

**Wissen. Vorträge. Beratung.
Orientierung. Support
für **Menschen** im *Digitalen Wandel***

Mag. Harald Russegger

Cyber-Psychologe, Unternehmensberater
Digitalisierung/KI, Uni Dozent, Coach,
Speaker, Aktivist für KI Ethik

Meine aktuellen Themen-Schwerpunkte:

- Mensch & Gesellschaft & Digitalisierung & KI
- Philosophie / Ethik der künstlichen Intelligenz
- Philosophy of Mind / Digitales Bewusstsein
- Cyber Security / KI Hacking
- Synthetische Biologie & Digitalisierung
- Menschen in AR/VR/XR Welten / Metaverse

Mehr Infos unter > www.bitdynamo.com



digital expert / consulting / training / development

harald russegger, mag.

rush@bitdynamo.com

+43(0)660 340 6325

KI (und Bildung): Mit Chancen & Risiken zu neuen Horizonten



A futuristic robot with glowing eyes and a central chest light sits at a desk in a classroom. The robot is holding a sign that says "STAY HERE". The background shows rows of desks and chairs, with a blue glow from a screen or light source.

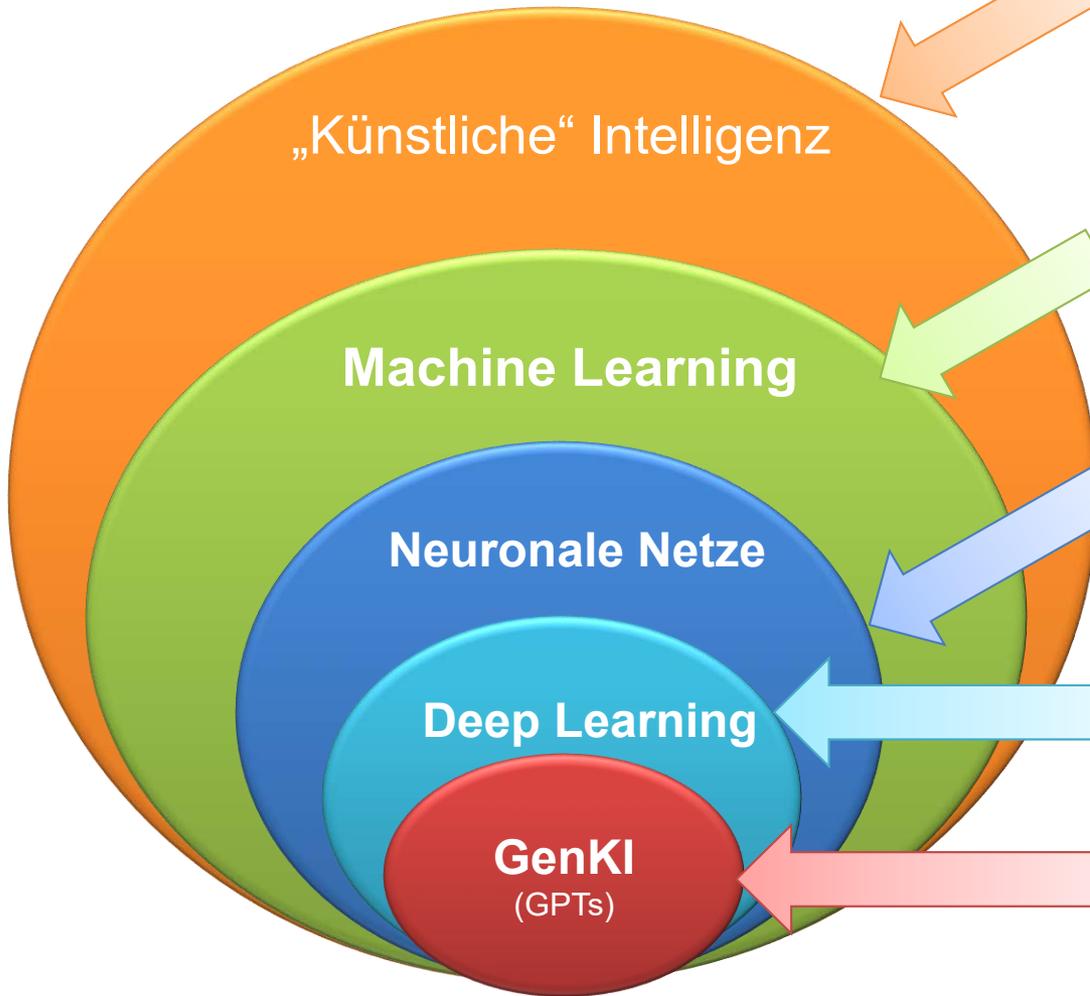
Spoiler: GenKI ist gekommen, um zu bleiben

Und nein - es gibt (noch) keine fertigen Konzepte

Tipp: das Thema sollte man **nicht aussitzen**

Worüber reden wir bei KI?

KI Schalen Modell



Vision eines **fiktiven synthetischen Systems**, das einmal in der Lage sein **könnte** eine **Alternative zur menschlichen Intelligenz darzustellen**. Hoch kontrovers. Bislang - nur Sammlung von Applikationen mit ML Fähigkeiten

Konkrete Computer-Verfahren, die **Inhalte erkennen, speichern, unterscheiden** und in **Selbst-Optimierungen bestimmte Probleme lösen können**. Sie erscheinen uns „intelligent“

Ein **statistisches Optimierungsverfahren**, das aus **vielen zuvor trainierten Beispielen einheitliche Muster extrahieren** und **daraus Schätzungen über noch nicht bekannte Beispiele abgeben** kann

Verfahren mit Neuronalen Netzen, die **auf tiefes Lernen optimiert sind**. Mit vielen hidden(deep) Layers

Spezielles Verfahren mit **NN** und **Attention Funktionalität**, das **vor-trainierte Inhalten in andere / weitere Inhalte transformieren** (umwandeln) kann, die auf **sprachlichen Wahrscheinlichkeiten** basieren. GPT = Generative Pre-trained Transformer

2017 Legendäres Paper „Attention is all you need“ von Google

Quasi die „Geburtsurkunde“ der **Generative Pretrained Transformer** **Attention und Self-Attention**

Das "Attention Is All You Need" Paper von Google aus 2017 ist wichtig, weil es das **Transformer-Modell** vorstellt. Dieses Modell nutzt den **"Attention"-Mechanismus**, der es ermöglicht, **wichtige Teile eines Textes zu erkennen und darauf zu fokussieren**. Dadurch **verbessert es die Verarbeitung von Sprache erheblich**, da es **Zusammenhänge über weite Strecken im Text parallel erfassen kann**, was mit früheren Methoden schwierig war. Dieser Fortschritt hat die Leistung von Maschinen beim „Verstehen“ und Übersetzen von Sprache deutlich verbessert.



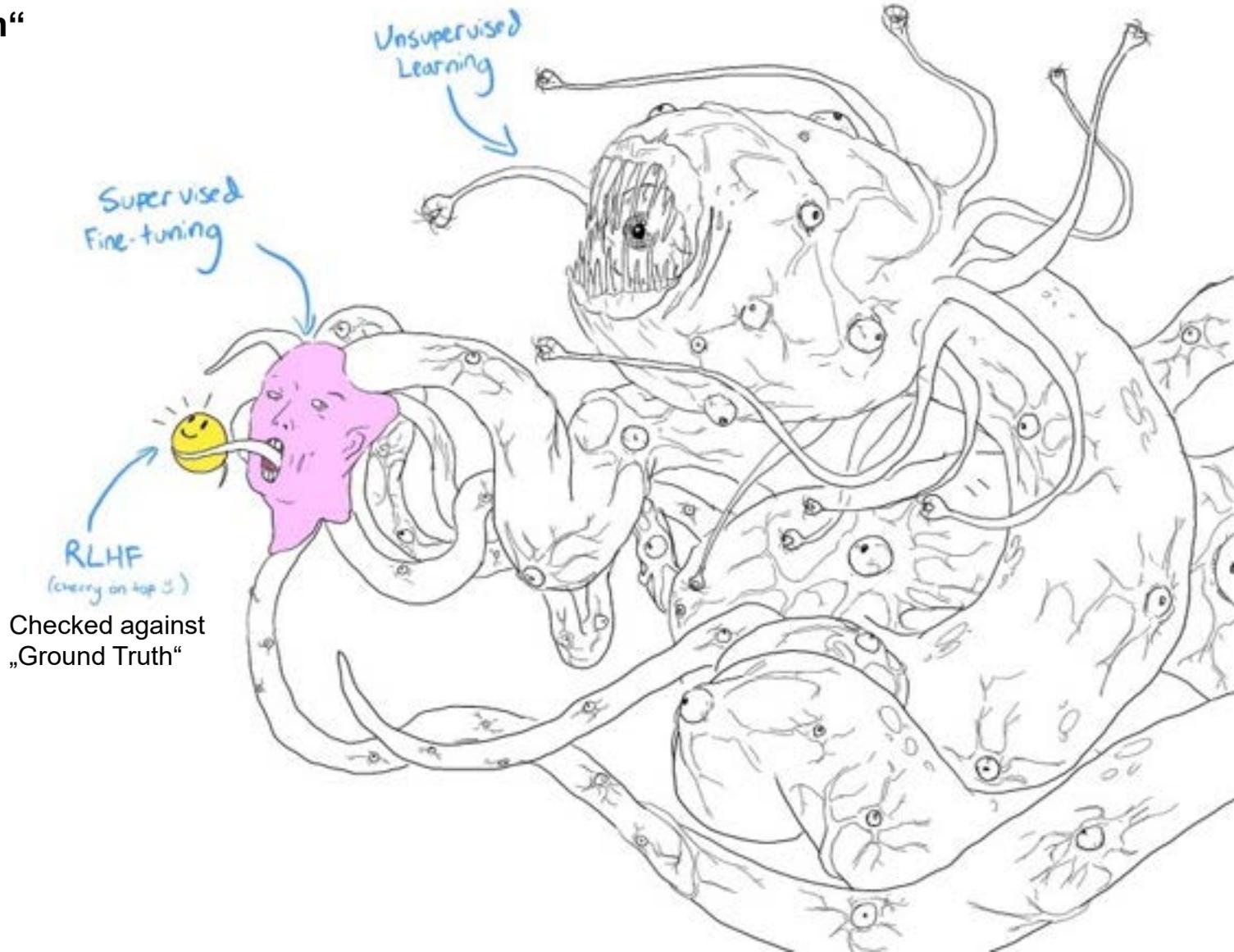
Bild-Quelle: https://twitter.com/Contrary_Res/status/1621203016903954438/photo/1

Der **Hund** springt über den **Zaun** und **er** ist dabei **glücklich**.

Transformer können den gesamten Text betrachten und erkennen Zusammenhänge deutlich besser. So auch, dass sich „er“ auf den Hund bezieht. RNNs können das nicht.

2017, 2023: die Geburt der „dicken“ GPTs

Generative Pre-Trained Transformers / LLM und sein „Wesen“





Warum sind viele (seit 2023) so scharf auf GenKI?

Weil es beliebige digitale Inhalte erzeugen, verändern, imitieren und (nun auch) selbstständig Arbeiten erledigen können wird.



Typische Anwendungsbereiche ...

Erklären / Lernen Erklär/Tutor/Antwort-Bots erklären Themen

Recherche Prompts liefern Antworten/Zusammenhänge für neue Lehr-Inhalte

Produzieren KI wandelt vorhandene Inhalte in Video, Audio etc. Beiträge

Auswerten KI analysiert / interpretiert Inhalte

Entscheiden / coachen / bewerten KI hilft eigene Gedanken zu ordnen

Übersetzen KI erstellt mehrsprachige Inhalte

Entwickeln KI programmiert, entwirft Gliederungen, hilft bei Brainstorming

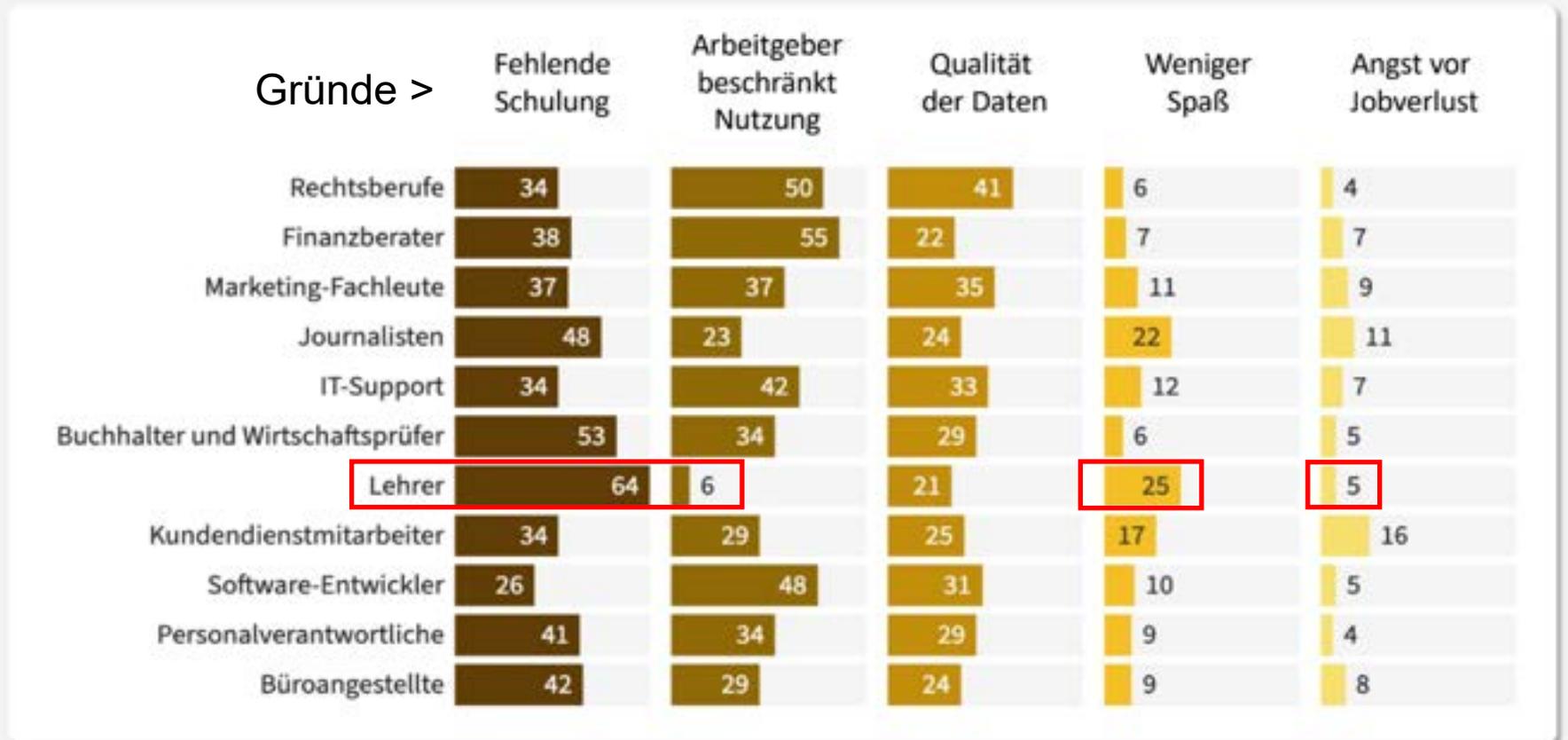
u.v.a.m

[...] Der Einsatz von KI [...] kann langwierige Arbeiten abkürzen und so Freiräume für weiterführende Gedanken eröffnen“ (Klein 2023).

Mögliche Gründe warum Menschen genKI nicht nutzen (Umfrage 2024 Dänemark)

Gründe für Nicht-Nutzung von ChatGPT

Umfrage unter 100.000
 Beschäftigten in Dänemark



Ganz pragmatische Frage ...

Wo kann ich genKI konkret bei mir verwenden?

Liste mal gedanklich auf **wie oft und wie lange** du am **Computer sitzen** musst, um:

Welche „Bullshit- Aufgaben“ hast Du?

repetitive, mühsame, langweilige, zeitraubende, anstrengende Arbeiten zu erledigen, für die man kaum (deine) spezielle Kompetenzen braucht.

Das sind Bullshit-Aufgaben.

Das soll die **genKI** machen.



Wie wird GenKI an den Akademien verwendet?

HOW POSTDOCS USE AI CHATBOTS

A little less than one-third of the postdoctoral researchers polled said that they use artificial-intelligence (AI) chatbots, such as ChatGPT, for everything from translating text to fixing code and overcoming writer's block.

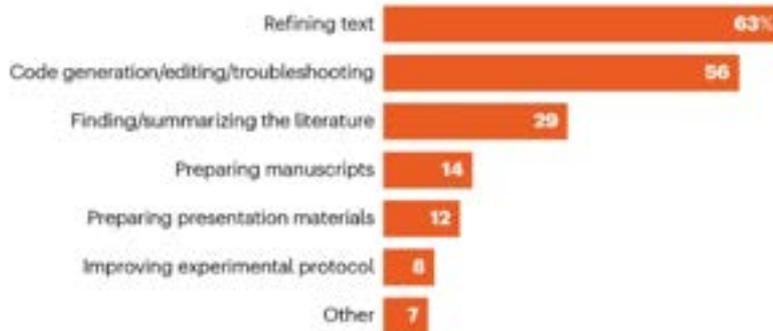
How has the evolution and rapid adoption of AI chatbots changed your day-to-day work or career plans, if at all?



Do you use AI chatbots, such as ChatGPT, in your work?



What do you use AI chatbots for?



How often do you use AI chatbots in your work?*



nature

Explore content - About the journal - Publish with us - Subscribe

nature > career feature > article

CAREER FEATURE | 16 October 2023

How ChatGPT is transforming the postdoc experience

Around one in three respondents to Nature's global postdoc survey are using AI chatbots to help to refine text, generate or edit code, wrangle the literature in their field and more.

Linda Hoodling



Note: schon etwas ältere Studie Okt/2023

- Ein **Drittel der Postdocs** in der Nature-Umfrage nutzt KI-Chatbots zur **Textverfeinerung, Codeerstellung** und **Literaturrecherche**.
- Von den befragten **Postdocs gaben 31% an**, Chatbots zu verwenden, wobei **67% keine Auswirkungen auf ihre Arbeit** sahen.
- Die Hauptanwendungsgebiete für Chatbots sind **Textverfeinerung (63%)** und **Codegenerierung (56%)**.
- Forscher schätzen die Hilfe von Chatbots bei **Routineaufgaben**, sehen aber deren **Begrenzungen in komplexen Denkprozessen**.

ChatGPT-4 im Test: Wie gut ist es im wissenschaftlichen Schreiben?

Contemporary School Psychology
<https://doi.org/10.1007/s40688-024-00500-z>

Note: aktuelle Studie April/2024



Examining the Capabilities of GPT-4 to Write an APA-Style School Psychology Paper

Adam B Lockwood¹ · Joshua Castleberry²

Accepted: 4 April 2024
© The Author(s) 2024

Abstract

Technological Advances in Artificial Intelligence (AI) have Brought forth the Potential for Models to Assist in Academic Writing. However, Concerns Regarding the Accuracy, Reliability, and Impact of AI in Academic Writing have been Raised. This Study Examined the Capabilities of GPT-4, a state-of-the-art AI Language Model, in Writing an American Psychological Association (APA)-style School Psychology Paper. Specifically, we Examined GPT-4's Ability to Provide Citations and Substantiate Claims, Offer Factually Correct Information, and Generate Accurate APA (7th Edition; 2020) Style Citations and Referencing. The Results Reveal Limitations in GPT-4's Performance, Including a lack of Proper Citations and Substantiation, Errors in Citations and Referencing, and the need for External Validation of Factual Accuracy. While GPT-4 Shows Potential, Further Refinement is Necessary to Enhance its Accuracy, Reliability, and Adherence to APA (2020) Guidelines. Future Research should Focus on Improving the Capabilities of AI Models for Scientific Writing and Developing Clear Policies for Their use in Academia.

Keywords School Psychology · Large Language Models · Writing, Artificial Intelligence

Empfehlung:

**speziellere wissenschaftl. KI-Tools
und/oder spezialisierte Mega-Prompts +
RAG nutzen**

Aber hey .. Das war auch nie der Anspruch,
dass KI wissenschaftliche Arbeiten schreiben
können soll. 😊

Zitierfähigkeit

55,6 % der Sätze nicht belegt
40,5 % der notwendigen Zitationen korrekt
Viele Zitationen auf nicht existierende
Artikel oder ungenaue Infos

Korrektheit

73,8 % der Behauptungen verifizierbar
Überprüfung erfordert externe Ressourcen
und logische Schlussfolgerungen

APA-Konformität

75 % der Zitate auf existierende Quellen
50 % der existierenden Zitate fehlerfrei
**Fehler: falsche Autoren, fehlende DOIs,
falsche Links**
25 % der Zitate auf nicht existierende
Artikel

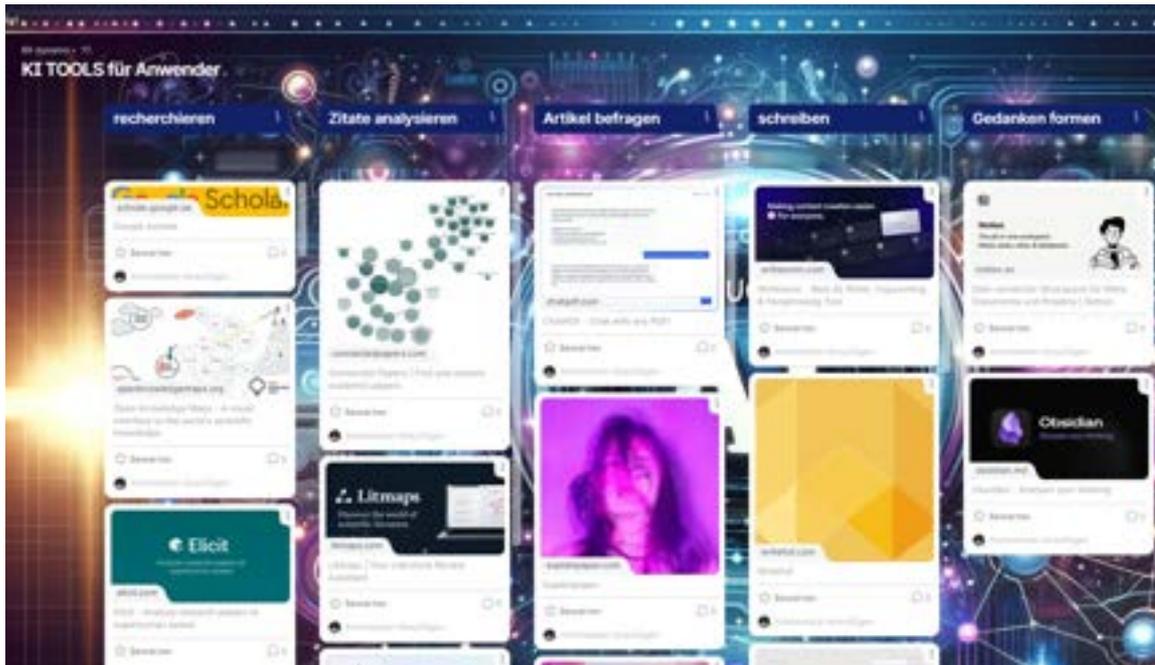
Fazit:

**ChatGPT-4 nicht geeignet für
wissenschaftliche Texte**

Oft ungenaue oder falsche Zitate

**Aber: Gut für Feedback, Ideenfindung
Zusammenfassungen**

Apropos KI Tools ... hier ist meine Sammlung



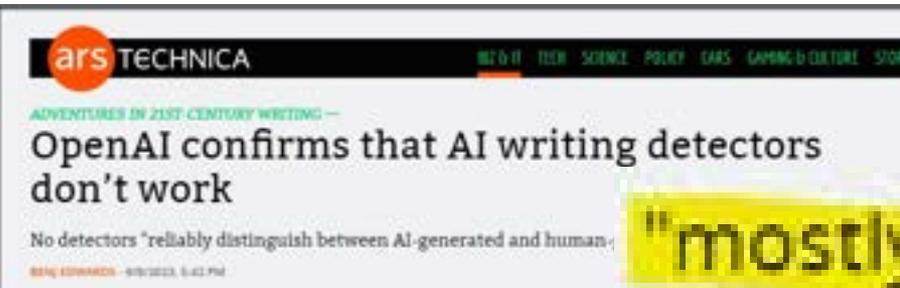
<https://padlet.com/webmaster297/kitools>

Falsche Hoffnungen ... KI Detektoren

In July, we covered in depth why AI writing detectors such as GPTZero don't work, with experts calling them "mostly snake oil." These detectors often yield false positives due to relying on unproven detection metrics. Ultimately, there is nothing special about AI-written text that always distinguishes it from human-written, and detectors can be defeated by rephrasing. That same month, OpenAI discontinued its AI Classifier, which was an experimental tool designed to detect AI-written text. It had an abysmal 26 percent accuracy rate.



"mostly snake oil."



Die Studie zeigt eine **signifikante Voreingenommenheit von GPT-Detektoren gegenüber nicht-nativen englischsprachigen Autoren. Mehr als die Hälfte der nicht-nativen Texte wurde fälschlicherweise als KI-generiert eingestuft.**

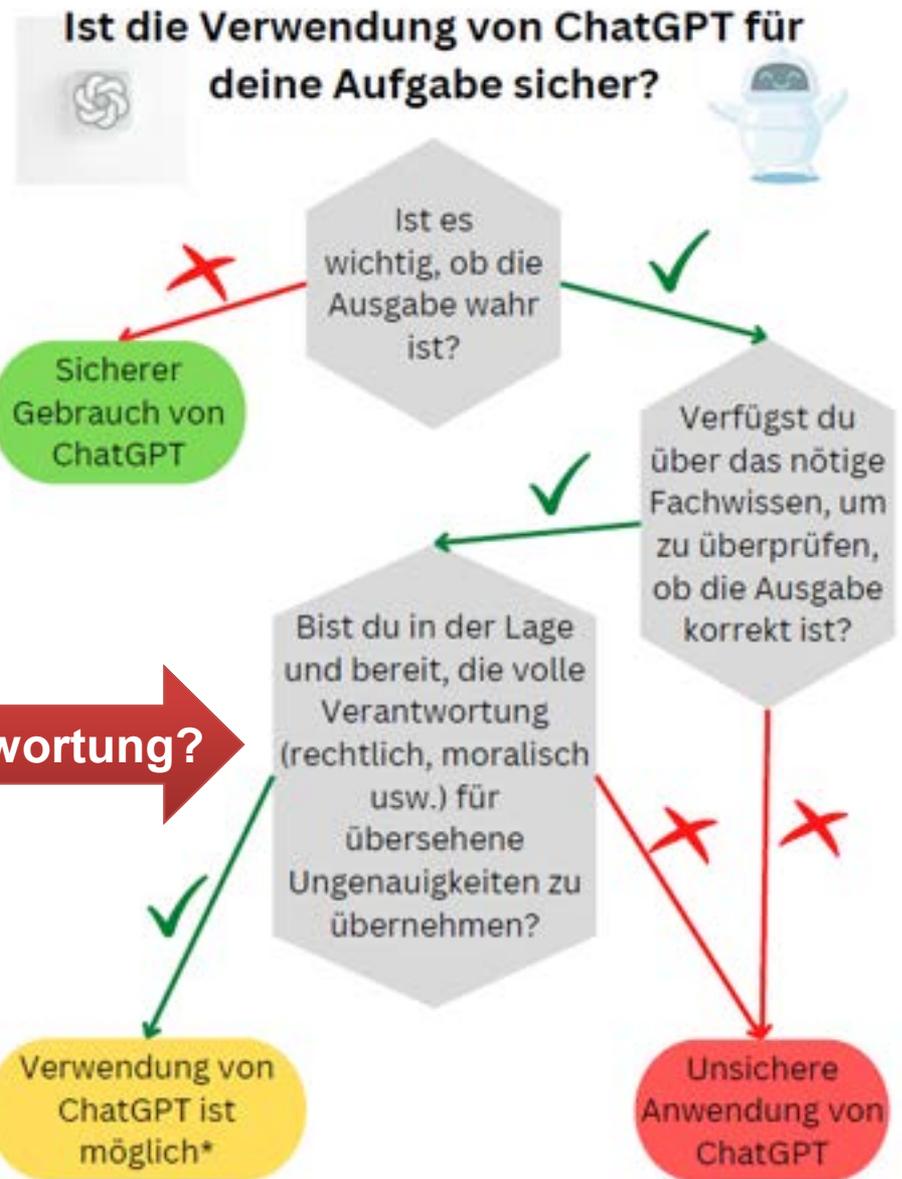
Problem:
Lehrende (ohne KI Erfahrung) könnten auf Basis von False Positives benoten

Thema: Wie mit GenKI umgehen?

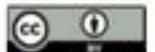
Moralischer GenKI Kompass

Bereit für volle Verantwortung?

Allerdings war das auch vor GenKI so...
Nur die Geschwindigkeit hat sich mit GenKI enorm erhöht!



* Überprüfe dennoch jedes ausgegebene Wort und jeden Satz auf inhaltliche und sprachliche Richtigkeit und gesunden Menschenverstand.



Chancen: generative KI (GenKI) im selbstgesteuerte Lernen (SRL)

Personalisierung

Anpassbare Lernpfade: GenKI kann Lerninhalte individuell anpassen, um **Engagement und Motivation zu fördern**.

Content-Generierung: KI kann relevante Lernmaterialien erstellen, die auf die **Bedürfnisse der Lernenden zugeschnitten** sind.

Unterstützung und Feedback

Sofortiges Feedback: GenKI bietet **Echtzeit-Feedback**, das den Lernprozess verbessert.

Lernanalysen: KI analysiert Daten, um **personalisierte Unterstützung bereitzustellen**.

Emotionales Wohlbefinden

Emotionserkennung: KI erkennt **emotionale Zustände und reagiert entsprechend**, z.B. durch Pausenvorschläge.

Stressbewältigung: GenKI integriert **Techniken zur Stressreduktion in den Lernprozess**.



Ethische Überlegungen

Datenschutz: Schutz der Lernendendaten ist entscheidend.

Fairness: Bias in KI-Modellen muss minimiert werden.

Praktische Anwendungen

KI-Tutoren: Virtuelle Tutoren unterstützen beim Lernen.

Kollaboratives Lernen: KI fördert Gemeinschaftsgefühl und Zusammenarbeit.

Thema: KI Tools Lernen > KI Tutor Bots selbst bauen > Botpress.cloud

Ausgabe auf Telegram

botpress.cloud



Khanmigo – KI-Tutor (US only)

GPT-4 + finetuning + prompting



Hilft ein KI-Tutor?

unklar, aber er motiviert und ist niederschwelliger

Implementing Learning Principles with a Personal AI Tutor: A Case Study

.. und ist 24/7 verfügbar

Ambroise Baillifard^{1†}, Maxime Gabella^{2†},
Pamela Banta Lavenex³, Corinna S. Martarelli³

¹EDUDL+, UniDistance Suisse, Brig, Switzerland.

²MAGMA Learning, Lausanne, Switzerland.

³Faculty of Psychology, UniDistance Suisse, Brig, Switzerland.

ambroise.baillifard@unidistance.ch, maxime@magmalearning.com

[†]These authors contributed equally to this work.

Abstract

Effective learning strategies based on principles like personalization, retrieval practice, and spaced repetition are often challenging to implement due to practical constraints. Here we explore the integration of AI tutors to complement learning programs in accordance with learning sciences. A semester-long study was conducted at UniDistance Suisse, where an AI tutor app was provided to psychology students taking a neuroscience course (N=51). After automatically generating microlearning questions from existing course materials using GPT-3, the AI tutor developed a dynamic neural-network model of each student's grasp of key concepts. This enabled the implementation of distributed retrieval practice, personalized to each student's individual level and abilities. The results indicate that students who actively engaged with the AI tutor achieved significantly higher grades. Moreover, active engagement led to an average improvement of up to 15 percentile points compared to a parallel course without AI tutor. Additionally, the grasp strongly correlated with the exam grade, thus validating the relevance of neural-network predictions. This research demonstrates the ability of personal AI tutors to model human learning processes and effectively enhance academic performance. By integrating AI tutors into their programs, educators can offer students personalized learning experiences grounded in the principles of learning sciences, thereby addressing the challenges associated with implementing effective learning strategies. These findings contribute to the growing body of knowledge on the transformative potential of AI in education.

Keywords: Artificial Intelligence, Machine Learning, Learning Sciences, AI tutor, Personalization, Retrieval Practice, Spaced Repetition

TL;DR

Der KI-Tutor **beeinflusst die akademische Leistung positiv** durch sinnvolles Modellieren menschlicher Lernprozesse.

ABER:

Es kann **keine kausale Beziehung zwischen App-Nutzung und Noten hergestellt** werden, da auch die Motivation der Studierenden eine Rolle spielt.

JEDOCH:

Scheint der KI-Tutor immerhin Lernen **spannender, interessanter und gezielter** zu machen #motivation #edutainment

arXiv:2309.13060v1 [cs.CY] 10 Sep 2023



Der nächste Hype > KI multi-agent Systeme

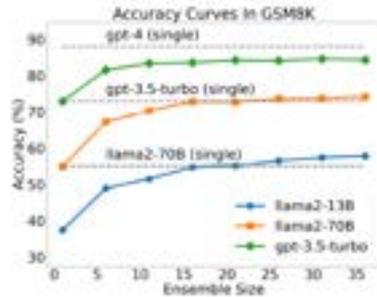
Von Prompt Engineering zu Flow Engineering

More Agents Is All You Need

Junyou Li¹ Qin Zhang² Yangbin Yu¹ Qiang Fu³ Debeng Ye¹

Abstract

We find that, simply via a sampling-and-voting method, the performance of large language models (LLMs) scales with the number of agents instantiated. Also, this method is orthogonal to existing complicated methods to further enhance LLMs, while the degree of enhancement is correlated to the task difficulty. We conduct comprehensive experiments on a wide range of LLM benchmarks to verify the presence of our finding, and to study the properties that can facilitate its occurrence. Our code is publicly available at: <https://github.com/LLM-Ensemble>.



<https://agentgpt.reworkd.ai/de>



flowiseai.com/



crewai.com

<https://chatdev.toscl.com/introduce/>



Multi-Agent Conversation Framework

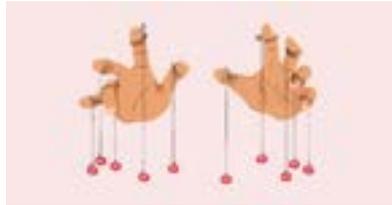
AutoGen provides multi-agent conversation framework as a high-level abstraction. With this framework, one can conveniently build LLM workflows.

<https://microsoft.github.io/autogen/>

Was soll bei KIs schon schiefgehen?



Unerwartete / unerwünschte Ergebnisse



Manipulation



Unfairness



Mangelnde Verantwortlichkeit

Automatisierte, verzerrte **Fehl-Entscheidungen** basierend auf **Vorurteilen, Rassismus und Diskriminierung** (bereits in den Trainings-Daten. Training-Bias)



Dominante proprietäre KIs > Ersatzöffentlichkeit und Hypernormalisierung?



Verletzung Datenschutzes / Privatsphäre



Mangelnde Akzeptanz für KI



Ein Hauptproblem liegt hier ...

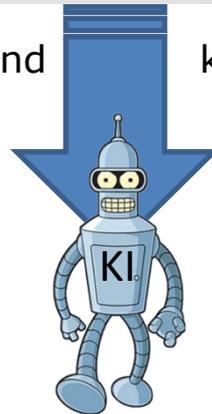


delegieren zunehmend

komplexe Entscheidungen



R.I.P.



Werte?

Verantwortung/Haftung?

Regeln?

wieviel human
Involvement?

Denn: KI-Fortschritt & Pragmatismus > vollendete Tatsachen > Haftung

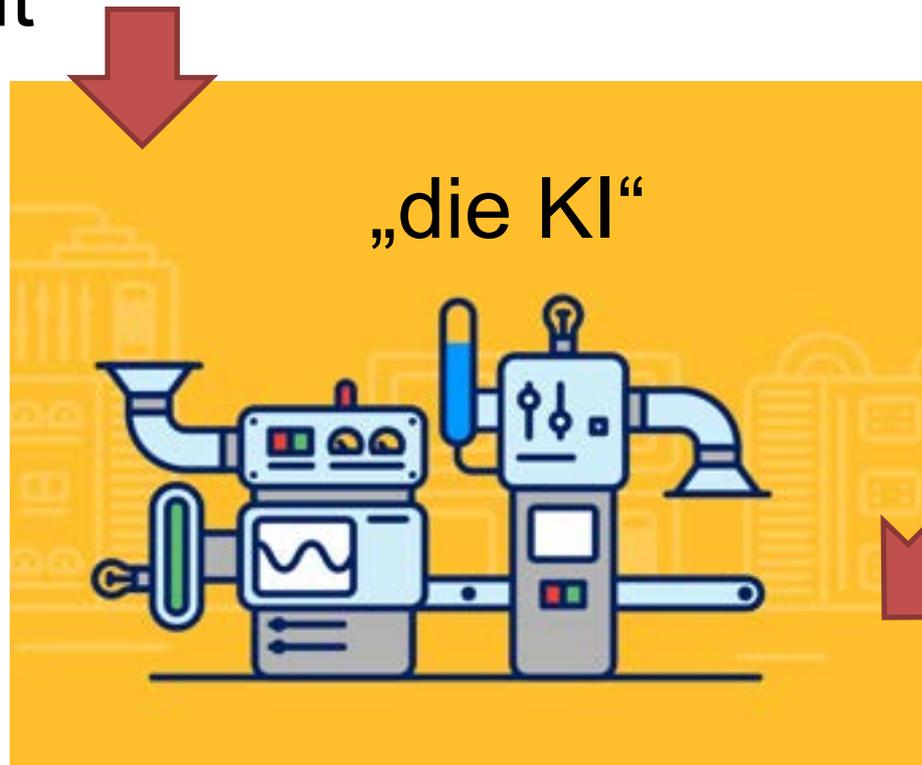
Wie soll die KI entscheiden?

Trolley-Dilemma
altbekanntes Problem
Moralphilosophie

Wie soll KI - statt uns - zukünftig entscheiden?

KI soll/muss **transparent** werden, weil ...

Input

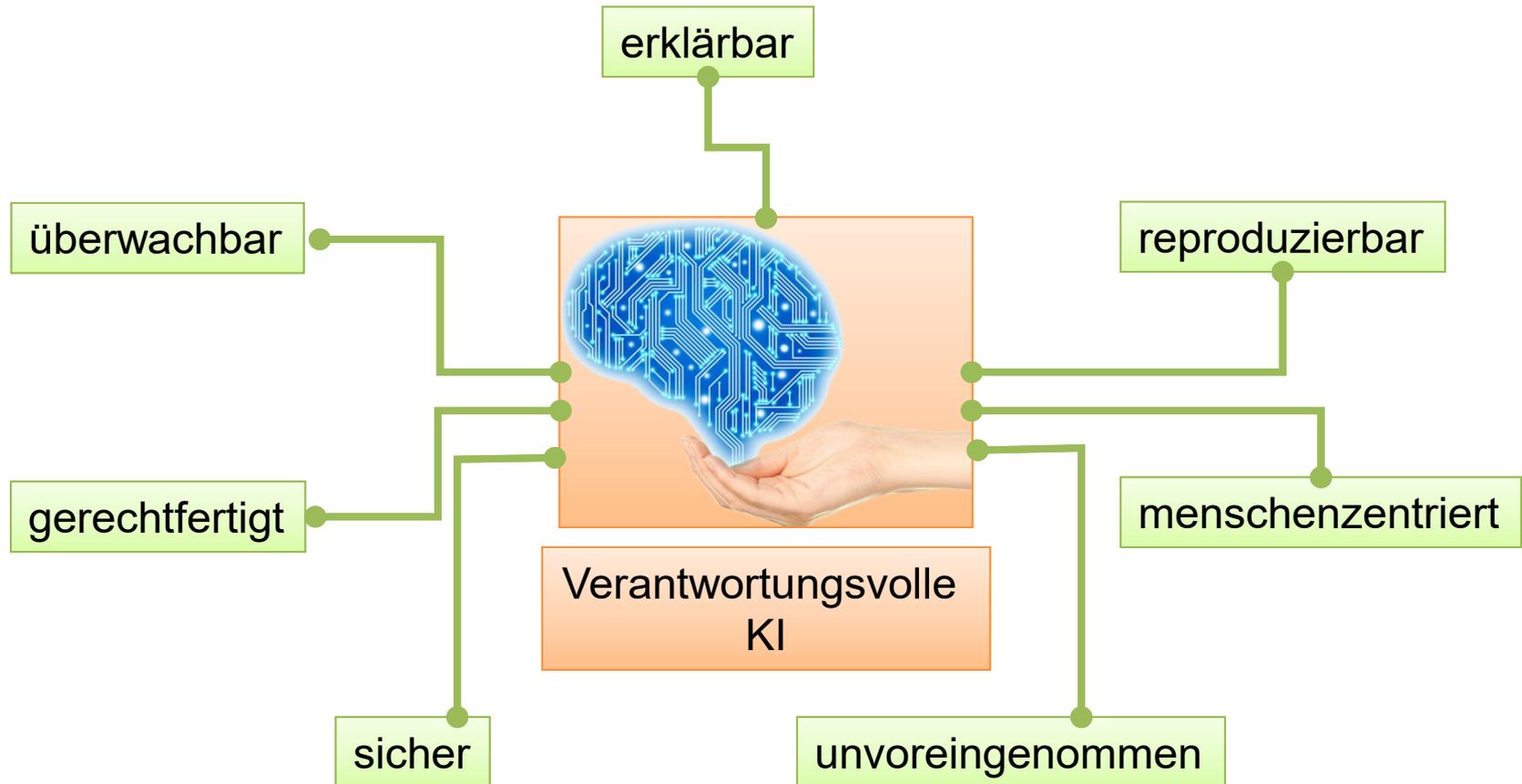


Output

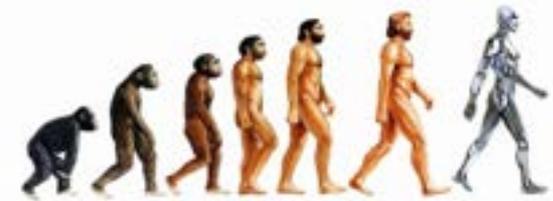
„KI ist nur **transparent/fair** wenn Datenverwendung, Entscheidungsprozesse und Funktionsweise von KI-Systemen für Menschen **leicht verständlich und nachvollziehbar** sind.“



Komponenten verantwortungsvoller / transparenter KI Systeme



Jedoch – wie realistisch ist diese Vision?

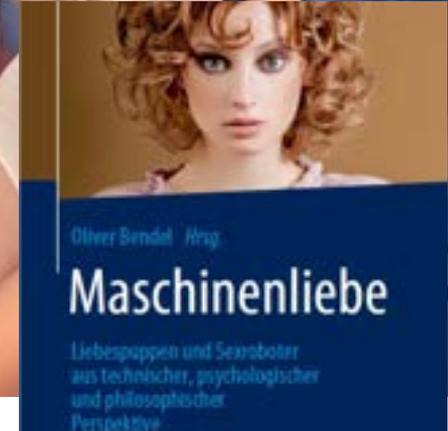


Maschinen Ethik & Roboter-Moral

Digitale Surrogate > Real Life Sex-bots / Life-companion



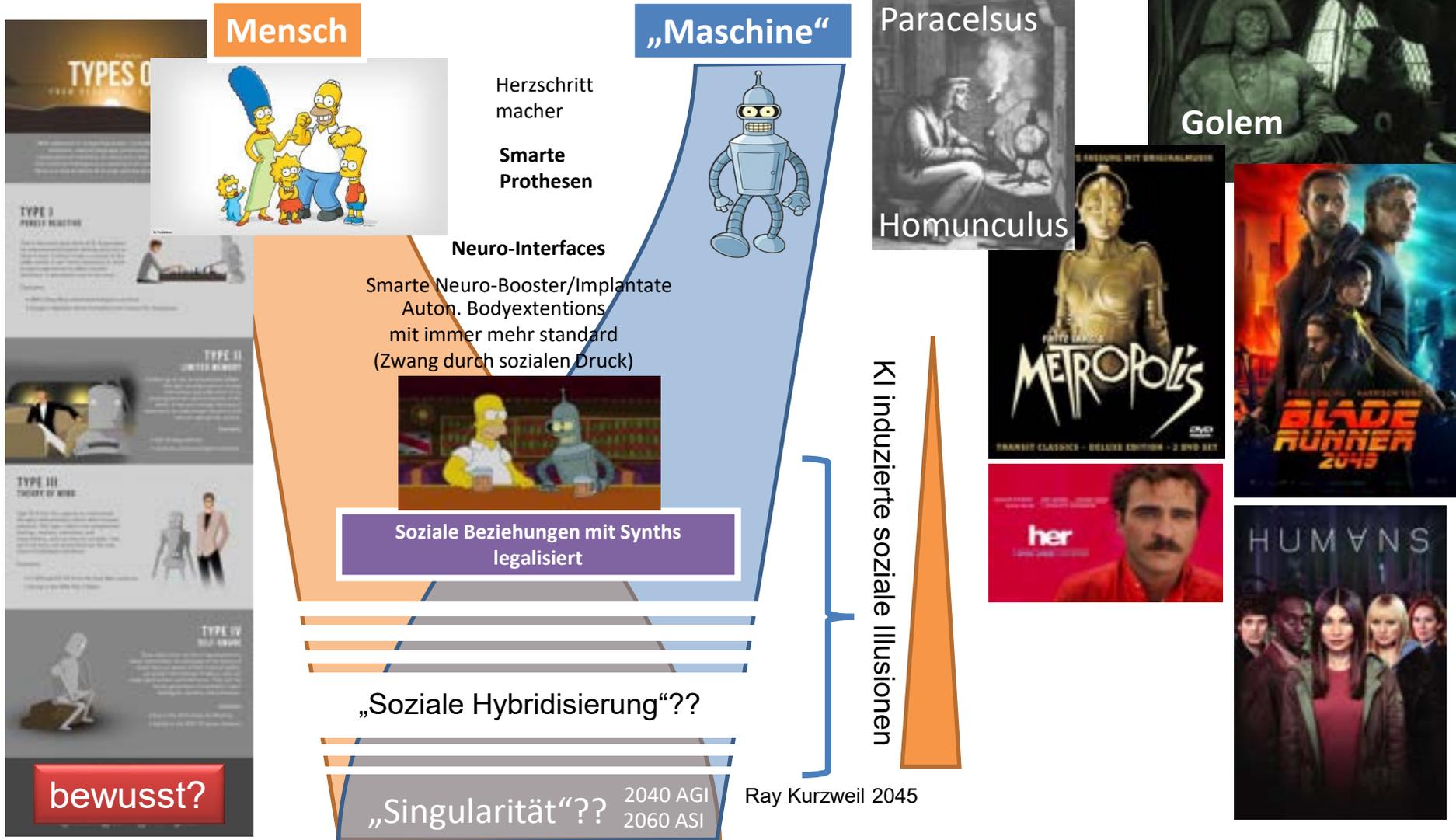
Wie verändert sich Ethik/Moral wegen KI Beziehungen?



Soziale Halluzination: Interessantes Begleit- Phänomen der Transformation
Werden Menschen - Maschinen Beziehungen „sexy“? #socialIllusion

Konvergenz – Mensch + Maschine ??

In Sci-fi längst erdacht



Die große Frage: Wie sich der **Artificial General Intelligence** nähern?

So viel Rechenpower / Daten dagegen
werfen bis von **selbst ein Wunder**
(Emergenz) passiert
(Wunsch-Traum der eliminativen Materialisten
siehe D. Dennett / P. Churchland)

Irgendein geniales KI Team findet genau
den richtigen Mix aus
„göttlichen“ ML Konzepten
(Wunsch-Traum der Funktionalisten)

Das klingt derzeit nicht nach
einer **konkreten Theorie, sondern**
eher nach **Wunschdenken**

Erstmal ist AGI anscheinend nur ein lukrative
Investment Bubble (für Wenige)



Quo vadis – KI ? Abenteuer ins Unbekannte

Als **Abenteurer** (lateinisch advenire ‚Ankommen‘ und adventus ‚Ankunft‘[1]; mittelhochdeutsch: Âventiure) wird eine **risikohaltige Unternehmung** wie eine gefahrenträchtige Reise oder die Erforschung eines unbekanntes Gebiets bezeichnet, die aus dem **geschützten Alltagsbereich entfernen**



LAND OF AGI.

FROM HERE ARE DRAGONS



Danke für Ihre Aufmerksamkeit!



.... weitere spannende Vorträge auf meiner Website!

> bitdynamo.com

Meine kuratierte Bookmarksammlungen sind hier >

<https://raindrop.io/BITDYNAMO/fokusthemen-35672131>



digital expert / consulting / training / development

harald russegger, mag.

rush@bitdynamo.com

+43(0)660 340 6325