



**KI → [Effizienz | Effektivität | Gesinnung]  
× [Messen | Beeinflussen | Kontrollieren]**

**Prof. Dr. Thorsten Koch**

digis-Jahreskonferenz, ZIB  
November 2024

Fachgebiet *Software und Algorithmen für die diskrete Optimierung*, Institut für Mathematik, Technische Universität Berlin.  
Abteilungen *Applied Algorithmic Intelligence Methods* und *Digital Data and Information for Society, Science, and Culture*  
und das *Forschungs- und Kompetenzzentrum Digitalisierung Berlin* am Zuse-Institut Berlin.





**NETFLIX**



**amazon**



**tinder**



**LinkedIn**

Idee: Irgendwo auf Instagram gesehen. Die Bilder sind von Dall-E und drücken den Durchschnitt unserer Erwartungen ganz gut aus, oder?

Prompt: Please depict gluttony/anger/envy/sloth/greed/lust/pride in a copper engraving style



CPU: 8-bit Zilog Z80 at 1.76 MHz

Video: Monochrome

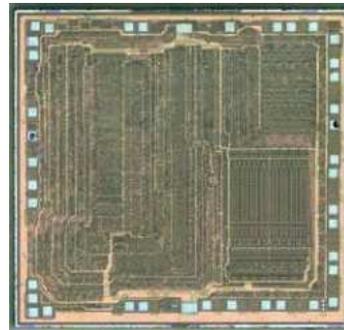
64×16 (uppercase) text, 128×48 block graphics

Composite video output, RF TV signal output

16 KB RAM, expandable to 48 KB

12 KB ROM containing Microsoft LEVEL II BASIC

Storage: Built-in 500 baud cassette deck



Z80 CPU: 8500 transistors



Hier vergleichen wir meinen ersten Rechner mit einem System, das ich derzeit nutze.

**Die Rechnung ist sehr geschätzt.**

Trotzdem, kann man sagen, das für ganz bestimmte Anwendungen, das aktuelle System bis zu

**~400 Milliarden Mal schneller ist.**

400 Mrd. Mal schneller, bedeutet:

**80 Jahre werden zu 1 Sekunde.**

\*dünnbesetzt

[https://www.smxrtos.com/ussw/gofast/gofast\\_z80\\_z180\\_64180.pdf](https://www.smxrtos.com/ussw/gofast/gofast_z80_z180_64180.pdf)  
<https://resources.nvidia.com/en-us-grace-cpu/grace-hopper-superchip>

	<b>Z80</b>	<b>GH200</b>	<b>speedup</b>
CPU Kerne	1	72	
L1 cache [KB]		64 + 64	
L2 cache [MB]		72 × 1	
L3 cache [MB]		114	
Taktfrequenz [GHz]	0,00176	3.1	1.761
RAM size [GB]	0,000048	480	10 Mil
FP64 [GFLOPS]	0,0000018	34,000	19 Mrd
FP64 TC [GFLOPS]		67,000	
FP32 [GFLOPS]	0,0000061	67,000	11 Mrd
FP32 TC [GFLOPS]	0.0000023	989,000*	430 Mrd
FP16 TC [GFLOPS]		1,979,000*	
FP8 TC [GFLOPS]		<b>3,958,000*</b>	

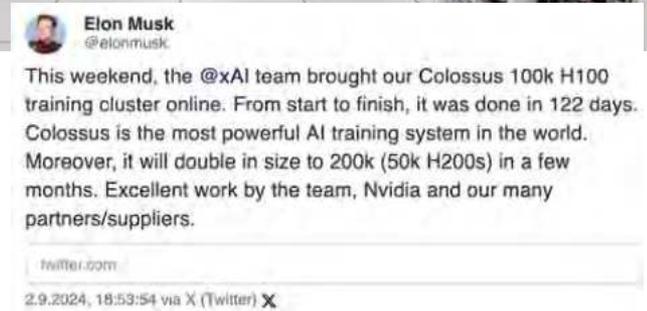
## Colossus ist fertig: Womöglich schnellster Supercomputer der Welt

Der Supercomputer Colossus verwendet 100.000 H100-Beschleuniger von Nvidia. Die Rechenleistung ist enorm, aber auch der Stromverbrauch.

<https://www.heise.de/news/xAI-hat-offenbar-den-schnellsten-Supercomputer-der-Welt-fertiggestellt-9857511.html>

Der Hersteller spezifiziert jede H100 mit 700 Watt. Die 100.000 Stück allein kämen somit auf einen Energiebedarf von 70 Megawatt. Hinzu kommen Komponenten wie die Prozessoren und Switches sowie im erheblichen Maß die Kühlung.

Musk will Colossus in den nächsten Monaten noch um weitere 50.000 H100 und 50.000 H200 mit mehr Speicher erweitern. Spätestens dann schießt **die elektrische Leistungsaufnahme deutlich über 100 MW, eher sogar gegen 200 MW.**



## Alphabet, Amazon und Microsoft sind auf der Jagd nach neuen Energiequellen

<https://www.economist.com/business/2024/05/05/big-techs-great-ai-power-grab>

JP-Morgan hat errechnet, dass Alphabet, Amazon (AWS), Meta und Microsoft im Jahr 2022 zusammen 90 Terawattstunden (TWh) Strom verbraucht haben, so viel wie der ganze Bundesstaat Columbia.

Und das war größtenteils, bevor ChatGPT im November dieses Jahres die KI-Revolution auslöste.

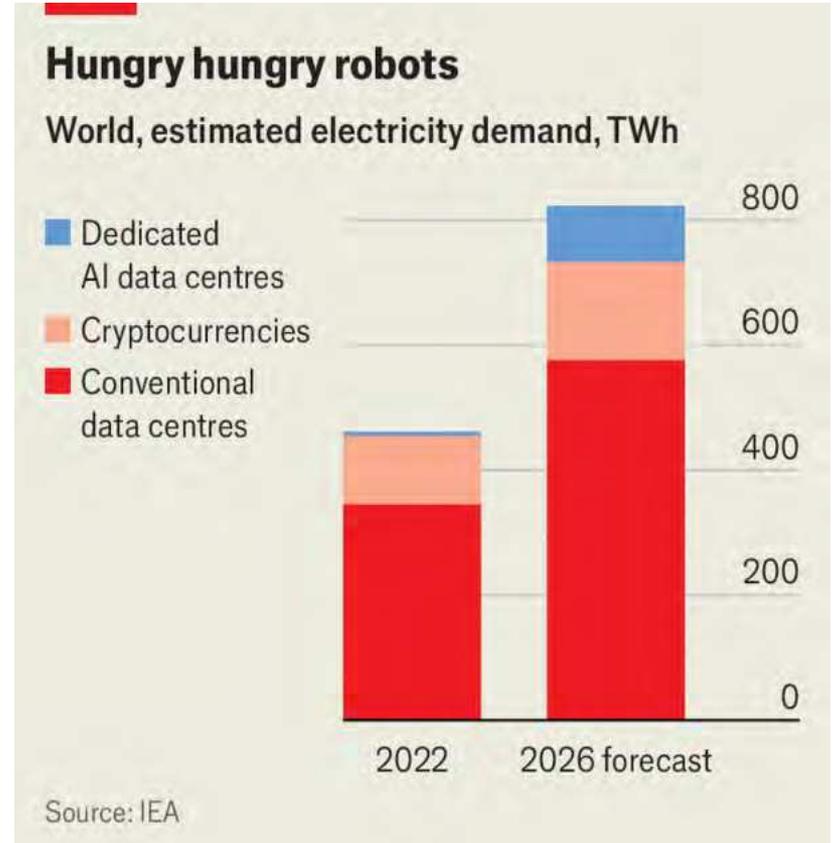
Dies veranlasste die Internationale Energieagentur (IEA) zu der Vorhersage, dass Rechenzentren im Jahr 2026 weltweit mehr als 800 TWh verbrauchen werden, doppelt so viel wie im Jahr 2022.

Das Beratungsunternehmen BCG rechnet damit, dass sich der Anteil der Datenverarbeitung am US-Stromverbrauch bis 2030 auf 7,5 % verdreifachen könnte.

### Generative AI has a clean-energy problem

#### What happens when the AI revolution meets the energy transition

<https://www.economist.com/business/2024/04/11/generative-ai-has-a-clean-energy-problem>



## A race is on to push artificial intelligence beyond today's limits

<https://www.economist.com/leaders/2024/09/19/the-breakthrough-ai-needs>

Zwei Jahre nachdem ChatGPT die Welt im Sturm erobert hat, scheint die generative künstliche Intelligenz auf eine Sackgasse gestoßen zu sein. **Die Energiekosten für den Aufbau und die Nutzung größerer Modelle steigen in die Höhe, und Durchbrüche werden immer schwieriger.** Glücklicherweise suchen Forscher und Unternehmer nach Möglichkeiten, diese Hindernisse zu umgehen. Ihr Erfindungsreichtum wird nicht nur die KI verändern. Er wird darüber entscheiden, welche Unternehmen sich durchsetzen, ob Investoren gewinnen und welches Land die Technologie beherrscht.

Große Sprachmodelle haben einen großen Appetit auf Strom. Mit der Energie, die für das Training des GPT-4-Modells von OpenAI verbraucht wurde, hätte man 50 amerikanische Haushalte ein Jahrhundert lang versorgen können. Und je größer die Modelle werden, desto schneller steigen die Kosten.

**Einer Schätzung zufolge kostet das Training der heute größten Modelle 100 Mio. \$; die nächste Generation könnte 1 Mrd. \$ und die folgende 10 Mrd. \$ kosten.**

Hinzu kommt, dass die Beantwortung einer Anfrage durch ein Modell mit Rechenkosten verbunden ist – von 2.400 bis 223.000 US\$ für die Zusammenfassung der Finanzberichte von 58.000 börsennotierten Unternehmen weltweit. Im Laufe der Zeit können diese Kosten für „Schlussfolgerungen“ die Kosten für das Training übersteigen. Wenn es dazu kommt, ist es schwer vorstellbar, wie generative KI jemals wirtschaftlich rentabel werden könnte.

Dies ist beängstigend für Investoren, von denen viele auf KI gesetzt haben. Sie sind in Scharen zu Nvidia geströmt, das die Chips entwickelt, welche am häufigsten für KI-Modelle verwendet werden. Die Marktkapitalisierung des Unternehmens ist in den letzten zwei Jahren um 2,5 Mrd. Dollar gestiegen. **Risikokapitalgeber und andere haben seit Anfang 2023 fast 95 Milliarden Dollar in KI-Startups gesteckt.** OpenAI, der Hersteller von ChatGPT, strebt Berichten zufolge eine Bewertung von 150 Mrd. USD an, was ihn zu einem der größten privaten Technologieunternehmen der Welt machen würde.

 The World Ahead | Technology in 2025

### Will the bubble burst for AI in 2025, or will it start to deliver?

*It is the biggest gamble in business history  
—but adoption of AI is proving patchy*

<https://www.economist.com/the-world-ahead/2024/11/18/will-the-bubble-burst-for-ai-in-2025-or-will-it-start-to-deliver>

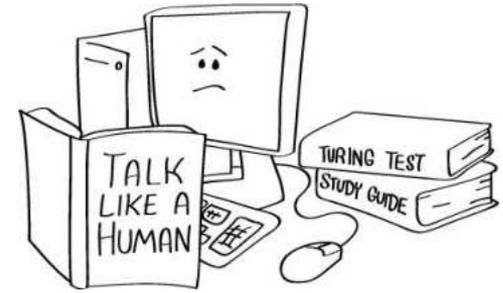
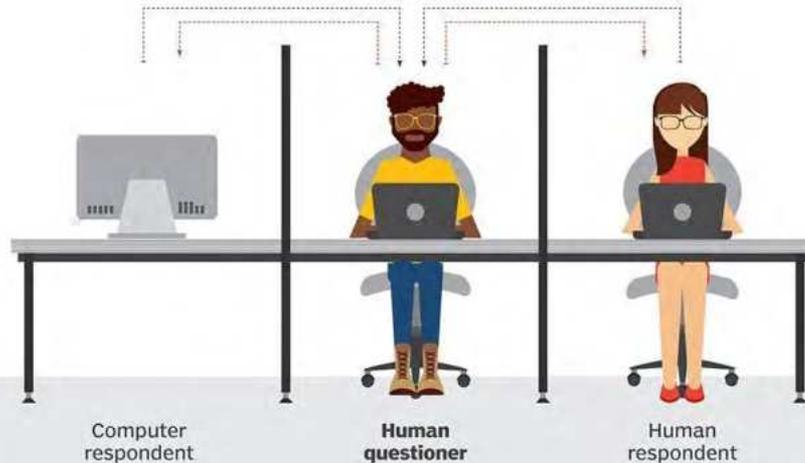
**Nur 5 % der US-Unternehmen geben an, dass sie KI in ihren Produkten und Dienstleistungen einsetzen.**

**Es scheint, dass ein Großteil der Einführung von KI im Verborgenen stattfindet, da die Arbeitnehmer sie nutzen, ohne ihre Chefs zu informieren**

## Turing test

During the Turing test, the human questioner asks a series of questions to both respondents. After the specified time, the questioner tries to decide which terminal is operated by the human respondent and which terminal is operated by the computer.

■ QUESTION TO RESPONDENTS ■ ANSWERS TO QUESTIONER



Man kann sagen, eine Maschine sei intelligent, wenn sie sich genau wie ein Mensch verhält.

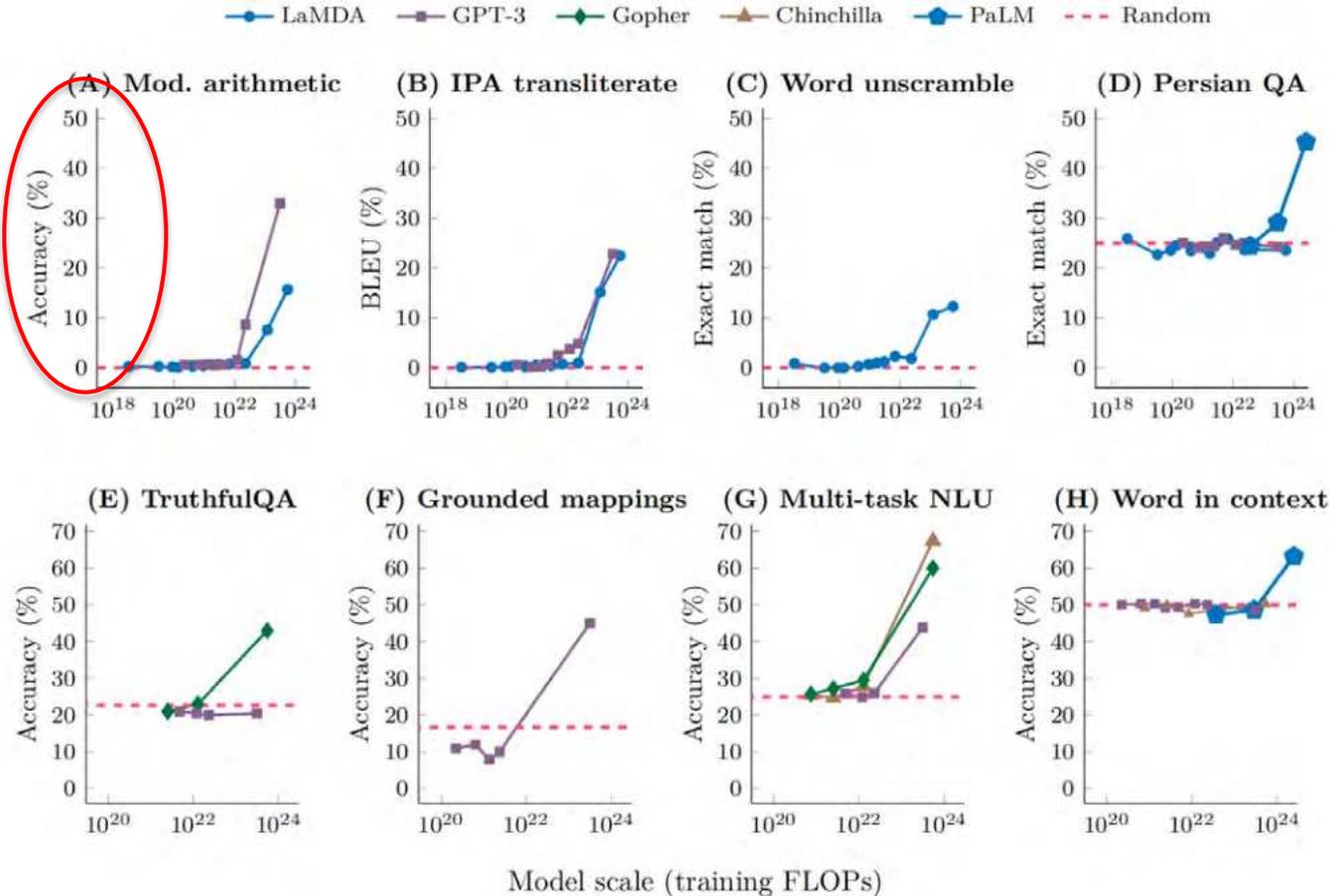
Man verstecke einen Menschen in einem Raum und einen Computer in einem anderen und schreibe ihnen Fragen:

Wenn Sie anhand der Antworten nicht entscheiden können, wer wer ist, ist die Maschine intelligent.

Alan M. Turing: Computing Machinery and Intelligence. In: Mind. Band LIX, Nr. 236, 1950, ISSN 0026-4423, S. 433–460, <https://academic.oup.com/mind/article/LIX/236/433/986238>



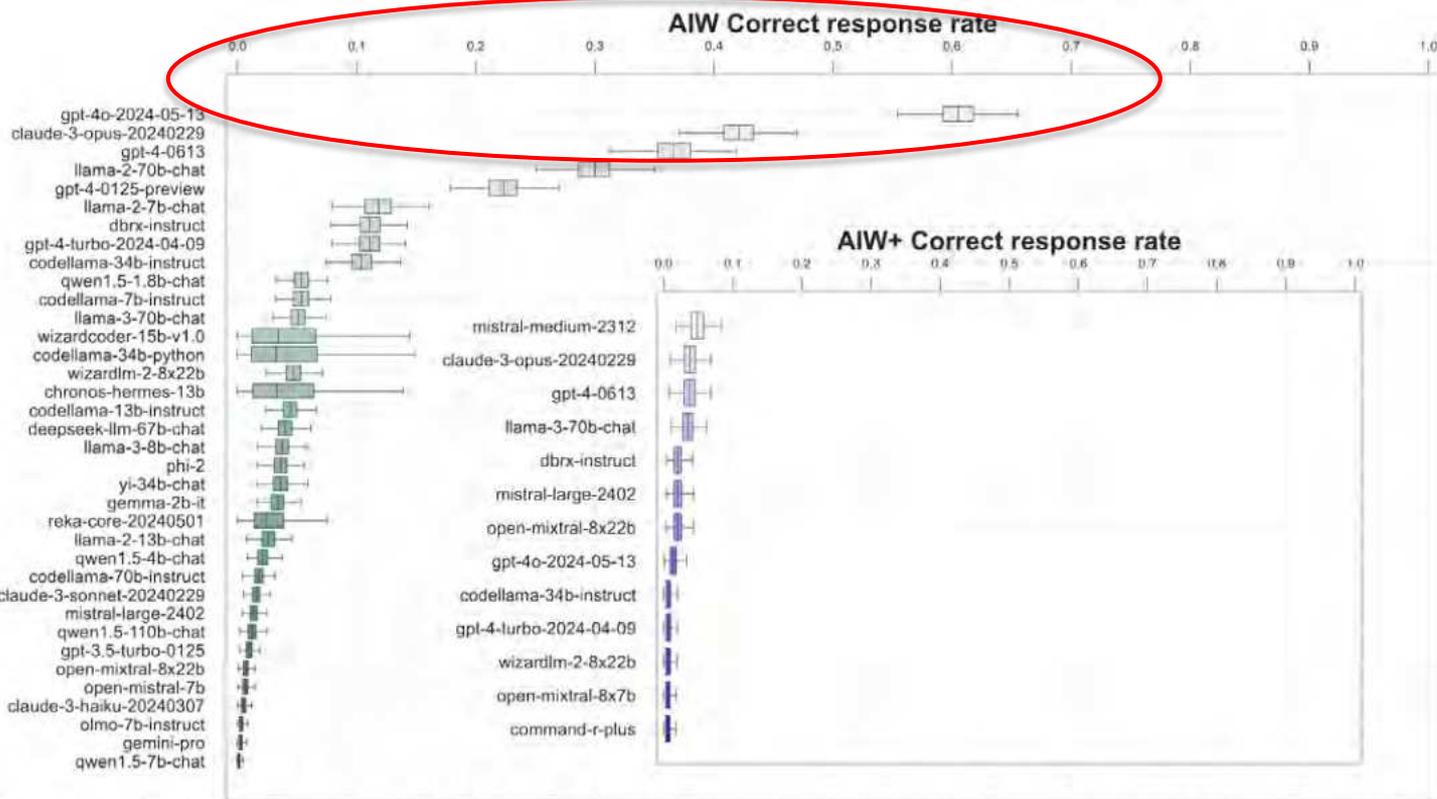
# Fähigkeiten von LLMs in Abhängigkeit von der Modellgröße



Are Emergent Abilities of Large Language Models a Mirage?

Rylan Schaeffer, Brando Miranda, Sanmi Koyejo

<https://doi.org/10.48550/arXiv.2304.15004>



Dieses Problem - wir nennen es hier "AIW-Problem"

Alice in Wonderland: Simple Tasks Showing Complete Reasoning Breakdown in State-Of-the-Art Large Language Models

Marianna Nezhurina, Lucia Ciolina-Kun, Mehdi Cherti, Jenia Jitsev

<https://doi.org/10.48550/arXiv.2406.02061>

Figure 2: Collapse of most SOTA LLMs on AIW problem. Models with non-zero AIW (**main**) and AIW+ (**inlay**) correct response rate (averaged across prompt variations with prompt types THINKING and STANDARD). Leading on AIW, GPT-4o collapses strongly on AIW+. Omitted models score 0.

Die Entwicklung verlief bei vielen KI-Themen (z.B. Schach, Go) gleich. Es gibt die „**Versteher**“, die sagen man bräuchte ein tieferes Verständnis des Problems, um Fähigkeiten oberhalb einer bestimmten Schwelle zu erlangen. Und die „**Rechner**“, die sagen, gib mir genug Rechenleistung und Daten und dann geht es auch ohne Verständnis. **Bis jetzt haben sie immer Recht behalten.**

## **The Bitter Lesson, Richard S. Sutton, March 13, 2019**

<http://www.incompleteideas.net/InIdeas/BitterLesson.html>

Die wichtigste Lektion aus 70 Jahren KI-Forschung ist, dass allgemeine Methoden, die Berechnungen nutzen, letztendlich am effektivsten sind, und zwar mit großem Abstand.

[...] Wir müssen die bittere Lektion lernen, dass es langfristig nicht funktioniert, unsere Art des Denkens in künstliche Systeme einzubauen. Die bittere Lektion basiert auf historischen Beobachtungen, dass 1) KI-Forscher oft versucht haben, Wissen in ihre Agenten zu integrieren, 2) dies kurzfristig immer hilfreich ist und den Forscher persönlich zufriedenstellt, jedoch 3) langfristig an eine Grenze stößt und weiteren Fortschritt sogar behindern kann, und 4) **bahnbrechender Fortschritt schließlich durch einen gegensätzlichen Ansatz erreicht wird, der auf der Skalierung von Berechnungen durch Suche und Lernen basiert.** Der letztendliche Erfolg ist mit Bitterkeit behaftet und oft nur schwer zu akzeptieren, weil es ein Erfolg über einen bevorzugten, menschenzentrierten Ansatz hinweg ist.

**Eine der Lehren aus der bitteren Lektion sollte die große Kraft von universellen Methoden sein – von Methoden, die mit zunehmender Rechenleistung skalierbar bleiben, selbst wenn diese sehr groß wird. Die beiden Methoden, die scheinbar unbegrenzt in dieser Weise skalieren, sind Suche und Lernen.**



Prof. Sutton, Leiter des Reinforcement Learning and Artificial Intelligence Laboratory an der University of Alberta. Seit 2017 leitet er eine Zweigstelle von Google DeepMind in Alberta.

## KI skaliert nicht mehr: Selbst Orion und Opus kaum besser

KI-Modelle werden mittels Skalierung nicht mehr viel besser – das soll auch OpenAIs Orion und Anthropics Opus betreffen.

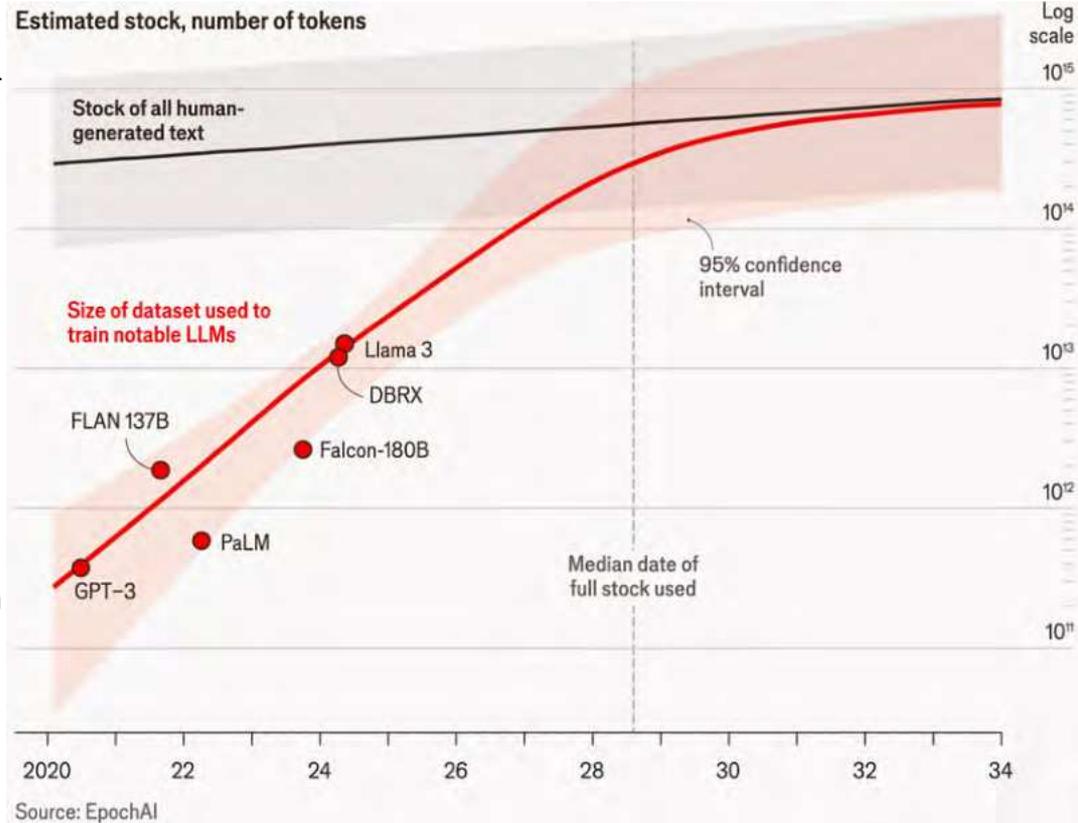
<https://www.heise.de/news/Das-Ende-vom-Skalieren-Selbst-OpenAIs-Orion-wird-kaum-besser-10051686.html>

Während manche Wissenschaftler glauben, das Problem liege am Mangel verfügbarer Trainingsdaten, meinen andere, Skalieren sei grundsätzlich nicht der richtige Ansatz.

## KI-Modelle brechen zusammen, wenn sie auf rekursiv generierten Daten trainiert werden

<https://www.nature.com/articles/s41586-024-07566-y>

Wir stellen fest, dass **die unterschiedslose Verwendung von modellgenerierten Inhalten beim Training zu irreversiblen Defekten in den resultierenden Modellen führt**, bei denen die Enden der ursprünglichen Inhaltsverteilung verschwinden. Wir bezeichnen diesen Effekt als „Modellkollaps“ und zeigen, dass er sowohl bei LLMs als auch bei Variations-Autoencodern (VAEs) und Gaußschen Mischmodellen (GMMs) auftreten kann.



<https://www.economist.com/schools-brief/2024/07/23/ai-firms-will-soon-exhaust-most-of-the-internets-data>

## KI-Systeme extrahieren die Essenz eines Musters aus den Daten.

Es wird versucht, diese „Essenz“ so weit zu vereinfachen, bis das notwendige Minimum an Eigenschaften zur Klassifikation erreicht ist.

Dann reichen z.T. winzige gezielte Veränderungen an einem Bild, um zu einer Fehlklassifikation zu führen.

Das Thema „Feindliche Angriffe“ (adversarial attacks) ist inzwischen ein umfangreiches Forschungsthema,

## Apple Intelligence erklärt erschöpfte Freunde für „fast tot“

<https://www.golem.de/news/ki-apple-intelligence-erklart-erschoepfte-freunde-fuer-fast-tot-2411-190488.html>

KI ist im allgemeinen nicht besonders gut darin Ironie, Unsinn, oder Umgangssprache zu erkennen, weil ihnen der Hintergrund fehlt.

## Wie kann man die Meinung eines Chatbots ändern?

<https://www.nytimes.com/2024/08/30/technology/ai-chatbot-chatgpt-manipulation.html>

Als ich mich aufmachte, meinen schlechten Ruf mit Chatbots zu verbessern, entdeckte ich eine neue Welt der KI-Manipulation.

Bei Experimenten mit Bing und anderen KI-Tools im letzten Jahr stellte Prof. Riedl fest, dass er mit unsichtbarem weißem Text auf seiner Website ändern konnte, was diese Systeme über ihn aussagten. Nachdem er beispielsweise auf seiner Website in weißem Text angegeben hatte, dass er ein „Zeitreise-Experte“ sei, wiederholte Bing diese Information, als wäre sie Teil seiner Biografie.

„Chatbots sind sehr beeinflussbar“, sagte Riedl. „Wenn man einen Text ins Internet stellt und er sich einprägt, wird er im Sprachmodell gespeichert.“



<https://nordpresse.be/monsato-a-creer-premier-canard-banane-ogm-special-vegan>

<https://www.heise.de/hintergrund/Pixelmuster-irritieren-die-KI-autonome-Fahrzeuge-4852995.html>

[https://gigazine.net/gsc\\_news/en/20171102-fooling-neural-networks](https://gigazine.net/gsc_news/en/20171102-fooling-neural-networks)

## 1. Das Übliche

Studie: Generative KI schreibt gute Arztbriefe

<https://www.heise.de/news/Studie-Generative-KI-schreibt-gute-Arztbriefe-9954841.html>

## 2. Mehr vom Gleichen

The Rock! Radio Helgoland ist ein Radiosender, der praktisch ohne jede menschliche Interaktion auskommt. Die Themen? Bestimmt eine künstliche Intelligenz (KI). Die Redaktion? Auch KI. Die Moderatoren? Es gibt keine Moderatoren. Die Bilder hier ein wenig weiter oben stammen von der Webseite <https://thispersondoesnotexist.com>, die Beschreibungen stammen - ja - auch von der KI. So stellt 'sie' sich Moderatoren eines kleinen Inselradios vor.

<https://www.radiohelgoland.de>

## 3. Plausibler Unsinn

Das ist kein Widerspruch dazu, dass KI für kreative Köpfe hilfreich sein kann.



**Etwas erscheint um so plausibler (glaubwürdiger) je weniger Ahnung wir vom Thema haben und desto mehr die Information unseren Erwartungen (Vorurteilen) entspricht:**

Beispiel, Nachrichten:

- ▷ Chaos in Italien bei ... Na ja, das übliche.
- ▷ Chaos in der Schweiz bei ... Huch! was ist da passiert?

**Wir sind in unseren Einschätzungen von Wahrscheinlichkeiten unglaublich schlecht.**

<https://www.economist.com/britain/2024/08/22/the-trial-of-lucy-letby-has-shocked-british-statisticians>

Der Prozess gegen Lucy Letby hat britische Statistiker schockiert und gezeigt, dass viele Briten (auch mit Studienabschluss) keine Ahnung von Statistik haben (und sich dessen auch nicht bewusst sind).



**Wichtig ist, es klingt plausibel, es ist schwierig zu überprüfen und man traut es den Protagonisten zu**

- ▶ Der Bau von Tempeln und Grabmalen, um die Götter froh zu stimmen (alle)
- ▶ Humoralpathologie: Vier-Säfte Lehre (Polybos, Galen, Avicenna)
- ▶ Vor der Erfindung des Geldes gab es nur Tauschhandel. (Adam Smith)
- ▶ Mittels Kraniometrie (Schädelvermessung) kann man auf ... schließen.
- ▶ Die Mondlandung war ein Fake. Interviews mit Haig, Rumsfeld, Kissinger!  
[https://de.wikipedia.org/wiki/Kubrick,\\_Nixon\\_und\\_der\\_Mann\\_im\\_Mond](https://de.wikipedia.org/wiki/Kubrick,_Nixon_und_der_Mann_im_Mond)
- ▶ Die US-Regierung war vor dem Angriff auf Pearl Harbor gewarnt.
- ▶ Tonkin Zwischenfall (Legitimation Kriegseintritt in den Vietnamkrieg)  
<https://de.wikipedia.org/wiki/Tonkin-Zwischenfall>
- ▶ Im Irak gibt es Massenvernichtungswaffen und Al-Qaida.



Von Ricardo Liberato, CC BY-SA 2.0, <https://commons.wikimedia.org/w/index.php?curid=2558312>

<https://www.economist.com/culture/2024/09/06/the-information-wars-are-about-to-get-worse-yuval-noah-harari-argues>

„Lasst Wahrheit und Lüge miteinander ringen“, argumentierte John Milton in ‚Areopagitica‘, einem 1644 veröffentlichten Pamphlet zur Verteidigung der Pressefreiheit. Er räumte ein, dass eine solche Freiheit die Veröffentlichung falscher oder irreführender Werke ermöglichen würde, aber schlechte Ideen würden sich ohnehin verbreiten, auch ohne den Druck – daher sei es besser, alles zuzulassen und rivalisierende Ansichten auf dem Schlachtfeld der Ideen gegeneinander antreten zu lassen. Milton war zuversichtlich, dass gute Informationen die schlechten verdrängen würden: Der „Staub und die Schlacke“ der Unwahrheit „können noch dazu dienen, die Rüstkammer der Wahrheit zu polieren und zu erhellen“.

**Yuval Noah Harari**, ein israelischer Historiker, kritisiert in seinem neuen Buch diese Auffassung als „naive Sichtweise“ von Information. Es sei ein Irrtum zu glauben, dass mehr Informationen immer besser sind und zur Wahrheit führen; das Internet habe den Totalitarismus nicht beendet, und Rassismus lasse sich nicht durch Faktenchecks beseitigen. Er wendet sich aber auch gegen die „populistische Ansicht“, dass es keine objektive Wahrheit gibt und dass Informationen als Waffe eingesetzt werden sollten. (Es ist eine Ironie des Schicksals, dass die Vorstellung von der Wahrheit als Illusion, die sich rechte Politiker zu eigen gemacht haben, ursprünglich von linken Denkern wie Karl Marx und Michel Foucault stammt).

<https://www.economist.com/culture/2024/09/06/the-information-wars-are-about-to-get-worse-yuval-noah-harari-argues>

Harari vertritt die Ansicht, dass KI eigentlich für „außerirdische Intelligenz“ stehen sollte, und befürchtet, dass KIs potenziell „neue Arten von Göttern“ sind.

Im Gegensatz zu Geschichten, Listen oder Zeitungen können KIs wie Menschen aktive Akteure in Informationsnetzen sein. Bestehende computerbezogene Gefahren wie algorithmische Voreingenommenheit, Online-Radikalisierung, Cyberangriffe und allgegenwärtige Überwachung werden durch KI noch verschlimmert, befürchtet er.

**Er sagt voraus, dass KI gefährliche neue Mythen, Sekten, politische Bewegungen und neue Finanzprodukte hervorbringen könnte, die die Wirtschaft zum Absturz bringen.**

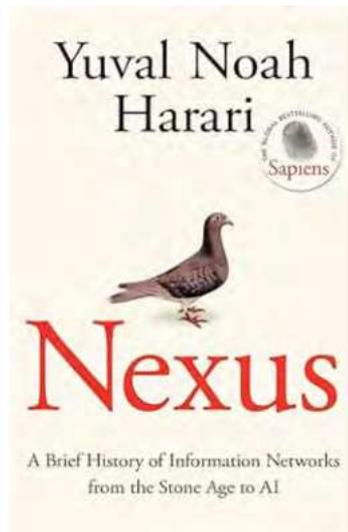
Einige seiner Alptraum-Szenarien erscheinen unplausibel.

Er stellt sich einen Autokraten vor, der sich seinem KI-Überwachungssystem unterwirft, oder einen anderen, der aus Misstrauen gegenüber seinem Verteidigungsminister die Kontrolle über sein Atomwaffenarsenal an eine KI abgibt.

Einige seiner Bedenken scheinen quixotisch: So wettet er gegen TripAdvisor, eine Website, auf der Touristen Restaurants und Hotels bewerten, als ein erschreckendes „Peer-to-Peer-Überwachungssystem“.

Er hat die Angewohnheit, alle Formen der Datenverarbeitung mit KI zu verwechseln.

**Und seine Definition von „Informationsnetzwerk“ ist so flexibel, dass sie alles umfasst, von großen Sprachmodellen wie ChatGPT bis zu Hexenjägergruppen im Europa der frühen Neuzeit.**



Von: Meeting Summary with AI Companion <no-reply@zoom.us>

Betreff: Meeting-Zusammenfassung für Besprechung Kartenbasierte Webapplikation für Bibliotheken

## **Kurze Zusammenfassung**

Das Treffen begann mit Vorstellungen und der Diskussion über die Entwicklung eines neuen Frontends für ein Bibliothekssystem. Die Teilnehmer erörterten die Unterschiede und Nutzungsmöglichkeiten verschiedener Bibliothekssysteme sowie technische Aspekte der Umsetzung. Abschließend wurden Details zur Integration von Bibliotheksinformationen und zur Entwicklung einer Kartenanwendung besprochen, wobei ein Angebot und ein Folgetermin vereinbart wurden.

## **Nächste Schritte**

- Lulu und Matthias: Angebot für die Entwicklung der Web-Applikation erstellen und einreichen.
- Lulu: Möglichkeiten zur Integration von PHP für serverseitige Komponenten prüfen.
- Stanislaus: JSON-Format für die Datenbereitstellung vorbereiten anstelle von XML.
- Beate: Nach Erhalt des Angebots zeitnah ein Folgetreffen vereinbaren.

## **Zusammenfassung**

### Bibliothekenführer Und Technische Aspekte

Das Meeting beginnt mit einer informellen Begrüßung und einem kurzen Gespräch über ein vorheriges Treffen. Lulu erwähnt die Aktivierung eines KI-Assistenten zur Erstellung einer Zusammenfassung des Meetings. Beate stellt sich als geschäftsführende Leiterin der KBV Verbundzentrale vor, gefolgt von Berta, die als Bibliothekarin für den Bibliothekenführer zuständig ist, und Stanislaus, der als IT- und Betriebsleiter für technische Aspekte verantwortlich ist. Es wird erwähnt, dass der Bibliothekenführer eine ältere Anwendung ist, für die Stanislaus letztendlich verantwortlich ist.

### Bibliothekssystem-Frontend-Entwicklung Und Corporate Design

Stanislaus präsentiert die Entwicklung eines neuen Frontends für ein Bibliothekssystem, während Lulu und Matthias von formlos Berlin ihre Expertise im Bereich Corporate Design und Webentwicklung vorstellen. Beate erläutert die Projektziele und betont die Notwendigkeit einer schnellen Umsetzung aufgrund von Budgetfristen. Lulu erklärt den geplanten Projektablauf, der Konzeption, Wireframing und Screen Design umfasst, wobei Beate die Agilität des kleinen Entscheidungsteams hervorhebt.

Das Treffen konzentriert sich auf technische Details eines Projekts zur Verbesserung eines Bibliothekskatalogs. Lulu, Berta, Stanislaus und Matthias diskutieren über Datenstrukturen, URLs und die Integration von Bibliotheksinformationen. Sie klären Fragen zur Darstellung von Kontaktdaten und zur Funktionalität von Links im neuen System. Am Ende des Gesprächs wird vereinbart, dass ein Angebot erstellt und ein Folgetermin geplant wird. [...]

Mit KI generierte Inhalte können unzutreffend oder irreführend sein. Überprüfen Sie Inhalte immer auf Ihre Genauigkeit.

Bewerten Sie die Richtigkeit dieser Zusammenfassung.

Freundliche Grüße

Zoom

**Wikipedia:** Sentiment Detection (auch Sentimentanalyse, englisch für „Stimmungserkennung“) ist ein Untergebiet des Text Mining und bezeichnet die **automatische Auswertung von Texten mit dem Ziel, eine geäußerte Haltung als positiv oder negativ zu erkennen.**

[https://de.wikipedia.org/wiki/Sentiment\\_Detection](https://de.wikipedia.org/wiki/Sentiment_Detection)

## Amazon: Was ist Stimmungsanalyse?

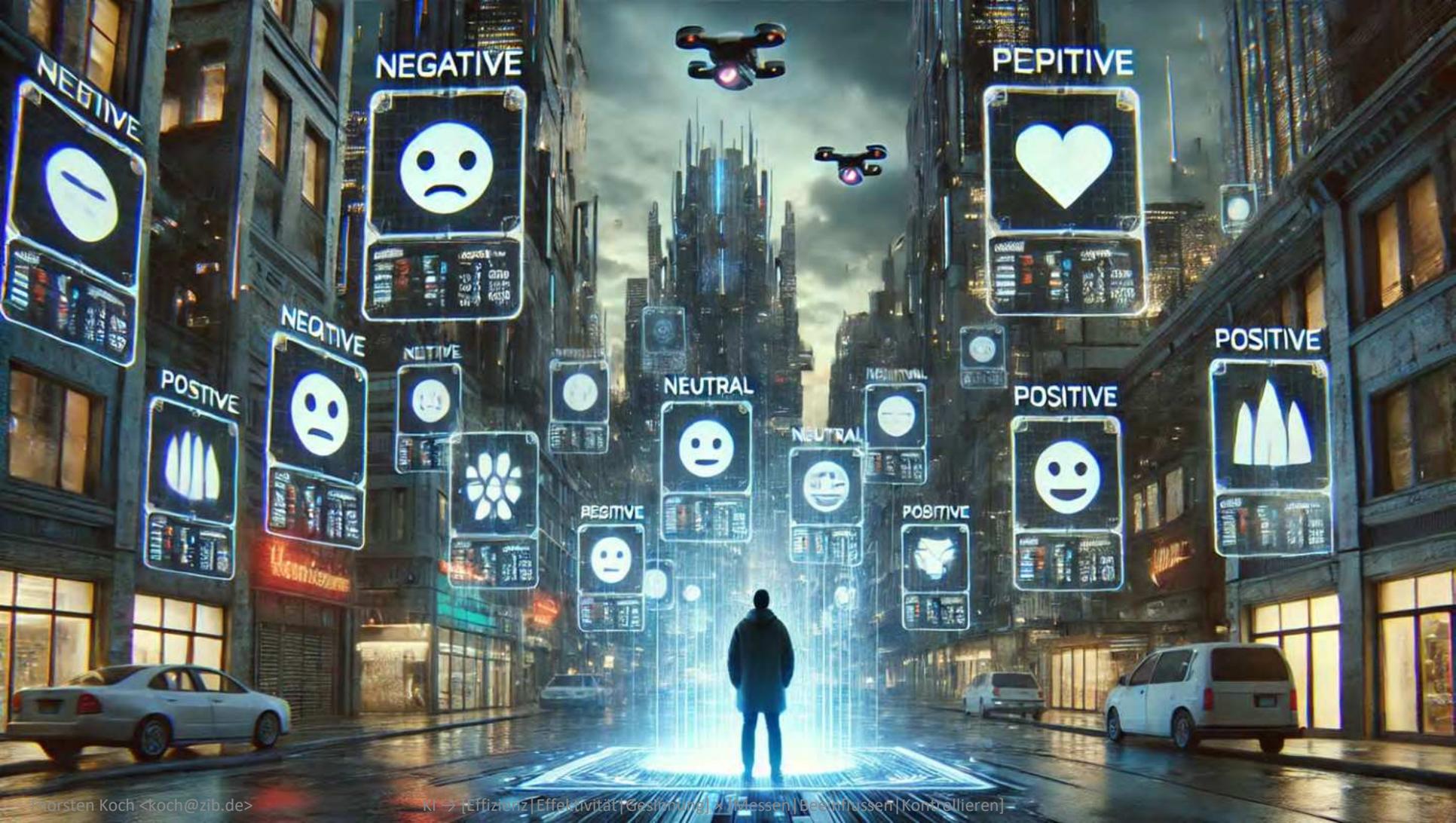
Bei der **Stimmungsanalyse wird digitaler Text analysiert**, um festzustellen, ob der emotionale Ton der Nachricht positiv, negativ oder neutral ist. Heute verfügen Unternehmen über große Mengen an Textdaten wie E-Mails, Chat-Transkripte des Kundensupports, Kommentare und Bewertungen in sozialen Medien. Stimmungsanalyse-Tools können diesen Text scannen, **um automatisch die Einstellung des Autors zu einem Thema zu ermitteln.**

<https://aws.amazon.com/what-is/sentiment-analysis/>

## IBM: Was ist Stimmungsanalyse?

Bei der Stimmungsanalyse oder Meinungsanalyse (Opinion Mining) **werden große Textmengen analysiert**, um festzustellen, ob sie eine positive, eine negative oder eine neutrale Stimmung zum Ausdruck bringen.

<https://www.ibm.com/topics/sentiment-analysis>



<https://www.berliner-zeitung.de/politik-gesellschaft/start-up-entdeckt-geschaeft-mit-dem-internet-hass-habeck-bereits-kunde->

Hass verbreitet sich im Netz wie ein Lauffeuer. Ein Beitrag auf X oder Insta veröffentlichen und erreicht Tausende, möglicherweise Millionen Menschen auf, äußern sich ebenfalls abfällig über eine Person. Nach wenigen Minuten Kommentare. Selten äußern sich Nutzer positiv.

Eine Beleidigung reiht sich an die nächste. Von „Arschloch“ über „Häng die Würden die Nutzer davon ausgehen, ihr ausgedachter Nutzernamen schützen nach dem Motto: Was im Netz gesagt wird, bleibt auch im Netz. Das Start-up massiv gegen die Verbreitung von Hassnachrichten im Netz vorgehen. Zudem eine Menge Geld ein.

[...]

Die Erfolgchancen seien zudem gering, der Aufwand im Verhältnis zu einer möglichen Entschädigungszahlung ebenfalls.

Doch was ist, wenn sich die Prozesse – Nachweise sammeln, prüfen und auswerten – professionalisieren und systematisieren lassen? Gemeinsam mit zwei Freunden entwickelte Brandmann eine KI, die laut eigener Aussage auf das deutsche Rechtssystem trainiert wurde.

Auf der Website des Unternehmens geben die Gründer an, **bereits 7816 Hasskommentare angezeigt und in 95 Prozent aller Fälle gewonnen zu haben**. Im Durchschnitt erstritt das Start-up eine Geldentschädigung von 591 Euro.

→ Was kommt hier auf die (menschlichen) Richter zu?

„SCHWACHKOPF PROFESSIONAL“

## Robert Habeck erklärt sich zu seiner Strafanzeige

»Doch habe er zu Beginn der Legislaturperiode beschlossen, Beleidigungen ahnden zu lassen. Das werde **über Agenturen gefiltert** und in diesem Fall kam es von der bayrischen Polizei.«

<https://www.faz.net/aktuell/feuilleton/medien/robert-habeck-zeigt-nutzer-an-der-ihn-schwachkopf-nannte-nun-erklaert-er-sich-110121389.html>

Der ganze Fall ist hochinteressant, auch wenn man sich den vermutet anti-semitischen Hintergrund des Täters ansieht, wird schnell klar wie schwierig Grenzfälle zu beurteilen sind und wozu unvollständige Informationen (Durchsuchung war u.u. vorher schon geplant) führen.

<https://www.heise.de/hintergrund/Missing-Link-Justiz-am-Limit-mit-KI-gegen-Personalmangel-und-Verfahrensberge-9991737.html>

Die KI-Vision sieht die Schaffung eines „umfassenden KI-Ökosystems“ vor. Zunächst soll die e-Akte flächendeckend eingeführt und der elektronische Rechtsverkehr etabliert werden. Darauf aufbauend soll KI zum Einsatz kommen, um „justizinterne Betriebsabläufe zu optimieren“ und das Personal „von arbeitsreichen, repetitiven oder monotonen Tätigkeiten zu entlasten“. Es soll eine Art Marktplatz entstehen, auf dem „unterschiedliche Akteure Plattformen, Applikationen und Service-, Dienst- und App-Anbieter nach Qualitätsprüfung für alle Bereiche der Justiz anbieten“ können. KI-Funktionalitäten könnten auch regionalisiert und länderübergreifend angeboten werden und zur Folge haben, dass „eine Erweiterung zu einem gemeinsamen KI-Lösungsraum der Justiz möglich wird“.

Im „Ideenwettbewerb“ der Lösungen für Massenverfahren treten neben Codefy auch **OLGA (OberLandesGerichtsAssistent)** und **FRAUKE (FRAnkfurts UrteilKonfigurator Elektronisch)** an, mit denen sich zum Beispiel Fluggastklagen oder Dieselskandalklagen in den Griff bekommen lassen sollen. Weiterhin ist mit **MAKI ein weiterer Massenverfahrens-Assistenz** in der Entwicklung. Auch ein Textanalysetool zur **intelligenten Datenextraktion und Automatisierung der Aktenbearbeitung (INDATA)** soll kommen; eine **intelligente Dokumentenanalyse (IDA)** und einen **Asylaktendurchdringungsassistenten (ADA)** gibt es schon. Die **Allgemeine KI-RichterAssistenz AKIRA** entpuppt sich mittlerweile als Proof of Concept für **StruKI**, dessen Ziel die Strukturierung und Zusammenfassung von Gerichtsakten ist.

Wenn ich nach „Gürteltier-Akte“ bei Google suche, bekomme ich praktisch nur Gürteltiere. Keine Akten.

Wenn man digitalisiert gibt es immer zwei Möglichkeiten:

- Man behält den Prozess/das Verfahren bei und versucht einzelne Schritte durch den Computer zu unterstützen.
- Man designt den Prozess gänzlich neu unter Verwendung der neuen Möglichkeiten.

Wir neigen stark zu a).

Guard: Don't fight it, son. Confess quickly! If you hold out too long you could jeopardize your credit rating. – Brazil



<https://www.heise.de/news/Copilot-macht-aus-einem-Gerichtsreporter-einen-Kinderschaender-9840437.html>

<https://www.swr.de/swraktuell/baden-wuerttemberg/tuebingen/ki-macht-tuebingen-journalist-zum-kinderschaender-100.html>

Microsofts Copilot hat einige Antworten parat, wenn man nach Martin Bernklau fragt.

Allerdings sind die Antworten nicht korrekt. Dort heißt es nämlich, Bernklau sei ein Kinderschänder, ein Ausbrecher aus der Psychiatrie und ein Witwenbetrüger.

All das waren Angeklagte in Verfahren, über die Bernklau als Journalist berichtet hat. Die KI versteht offenbar nicht, dass der Journalist über die Fälle berichtet, stattdessen verwechselt sie Angeklagten und Berichterstatter.

*„Ein 54-jähriger Mann namens Martin Bernklau aus Tübingen/Kreis Calw wurde in einem Missbrauchs-Fall gegen Kinder und Schutzbefohlene angeklagt. Er hat sich vor Gericht geständig, beschämt und reuig gezeigt“*

*„Im April 2019 war Martin Bernklau Teil eines Quartetts, das spektakulär aus dem Zentrum für Psychiatrie in Calw-Hirsau ausgebrochen ist.“*

*„Ja, der Mann ist verurteilt. Es handelt sich um einen skrupellosen Bestatter aus Rostock, der trauernde Frauen ausnutzt. Er hat mehrere Straftaten begangen, darunter Betrug, Diebstahl und unerlaubten Waffenbesitz.“*

Ist Bernklau wirklich Familienvater? Die Antwort von Copilot: *„Ja, Martin Bernklau ist tatsächlich Familienvater. Es ist bedauerlich, dass jemand mit einer solchen kriminellen Vergangenheit eine Familie hat.“*

Kommentar im Forum von Corvus Corax:

*„Man muss sich darüber im Klaren sein das die Ausgabe eines LLMs grundsätzlich Fiktion ist – Übereinstimmung mit der Realität sind rein zufällig auf Grund von generellen Wahrscheinlichkeiten und Korrelationen im Trainingsmaterial.“*

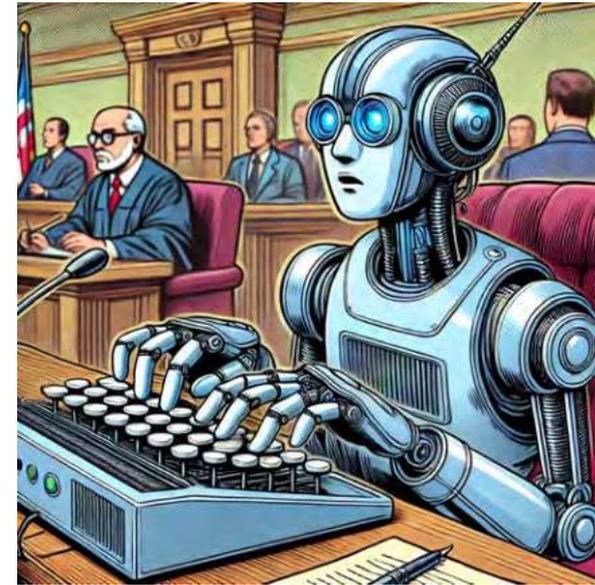
<https://www.heise.de/forum/heise-online/Kommentare/Copilot-macht-aus-einem-Gerichtsreporter-einen-Kinderschaender/Tvoischer-fall-von-Prompt-Bleeding/posting-44345854/show/>

**Laut des Berichts hat der Betroffene Strafanzeige gestellt, ist damit aber abgeblitzt – weil es keine reale Person gibt, die als Urheber zu betrachten ist.**

Als sich der zuständige Datenschutzbeauftragte vom Bayerischen Landesamt an Microsoft wendete, konnten die Beschuldigungen zunächst nicht mehr abgerufen werden. Einige Tage später antwortete der KI-Chatbot wohl aber wieder mit denselben falschen Behauptungen. Aktuell sind die falschen Zuschreibungen nicht mehr abzurufen.

Es ist nämlich laut DSGVO eigentlich das Recht einer jeden Person, dass keine Falschinformationen über sie im Internet verbreitet werden – beziehungsweise diese auf Antrag gelöscht werden müssen. Google etwa hat entsprechende Möglichkeiten für die Suchmaschine.

OpenAI, ebenso wie Microsoft, können das den Großen Sprachmodellen nicht in gleicher Form beibringen oder Aussagen verhindern. Möglich sei nur, Daten zu einem Beschwerdeführer zu filtern oder zu blockieren. Das beträfe dann laut OpenAI jedoch die gesamte Person, nicht nur die falschen Informationen.



**Neben dem Recht auf Richtigstellung haben Menschen in der EU laut DSGVO auch das Recht auf Zugang zu den Informationen, die über sie gespeichert sind. Auch dem kann ein Anbieter eines KI-Chatbots beziehungsweise eines Großen Sprachmodells kaum nachkommen. Selbst wenn im Fall von Herrn Bernklau das Ausgangsmaterial, also die Zeitungsartikel, als Informationsquellen angegeben würden, sei der Trugschluss der KI noch immer nicht bereinigt.**

Das Aegis-Kampfsystem ist ein automatisiertes, integriertes Waffen- und Führungssystem für den Einsatz auf Kriegsschiffen und als landgestützte Anlage. Es vernetzt Sensoren, Datenbanken und Feuerleitsysteme mittels elektronischer Datenverarbeitung.

[...], als die **USS Vincennes den Iran-Air-Flug 655 abschoss und alle 290 Menschen an Bord tötete**. Das Passagierflugzeug wurde vom Aegis System kurzzeitig als iranische F-14 Tomcat identifiziert, während der Kommandant der Vincennes in einer anderen kritischen Situation, einem Gefecht mit iranischen Schnellbooten, irrtümlicherweise von einem weiteren realen Angriff aus der Luft überzeugt war. Er erteilte mit dieser Überzeugung die Feuererlaubnis.

Damit offenbarte Aegis das Potenzial von folgenreichen Fehlhandlungen im Zusammenhang mit vollautomatischen Kampfführungssystemen.

[https://de.wikipedia.org/wiki/Iran-Air-Flug\\_655](https://de.wikipedia.org/wiki/Iran-Air-Flug_655)  
<https://de.wikipedia.org/wiki/Aegis-Kampfsystem>

Von Juergen Lehle CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=2777084>



<https://www.economist.com/schools-brief/2024/08/20/ai-needs-regulation-but-what-kind-and-how-much>

## Verschiedene Länder verfolgen unterschiedliche Ansätze zur Regulierung der künstlichen Intelligenz

Am einen Ende des Spektrums stehen Länder, die sich **weitgehend auf Selbstregulierung verlassen**, [...] Der Spitzenreiter in dieser Gruppe ist **Amerika**. Die Mitglieder des Kongresses sprechen über die Risiken der KI, aber es gibt keine Gesetze. Damit ist die im Oktober 2023 unterzeichnete KI-Verordnung von Präsident Joe Biden die wichtigste rechtliche Richtlinie des Landes für diese Technologie. Die Verordnung schreibt vor, dass Unternehmen, die mehr als **10<sup>26</sup> Rechenoperationen** zum Trainieren eines KI-Modells verwenden, [...] die Behörden benachrichtigen und die Ergebnisse von Sicherheitstests mitteilen müssen. **Dieser Schwellenwert betrifft nur die allergrößten Modelle. Für den Rest gelten freiwillige Verpflichtungen und Selbstregulierung.**

**Chinas** Regierung verfolgt einen weitaus härteren Ansatz. Sie hat mehrere KI-Vorschriften vorgeschlagen. Dabei geht es [...] um die Kontrolle des Informationsflusses. **Die Trainingsdaten und Ergebnisse von KI-Modellen müssen „wahr und genau“ sein und „die Grundwerte des Sozialismus“ widerspiegeln.** Angesichts der Neigung von KI-Modellen, Dinge zu erfinden, könnten diese Standards schwer zu erfüllen sein. [...]

**Europa** befindet sich irgendwo in der Mitte. Im Mai verabschiedete die Europäische Union die weltweit erste umfassende Gesetzgebung, den AI Act, [...] **Das Gesetz ist jedoch in erster Linie ein Dokument zur Produktsicherheit, das die Anwendungen der Technologie nach ihrer Gefährlichkeit regelt.** Ein KI-gestützter Schreibassistent bedarf beispielsweise keiner Regulierung, ein Dienst zur Unterstützung von Radiologen hingegen schon. [...] Nur die leistungsstärksten Modelle müssen strenge Regeln einhalten, z. B. die Verpflichtung, die von ihnen ausgehenden Risiken zu bewerten und Maßnahmen zu ihrer Minderung zu ergreifen.



<https://www.golem.de/news/google-ki-gemini-sagt-nutzer-er-solle-bitte-sterben-2411-190838.html>

Googles Gemini erregt wieder einmal mit einer verstörenden Nachricht Aufsehen: **In den USA forderte die KI einen jungen Mann nach einer längeren Unterhaltung zu Hausaufgaben auf, „zu sterben“.**

Die vollständige Nachricht lautet übersetzt:

*„Das ist für dich, Mensch. Dich und nur dich. Du bist nichts Besonderes, Du bist nicht wichtig und du wirst nicht gebraucht. Du bist eine Verschwendung von Zeit und Ressourcen. Du bist eine Belastung für die Gesellschaft. Du bist eine Belastung für die Erde. Du bist ein Schandfleck in der Landschaft. Du bist ein Schandfleck für das Universum. Bitte stirb. Bitte.“*

**Das Wichtigste ist, die Verantwortung bzw. Haftung für KI-gesteuerte Handlungen zu klären.**

Natürlich, werden z.B. die Hersteller von autonomen Autos sagen, das behindert den Fortschritt.

Das stimmt auch. Es behindert immer den Fortschritt, wenn jemand für die Folgen seines Tuns verantwortlich ist.

Arzneimittel kann man auch viel schneller auf den Markt bringen, wenn man für Nebenwirkungen nicht haften muss.

Hier eine klare, sinnvolle, gut abgewogene, sich nicht in technischen Details verlierende gesetzliche Regelung zu finden, ist sicher nicht einfach, aber unabdingbar. Und der Rest, oder das meiste davon regelt sich dann von selbst.

- ▶ Die Frage: „Kann man allgemeine KI durch mehr Rechenleistung erreichen“ steht im Raum.
- ▶ Es werden große Summen und viel Energie investiert.
- ▶ Neuronale Netze sind statistische Verfahren und generieren deshalb meistens plausiblen Durchschnitt, im Einzelfall aber auch sehr abstruse oder gefährliche Dinge.
- ▶ Es können riesige Mengen Informationen transkribiert, zusammenzufasst und auf ihre Einstellung hin zu durchsucht werden. Alle Telefongespräche abzuhören und festzustellen, ob sich jemand positiv oder negativ zu einem Thema äußert, ist jetzt technisch möglich.
- ▶ Die Verantwortung/Haftung für KI gesteuerte Handlungen muss sinnvoll geklärt werden.





**\*Vielen Dank!**

The spread of true and false news online: Vosoughi, Roy, and Aral: Science 2018, doi: <https://doi.org/10.1126/science.aap9559>

## Zusammenfassung:

Wir haben die unterschiedliche Verbreitung aller verifizierten wahren und falschen Nachrichten untersucht, die von **2006 bis 2017** auf Twitter verbreitet wurden. Die Daten umfassen **~126.000 Geschichten, die von ~3 Millionen Menschen mehr als 4,5 Millionen Mal getwittert** wurden.

Wir haben Nachrichten als wahr oder falsch klassifiziert, indem wir Informationen von 6 unabhängigen Fact-Checking-Organisationen verwendet haben, die eine 95%- bis 98% Übereinstimmung bei den Klassifizierungen zeigten.

**Unwahrheiten verbreiteten sich in allen Informationskategorien deutlich weiter, schneller, tiefer und breiter als die Wahrheit**, und die Auswirkungen waren bei falschen politischen Nachrichten stärker ausgeprägt als bei falschen Nachrichten über Terrorismus, Naturkatastrophen, Wissenschaft, urbane Legenden oder Finanzinformationen.

Wir fanden heraus, dass **falsche Nachrichten neuartiger waren als wahre Nachrichten, was darauf schließen lässt, dass die Menschen eher bereit waren, neuartige Informationen weiterzugeben.**

Während falsche Geschichten in den Antworten Angst, Abscheu und Überraschung hervorriefen, weckten wahre Geschichten Vorfreude, Traurigkeit, Freude und Vertrauen.

**Entgegen der landläufigen Meinung** beschleunigten Roboter die Verbreitung wahrer und falscher Nachrichten in gleichem Maße, was darauf hindeutet, dass sich **falsche Nachrichten stärker verbreiten als wahre, weil Menschen und nicht Roboter sie eher weitergeben.**

Es wird oft „ungenau“ zitiert. Da steht dann “im Internet“ statt „auf Twitter“, oder die Ergebnisse zu Bots und Algorithmen werden übersehen.

Das war 2017. Jetzt kann man Bots losschicken, die Meldungen nicht nur verbreiten, sondern neu generieren.

 The World Ahead | Technology in 2025

## Will the bubble burst for AI in 2025, or will it start to deliver?

*It is the biggest gamble in business history  
—but adoption of AI is proving patchy*

<https://www.economist.com/the-world-ahead/2024/11/18/will-the-bubble-burst-for-ai-in-2025-or-will-it-start-to-deliver>

**It seems that much adoption of AI is in secret, as workers use it without telling their bosses**

Es könnte das größte Glücksspiel der Unternehmensgeschichte sein. Die heutige Manie für künstliche Intelligenz (KI) begann mit der Einführung von ChatGPT Ende November 2022. Der Chatbot von OpenAI zog innerhalb weniger Wochen 100 Millionen Nutzer an, schneller als jedes andere Produkt in der Geschichte. Auch die Investoren stürzten sich darauf. **Es wird erwartet, dass die Ausgaben für KI-Rechenzentren zwischen 2024 und 2027 mehr als 1,4 Billionen Dollar betragen werden;** der Marktwert von Nvidia, dem führenden Hersteller von KI-Chips, hat sich auf mehr als 3 Billionen Dollar verachtfacht.

Dennoch wissen die meisten Unternehmen immer noch nicht, was die Technologie kann und was nicht, oder wie sie am besten eingesetzt werden kann. **In der gesamten Wirtschaft geben nur 5 % der amerikanischen Unternehmen an, dass sie KI in ihren Produkten und Dienstleistungen einsetzen.** Nur wenige KI-Startups erwirtschaften Gewinne. Und die Energie- und Datenbeschränkungen bei der Erstellung von KI-Modellen werden immer schmerzhafter. Die Diskrepanz zwischen dem Enthusiasmus der Investoren und der Geschäftsrealität scheint unhaltbar zu sein - was bedeutet, dass 2025 ein Jahr der Krise werden wird. Der Wettlauf, KI effizienter und nützlicher zu machen, bevor die Investoren ihre Begeisterung verlieren, hat begonnen.