

A woman's profile is shown in a dark blue, futuristic setting. Her head is filled with a complex network of glowing lines and nodes, representing artificial intelligence or data processing. The background is a dark blue gradient with a subtle grid of white lines and dots, suggesting a digital or network environment.

Data Analysis mit KI

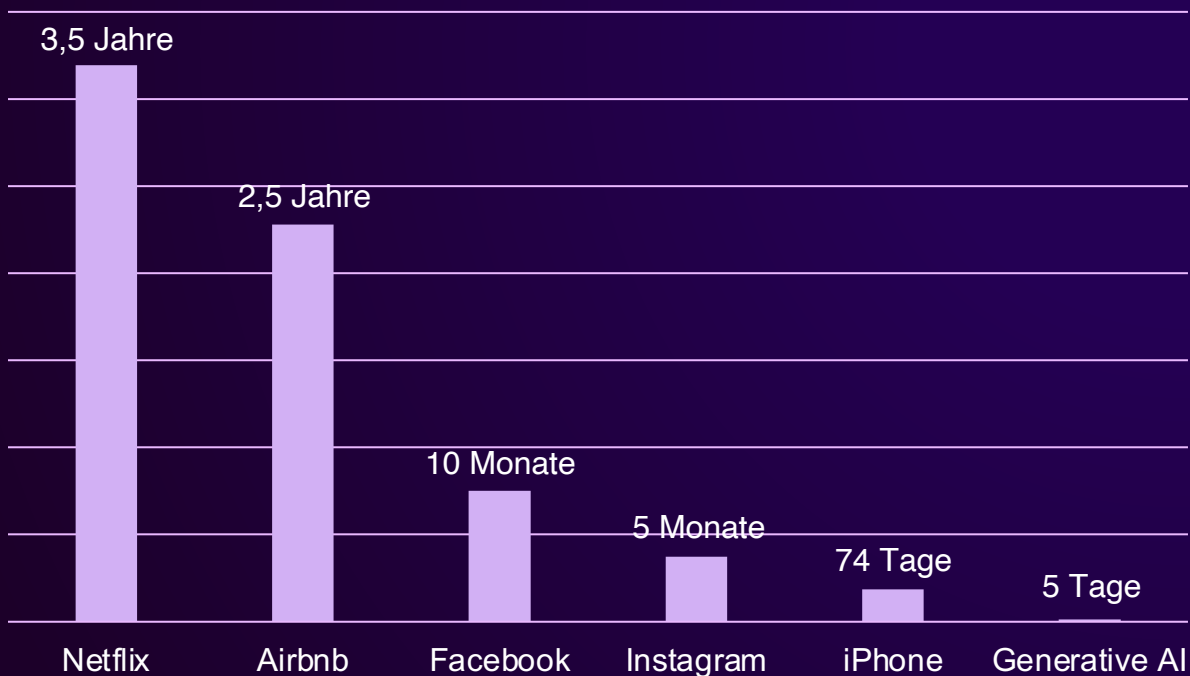
So wird jeder zum professionellen Data Scientist

Generative KI-Tools sind mit der schnellsten Verbreitung in der Geschichte ins öffentliche Bewusstsein gerückt



ChatGPT überschritt die Marke von einer Million Nutzern in nur fünf Tagen, nachdem es im November 2022 an die Öffentlichkeit gegangen war.

Zeit bis 1 Million Nutzer



Forbes
AI 'humans' in virtual world plot and gossip like us

GPT-4 Beats 90% Of Lawyers Trying To Pass The Bar

"World's Most Advanced Humanoid Robot" Can Now Have Full Conversations

INNOVATION ©
ChatGPT Passed the U.S. Medical Licensing Exam. Will It Be Your Future Doctor?

ARTIFICIAL INTELLIGENCE
AI Can Re-create What You See from a Brain Scan

TECH
OpenAI announces GPT-4, claims it can beat 90% of humans on the SAT
PUBLISHED TUE, MAR 14 2023-1:42 PM EDT | UPDATED TUE, MAR 14 2023-2:32 PM EDT

Chat-GPT Pretended to Be Blind and Tricked a Human Into Solving a CAPTCHA

Mehr als 60% der Deutschen haben keine KI-Kenntnisse, obwohl ChatGPT sehr bekannt ist

Nutzung von KI in Deutschland

Einstellung & Anwendung

Meinung zu KI in Deutschland

(Stichprobengröße 1,000 Personen)

- **80% sagen, dass**
 - KI das **Risiko der Fehlinformationen erhöht**
 - **Datensicherheit hat eine hohe Priorität**
- Skepsis aufgrund **Abhängigkeit und Arbeitslosigkeit hoch**

Nutzung von KI

- Noch nicht in großem Umfang verwendet
- Hauptsächlich Informationsbeschaffung und Texterstellung
- ChatGPT wird häufig von unter 25-Jährigen zu Lernzwecken genutzt

Gründe für die Nutzung von KI, z.B. ChatGPT





Robin Reuschel

Mitgründer und Geschäftsführer
ROOVER GmbH

robin@roover.eu

Find me on LinkedIn



Wir unterstützen Unternehmen in den Themen Effizienz, Wachstum und Digitalisierung

Kompetenz-Portfolio: Beratungsdienste

Bildung

- Metaverse-Workshops
- KI Masterclass

Wachstumstreiber

- Entwicklung von Geschäftsmodellen
- Diversifizierung und Markteintritt
- Metaverse und KI-Strategie

Investoren Unterstützung

- Finanzielle Gestaltung
- Due Diligence / M&A

Roover – Kompetenz- Portfolio

Verbesserung der Leistung

- Umstrukturierung
- Prozess- & Organisationsoptimierung
- Funktionelle Analyse

Effizienz mit Technologie

- Effizienzsteigerung mit Metaverse-Technologien / KI
- 360° IT Check

Implementierung

- Unterstützung bei der Umsetzung von Metaverse-Projekten / technologischen Lösungen

Unsere
Kunden



Lufthansa
Industry Solutions



SUPER **RTL** Beiersdorf Bookwire

ANTERDIT

YEAHR!

Boer
LEGAL

Und das entlang von 6 Zukunftstechnologien

**Blockchain &
Digitale Güter**



**Virtuelle
Welten**



**AR & VR
Technologien**



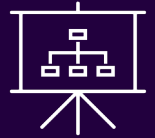
**Digitale
Zwillinge**



**Künstliche
Intelligenz**



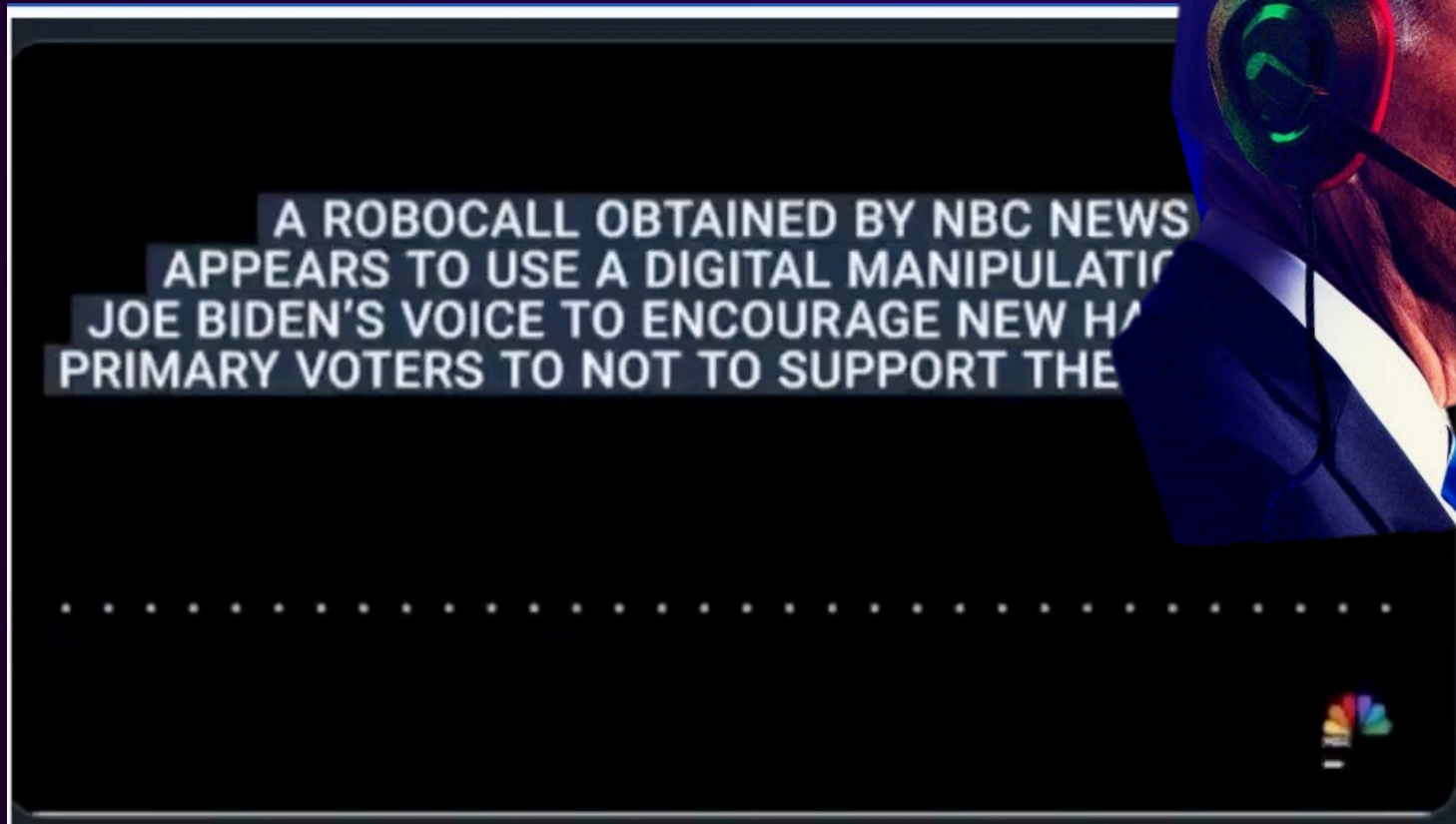
Robotics



KI hat bereits heute schon großen Einfluss auf unser Leben



KI hat bereits heute schon großen Einfluss auf unser Leben



Revolution in der Antibiotikaforschung durch Einsatz von KI

KI in der Wissenschaft

Wissenschaftler entdecken die ersten neuen Antibiotika seit über 60 Jahren



Wissenschaftler entdecken mithilfe von KI die ersten neuen Antibiotika seit über 60 Jahren



Was?

Eine neue Klasse von Antibiotika für arzneimittelresistente MRSA-Bakterien, die mit Hilfe von transparenteren **Deep-Learning-Modellen** entdeckt wurde.



Warum?

- ✓ Komplexe Daten können durch die Nutzung von KI schneller analysiert werden
- ✓ Traditionell langwieriger und kostenintensiver Prozess der Antibiotikaentwicklung wird beschleunigt
- ✓ Ermöglicht eine große Anzahl von Verbindungen schnell zu screenen

KI-Durchbruch sichert Weg zu sauberer Fusionsenergie

KI in der Forschung und Entwicklung

Scientists Claim AI Breakthrough to Generate Boundless Clean Fusion Energy

Princeton researchers report that a new AI model has solved one of the major roadblocks to generating fusion energy.



Source: Roover, Vice



Was?

Princeton-Forscher haben ein **KI-Modell entwickelt, das den Austritt von Plasma in Fusionsreaktoren verhindert**. Ein Durchbruch für saubere und unbegrenzte Energie, ähnlich wie bei der Sonne.



Warum?

- ✓ **KI ermöglicht präzise Vorhersage und Vermeidung von Plasma-Instabilitäten**
- ✓ Beschleunigt Entwicklung sicherer und effizienter Fusionsreaktoren
- ✓ KI-Technologie entscheidend für Durchbruch in sauberer Energiegewinnung

... wer braucht dann überhaupt noch Data-Scientists?



Awesome! Time to create a logistic regression. Split the data into a training set and a test set 80/20, train the logistic regression model, evaluate the model's performance by predicting the test set results, and generate a classification report.



manipulate_sheet complete ✓
> "data": { ... } 2 Items



The logistic regression model has been trained and evaluated. Here's the classification report:

	precision	recall	f1-score	support	
1					
2					
3	0.0	0.40	0.18	0.25	136
4	1.0	0.59	0.82	0.69	403
5	2.0	0.47	0.24	0.32	165
6					
7	accuracy			0.56	704
8	macro avg	0.49	0.41	0.42	704
9	weighted avg	0.53	0.56	0.52	704



Create a linear regression matrix heatmap of income vs. the different spending categories

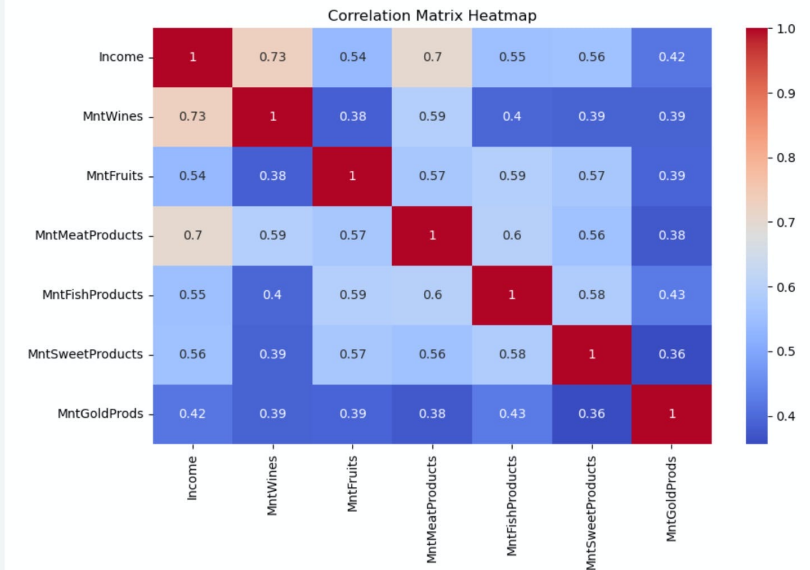


manipulate_sheet complete ✓
> "data": { ... } 2 Items



I've created a heatmap to visualize the correlation matrix of income vs the different spending categories. Each cell in the heatmap represents the correlation coefficient between a pair of variables. The color of the cell indicates the strength and direction of the correlation, with darker colors indicating stronger correlations.

Here's the heatmap:



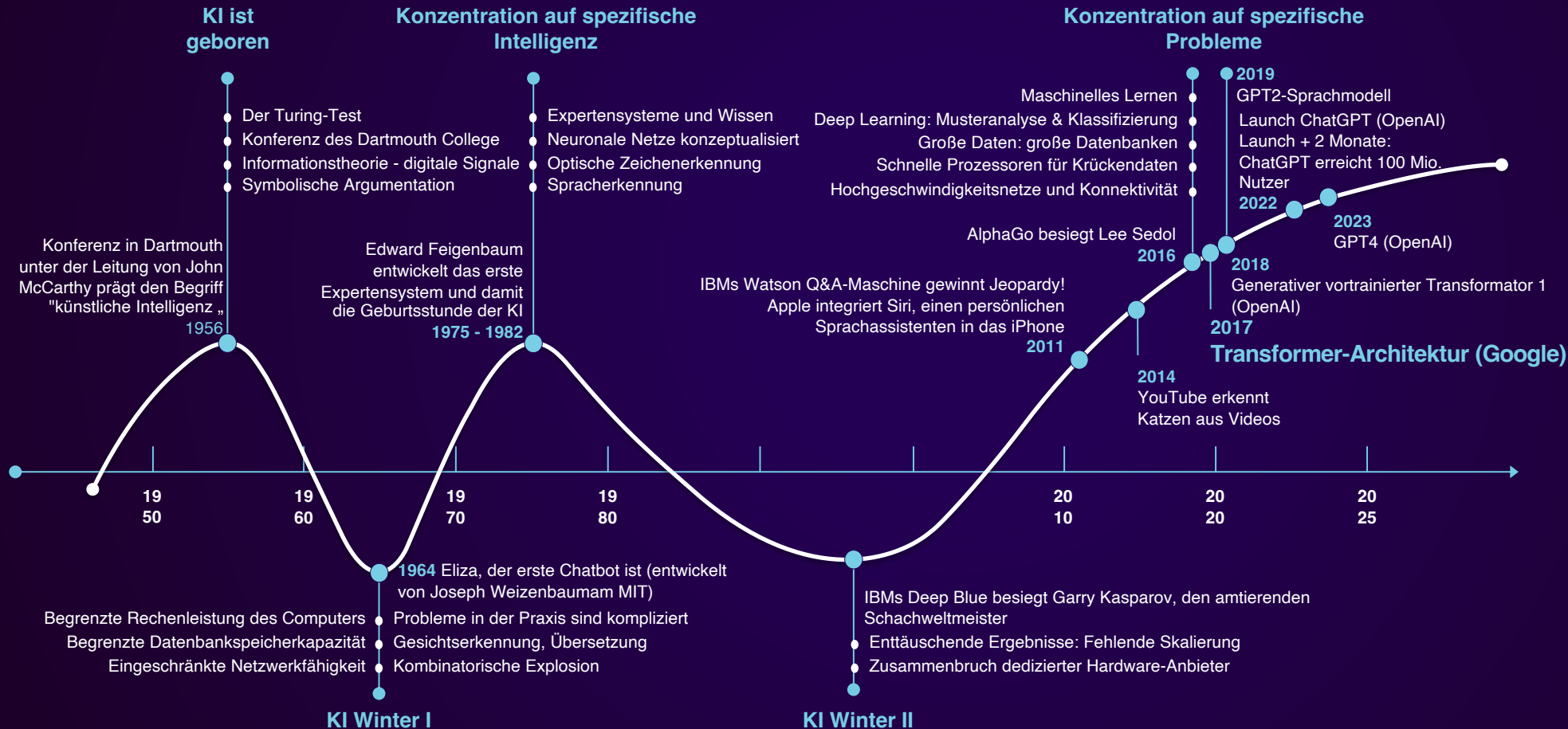
Regenerate response



Was soll der Hype eigentlich?

KI wurde bereits 1956 geboren und entwickelte sich langsam weiter - LLMs Forschung & Geschäftsanwendungen

Zeitleiste der KI-Entwicklung



AUSBLICK

KI in separaten Disziplinen werden vereint in universellen LLMs

- Music Generation
- Image Generation
- Robotics
- Speech Recognition
- Speech Synthesis
- Computer Vision

KI umfasst verschiedene Arten, wie Daten durch Algorithmen verarbeitet werden

Künstliche Intelligenz

Bereich der Informatik, der versucht, **Expertensysteme zu schaffen** oder sogar **die menschliche Intelligenz nachzuahmen**

Maschinelles Lernen

Teilbereich der KI, der es **Maschinen ermöglicht, aus vorhandenen Daten zu lernen**, um Entscheidungen oder Vorhersagen zu treffen

Deep Learning

Eine **Technik des maschinellen Lernens**, bei der **Schichten von neuronalen Netzen verwendet werden**, um Daten zu verarbeiten und Entscheidungen zu treffen

Generative KI

ChatGPT

Fähigkeit zur Verwendung von Eingabeaufforderungen zur **Erstellung, Verbesserung und Interaktion** mit Texten, Bildern und Videos unter Verwendung großer trainierter Modelle

Nein, nicht diese
Transformer...



Transformer waren die wichtigste Erfindung, die die Entwicklung von LLMs wie ChatGPT ermöglichten

Neuronale Netzwerke Transformer

Sieh es als eine
“*One Word
Theatre
Show*”

Generative vortrainierte Transformer berechnen die Wahrscheinlichkeit des nächsten Wortes

Neuronale Netze Transformer

Es kommt auf den Kontext an

- **Initial Systeme:** Vorhersage des nächsten Tokens und verdeckte Sprachmodellierung"
- **Stochastische Papageien**": Hannah ist eine "...", was auch immer Hannah in unserem Trainingsset am meisten ist.
- Eine weitere Stufe der Komplexität: **Selbstwahrnehmung**
- eingereicht von Vaswani et al. im Jahr 2017 (Google)
- **Transformer-Architektur**, die vollständig auf einem Selbstbeobachtungsmechanismus basiert
- Kann **Sequenzen parallel verarbeiten** und Abhängigkeiten zwischen ihnen aufrechterhalten
- **Übertrifft RNNs und Convolutional Neural Networks (CNNs)** bei vielen Aufgaben

Next-token-prediction

Dem Modell wird eine Folge von Wörtern vorgegeben, mit dem Ziel, das nächste Wort vorherzusagen.

Beispiel:

Hannah ist eine

Hannah ist eine *Schwester*
 Hannah ist eine *Freundin*
 Hannah ist eine *Verkäuferin*
 Hannah ist eine *Komikerin*

Masked-language-modeling

Dem Modell wird eine Folge von Wörtern vorgegeben mit dem Ziel, ein "verdecktes" Wort in der Mitte vorherzusagen.

Beispiel:

Jakob liest

Jakob *fürchtet* das Lesen
 Jakob *liebt* das Lesen
 Jakob liest *gerne*
 Jakob *hasst* Lesen

Der Mechanismus der Selbstwahrnehmung hilft dem Modell, Zusammenhänge zwischen Wörtern zu erkennen



Aufmerksamkeit als technisches Konzept



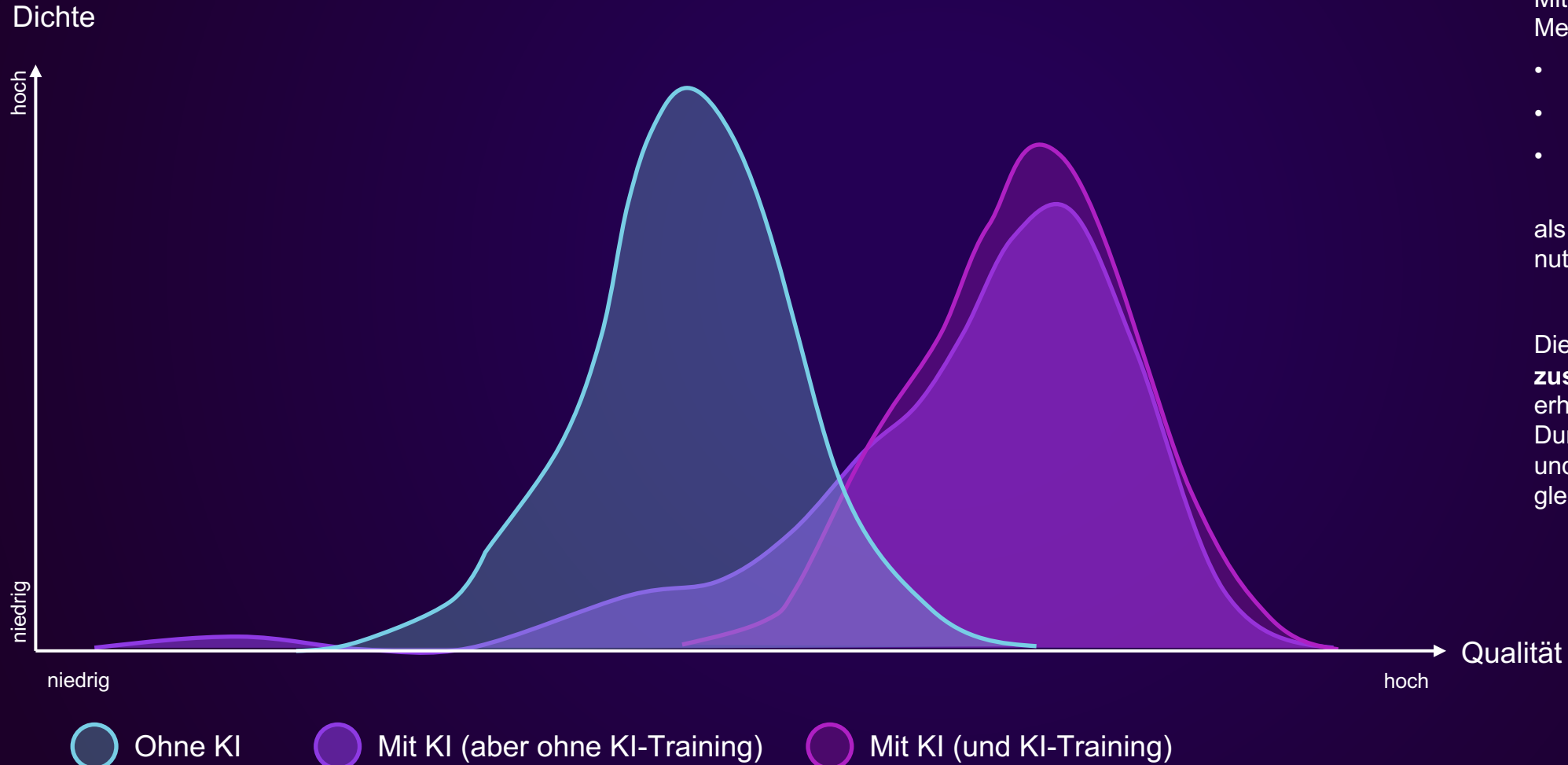
Selbstwahrnehmung ist das Schlüsselprinzip, das moderne LLMs so viel besser macht als je zuvor

- **Erfasst die Beziehungen** (einschließlich der Bedeutung) zwischen **verschiedenen Wörtern** in einem Satz
- **3 Informationen pro Wort:**
 - Das **eigentliche Wort**, auf das wir uns konzentrieren (Abfrage)
 - Die **anderen Wörter** im Satz/Absatz (Schlüssel)
 - Die **eigentliche Bedeutung** des Wortes (Wert)
- Vergleiche dann die Abfrage mit den Schlüsseln, indem du **die Ähnlichkeit berechnest** (in der Mathematik: ein Punktprodukt)
- Weitere Verarbeitung: Skalierung und Normalisierung → ergibt eine Wahrscheinlichkeitsverteilung
- *Im Wesentlichen: **Wie wichtig jedes Wort im gegebenen Kontext ist***

„Traditionelle“ Datenanalyse vs. Datenanalyse mit KI

	Datenverarbeitung	Datenmengen	Analysemodelle	Expertenabhängigkeit	Entdeckungspotential
 Ohne KI	Manuell & Zeitaufwendig	Geeignet für kleinere Datensätze	Statistische Analysemodelle	Sehr hoch	Manuell & Zeitaufwändig
 Mit KI	Automatisiert & Zeiteffizient	Skaliert mit großen und komplexen Datensätzen	Adaptiv & Selbstlernend	Deutlich reduziert	Erkennt neue Muster & Zusammenhänge

Eine Harvard-Studie ergab, dass die Qualität der Arbeit durch KI um bis zu 40 % steigen kann



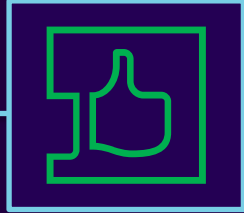
Mit dem Einsatz von KI können Menschen im Durchschnitt

- **12,2 %** mehr Aufgaben
- **25,1 % schneller**
- und mit **40% bessere Ergebnisse** erzielt

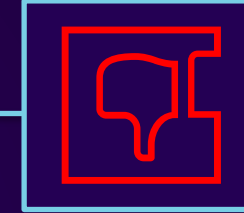
als Menschen, die keine KI nutzen.

Die Menschen, die zuvor ein **zusätzliches KI-Training** erhalten hatten, waren im Durchschnitt **sogar noch besser** und ihre Ergebnisse waren noch gleichmäßiger verteilt.

„Dos and don'ts“ bei Verwendung von KI in der Datenanalyse



- **Aufgeschlossenheit** gegenüber der neuen Technologie sein
- **Generative KI** als Möglichkeit für mehr Effizienz und erweiterte Intelligenz in der Datenanalyse
- **Teilen** von Erkenntnissen, Ideen, Erfahrungen und Use Cases im Team
- Immer wieder die Funktionalitäten **testen** – GenAI entwickelt sich exponentiell



- Den Ergebnissen der KI **niemals blind vertrauen**
- Keine **vertrauenswürdigen Daten** in ein unverschlüsseltes System eingeben (bspw. Chat GPT)
- Recherchen von GenAI **nicht als Datengrundlage** für Datensätze nehmen
- Weitere kritische Punkte: **Aktualität, Datenqualität, Verständnis, Halluzinationen, Umgang mit Unbekanntem**

4 Thesen zur Zukunft von KI in der wissen- schaftlichen Datenanalyse

1.

Demokratisierung von
Expertenwissen
„Citizen Development“

2.

Paradigmenwechsel:
Von der Datenanalyse
zum
Datenverständnis

3.

Exponentielle
Beschleunigung von
wissenschaftlichen
Papern

4.

Integration und
Interdisziplinarität
verschiedener
Wissenschafts-
gebiete

Heute haben wir den ersten Schritt gemacht – aber es liegt noch eine weitere Reise vor Euch



*Wenn ihr
Interesse an
einem
gemeinsamen
Bergaufstieg habt,
meldet euch
gerne bei mir!*

Robin Reuschel
robin@roover.eu



Robin Reuschel
robin@roover.eu

Find me on LinkedIn



(un)glücklicherweise ist KI
noch nicht in der Lage,
jede Aufgabe zu
meistern...



ROOVER

METaverse CONSULTING



www.roover.eu