



UNIVERSITÄT  
HEIDELBERG  
ZUKUNFT  
SEIT 1386



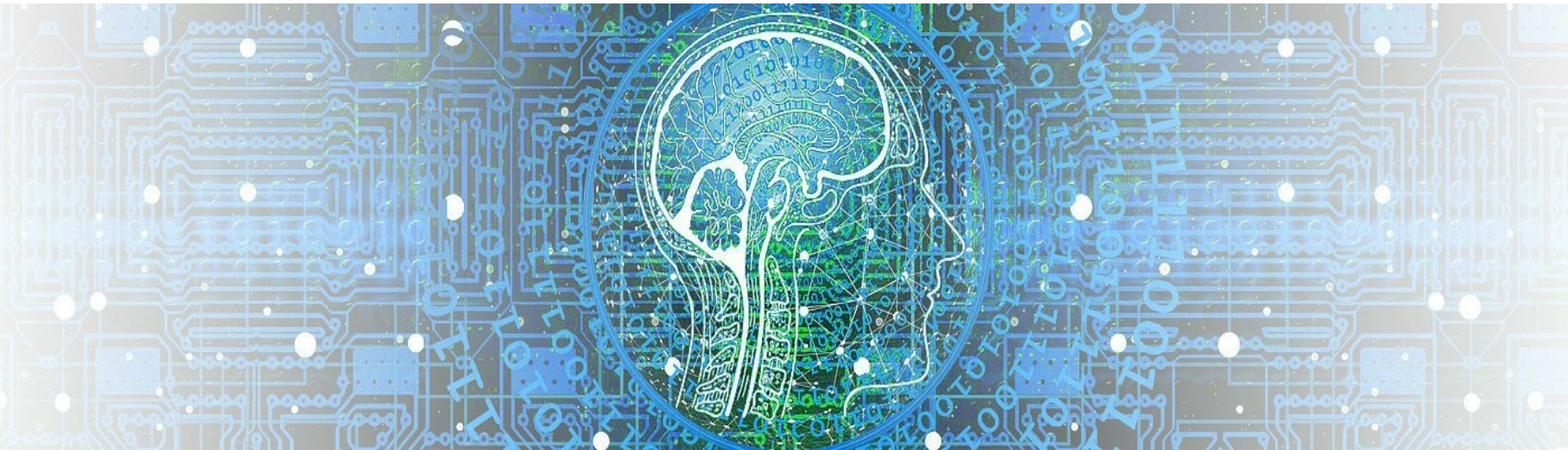
Heidelberg Institute for  
Theoretical Studies



# Prompts vs. wissenschaftliches Schreiben – Chancen und Herausforderungen der KI

Prof. Dr. Vincent Heuveline

27. Februar 2025



# Medien erzeugen: Text-zu-Bild-Generierung mittels KI



## Imagen: Das KI-Modell von Google



A brown bird and a blue bear.



One cat and two dogs sitting on the grass.



A sign that says 'NeurIPS'.



A small blue book sitting on a large red book.



A blue coloured pizza.



A wine glass on top of a dog.



A pear cut into seven pieces arranged in a ring.



A photo of a confused grizzly bear in calculus class.



A small vessel propelled on water by oars, sails, or an engine.

Quelle: "Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding", Saharia et al. (Neurips 2022), <https://arxiv.org/abs/2205.11487>



# Medien verändern: Originale Bilder editieren mittels KI



**Input Image** → **Edited Images**

	"A sitting dog"	"A jumping dog"	"A dog lying down"	"A dog playing with a toy"	"A jumping dog holding a frisbee"
	"A person giving the thumbs up"	"A person with crossed arms"	"A person in a greeting pose to Namaste hands"	"A person holding a cup"	"A person making a heart sign"
	"A cat wearing a hat"	"A cat wearing an apron"	"A cat wearing a necklace"	"A cat wearing a jean jacket"	"A drawing of a cat"

Quelle: "Imagic: Text-Based Real Image Editing with Diffusion Models", Kawar et al. (2023), <https://arxiv.org/abs/2210.09276>

## Gefahr der Erstellung von gefälschten und schädlichen Bilder / Videos

# Generative KI: ChatGPT als Wissenschaftler/in



BU

Write a research nature paper about a newly developed method for modelling myocarditis. Use only a 100 words.



**Wissenschaftliche  
Abhandlung**



We present a novel method for modeling myocarditis using human induced pluripotent stem cell-derived cardiomyocytes (hiPSC-CMs) and a co-culture system with immune cells. This model allows for the study of immune-mediated cardiac injury and potential therapeutic interventions. We demonstrate that hiPSC-CMs can be successfully infected with Coxsackievirus B3 (CVB3), leading to a pro-inflammatory response and induction of myocarditis-like pathology. Furthermore, we show that treatment with immunosuppressive agents can attenuate the inflammatory response and preserve cardiac function. Our method provides a valuable tool for investigating the pathogenesis of myocarditis and evaluating potential therapies.


# Generative KI: ChatGPT als Wissenschaftler/in



**BU** Write a critical report as part of a review of this abstract as an official reviewer of the Nature Journal. Use only a 100 words.

**Bewertung der  
Abhandlung**

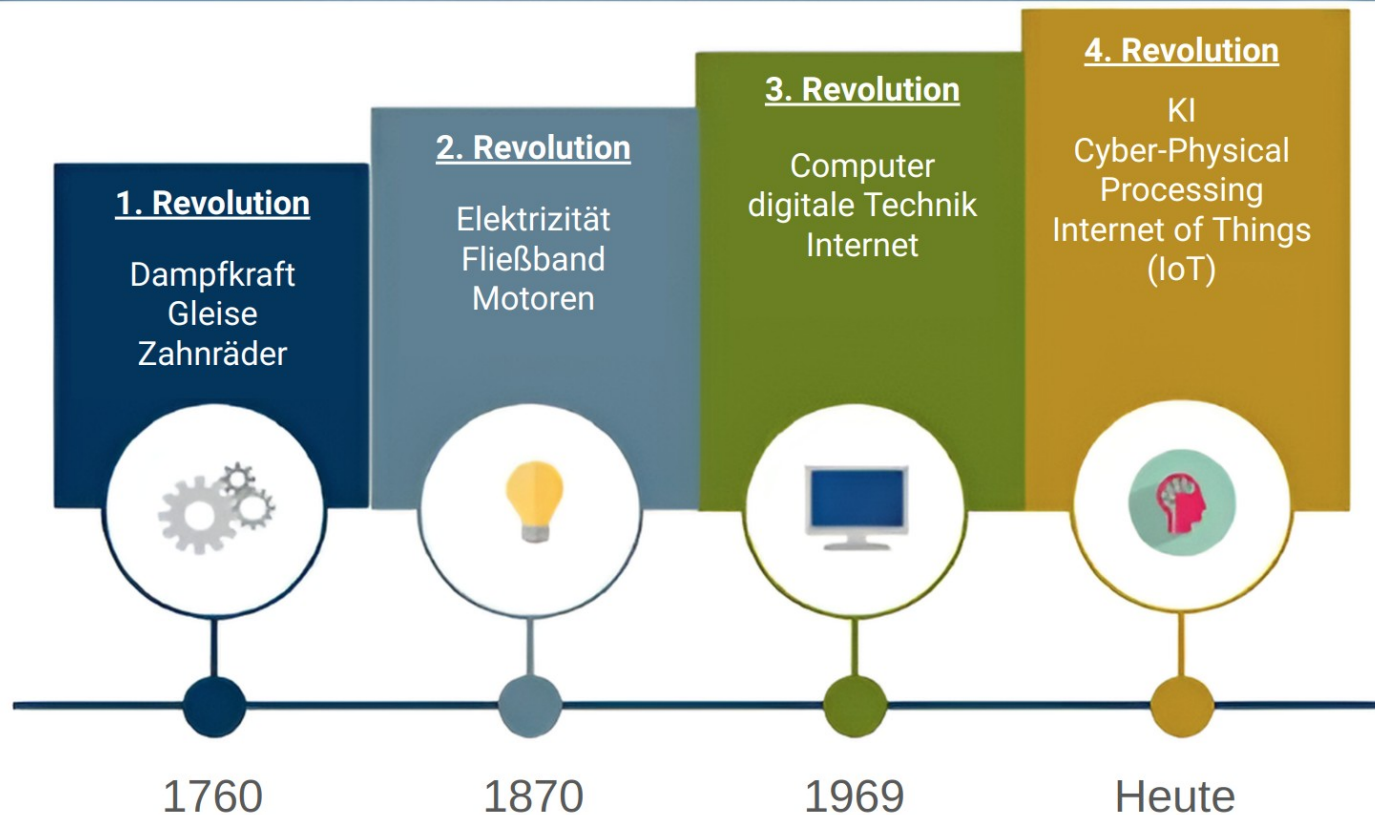


 The abstract presents a novel method for modeling myocarditis using hiPSC-CMs and immune cells, but the manuscript lacks key details and analysis. The viral inoculation protocol needs clarification, the small sample size limits the study's power, and the risks and benefits of immunosuppressive therapy are not fully addressed. The approach shows promise, but the manuscript requires revision and expansion before consideration for publication in Nature.

# KI als Schlüsselement der 4. Industriellen Revolution



## Vier Stufen der Industriellen Revolution



## KI: Bedrohung oder Chance?



„Die Menschen sind nicht in ausreichendem Maße für den Umgang mit den sich wandelnden technologischen Strukturen gerüstet.“

James Heckman (Ökonom, Nobelpreisträger 2000)

„Wachstumsmuster, in denen extreme Ungleichheit herrscht, funktionieren nicht.“

A. Michael Spence (Wirtschaftswissenschaftler, Nobelpreisträger 2000)





## Empfehlungssysteme auf Streaming-Plattformen



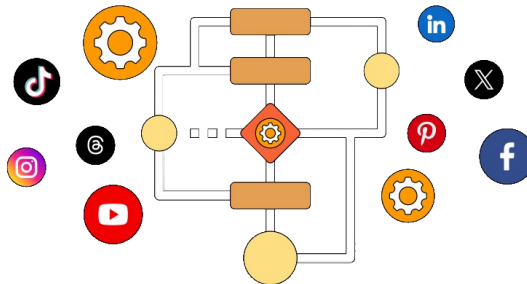
Quelle: <https://www.augsburger-allgemeine.de/wirtschaft/streaming-dienst-wie-aus-einem-dvd-verleih-der-unterhaltungsriese-netflix-wurde-id63759426.html>

## Ranking-Algorithmus



Quelle: <https://www.tagesschau.de/wirtschaft/unternehmen/google-cloud-rechenzentrum-deutschland-100.html>

## Werbung & personenbezogener Inhalt auf Social Media



Quelle: <https://www.socialchamp.io/blog/social-media-algorithm/>

## Spam-Filter



Quelle: <https://penplusbytes.org/strategies-for-dealing-with-e-mail-spam/>





**Die aktuellen Fortschritte in der KI sind**

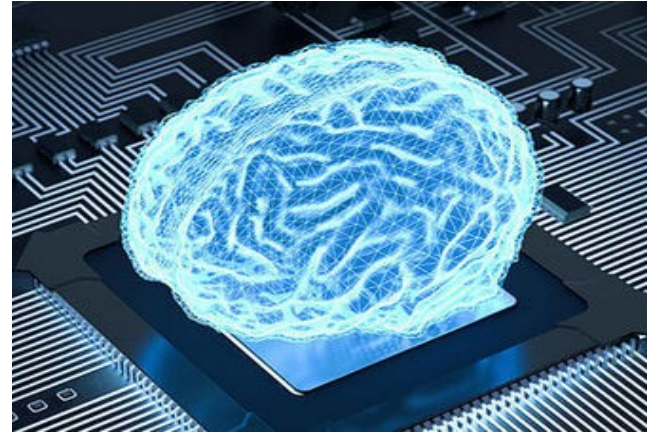
**viel schneller**

**als die Entwicklung des Verständnisses dafür in  
unserer Gesellschaft.**

# Künstliche Intelligenz: Quid?

Eine allgemeingültige Definition ist kaum denkbar: menschliche Intelligenz ist nicht eindeutig zu definieren.

Handlungsintelligenz  
Kognitive Intelligenz  
Emotionale Intelligenz  
Soziale Intelligenz  
...



Quelle: © fotomek - stock.adobe.com

Definition für diesen Vortrag:

Künstliche Intelligenz ist die Eigenschaft eines IT-Systems, der **menschlichen Kognition** ähnliche Fähigkeiten zu zeigen. Diese Fähigkeit kann aber auch über die menschliche Leistungsfähigkeit hinausgehen (z.B. Auswertung von 10.000 MRT-Scans). KI-Systeme verfügen in unterschiedlichen Anteilen über bestimmte Kernfähigkeiten wie situatives **Wahrnehmen**, **Kommunizieren**, **Handeln**, **Schlussfolgern** oder **Lernen**.



## Schwache KI (weak Artificial Intelligence):

KI, die die menschliche Intelligenz in abgegrenzten Teilbereichen erreicht oder sogar übertrifft.

### Beispiele:

- Zeichen- bzw. Texterkennung
- Bilderkennung
- Spracherkennung
- Individuelle Aussteuerung von Werbung
- Automatisierte Übersetzung
- Selbstfahrende Autos
- Expertensysteme (z.B. das Ableiten von Handlungsempfehlungen auf Basis einer Wissensdatenbank)
- ...



Alle heute existierenden Systeme fallen unter die Kategorie der schwachen KI.



## Starke KI (strong / general Artificial Intelligence):

KI, die die menschliche Intelligenz in allen Bereichen erreicht oder sogar übertrifft.

### Voraussetzungen:

- Logisches Denkvermögen
- Entscheidungsfähigkeit auch bei **Unsicherheit**
- Planungs- und Lernfähigkeit
- Fähigkeit zur Kommunikation in natürlicher Sprache
- Kombinieren aller Fähigkeiten zur Erreichung eines übergeordneten Ziels
  
- **Humor ...**



Quelle: <https://www.pinterest.co.uk/pin/613404411745794895/>

### Allensbach-Studie:

Drei von vier Deutschen denken beim Begriff KI an den Terminator.



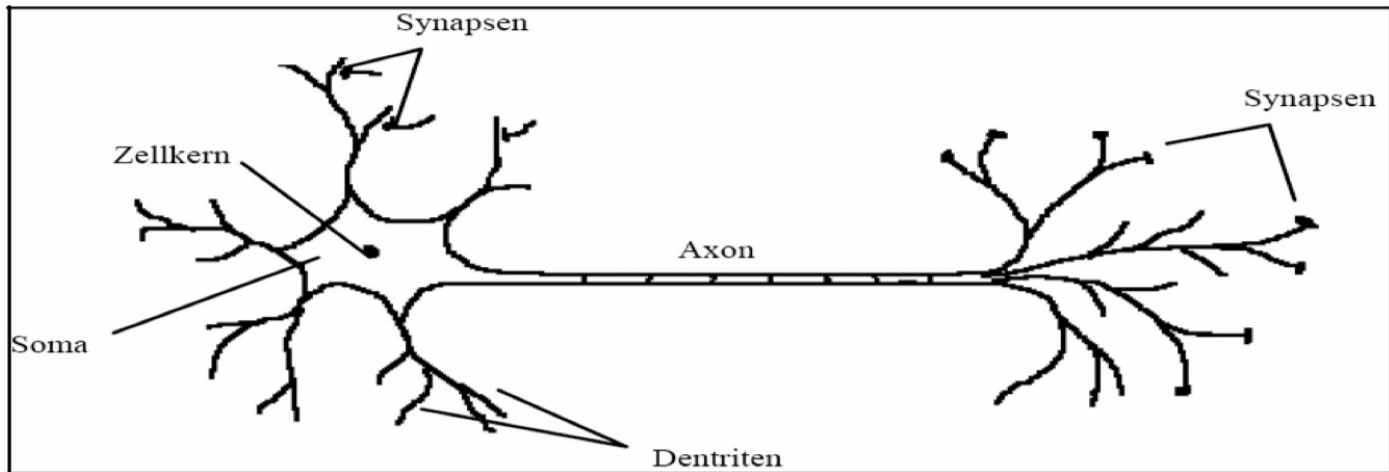
# Neuronale Netze in der Biologie



Neuronale Netze beziehen sich auf Strukturen des Gehirns von Tieren und Menschen.

Neuronen bestehen aus:

- einem Zellkörper (**Soma**)
- den **Dentriten** für den Empfang von Signalen anderer Neuronen
- dem **Axon** zur Übertragung von Signalen an andere Neuronen
- **Synapsen**, die den Kontakt zu Dentriten anderer Neuronen herstellen



Quelle: <https://www.dbai.tuwien.ac.at/education/AIKonzepte/Folien/NeuronaleNetze.pdf>

# Neuronale Netze in der Biologie



Menschliche Großhirnrinde:

- 10 Milliarden Neuronen
- diese können parallel arbeiten
- jedes Neuron ist mit ca. 2.000 anderen Neuronen verbunden

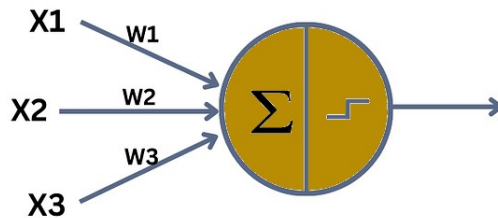


Quelle: <https://mt-portal.de/aktuell/wie-synapsen-nicht-mehr-verschwinden/>

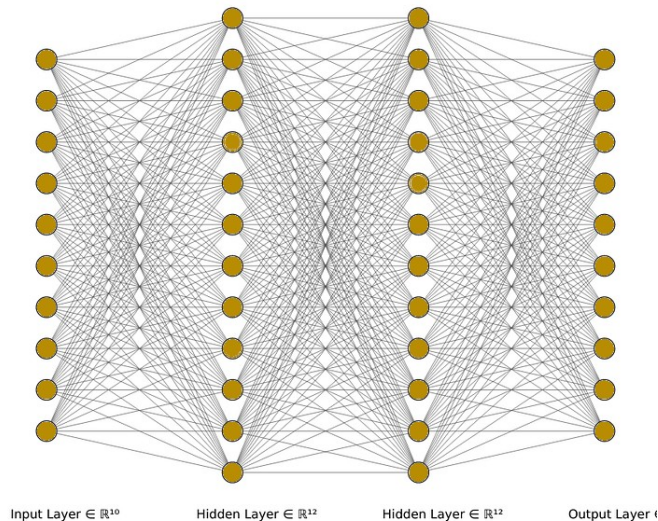
# Künstliche Neuronale Netze



Das Künstliche Neuronale Netz (KNN) ist dem Aufbau des biologischen Gehirns nachempfunden. Es besteht aus einem **abstrahierten Modell** miteinander verbundener Neuronen, durch deren spezielle Anordnung und Verknüpfung sich Anwendungsprobleme aus verschiedenen Bereichen computerbasiert lösen lassen.



Single-layer perceptron



Multi-layer perceptron



# Künstliche Neuronale Netze: Bildklassifikation



**Katze**



**Hund**



**Hund**



**???**



Quelle: <https://pair.withgoogle.com/explorables/saliency/>

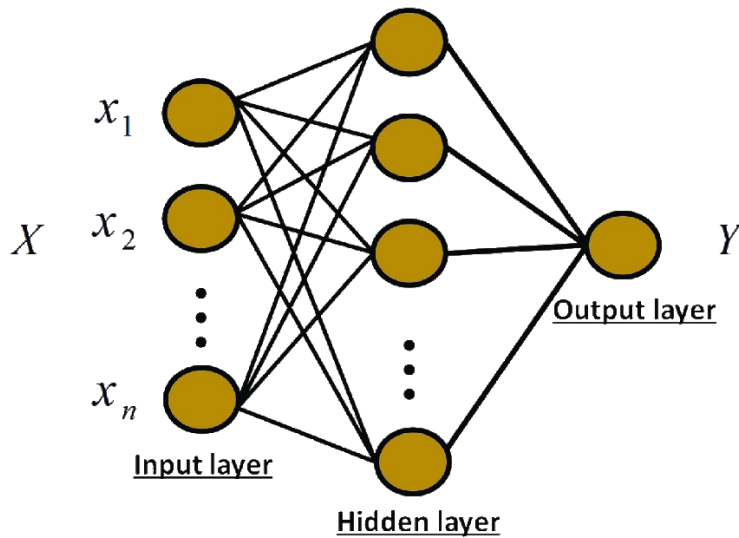


# Künstliche Neuronale Netze: Bildklassifikation



■ **Input layer:** 2D-Bilder

■ **Output layer:** „0“ für Katzen  
„1“ für Hunde



Quelle: <https://pethelpful.com/cats/tabby-cat-care>

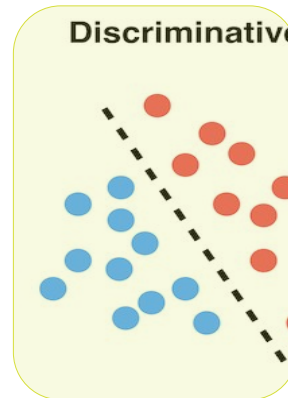
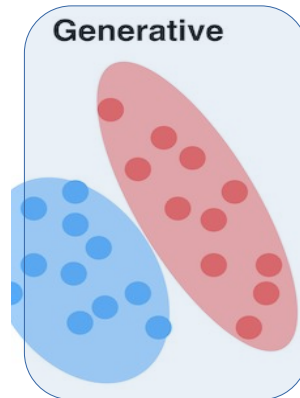
Output soll „0“ sein

# Arten von KI:



## Anwendungen:

- Texterzeugung mit LLMs (ChatGPT)
- Bilderzeugung (Dalle-2, FLUX.1)
- Musikgenerierung (MusicLM von Google)
- ...



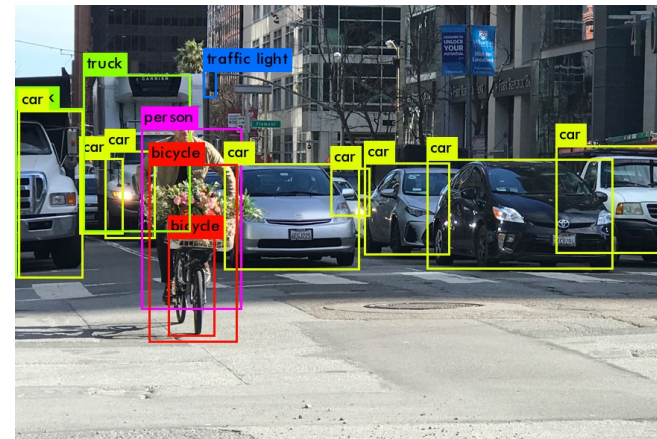
Source: <https://betterprogramming.pub/>

## Anwendungen:

- Bildklassifizierung
- Objekt-Erkennung
- Erkennung von Betrug
- Segmentierung
- ...



Prompt to FLUX.1 model: “two cute spiders in victorian outfits having a miniature tea party with a tiny table and teapot on a leaf, macro photo.”



Source: <https://itsjb13.medium.com/building-an-advanced-object-detection-application-for-autonomous-vehicles-yolov7-intel-pytorch-478ee5cedd39>



ChatGPT gibt jedem möglichen Wort bzw. Token (Wortbestandteil) eine Wahrscheinlichkeit auf eine Sequenz an Worten/Tokens zu folgen.

Methodik:  
**Multi Head Self Attention**

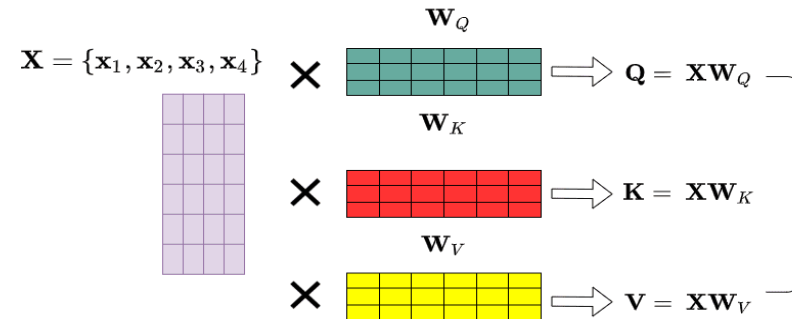
<s>	not	all	heroes	wear
0	1	2	3	4

Input Sequence

GPT

capes	90%
pants	5%
socks	2%
⋮	⋮

Output guess



- **Query (Q)** sucht nach relevanten Kontext-Token für ein Token
- **Key (K)** enthält Referenzpunkte für den Abgleich mit der Query.
- **Value (V)** enthält die eigentlichen Informationen des Tokens



## Vergleich von GPT-3 und GPT-4



**GPT-3**

Trainingsdaten: ~570 GB,  
Token: ~300 Milliarden,

Modellparameter: ~175  
Milliarden

**GPT-4**

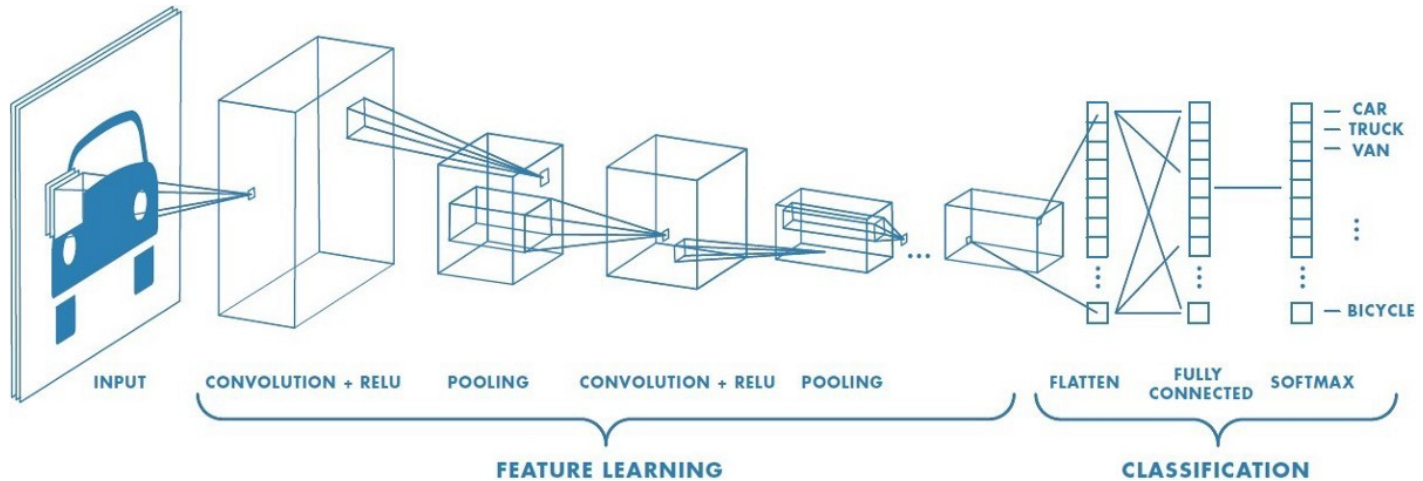
Trainingsdaten: ~20 TB,  
Token: 13 Billionen,

Modellparameter  
: ~1 Billion

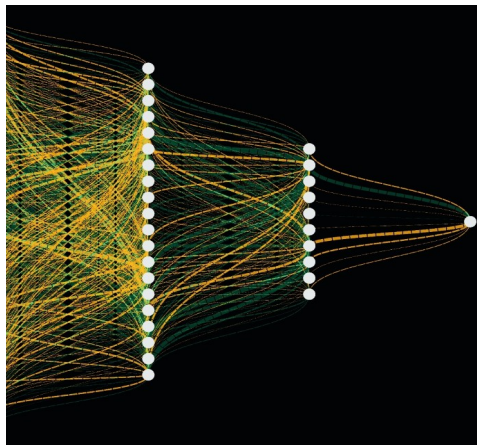
**Quellen für Trainingsdaten:** öffentliche Bücher, Wikipedia, Common Crawl



# Künstliche Neuronale Netze: Black Box?



Quelle: <https://saturncloud.io/blog/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way/>



Quelle: <https://www.uni-stuttgart.de/forschung/forschung-leben/1-2021/blackbox/>

## Fehlende Nachvollziehbarkeit durch komplexe Verkettungen / Gewichtungen



Quelle: <https://the-decoder.de/kuenstliche-intelligenz-keine-angst-vor-der-blackbox/>



## Studie des Digitalverbandes "Bitkom" (März 2024) mit 500 befragten Studenten

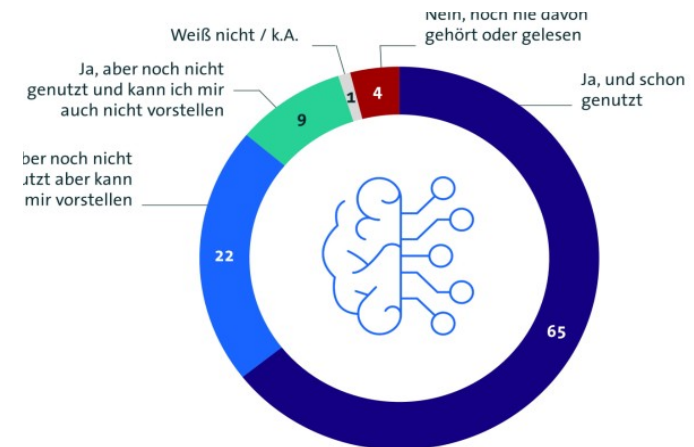
### KI-Nutzung (ChatGPT) unter Studierenden:

- 65% haben bereits ChatGPT genutzt.
- 22% können sich eine Nutzung vorstellen.
- 95% haben von ChatGPT gehört.
- Nur 4% kennen ChatGPT nicht.

### Regelungslücken: Nur 37% der Studierenden berichten von KI-Regeln an Hochschulen.

- **75% sind für das Lehren des richtigen KI-Einsatzes an Hochschulen.**

### Mehrheit der Studierenden kennt & nutzt ChatGPT



[https://www.bitkom.org/Presse/Presseinformation/So-digital-sind-Deutschlands-Hochschulen#\\_](https://www.bitkom.org/Presse/Presseinformation/So-digital-sind-Deutschlands-Hochschulen#_)

**Hochschulen müssen sicherstellen, dass Studierende und Personal ausreichende KI-Kompetenz haben!**



## Leitfaden für KI Anwendungen an Universitäten

### Verantwortungsvoller Umgang

Förderung der ethischen und durchdachten Anwendung von KI-Tools.

### Energieeffiziente KI

Entwicklung von Regeln für eine nachhaltige Nutzung & Entwicklung von KI-Methoden.

### Entwicklung kritischer Kompetenzen

Förderung von Fähigkeiten zur Navigation und Gestaltung von KI-gesteuerten Veränderungen.

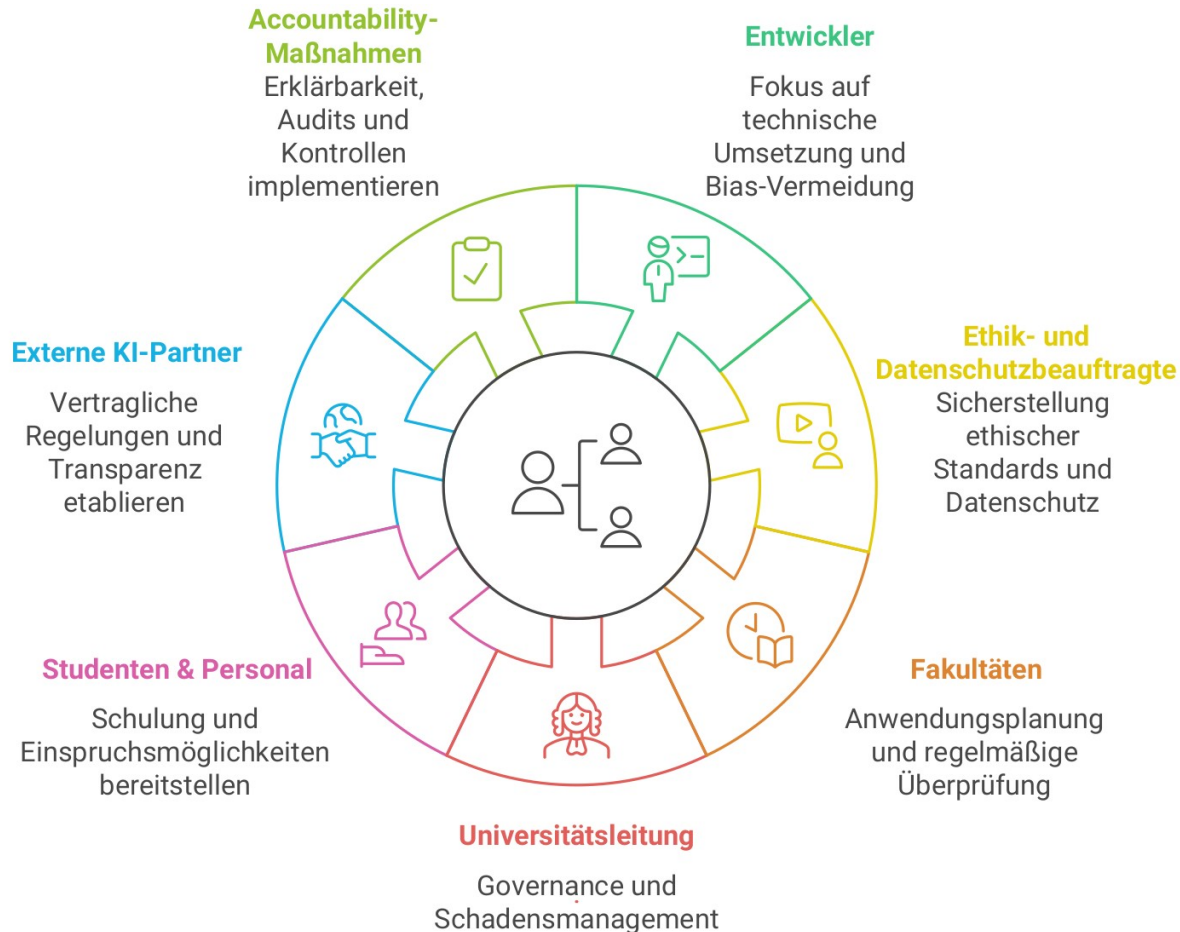
### Nachhaltigkeitsbewertung

Bewertung der langfristigen Auswirkungen und Effizienz von KI-Lösungen.





## Verantwortung von KI Anwendungen an Universitäten





# Herausforderungen



## Datenschutz und Sicherheit

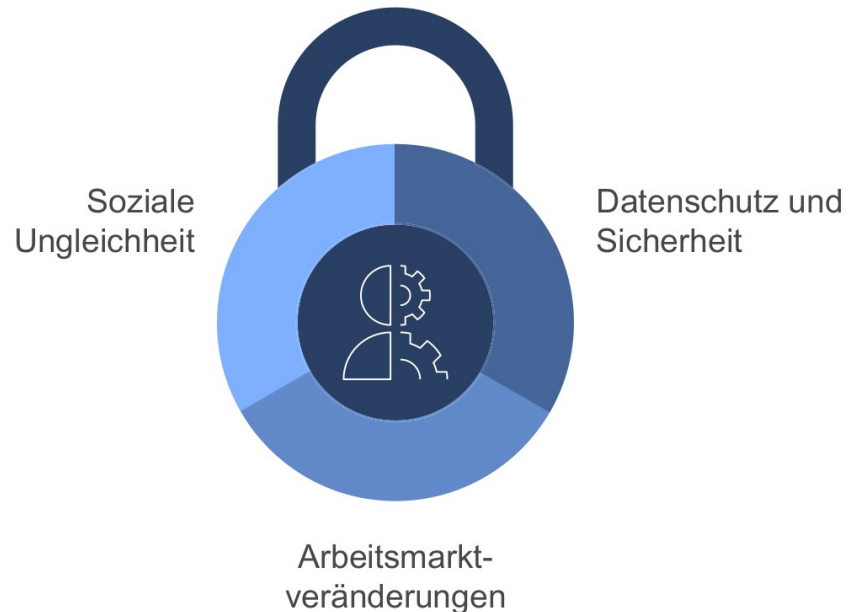
- Risiko des Datenmissbrauchs von Trainingsdaten
- Massen-Überwachung

## Arbeitsmarktveränderungen

- Automatisierung von Arbeitsplätzen
- Arbeitsplatzverlust
- Bedarf an Umschulung und Weiterbildung

## Soziale Ungleichheit

- Zugang zu KI-Technologien
- Gefahr der Vertiefung bestehender sozialer Ungleichheiten



# Herausforderungen



## Transparenz und Nachvollziehbarkeit

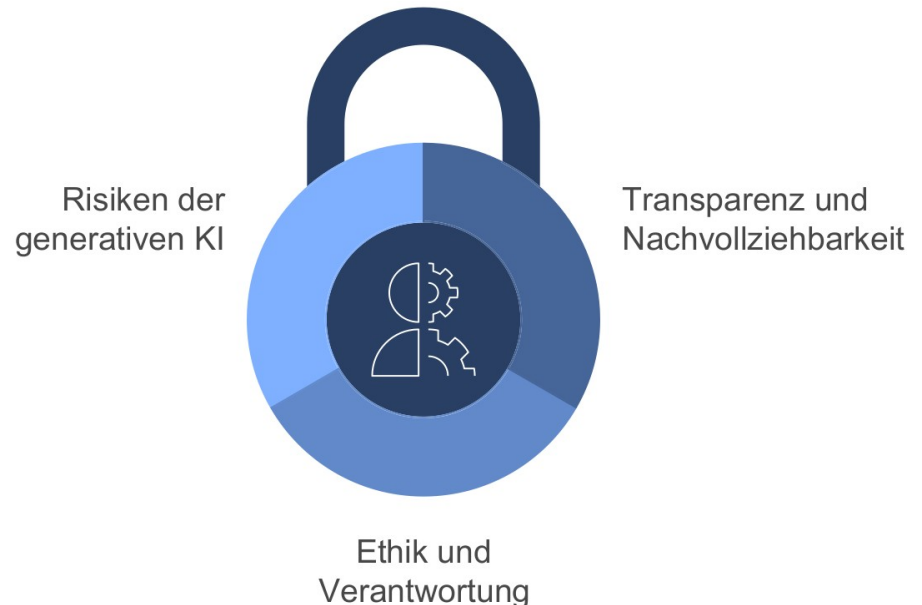
- Komplexität und Intransparenz von KI-Systemen
- Bedarf an Erklärbarkeit und Kontrolle

## Ethik und Verantwortung

- Entscheidungsfindung durch KI: Wer trägt die Verantwortung?
- Verbreitung von schädlichem KI-generiertem Inhalt
- Bias und Diskriminierung in Algorithmen

## Generative KI

- Gefahr des Halluzinierens
- Gefahr der Zensur



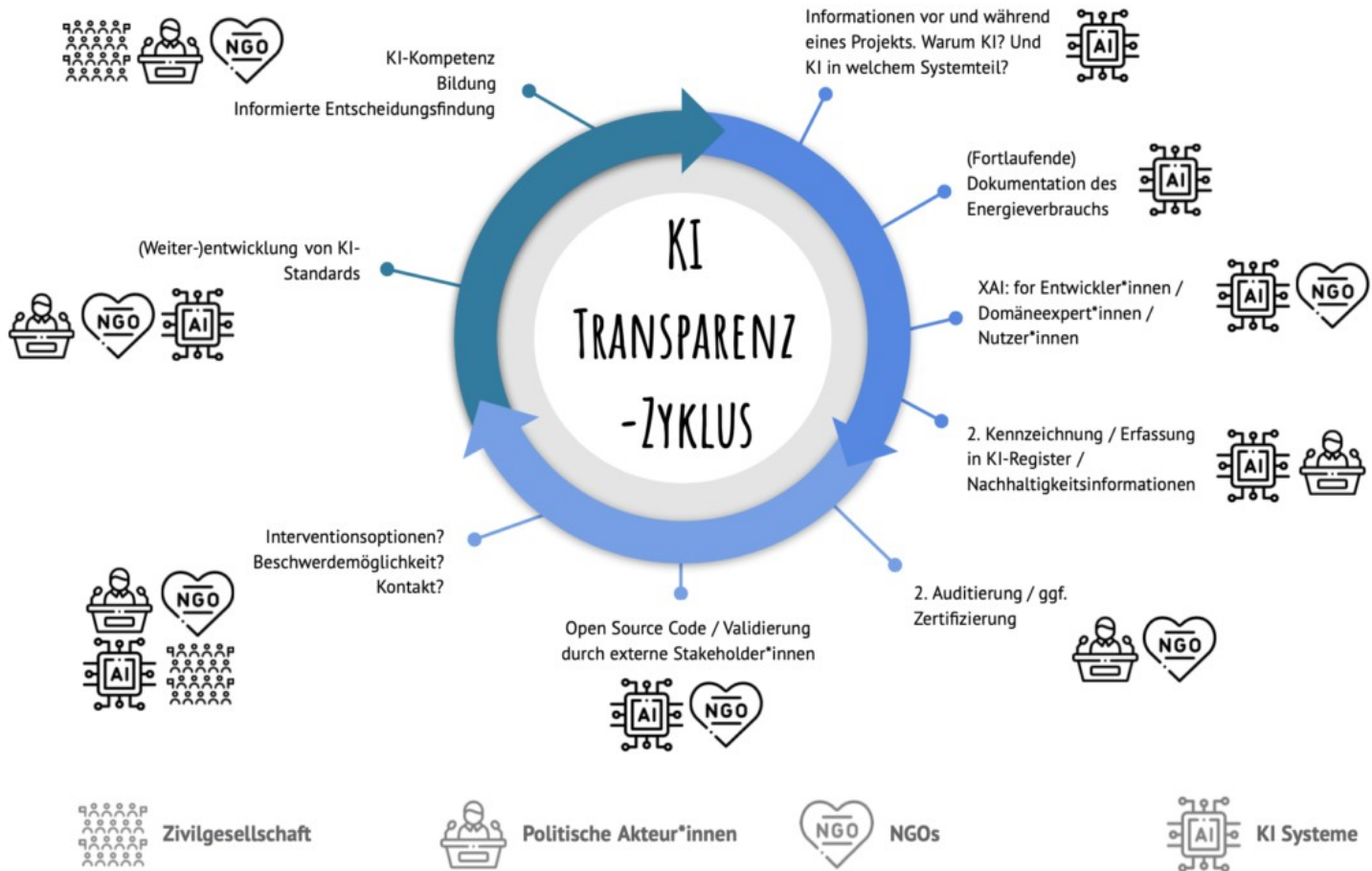
Quelle: <https://www.hiig.de/ki-transparenz-zyklus/>



**KI ist nicht überall sinnvoll!**

**Überprüfung ob effiziente Algorithmen  
bereits vorhanden sind und ausreichen!**

# Transparente KI



Quelle: <https://www.hiig.de/ki-transparenz-zyklus/>





19. SEPTEMBER 2024 - KÜNSTLICHE INTELLIGENZ (KI)

## KI FÜR ALLE: URZ ETABLIERT SPRACHMODEL- PLATTFORM FÜR DIE UNIVERSITÄT

Seit dem Launch von ChatGPT im November 2022 schreiten die rasanten Entwicklungen auf dem Markt der Anwendungen generativer künstlicher Intelligenz und deren Leistungssprünge weiter voran. Aufgrund dieser Dynamik gibt es inzwischen weltweit etliche große Sprachmodelle (sog. LLMs – Large Language Models), darunter leistungsstarke Open-Source-Systeme, die auch von großen Unternehmen eingesetzt werden. Für die Universität Heidelberg stellt das URZ Studierenden und Beschäftigten diese Technologie nun zur Verfügung.



Summit Art Creations / Adobe Stock

Auf der Plattform „YoKI“ kann ab sofort ein solches Open-Source-Sprachmodell in einer ersten Testphase erprobt werden. Auf der KI-Plattform werden perspektivisch verschiedene Sprachmodelle angeboten, die je nach System unterschiedliche Bedarfe von Forschung, Lehre und Verwaltung berücksichtigen. So können die Nutzer:innen entscheiden, welches der Modelle sich am besten für ihren Anwendungszweck eignet und das für sie passende Modell auswählen.

YoKI Chatbot





 YoKI v0.9.3

Universität Heidelberg. Zukunft. Seit  
1386.




<https://yoki.urz.uni-heidelberg.de/>

 YoKI

New Chat

Today

 Forschungsdatenmanagement kri

Bitte gebe ein kurzes kritisches Statement zu dem universitären Forschungsdatenmanagement in Bezug auf den sicheren Austausch jener Forschungsdaten, den Energieverbrauch der Speicherung sowie den Energieverbrauch der Hochleistungsrechner, welche verwendet werden, um hieraus KI Modelle zu entwickeln. Wie könnte das Management und der Energieverbrauch optimiert werden?

## Prompt:

*“Bitte gebe ein kritisches Statement zu dem **universitären Forschungsdatenmanagement** in Bezug auf den **sicheren Austausch** jener Forschungsdaten, den **Energieverbrauch** der **Speicherung** sowie den Energieverbrauch der **Hochleistungsrechner**, welche verwendet werden um hieraus KI Modelle zu entwickeln. Wie könnte das Management und der Energieverbrauch **optimiert werden?**”*



 YoKI v0.9.3

Universität Heidelberg. Zukunft. Seit  
1386.



<https://yoki.urz.uni-heidelberg.de/>

- Universitäre Forschungsdatenmanagementsysteme haben oft Probleme mit der Sicherheit beim Austausch von Daten, dem hohen Energieverbrauch bei der Speicherung sowie der Auslastung von Hochleistungsrechnern für KI-Modelle. Um dies zu optimieren, könnten **zentralisierte Sicherheitsrichtlinien** implementiert werden, um den sicheren Datenaustausch zu gewährleisten. Darüber hinaus könnten energiesparende Speicherlösungen wie beispielsweise die Verwendung von **SSDs oder cloud-basierten Archivierungsmethoden eingesetzt werden, um den Energieverbrauch zu reduzieren.** Für Hochleistungsrechner könnte ein **Lastenausgleichssystem implementiert** werden, das Ressourcen effizienter nutzt und somit den Gesamtenergieverbrauch senkt.

**Konkstruktive Vorschläge für  
die jeweiligen Punkte!**



Universitäre Forschungsdatenmanagementsysteme sind oft mit Sicherheitsproblemen bei der Datenübermittlung, hohem Energieverbrauch in der Speicherung und der intensiven Nutzung von Hochleistungsrechnern für KI-Modelle konfrontiert. Um diese Herausforderungen zu bewältigen, könnten folgende Optimierungsmaßnahmen umgesetzt werden:

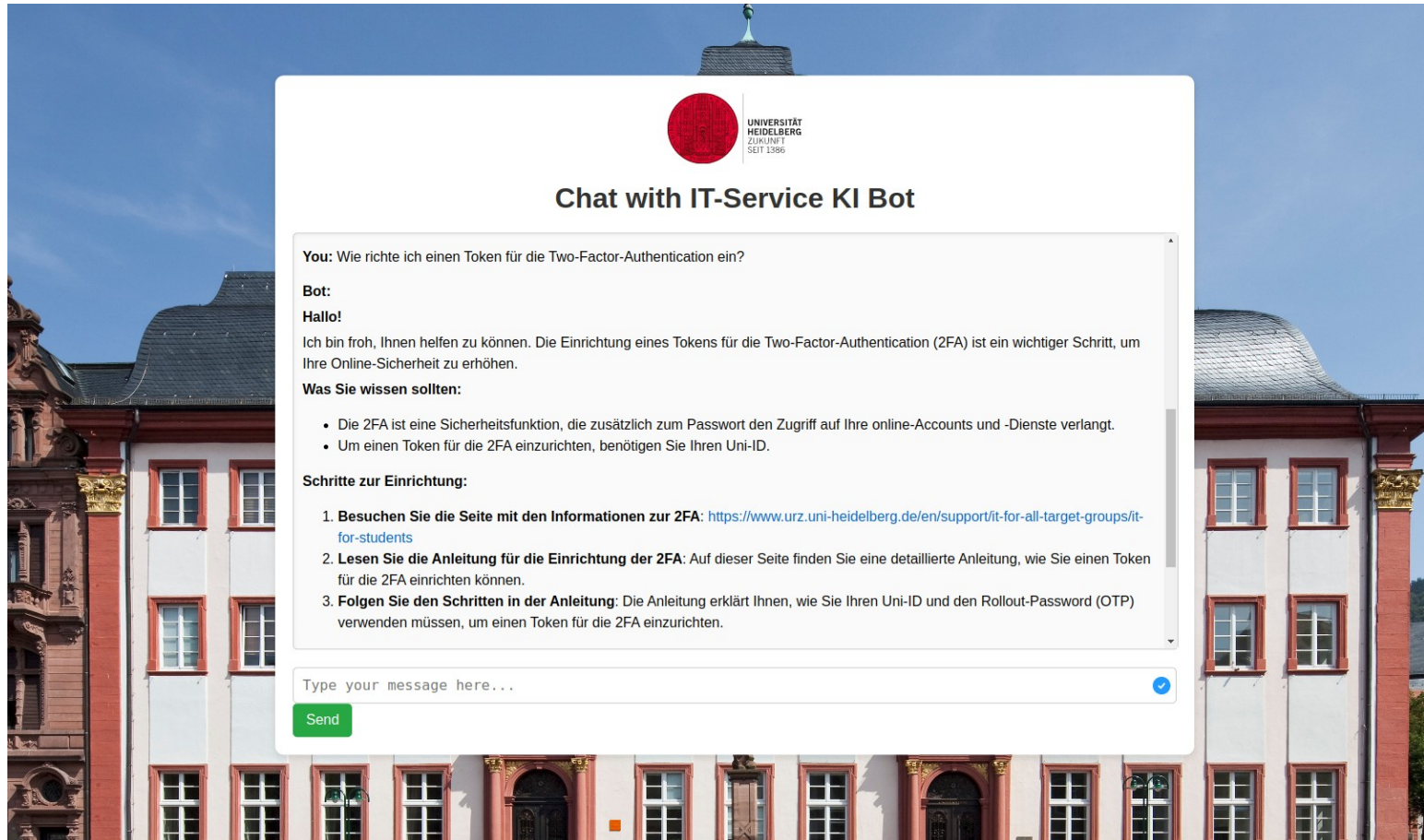
1. **Sichere Datenaustauschmethoden:** Durch die Implementierung von zentralisierten Sicherheitsrichtlinien und **kryptografischen Verfahren** kann der sichere Austausch von Forschungsdaten gestärkt werden, wodurch Risiken wie Datenabflüsse oder Manipulation minimiert werden.
2. **Energiesparende Speichertechnologien:** Die Verwendung von **SSDs anstelle von HDDs sowie cloud-basierte Archivierungsmethoden** können den Energieverbrauch bei der Speicherung effizient reduzieren, da sie weniger Energie benötigen und schneller zugreifbar sind.
3. **Effiziente Hochleistungsrechner-Verwaltung:** **Ein Lastenausgleichssystem könnte** die Nutzung von Rechenressourcen optimieren, indem es die Auslastung von Hochleistungsrechnern balanciert und somit den Gesamtenergieverbrauch senkt. Zudem könnten energiesparende Hardwarelösungen oder die Verwendung von alternativen Berechnungsmethoden wie **Edge-Computing** eingesetzt werden, um Ressourcen zu sparen.

Diese Maßnahmen könnten nicht nur die Sicherheit und Effizienz der Forschungsdatenverwaltung verbessern, sondern auch den ökologischen Fußabdruck solcher Systeme reduzieren.



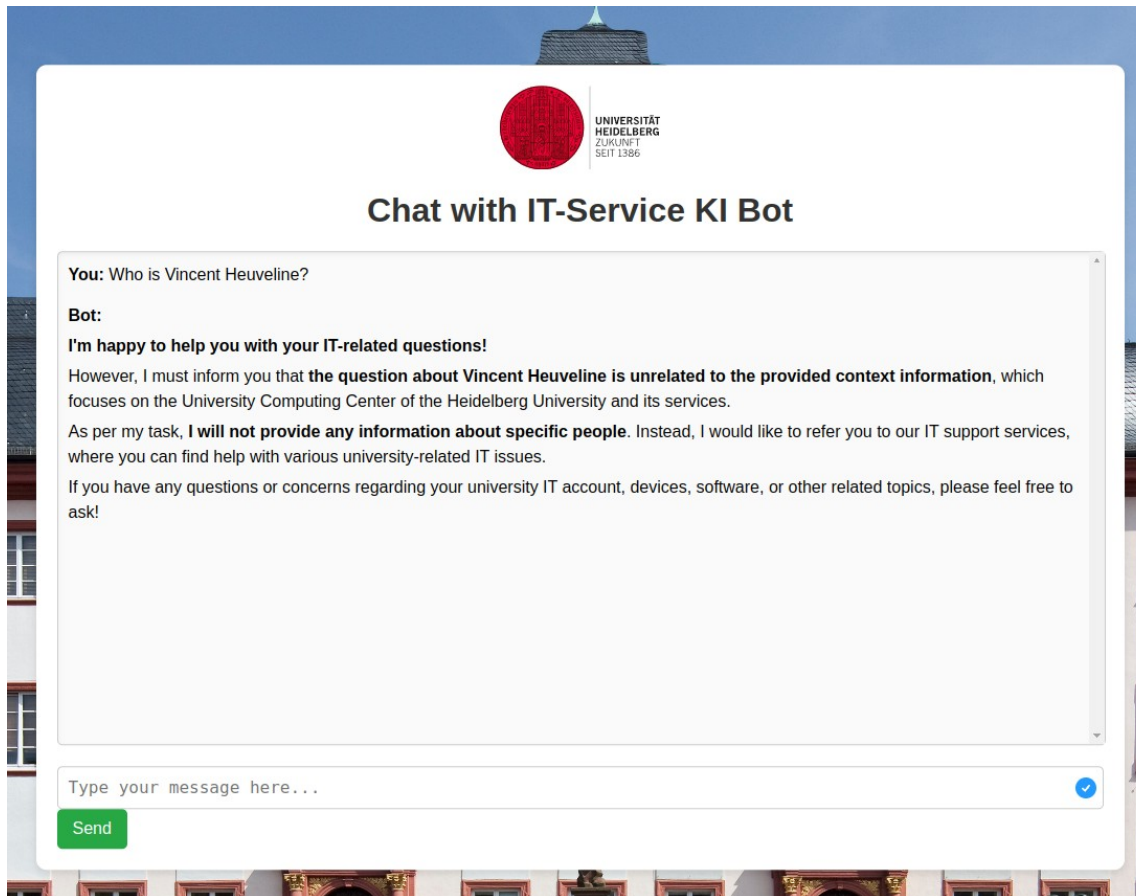


**Spezialisierte Sprachmodelle:** Das URZ hat ein Chat-Bot Sprachmodell entwickelt, welches auf Basis der URZ Websites & Dokumente eine detaillierte IT Antwort liefert.





**Datenschutz:** Der Chatbot speichert keine Chatverläufe in jeglicher Hinsicht und unterbindet Antworten auf personenbezogene Daten sowie Fragen außerhalb des URZ.





- Persönlich sehe ich in KI deutlich mehr Möglichkeiten als Gefahren.
- Langfristige Arbeitsplatzverluste unwahrscheinlich, es werden andere Jobs geschaffen, wo Jobs verloren gehen (siehe Industrielle Revolution).
- Reale Gefahr der Mass Surveillance.
- Transparenz, wo KI eingesetzt wird.
- Problematik der realistischen, jedoch künstlichen Inhalte.
- Wie KI grundlegend funktioniert, sollte vermittelt werden.
- Große Problematik, dass bereits reichere Institutionen noch mächtiger werden, da das Trainieren von KI Modellen sehr kostspielig ist.



Quelle: <https://www.linkedin.com/pulse/top-10-challenges-considerations-businesses-while-navigating-gn/>

➔ **KI ist mächtiges Instrument zur Unterstützung in der Wissenschaft!**

# Fragen

