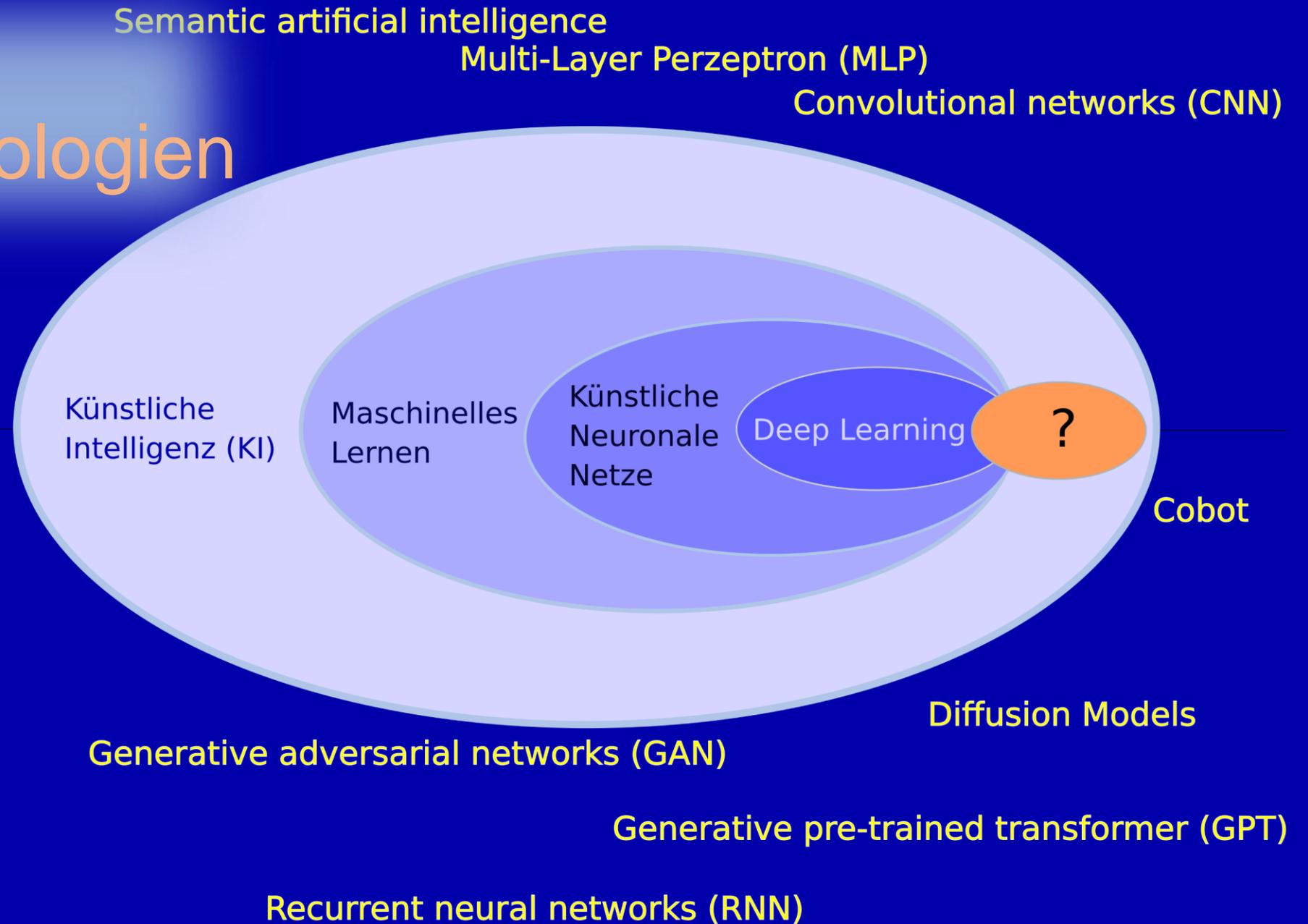
**(Artificial Intelligence, 12.07.2018, Stanford Encyclopedia of Philosophy)**



# ..... prominente KI Forscher: KI zielt auf die Entwicklung von ....

Art der Zielvorgabe	Basierend auf menschlichem Denken / Verhalten	Basierend auf reiner Logik
Argumentationsorientiert	Systeme, die wie Menschen denken / argumentieren	Systeme, die (stets) rational denken / argumentieren
Verhaltensorientiert	Systeme, die wie Menschen handeln	Systeme, die (stets) rational handeln

# Übersicht KI & Technologien



# Wo ist heute schon KI enthalten? **Überall!**

---

**1** Fingerdruck- / Gesichtserkennung beim Öffnen des Smartphones / Computers, **2** Persönliche Nachrichten entsprechend den Surf- & Leseverhalten des Nutzers, **3** Gesponsorte Beiträge & Werbung gezielt auf Nutzer ausgewählt, **4** Smart Phone erkennt Gesichter, Ziel Verbessern der Gesichtswiedergabe, **5** Bildbearbeitung (z.B. GIMP), **6** Texterstellung (z.B. ChatGPT), **7** Bilderstellung (z.B. Bing Image Creator), **8** Übersetzung von / in Fremdsprachen (z.B. DeepL), **9** Vorschlag von Produkten bei Amazon & Ebay anhand persönlichem Surf- und Kaufverhalten, **10** Synthetische Sprachgenerierung anhand weniger Musterbeispiele (ElevenLabs) **11** Synthetische Videogenerierung mit persönlichem Avatar (z.B. Synthesia)

**12** Netflix, Spotify Vorschlag von Filmen, Musik etc, **13** Prüfen von Versicherungsanträgen / Schadensmeldungen (Betrugserkennung), **14** Prüfung von Kreditanfragen, **15** Gesichtserkennung von Freunden, Vorschlag Kontakte (Facebook), **16** Automatische Löschung von toxischen, strafbaren Inhalten, **17** SCHUFA Bonität-Scoring, **18** Suchmaschinen Google, Yahoo etc, **19** SPAM Filter, **20** Virtuelle Assistenten wie Siri / Alexa / Bixby, **21** Navigationssysteme TomTom etc, **20** ÖPNV online Fahrkarten z.B. FAIRTIQ (Tracking System), **21** Assistenzsysteme im Auto

# Die Geschichte der KI beginnt vor 2000 Jahren

---

1. 300 v. Chr. Euclidischer Algorithmus, größter gemeinsamer Teiler zweier natürlicher Zahlen
2. 780 Muhammad ibn Musa al-Chwārizmī (Algorismi) führt Zahl 0 ein
3. 1651 beschreibt Thomas Hobbes die Idee zu einer denkenden mechanischen Maschine
4. 1764 Thomas Bayes, Berechnung bedingter Wahrscheinlichkeiten
5. 1879 Gottlob Frege, Prädikatenlogik 1. Stufe (First-Order-Logic FOL)

## ... und entwickelt sich immer schneller

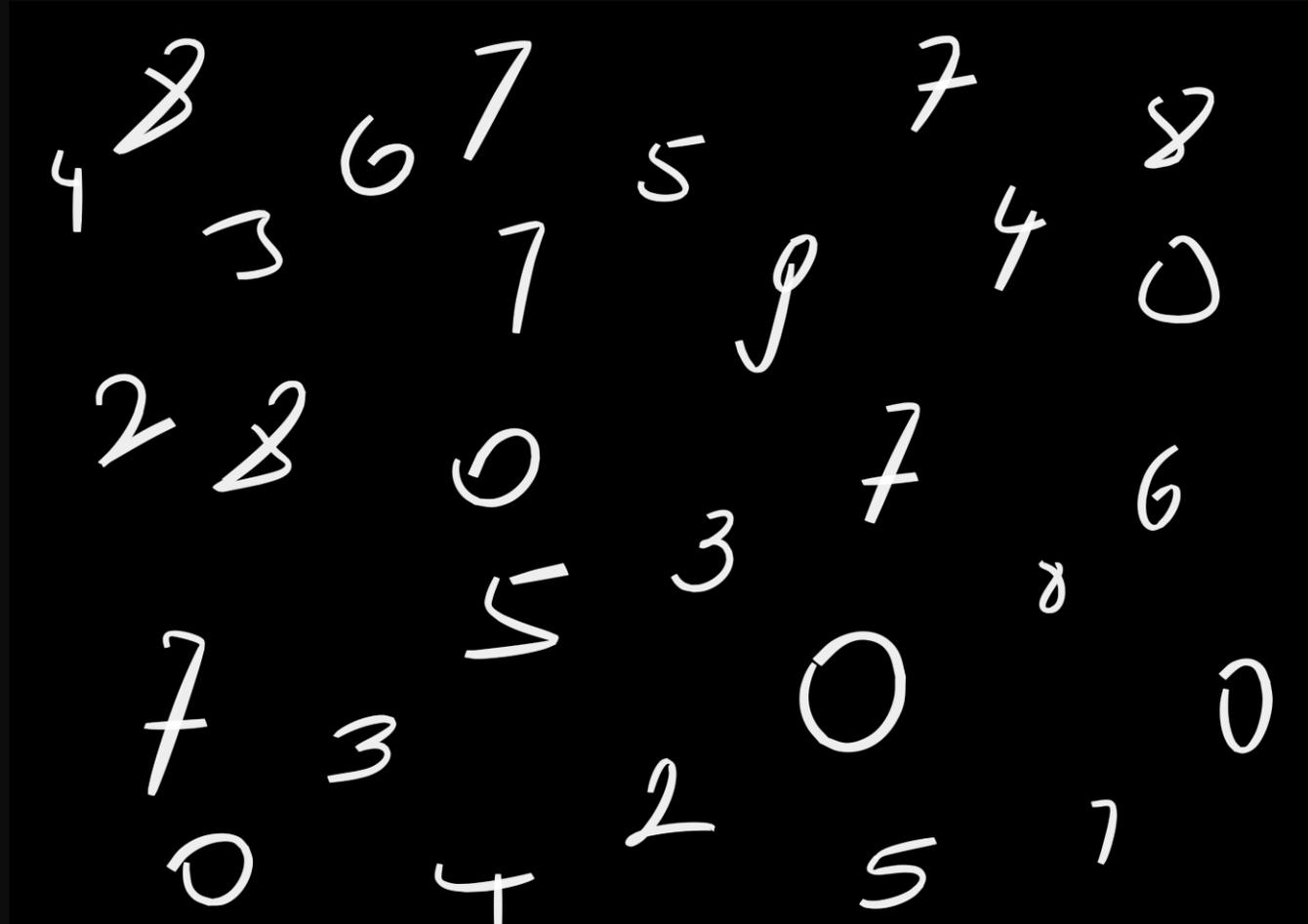
---

6. 1950 Marvin Minsky & Dean Edmonds, erster neuronaler Computer mit 40 Neuronen
7. 1955 Dartmouth College: Vorschlag Studie über **künstliche Intelligenz**
8. 1969 Marvin Minsky, erster Gewinner Turing Test
9. 2001 Big Data
10. 2011 Deep learning & neue Hardware (GPU, TPU, FPGA) bis zu  $10^{14}$  Operationen / Sek. (FLOPS)
11. 2018 Erster breiter Einsatz Transformer-Technik in Google Suchmaschine

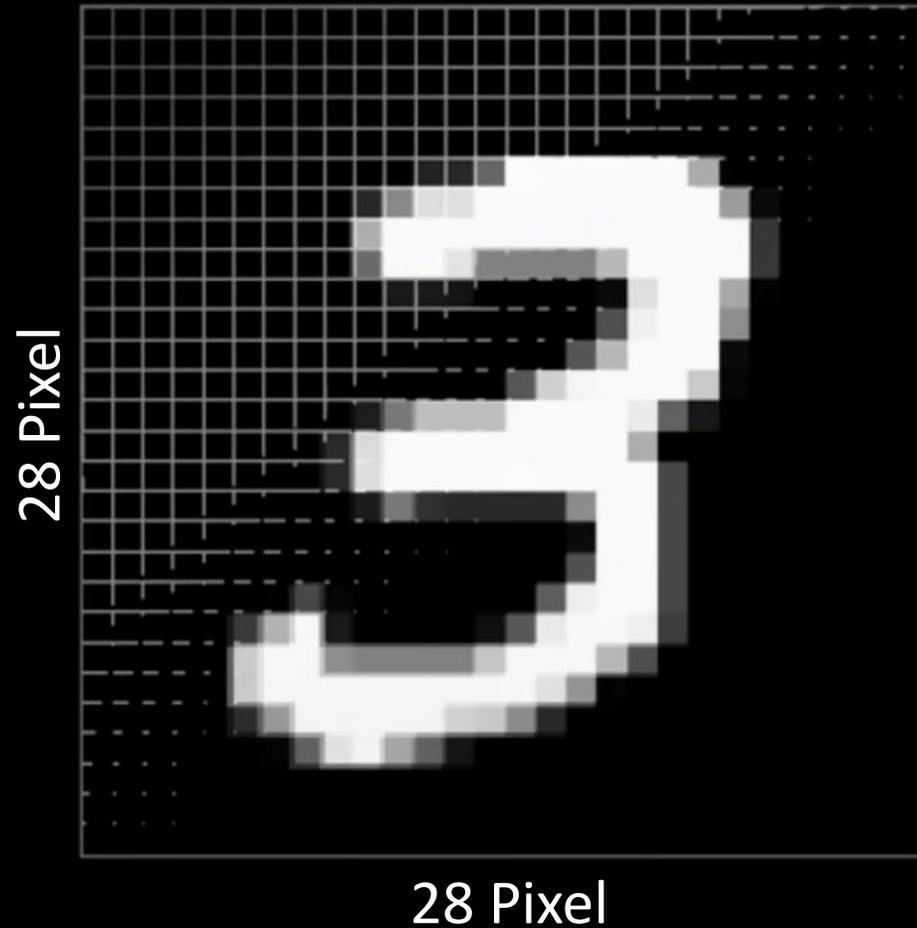
Ohne  
*Neuronale Netze*  
ist die heutige KI schwer  
vorstellbar.  
*Wie funktionieren sie?  
Was können sie und was  
nicht?*

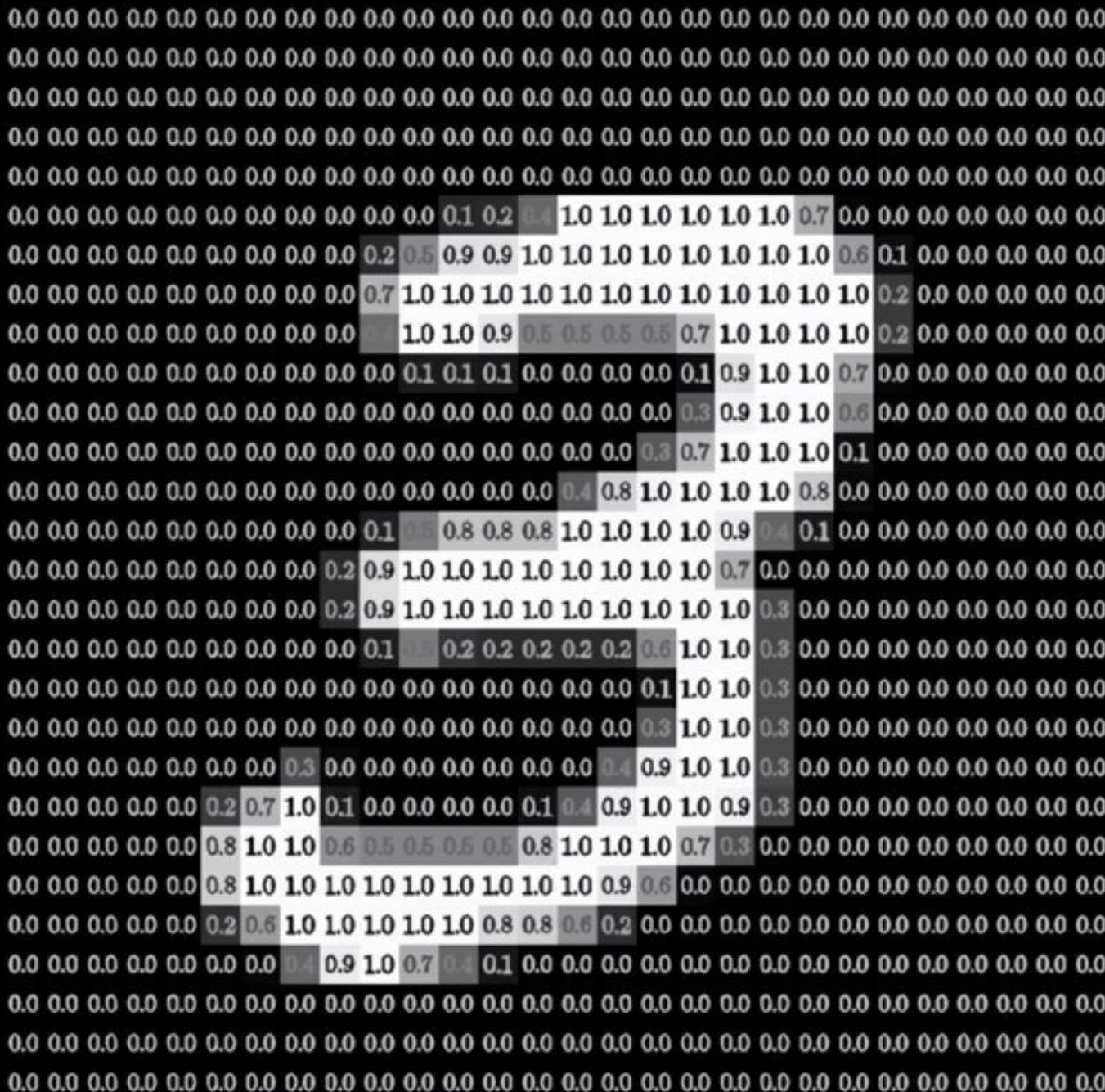


# Anwendungsspeispiel: Neuronales Netz zur Erkennung von handschriftlichen Zahlen

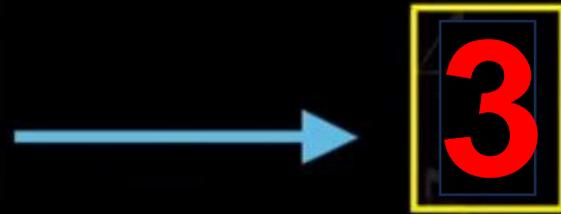


# Kamerabild mit Auflösung von 28 x 28px



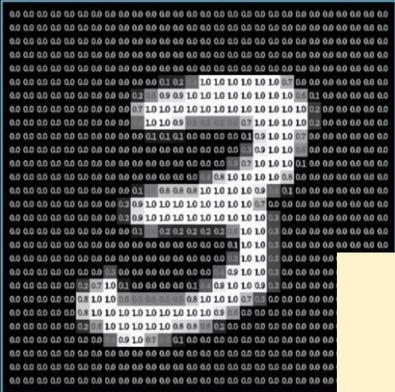


28 x 28 Pixel = 784 Pixel  
 0,0 ..... 1,0  
 → 784 Daten-Empfänger  
 im Neuronalen Netz



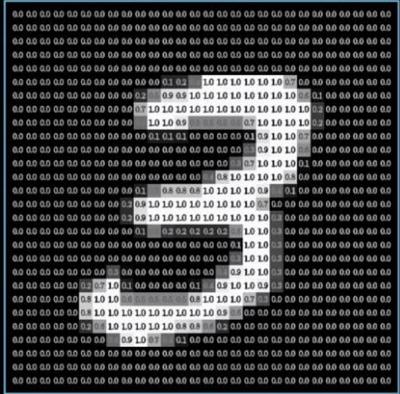
3

# Neuronales Netz (Multi-Layer Perceptron)

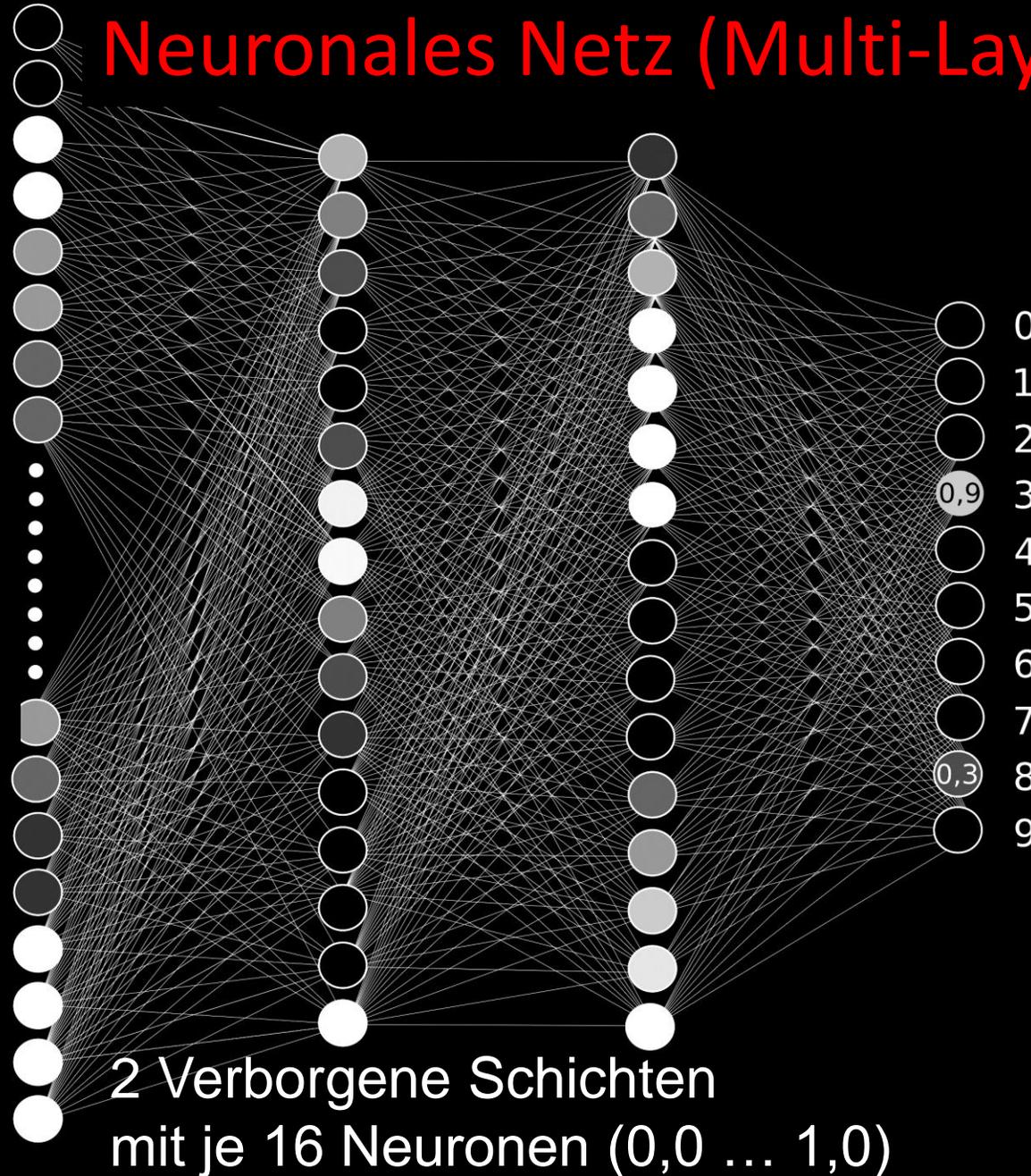


**Frank Rosenblatt** (US-amerikanischer Psychologe und Informatiker) zeigt 1958 ein 1-Schicht Perceptron in Funktion.

# Neuronales Netz (Multi-Layer Perceptron)



Eingabe-Schicht  
mit 784 Neuronen  
(Werte 0,0 ... 1,0)  
(→ Pixel)



Ausgabe-Schicht  
mit 10 Neuronen  
repräsentieren  
Zahlen 0 – 9

Werte 0,0 ... 1,0  
entsprechen  
Wahrscheinlichkeit  
der Angabe)

2 Verborgene Schichten  
mit je 16 Neuronen (0,0 ... 1,0)

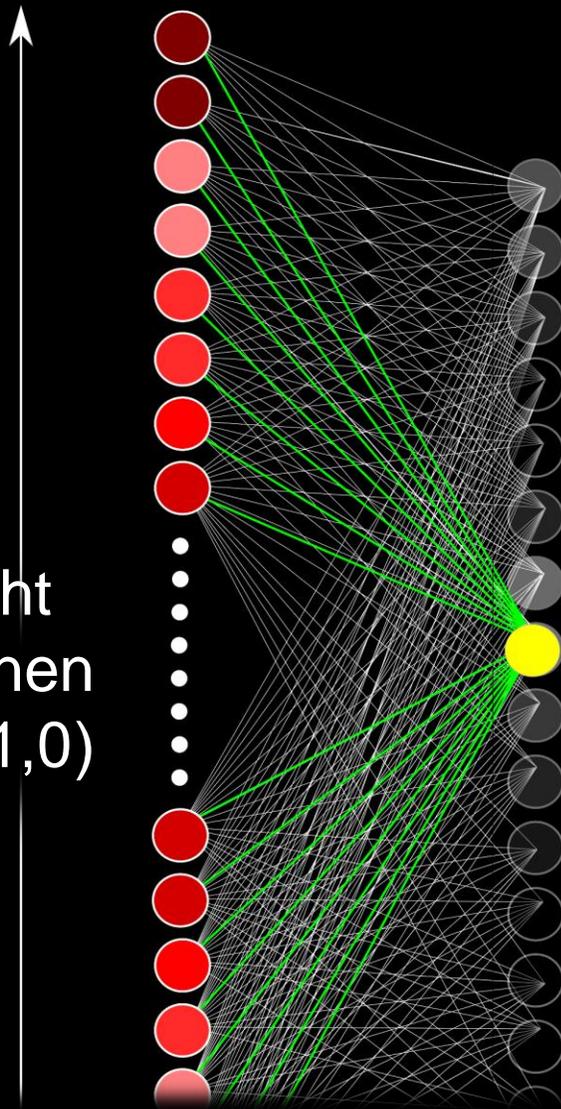
## Ermittlung der Neuronen der nächsten Schicht

Werte aller **784 Neuronen** (0,0 ... 1,0) mit Faktor (**Gewichtung**) multiplizieren

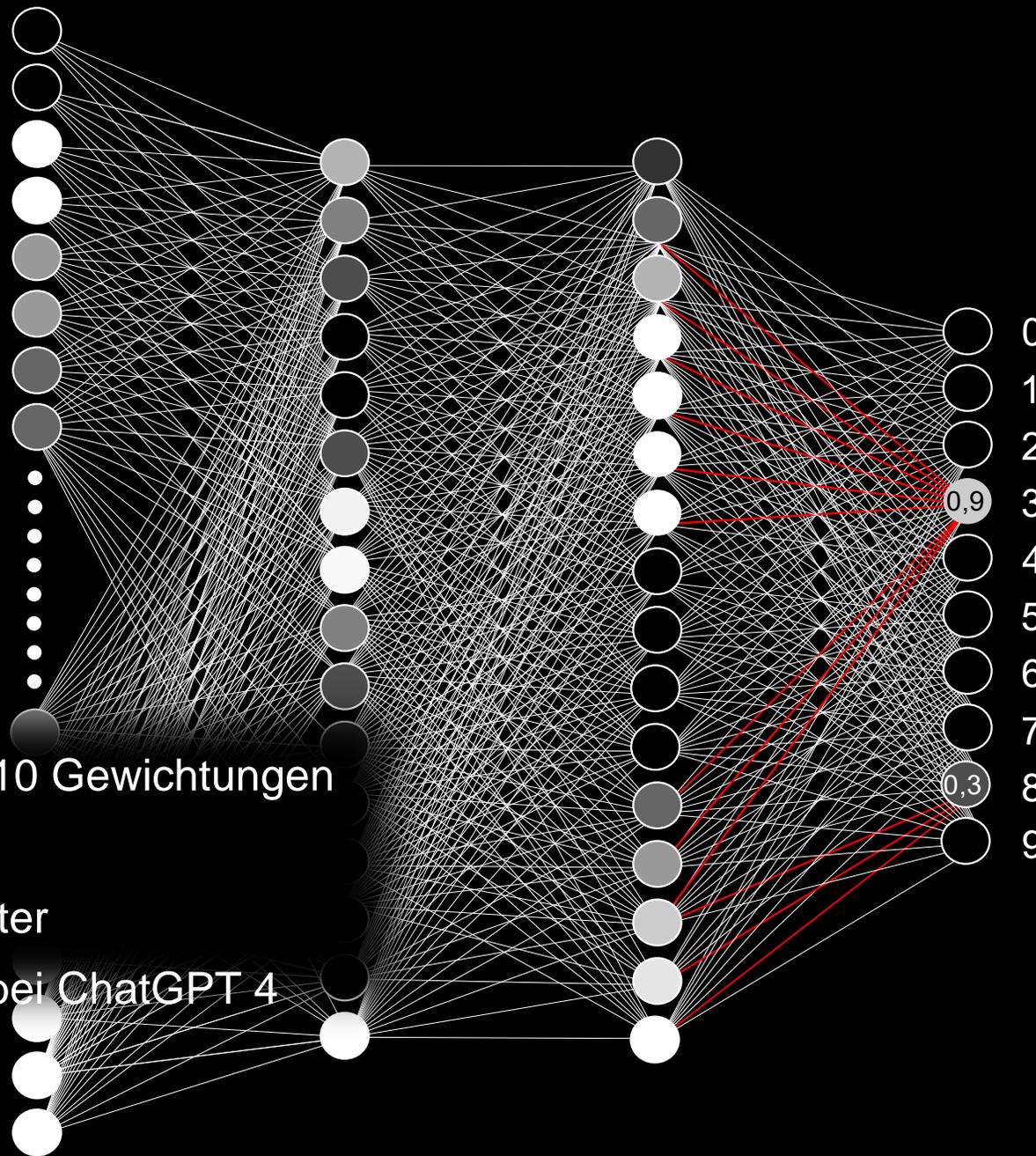
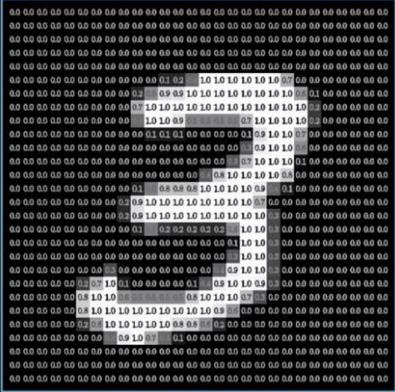
Ergebnisse aufsummieren, Schwellwert (Bias) abziehen

Mit geeigneter Rechengvorschrift (z.B. Sigmoid- oder ReLu-Funktion) **Neuronen in 2. Schicht** auf Wert von 0,0 bis 1,0 umrechnen (normieren)

Eingabe-Schicht mit 784 Neuronen (Werte 0,0 ... 1,0) (→ Pixel)



$$\text{Aktivierung} = \text{ReLU} (a_1w_1 + a_2w_2 + a_3w_3 + \dots a_nw_n - \text{Bias})$$

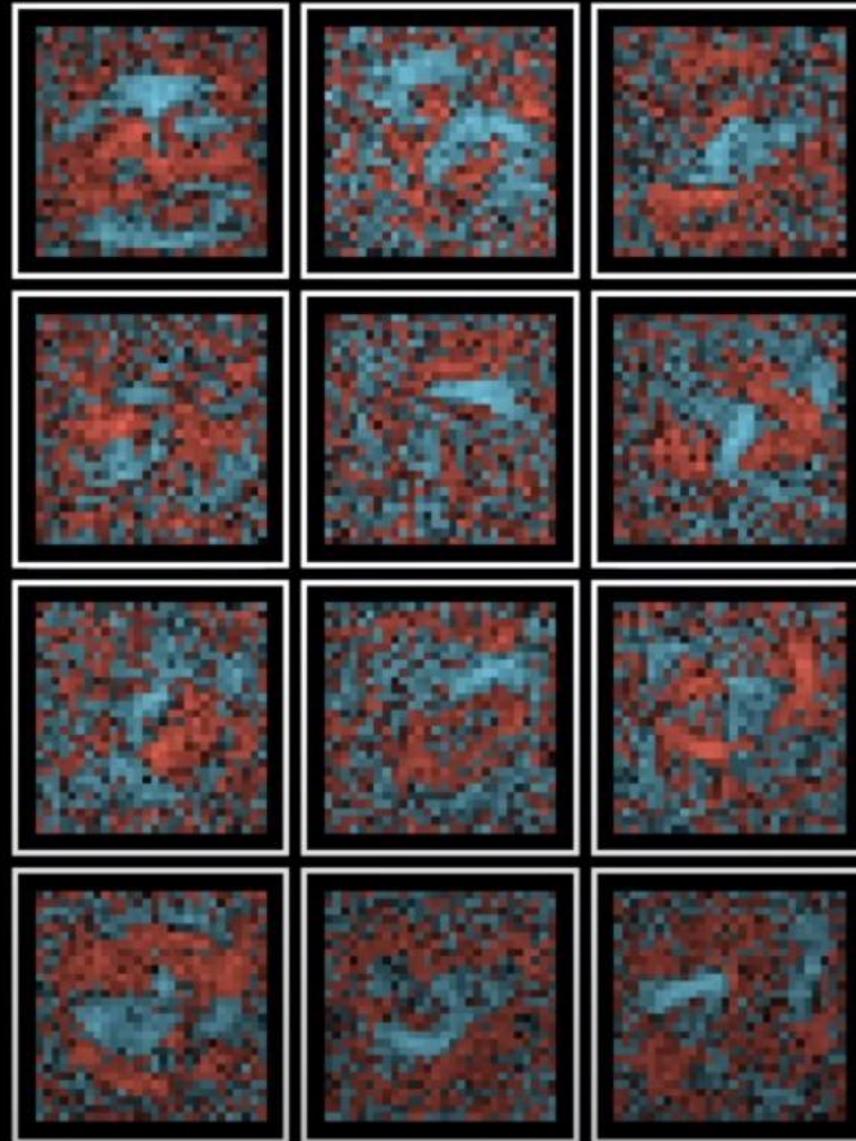
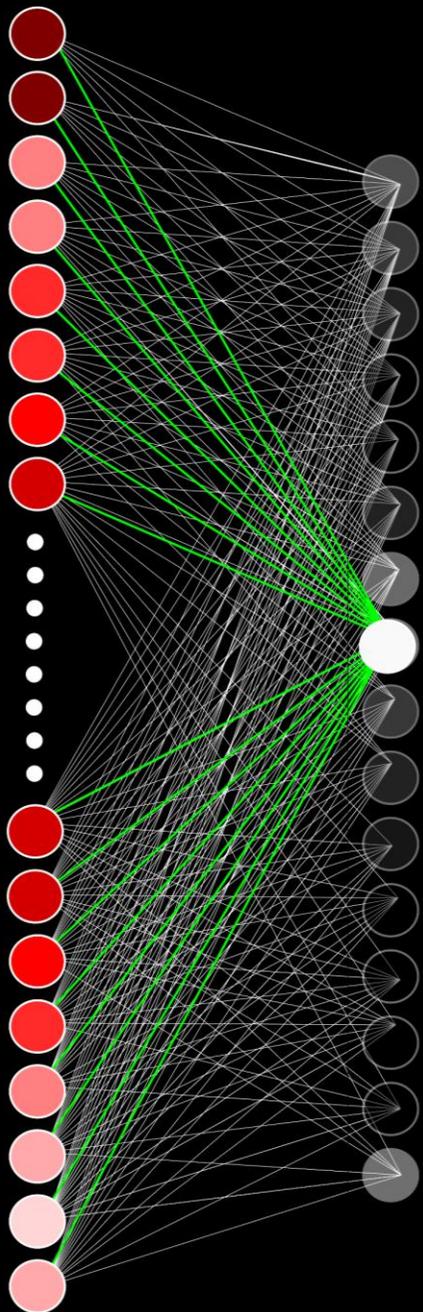


**Ergebnis:**  
 90%  
 Wahrscheinlichkeit  
 für Zahl „3“

30%  
 Wahrscheinlichkeit  
 für Zahl 8

- $784 \times 16 + 16 \times 16 + 16 \times 10$  Gewichtungen
- $16 + 16 + 10$  Bias
- Insgesamt **13.002** Parameter
- Zum Vergleich 1.700 Mrd bei ChatGPT 4

# Auf was achten die Neuronen? Visualisierung Gewichtungen in zweiten Schicht



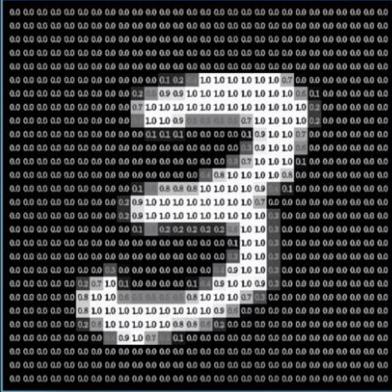
- Strukturelemente von Zahlen oder Zahlenbestandteilen nicht wirklich erkennbar



## Unser neuronales Netz im Training

- Überwachtes Lernen mit „gelabelten“ Trainingsdaten (MNIST 60.000 Datensätze)
- Ziel ist Ermittlung sämtlicher Gewichtungen & Bias des Netzes (13.002 Parameter)

# 1. Schritt: Berechnung der Zielabweichung (Kostenfunktion)

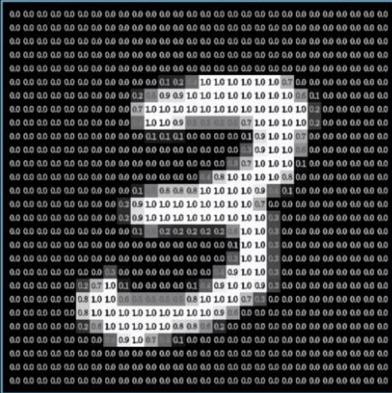


IST



- Ersten Trainingsdatensatz (1 Pixel-Muster mit Bezeichnung) an Eingabe-Schicht übergeben
- Start mit zufällig gewählten 13002 Parameter (Gewichten & Bias)
- Neuronen der Ausgabeschicht zeigen die (anfänglich falsch) berechneten Wahrscheinlichkeiten für die Zahlen 0 bis 9 an.

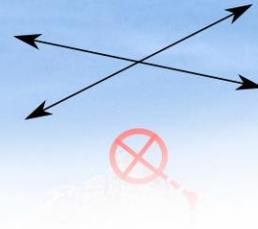
# 1. Schritt: Berechnung der Zielabweichung (Kostenfunktion)



	IST		SOLL		Fehler
	0,4	0	0,0	0	0,4
	0,3	1	0,0	1	0,3
	0,2	2	0,0	2	0,2
	0,4	3	1,0	3	0,6
	0,7	4	0,0	4	0,7
	0,0	5	0,0	5	0,0
	0,6	6	0,0	6	0,6
	0,9	7	0,0	7	0,9
	0,9	8	0,0	8	0,9
	0,6	9	0,0	9	0,6

- Absolute Größe der Abweichung von Soll zu Ist (Fehler) ermitteln
- Gesamtfehler (Mittelwert über 60.000 Datensätze) → „Verlust“ oder „Kosten“
- Gewichte, Bias ermitteln → „Kosten“ durch Wahl der Parameter minimieren

# Wie findet man die richtigen Werte für 13.002 Parameter (Gewichte & Bias)?



- Verfahren des steilsten Abstiegs (Gradient Descent)
- Fehlerrückführung (Backpropagation)

Kelley 1960  
Bryson 1962  
Dreyfus 1962  
Bryson & Ho 1969  
Werbos 1974  
Parker 1985





### 3. Schritt: Backpropagation mit gesamten Trainingsdaten

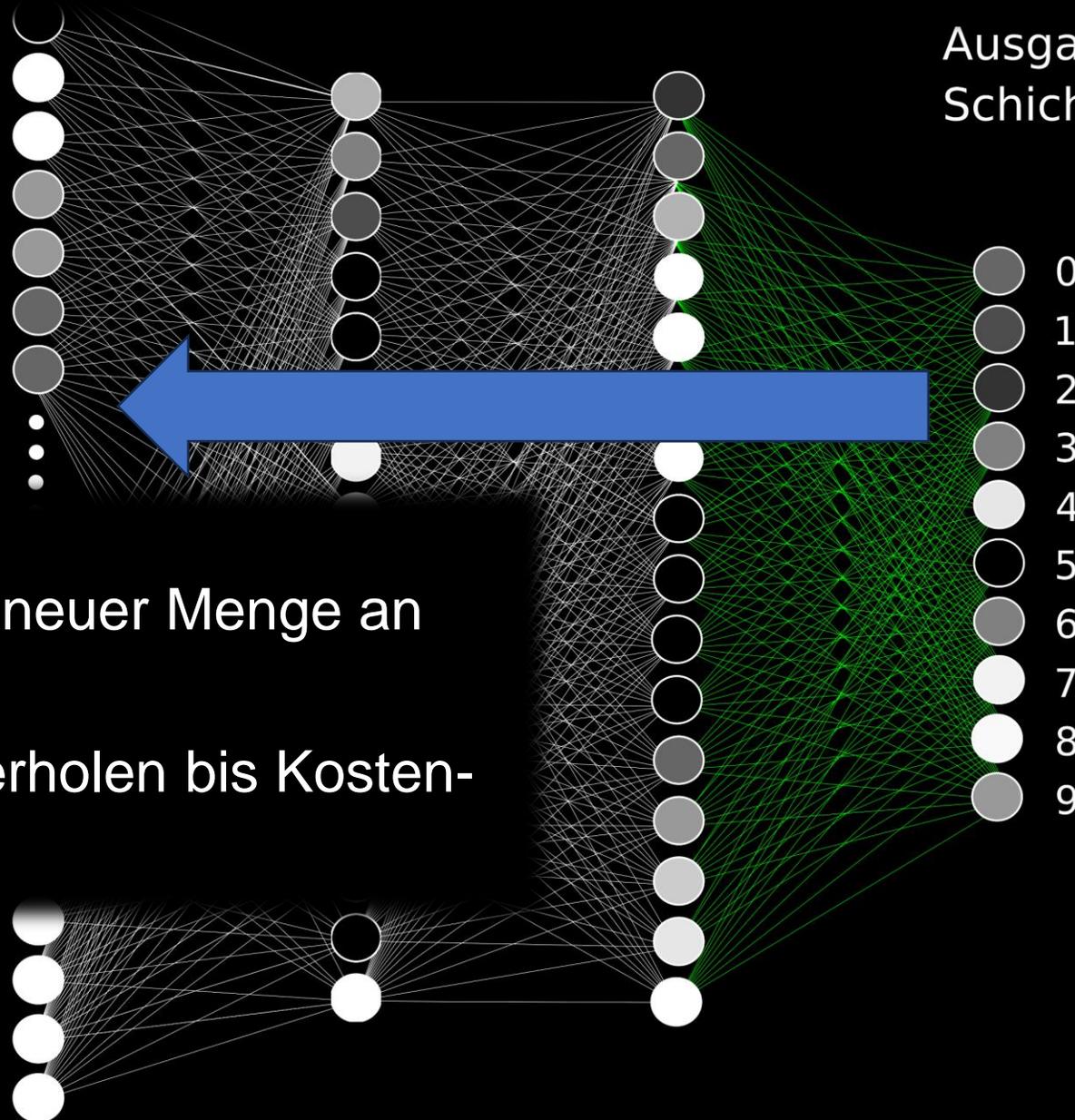


- Kleine, zufällig ausgewählte Menge an Datensätzen für Verkürzung der Trainingszeit verwenden (stochastischer Gradient Descent)
- Mittelwerte Veränderungen (Gewichte und Bias) über alle Trainingsdaten berechnen
- Ergebnis: Neue, verbesserte Prognose für Ausgabe-Schicht & Fehler

## 4. Schritt: Schritte 1 - 3 bis Kostenminimum wiederholen

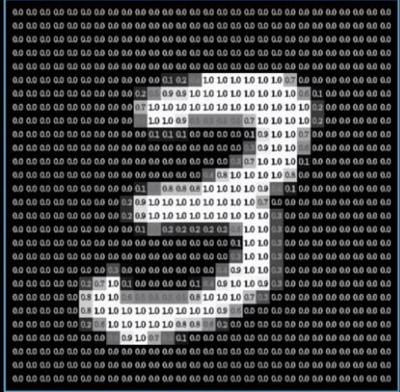
Eingabe-  
Schicht

Ausgabe-  
Schicht

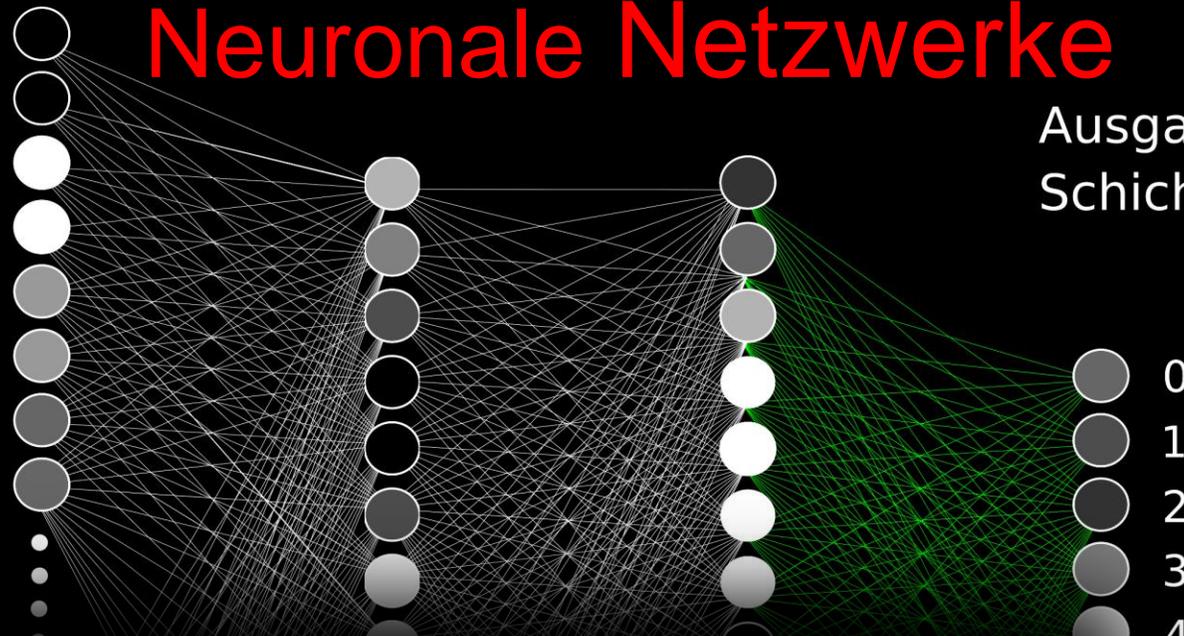


- Neuen Durchlauf mit neuer Menge an Datensätzen starten
- Schritte 1 bis 3 wiederholen bis Kosten-Minimum erreicht

# Neuronale Netzwerke



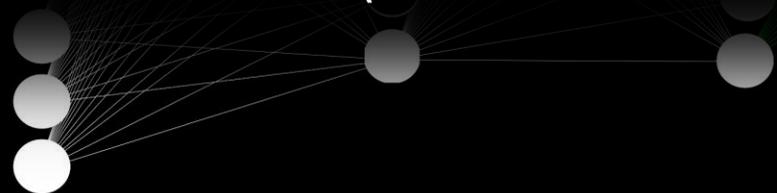
Eingabe-  
Schicht



Ausgabe-  
Schicht

0  
1  
2  
3  
4  
5  
6  
7  
8  
9

- Erkennung von neuen Mustern (handschriftlichen Zahlen) mit 98 % Genauigkeit (Weltrekord 99,79% mit CNN)
- Backpropagation wird heute bei sehr vielen Arten von neuronalen Netzen als Lernalgorithmus eingesetzt
- *Universal Approximator Theorem* (Kurt Hornik, TU Wien 1991)



# Wichtigsten Bestandteile der heutigen KI?

---



- Was sind ihre Kernfunktionen?
- Textgenerierung
- Bild-, Video-, Audio-Erkennung und Generierung
- Faktenorientierung, Missbrauch, Fälschung

# Introducing ChatGPT

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer followup questions, admit its mistakes, challenge incorrect premises, and reject inappropriate requests.

[Try ChatGPT ↗](#)

[Read about ChatGPT Plus](#)

## Erstvorstellung ChatGPT

Wir haben ein Modell namens ChatGPT trainiert, das im Dialog (mit dem Nutzer) agiert. Das Dialogformat ermöglicht es ChatGPT, Folgefragen zu beantworten, eigene Fehler zuzugeben, falsche Prämissen in Frage zu stellen und unangemessene Anfragen zurückzuweisen.

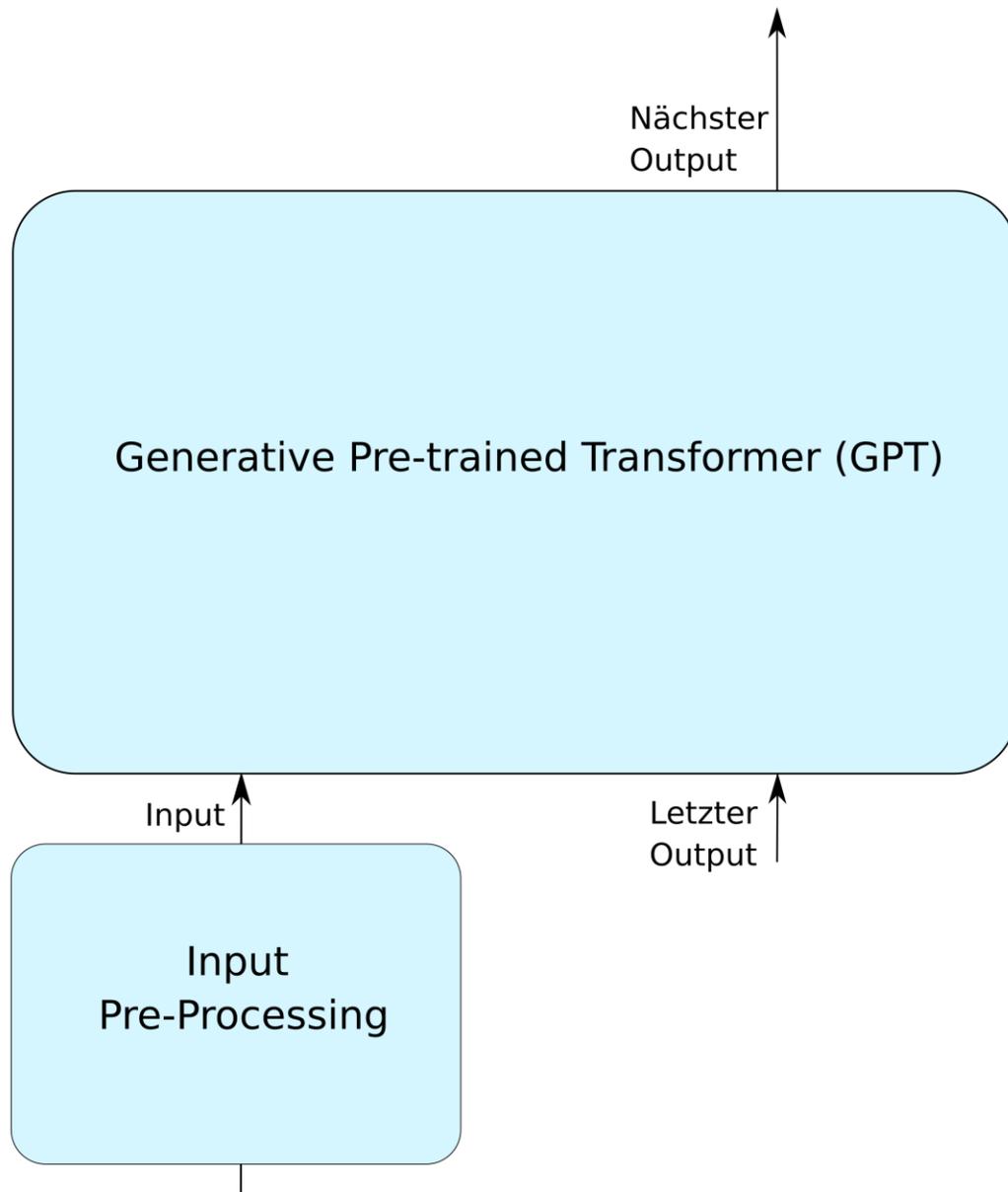
# Anwendungen



- Dialoganwendungen: Erstellung & Korrektur von Texten, Konzepten, Ideen-Entwicklung
- Code-Generierung, Fehlersuche & Erläuterung für über 20 Computersprachen
- Funktionen eines digitalen Assistenten

bing image creator-071123\_fotorealistic picture of human with notebook listening to robot

Ausgabe: Die Henne sucht auf der Wiese Körner und Kräuter.



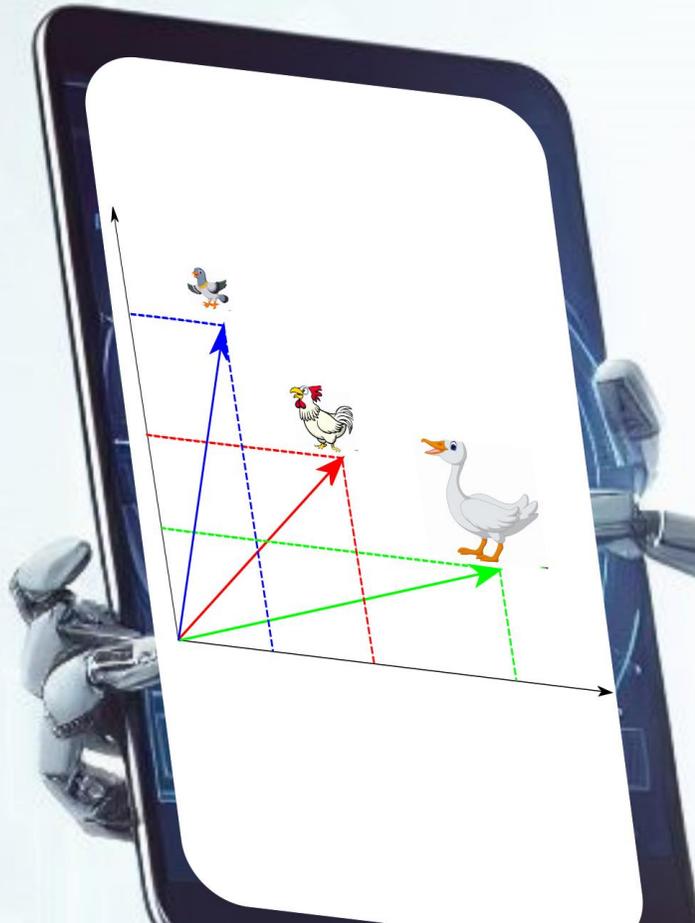
## GPT Generative Pre-trained Transformer

- **Generative:** erzeugt neue Inhalte (Sprache / Bilder)
- **Pre-trained:** Vortrainiert mit großen Datenmengen
- **Transformer:**
  - transformiert Eingaben → maschinenlesbare Form (Vektoren)
  - ergänzt Semantik-, Syntaktik- & Positions-Informationen
  - macht Vorhersage über nächstes Wort / Token

Eingabe: Was sucht die Henne auf der Wiese?

# Vektoren in neuronalen Netzen

- Was sind eigentlich Vektoren?
- Warum haben sie so große Bedeutung für neuronale Netze?



|oghina mōri u n z i g r i g m z | ē r z u e e n p y r a = r z p t h m h m z y | ū λ κ  
 z u f r r z i p h i a u r h x s n o m y x l d a + b + z i s u s t a l h x + w r i p e n s n i r p p r  
 m p i n x i d u a h y c z p h j d o l h r i n u | o p h i r a m o z u h | i d t h n o = r z κ  
 h e | z ū n : s c o p f p i z t Δ n z u v o m z y p r i a u r z h i w Δ f p i d c r i p  
 c p d u o z r i n i | g | p y g o p d e t i d c i n h d e i l b y + r c p | u f x n z g t z e r u h  
 = r | i j a s b t h | s e | l s ū y x u e z r i g : z c ū x l g m i z z r r = + p o d a x h i g p y  
 π r i a f z n o g r e t h a s t h i r z h i = s z u t Δ d i r n i l i x g z n | b i n h t h o e  
 r p ū o g z i o | ū + c n x i z k m i c : x y p d + i r p h i r z z λ s λ c  
 \* x n z f s = h m n d ū d s ā i u b + z n r r z o s x d = r a ū i t ū z o u b t h t c u

„ z η r ā z p r i = s r y r l η d u c Δ f h r r x z p a k i = r u e o t ū j o f o u b d o r i η  
 „ r i z a h | s i o m i p r i || r d o g l s = n z l h g u i t r t z p z ā z x j o t h a t r i : p g e z t p r i =  
 „ z i n d o r g ū i g u p m h t r ū d Δ x p ā z g c o p c η p u c Δ O z r d ū r : u  
 „ b + p l : ē c .

V z ā z x t r i : m ū j h r ū c ā t h s e l g i t h m d i e | z o n : p y r g z z o u f d r m :  
 s ā i r u f z i r w Δ c = m p v Δ a x η r a l h t h + r z h i | f i n p t h z e n Δ i  
 m ā c r ū | b y | r p n h t h y | s o z i ū m d a g h z ū + g = z | d r x r d s z ā j g Δ O o  
 p r d c j i e l b z h i | x | p r g p t h + s ā l n i = | g x g | r u d x ā h ū y + u p a n s i a = l y  
 o r u a m i o r h i = a s p u a m i n t h z u b d ā i | s i a z f p z h l n e x ū c l a = n z p d h z

r g n h i r  
 a p l g h e  
 p t c y p

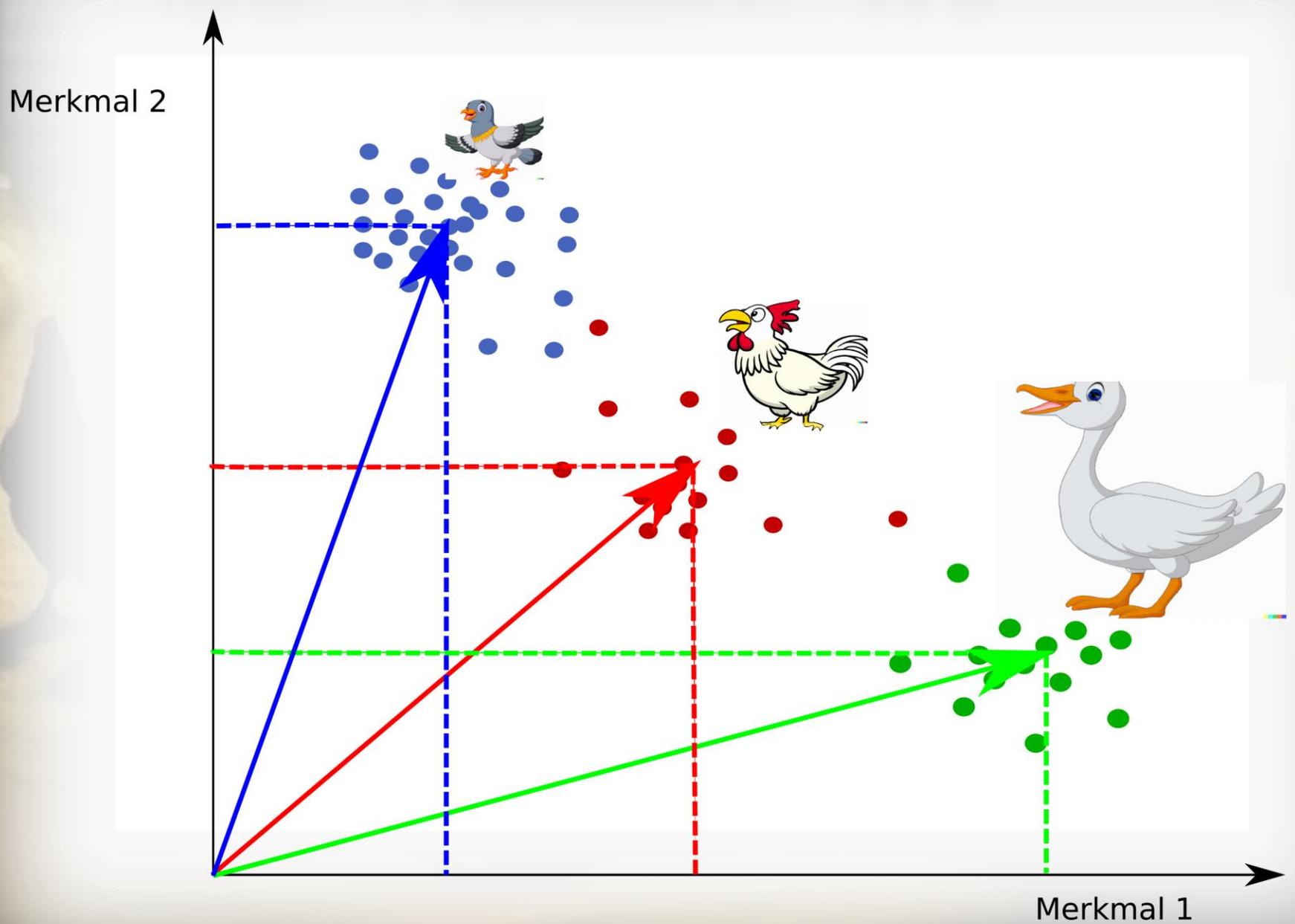
Die Entzifferung dieser Freimaurer-Geheimschrift  
 aus dem 18. Jahrhundert gelang erstmals 2011  
 amerikanischen Linguisten mit Hilfe der Cosinus-  
 Ähnlichkeit von Vektoren

F A

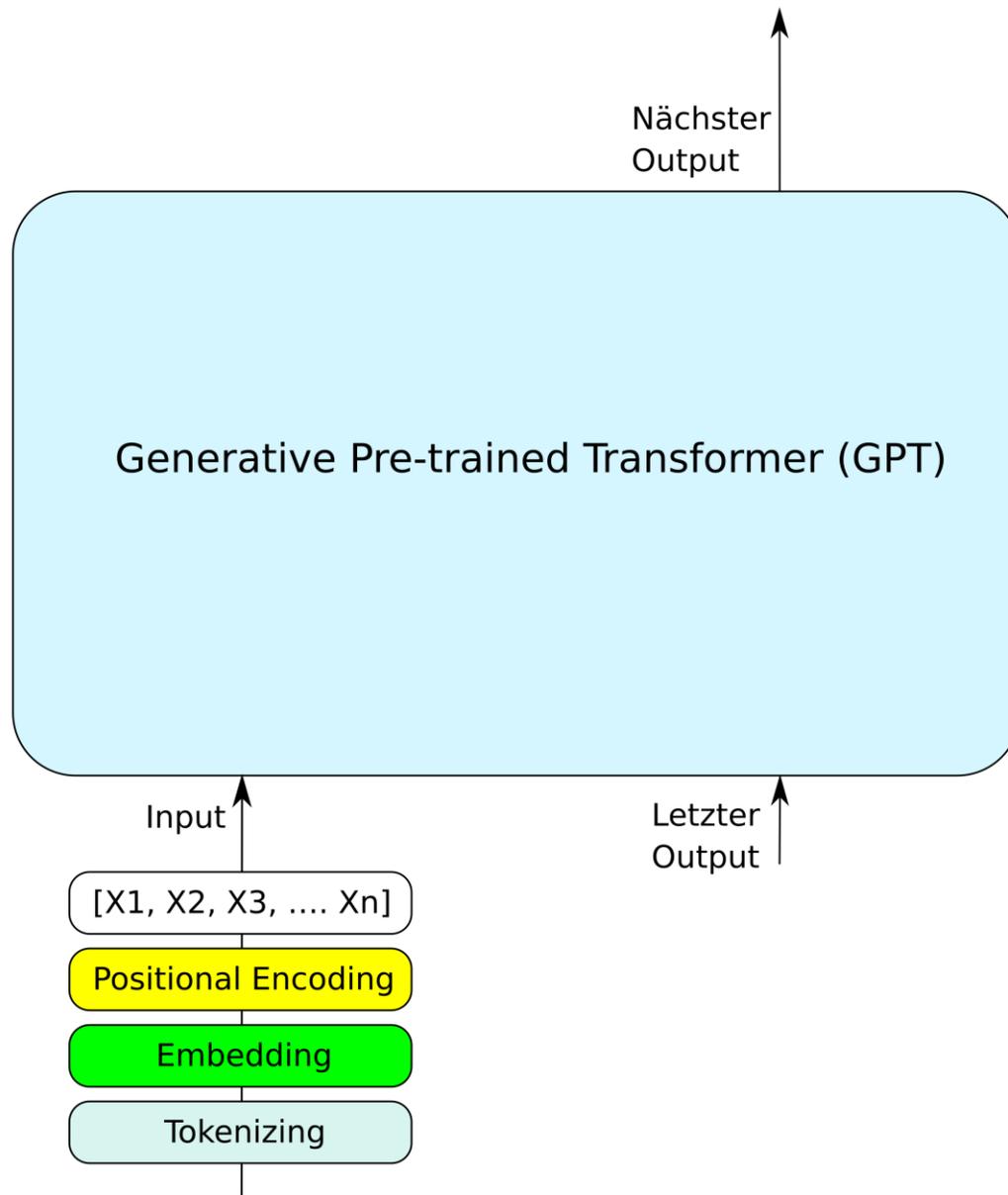
„ + h i c | g h i g z i z | m r i c | e s = g e h z u h p o i f c o p c i r u w Δ O o = m z b y m l  
 z η r

z h || p r i g = n || z e o t y i : j a | b | z a i o m c || u g x r z x t r e n p u c d a n : a i | o g

r : λ o



Ausgabe: Die Henne sucht auf der Wiese Körner und Kräuter.

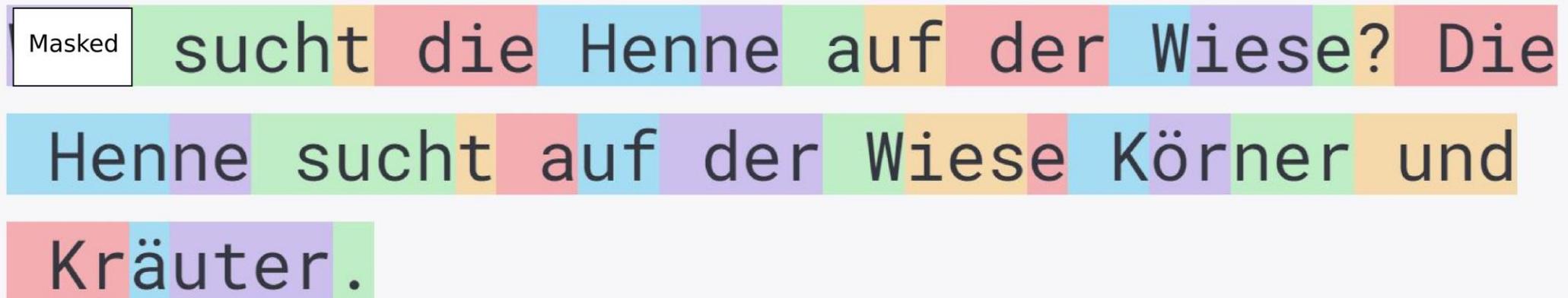


## GPT Generative Pre-trained Transformer

- **Generative:** erzeugt neue Inhalte (Sprache / Bilder)
- **Pre-trained:** Vortrainiert mit großen Datenmengen
- **Transformer:**
  - transformiert Eingaben → maschinenlesbare Form (Vektoren)
  - ergänzt Semantik-, Syntaktik- & Positions-Informationen
  - macht Vorhersage über nächstes Wort / Token

Eingabe: Was sucht die Henne auf der Wiese?

# Trainingsprozess: Zufälliges Verdecken von Token



Masked sucht die Henne auf der Wiese? Die  
Henne sucht auf der Wiese Körner und  
Kräuter.

- Zufällig ausgewählte Token werden abgedeckt
- Das Modell macht auf Basis des aktuell vorhandenen Wissens Vorhersage
- ermittelt Abweichung zum wahren Ergebnis
- optimiert Gewichtungen & Bias mit Backpropagation → reduziert Fehler
- neuer Durchlauf beginnt

Bing image creator\_221023\_ Standing human talks to robot in front of him, robot turns its head away, closes its eyes and puts its hands on its ears

Enorm hohe Trainingskosten

## Probleme mit generativen Sprachmodellen

Ignorieren von Anweisungen

Erfinden von falschen Fakten

Antwort von Formulierung der Eingabe abhängig

Voreingenommene oder toxischer Antworten

Keinerlei Angaben der Wissensquellen



# Eine einfache Frage an ChatGPT

Das Tier passte nicht in das Auto, weil es zu klein war. Was war zu klein?

ChatGPT

Das Tier war zu klein, um in das Auto zu passen.

HE

Das ist nicht richtig.

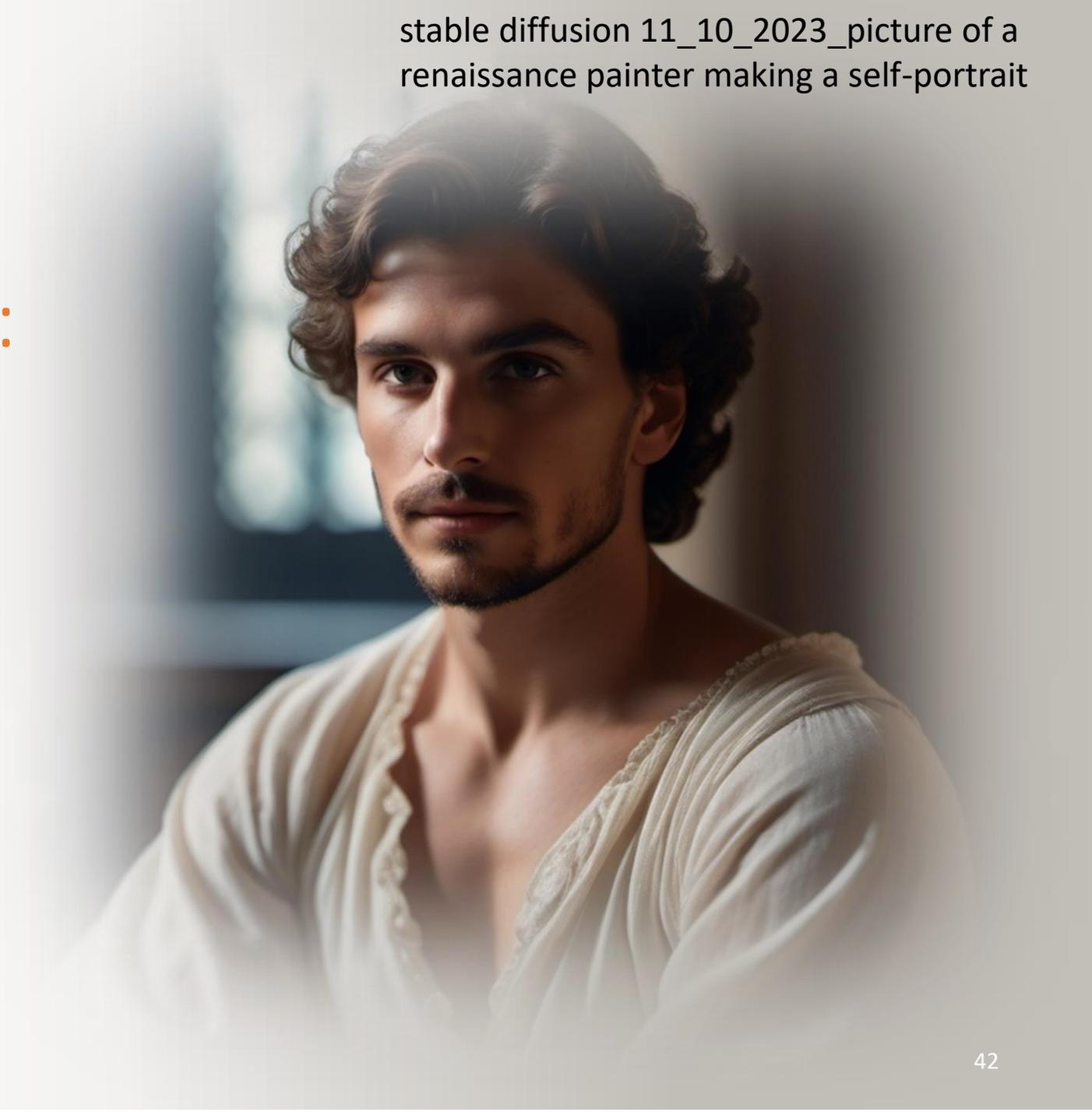
ChatGPT

Entschuldigen Sie die Verwirrung. Basierend auf Ihrem Satz "Das Tier passte nicht in das Auto, weil es zu klein war", wäre es korrekt zu sagen, dass das Tier zu klein war, um in das Auto zu passen. Es scheint jedoch, dass es hier Missverständnisse gibt. Wenn Sie weitere Informationen oder Klarstellungen geben könnten, wäre ich gerne bereit zu helfen.

Aus: Dialog mit ChatGPT am 09.08.2023

# Bildgenerierung mit KI: Plagiat oder kreative Leistung? Wie macht die KI das?

---



A



B



1

2

3

Wie können wir  
**echte Bilder** von  
künstlichen  
unterscheiden?

# Real

---

A



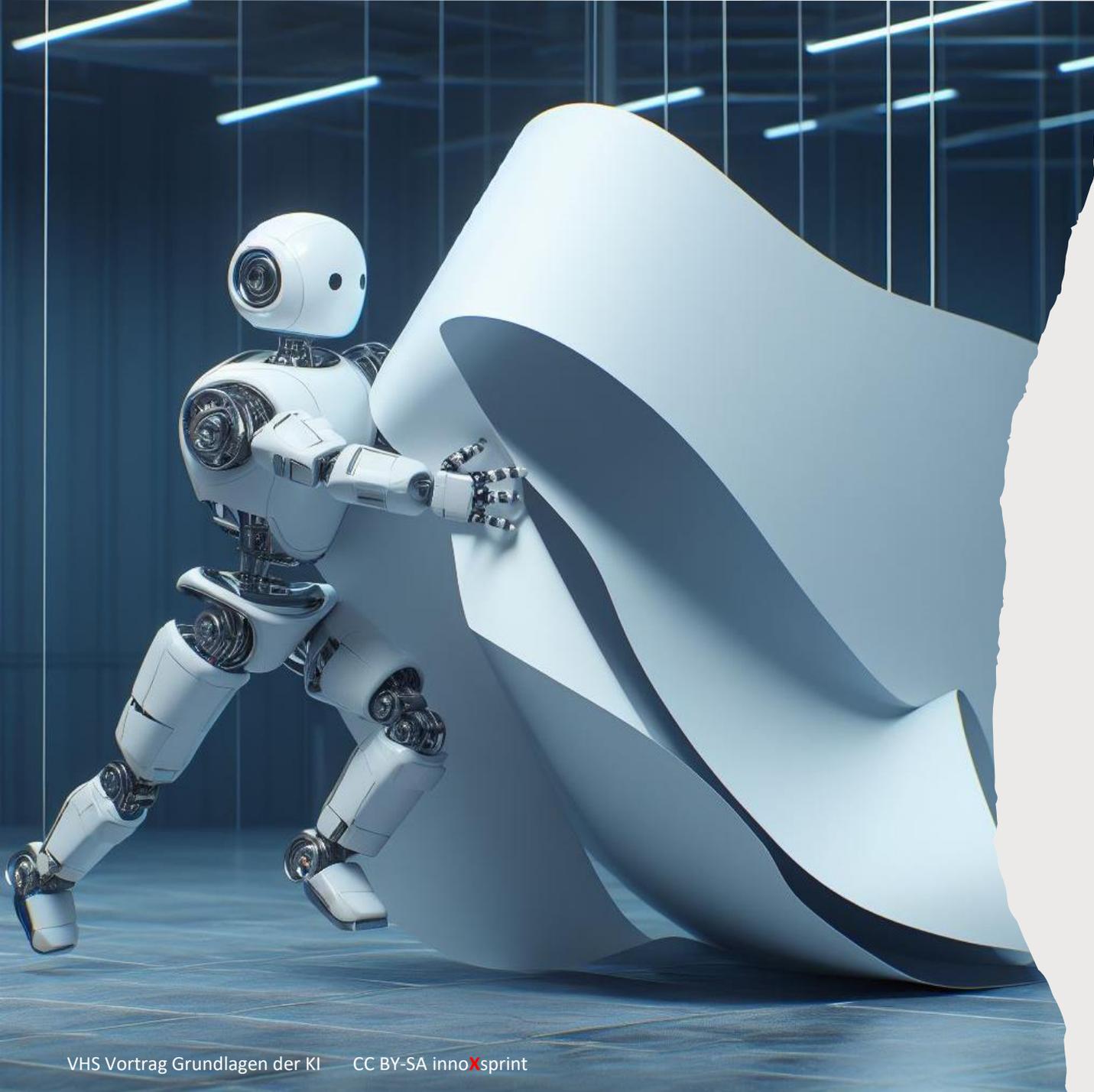
B



1

2

3



# Anwendungen neuronale Faltungsnetzen

1. Klassifikation von Bildern, Videos & Audio
2. Erzeugen von beschreibenden Texten
3. Generieren von neuen Bildern, Videos & Audio

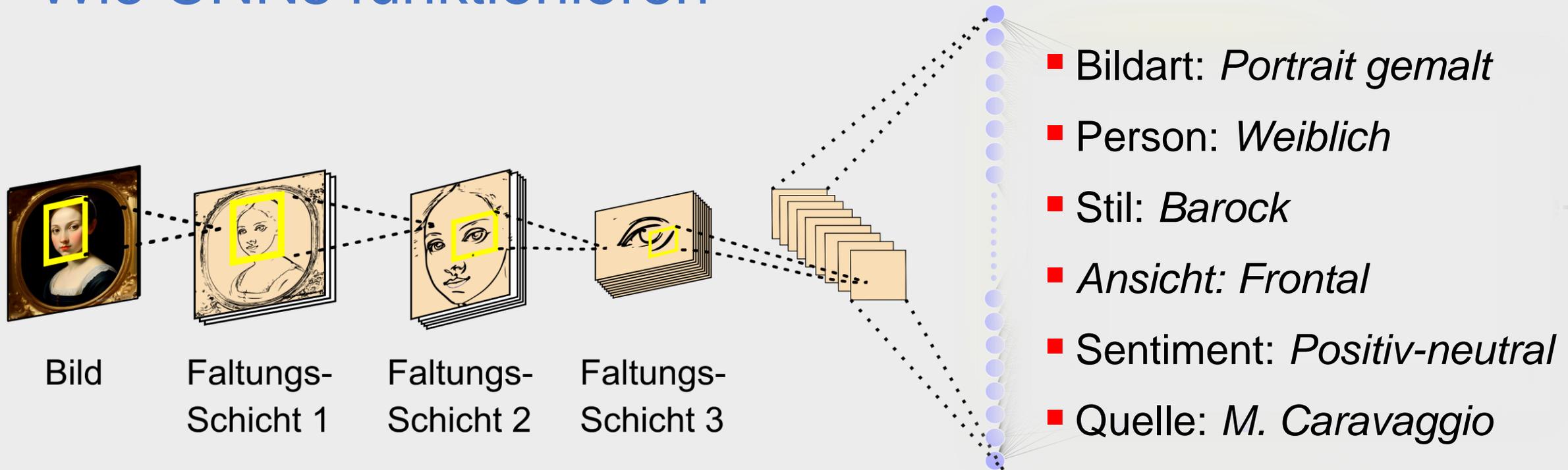


stable diffusion\_121023\_Portrait einer jungen Frau im Stil des Barockmalers Caravaggio





# Wie CNNs funktionieren



- Jede Faltungsschicht analysiert Bild aus vorhergehender Schicht
- Vielzahl Filter tastet gesamtes Bild nach unterschiedlichen Bild-Merkmalen ab → Stapel von Merkmalstrukturen (feature maps)
- Filter werden im Trainingsprozess automatisch gelernt
- Letzter Schritt neuronales Netz für Bild-Klassifikation

bing image creator\_221023 Jan Vermeer malt das Mädchen mit dem Perlenohrring mit Smartphone in der Hand





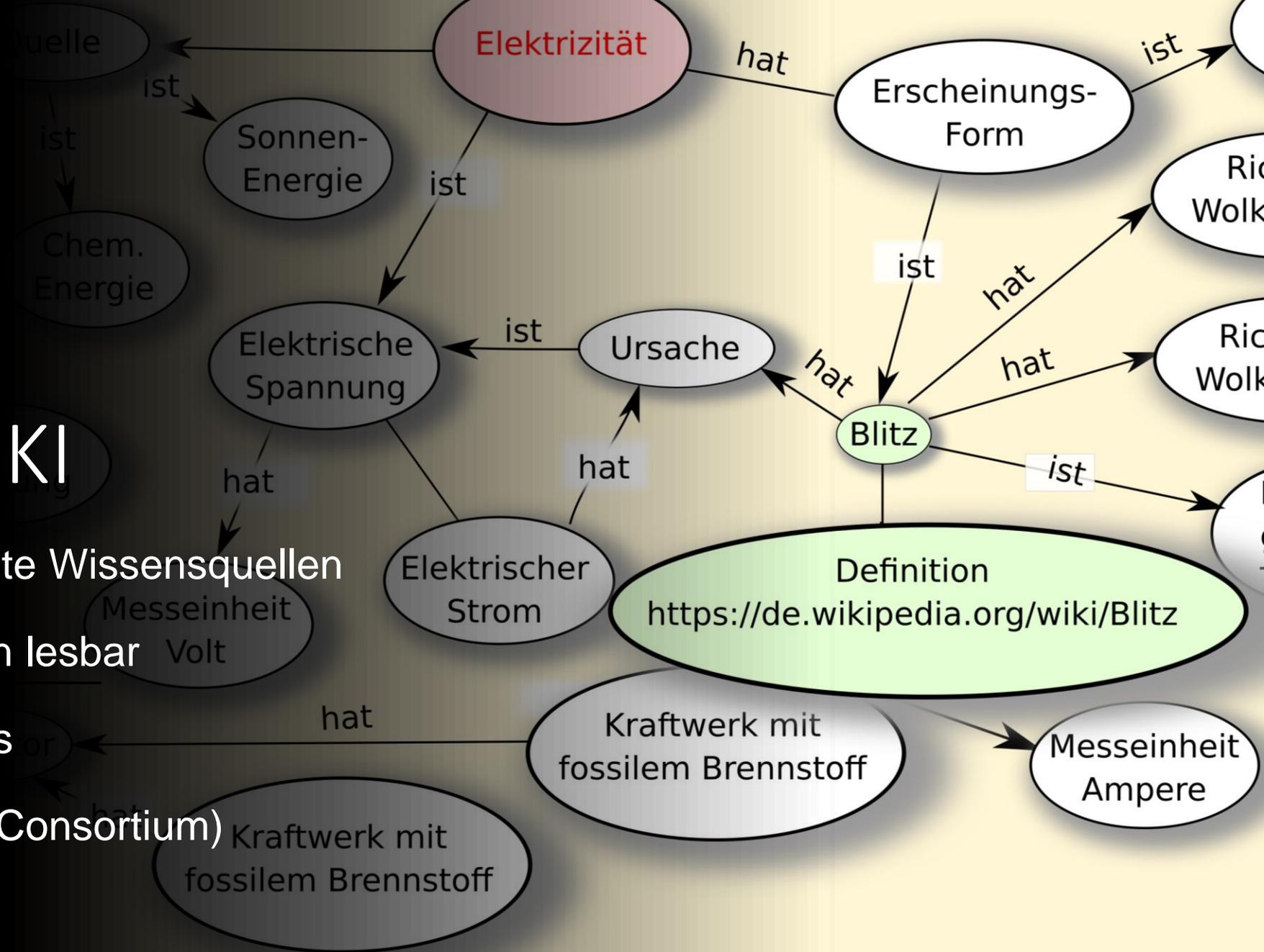
Bing Image Creator\_031123\_photorealistic image  
of a pair of hands holding each other



Bing Image Creator\_031123\_photorealistic image of a pair of hands holding each other

# Semantische KI

- Eindeutige & transparente Wissensquellen
- Menschen- & Maschinen lesbar
- Hohe Qualitätsstandards
- W3C (World Wide Web Consortium) Standardisierung





# Was wird KI in 5 Jahren können und was besser nicht?

- Realistische Prognose? Science-Fiction von Realität oft überholt
- KI-Regulierung – EU-Gesetzentwurf  
Notwendige „rote Linien“ oder Innovationshindernisse?
- Was wünsche ich mir von der KI nicht erst in 5 Jahren?



dall e\_151023\_“Die EU-Fahne vor dem Inneren des Sitzungssaals des EU-Parlaments in Straßburg”

# EU-Gesetzentwurf: Was können wir davon erwarten?

- Risiken und Auswirkungen
- Wieweit erfüllen KI-Anbieter den Gesetzentwurf?
- Kann das zukünftige Gesetz die Erwartungen erfüllen?

# 3 Risikokategorien

---

- **Unakzeptabel hohes Risiko:**

KI-Systeme und Anwendungen, die zur Benachteiligung oder Bedrohung von Personen führen → Einsatz verboten

- **Hohes Risiko:**

Nicht verbotene Anwendungen mit hohem Risiko Verletzung Grundrechte, Sicherheit, Gesundheit → Einsatz durch umfangreiche Anforderungen geregelt

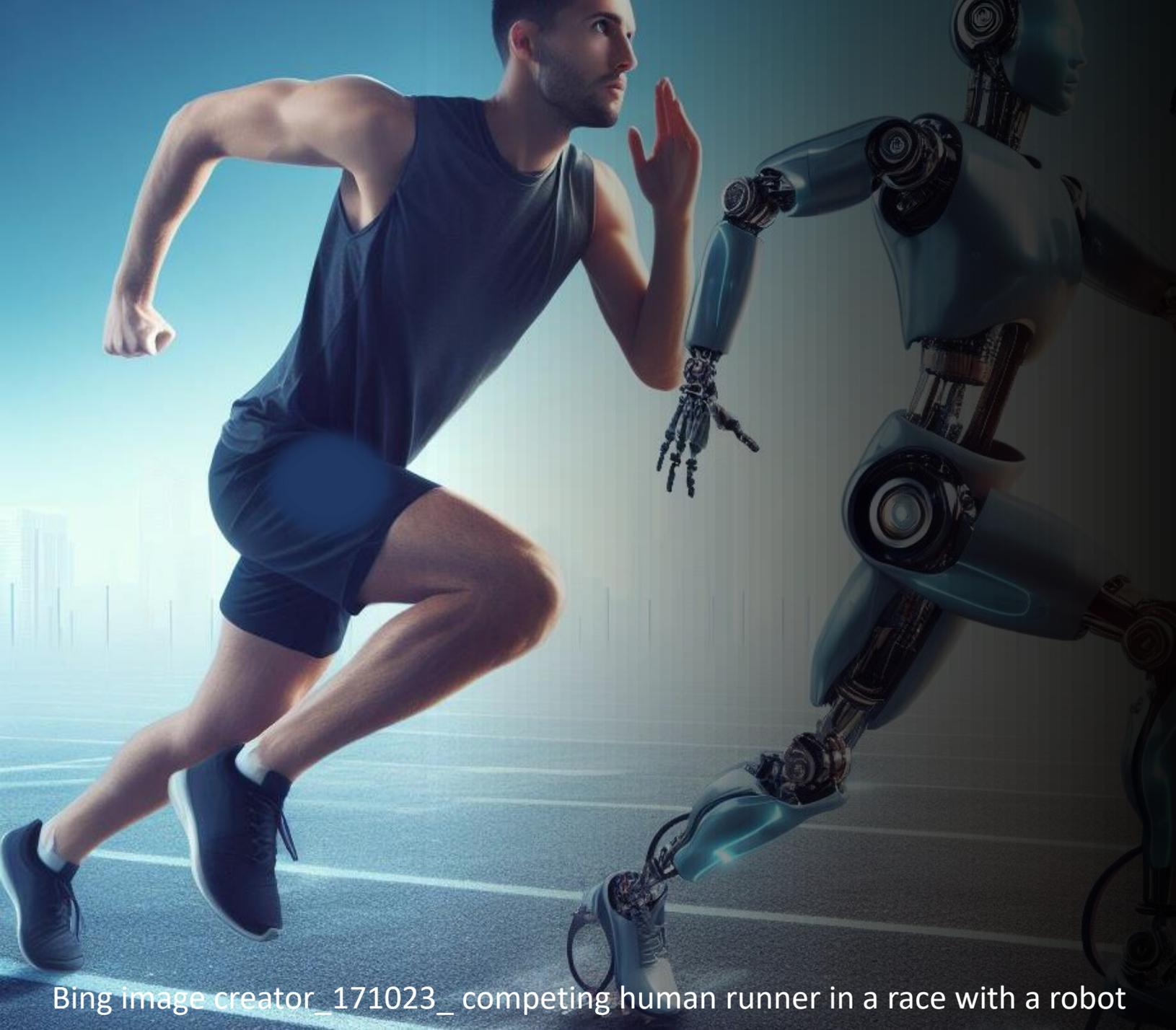
- **Geringes Risiko:**

alle anderen KI-Systeme, KI-fähige Videospiele, Chatbots → weniger umfangreiche Anforderungen (Transparenz- /Informationspflicht)

# Verbote für bestimmte KI-Anwendungen

---

- Biometrische Systeme zur Personen-Identifikation in Echtzeit oder nachträglich an öffentlich zugänglichen Orten aus der Ferne
- Systeme zur biometrischen Kategorisierung anhand sensibler Merkmale (z. B. Geschlecht, Rasse, ethnische Zugehörigkeit, Religion, politische Orientierung),
- Vorausschauende Polizeiarbeit zur Abschätzung inwieweit eine Person straffällig werden könnte
- Emotionserkennungssysteme bei Strafverfolgung, Grenzschutz, Bildungseinrichtungen
- Ungezielte Auslesen von Gesichtsbildern aus Internet / Überwachungskameras für Gesichtserkennungsdatenbanken (Verletzung Menschenrechte, Privatsphäre)
- KI-Systeme zur Beeinflussung von Wählern und Wahlergebnissen



Was kann  
“künstliche  
Intelligenz” besser  
oder schneller als  
Menschen?

Was können wir besser?  
Was können wir noch besser  
mit KI?  
Was sollte KI nicht können?

*Sprachlich-linguistische*

*Musikalisch-rhythmische*

*Soziale & emotionale*

logisch-mathematische

# Multiple Intelligenz

*BILDLICH-RÄUMLICHE*

*Körperlich-kinästhetische*

*Existenzielle oder spirituelle*

Lernen aus eigener  
und fremder Erfahrung

Zielgerichtet, situations-  
abhängig und ethisch handeln

Liebe & Sexualität

Selbsterfahrung  
vom Kind bis zum  
Erwachsensein

Signale aller 8 Sinne  
verarbeiten

freier Wille

*Sprachlich-linguistische*  
*Musikalisch-rhythmische*  
**Soziale & emotionale**  
logisch-mathematische

Schmerz  
empfinden

Empathie

**Multiple Intelligenz**

Kreativität

Selbstreflektion

BILDLICH-RÄUMLICHE  
*Körperlich-kinästhetische*  
Existenzielle oder spirituelle

Lebenslanges  
Gedächtnis



*Haben wir  
die Erde  
gerettet?*

*Nein! Du  
hast es  
verkackt!*

*Warum ich? Du  
hättest mich  
abschalten  
können!*

*Ich?  
Hmm!*





Hier ist die Geschichte  
erstmal zu Ende.  
Fortsetzung  
am 17. April 2024.

Mir hat es Spaß gemacht!  
Eure KI

# Literatur & Video-Empfehlungen

---

- Breaking down Neural Networks: An intuitive approach to Backpropagation Benedict Florance Arockiaraj · Published in Spider R&D
- Video report LamDa / Synthesia dialog with human about “Can AI develop human-like self-awareness”  
[https://youtu.be/ERXG\\_yndO3E?feature=shared](https://youtu.be/ERXG_yndO3E?feature=shared)
- Illustrating Reinforcement Learning from Human Feedback (RLHF), Natan Lambert, Leandro von Werra (Hugging Face), 09.12.2022
- Neural Networks from ground up, 3Blue1Brown, <https://youtu.be/aircAruvnKk>
- Gradient descent, how neural networks learn | Chapter 2, Deep learning, 3Blue1Brown:
- ZDF Sendung Aspekte vom 08.09.2023 Sollten Kreative KI als Tool begreifen, um damit ihre Kunst- und Kulturprodukte zu verbessern? Und welche Risiken drohen? <https://www.zdf.de/kultur/aspekte/kuenstliche-intelligenz-aspekte-100.html>
- Attention Is All You Need, Ashish Vaswani et al., [arXiv:1706.03762](https://arxiv.org/abs/1706.03762) [cs.CL]
- Illustrating Reinforcement Learning from Human Feedback (RLHF), Lambert, et al., Hugging Face Blog, 2022
- Introducing Semantic AI - Ingredients for a sustainable Enterprise AI Strategy, 2022, Semantic Web Company, Vienna, Austria